

DIPLOMARBEIT

Anwendung mathematischer Modelle zur Allokationsplanung
für eine verteilte Datenbank der Bundesanstalt für Arbeit

Verfasserin: **Dagmar Lück-Schneider**
 Richard-Wagner- Str. 41
 19059 Schwerin

Betreuer: **Prof. Dr. Schlageter**

Eidesstattliche Erklärung

Die vorliegende Diplomarbeit wurde von mir selbständig verfaßt. Alle von mir benutzten Quellen sind im Text belegt und im beigefügten Literaturverzeichnis aufgeführt. Fremde Hilfe habe ich darüber hinaus nicht in Anspruch genommen.

Datum: 22.01.1999 **Unterschrift:** Online-Version, Unt. im Original vorhanden

Vorwort

Heute findet man in vielen Bereichen Datenbanken, deren Daten auf mehrere Rechner verteilt sind, die aber dennoch gemeinsam verwaltet werden oder zwischen denen zumindest Zusammenhänge bestehen. Soll eine solche Situation kein Zufallsprodukt sein, so sind Überlegungen anzustellen, welche Daten „am günstigsten“ auf welchen Rechnern (mit der zugehörigen Peripherie) unterzubringen sind. Dabei können Kostenerwägungen sowie Anforderungen an Zugriffszeiten oder an die Ausfallsicherheit bestimmende Faktoren für die gewählte Allokation sein.

Die vorliegende Arbeit befaßt sich zum einen mit mathematischen Modellen, die eine solche Allokationsplanung zum Gegenstand haben. Zum anderen soll exemplarisch die Anwendung eines solchen Modells auf eine reale Allokationsplanung, wie sie innerhalb der Bundesanstalt für Arbeit¹ (BA) stattgefunden hat, aufgezeigt werden.

Eine der Hauptaufgaben² der BA ist die Unterstützung der Stellen- und Bewerbervermittlung innerhalb der Bundesrepublik Deutschland.

In den 70er Jahren wurde diese Tätigkeit durch die Bediensteten noch über Handkarteikartensysteme erledigt. Mit dem Jahr 1980 wurde die erste durch elektronische Datenverarbeitung (EDV) unterstützte Lösung eingeführt (vgl. im Literaturverzeichnis unter [HR94], S. 183). Diese Software wurde bis heute - natürlich in ständig angepaßter Form - beibehalten. Der Wunsch nach einer Neulösung wurde immer drängender. Gefordert wurden nicht nur neue Benutzeroberflächen, sondern vor allem auch eine gemeinsame Datenbank für die verschiedenen, zum Teil getrennt entwickelten EDV-Lösungen der unterschiedlichen Aufgabenbereiche der BA. Darüber hinaus steckt in der gemeinsamen Entwicklung ähnlicher funktionaler Aufgaben ein hohes Potential für weitere Synergieeffekte. So ähneln sich beispielsweise die Stellen- und Ausbildungsstellenvermittlung - bisher durch getrennt realisierte Programme unterstützt - in hohem Maße. Und schließlich erfordert die gegenwärtige Neuorganisation der Aufgabenerledigung in der Bundesanstalt³ eine diesem Schritt angepaßte EDV-Unterstützung.

Im Laufe des Jahres 1995 wurde daher in der BA ein Projekt namens PlaIT2000⁴ ins Leben gerufen, das unter Einbeziehung externer Firmen die Planung der Gestaltung und Entwicklung einer neuen Hard- und Softwaregeneration, kurz IT2000 genannt, zur Aufgabe hatte. Im Frühjahr 1997 hatte ich⁵ die Möglichkeit, in einer diesem Projekt angehörenden Arbeitsgruppe für kurze Zeit mitzuarbeiten. Dabei stieß ich auf Unterlagen zur Planung der Verteilung der umfangreichen Daten im Bereich der Stellen- und Bewerbervermittlung mittels mathematischer Modelle. Hierzu war von Seiten der Firma Siemens Nixdorf Informationssysteme (SNI) im Rahmen des Projektes ein Vortrag gehalten worden (vgl. [SNI96] / Anlage 1).

Diese Planungsüberlegungen zur Datenallokation weckten mein Interesse, waren mir doch gerade erst in meinem Studium mathematische Modelle im Kontext der Datenverteilung in verteilten Daten-

¹ Die Bundesanstalt für Arbeit (BA) mit Sitz in Nürnberg, der Rechtsform nach eine Körperschaft des öffentlichen Rechts, wurde am 1. Mai 1952 errichtet. In ihren Funktionen knüpft sie an die 1927 gegründete Reichsanstalt für Arbeitsvermittlung und Arbeitslosenversicherung an. Sie umfaßt neben der Hauptstelle weitere Dienststellen (Landesarbeitsämter, Arbeitsämter und besondere Dienststellen) im gesamten Bundesgebiet.

² Der BA sind per Arbeitsförderungsgesetz u. a. die Arbeitsvermittlung, die Berufsberatung und Vermittlung beruflicher Ausbildungsstellen, die Förderung der beruflichen Bildung sowie die Arbeitslosenversicherung übertragen.

³ Unter dem Stichwort Arbeitsamt 2000 (AA2000) werden gegenwärtig in den Arbeitsämtern umfangreiche Umstrukturierungen der Arbeitsorganisation vorgenommen. Die Neuorganisation von Arbeitsprozessen und Verantwortlichkeiten geht aber über die Arbeitsämter hinaus; sie hat auch die Landesarbeitsämter und die Hauptstelle erfaßt.

⁴ Kunstwort, entstanden aus „Planung der Informationstechnik- (IT-) Anwendungen 2000“

⁵ Die Autorin ist in einer der besonderen Dienststellen hauptberufliche EDV-Dozentin.

banken begegnet. So entstand der Gedanke, dieses Thema und seine konkrete Anwendung in der Bundesanstalt für Arbeit zum Gegenstand dieser Diplomarbeit zu machen.

Dafür, daß dies möglich war, möchte ich allen hieran Beteiligten meinen Dank aussprechen, zuerst Herrn Dr. Mühlbach, zuständig für Wirtschaftlichkeitsbetrachtungen und Berichtswesen im Projekt PlaIT2000, der mir von Seiten der BA zusagte, über das Thema schreiben zu dürfen sowie Herrn Kronewald, Leiter des Business Support Team for System Design bei Siemens Business Services, ehemals SNI in Nürnberg, der mir gestattete, die Studie zu verwenden. Darüber hinaus halfen sie mir wie auch insbesondere die Leitung und Mitarbeiter des Projektteams IT2000 der BA mit Informationen oder ergänzenden Hinweisen weiter. Bedanken möchte ich mich vor allem aber bei Herrn Professor Dr. Schlageter an der Fernuniversität Hagen für die Betreuung dieser Arbeit und das reichhaltige Angebot an interessanten Vorlesungen und Seminaren zum Thema Datenbanken in seinem Lehrgebiet, das weite Teile meines Hauptstudiums prägte und mit ausschlaggebend für mein Interesse an dem Thema dieser Arbeit war. Ganz besonders gilt der Dank auch seinem wissenschaftlichen Mitarbeiter, Diplominformatiker Herrn Laskowski, der den Themenwunsch unterstützte und auf diese Weise die Themenvergabe von Seiten der Fernuniversität möglich machte. Bedanken möchte ich mich bei ihm auch für seine engagierte und fachlich exzellente Betreuung der Arbeit. Ebenso erfuhr ich hervorragende Unterstützung bei schwierigen Literaturbeschaffungen durch Frau Ines Apel, Leiterin der Bibliothek der Dépendance der Fachhochschule des Bundes, Fachbereich Arbeitsverwaltung und ihre Mitarbeiterinnen in Schwerin sowie durch die Beschäftigten der Bibliothek der Fernuniversität Hagen. Und schließlich richtet sich ein ganz besonders herzlicher Dank an meinen Ehemann, der mein zahlreiche Jahre dauerndes nebenberufliches Fernstudium stets unterstützte, immer wieder Rücksicht nahm und viel Verständnis für die umfangreichen Stunden, die dem Studium vorbehalten waren, aufbrachte.

Hinweise zum Schriftbild und Aufbau der Arbeit

Im Rahmen der geltenden Übergangsfristen für die Umstellung der deutschen Rechtschreibregeln wurde die Arbeit noch nach den alten Rechtschreibregeln verfaßt.

Um die in den Literaturangaben genannte Literatur bequem und schnell auffinden zu können, wurde das Literaturverzeichnis in vier Bereiche untergliedert: Buchangaben, Zeitschriftenartikel, Beiträge aus Tagungsbänden sowie aus Schriften oder Erlassen der Bundesanstalt für Arbeit. Die Verweise im Text, auch bereits hier in diesem Vorwort, beinhalten in eckigen Klammern eine Abkürzung, die zu der Literaturangabe führt. In der Regel ist diese aus den Autoren- oder Titelanfangsbuchstaben gebildet. Sind sie in **fettem Schriftbild**, handelt es sich um eine Buchangabe, bei *kursivem Schriftzug* um einen Artikel, sind sie **kursiv und fett** dargestellt, um einen Tagungsbandbeitrag und wenn sie ohne Hervorhebung erscheinen, so wird auf eine BA-interne Schrift verwiesen. In letztem Falle kann auch ein Verweis auf ein stattgefundenes Gespräch vorkommen.

Als Anlage schließlich finden sich Literaturgrundlagen, die allgemein nicht zugänglich, für die Arbeit aber von grundlegender Bedeutung sind, ebenso herausnehmbare Übersichten zu Variablenbezeichnern aus den Kapiteln 3 und 4 sowie eine Diskette zum Ausprobieren der diversen Kalkulationen zu Kapitel 4.

Zum Verzeichnisteil am Ende der Arbeit gehört auch ein Abkürzungsverzeichnis für häufig verwendete Abkürzungen, das ein punktuelles Lesen der Arbeit erleichtern soll. BA-spezifische Abkürzungen sind gesondert aufgeführt. Meist gilt aber, daß einer Abkürzung im Text, die nicht Bestandteil eines Zitates ist, eine voll ausgeschriebene Nennung oder eine Erklärung vorausgeht. Auf die Seitenzahl dieser Angabe wird im Abkürzungsverzeichnis ebenfalls hingewiesen. Als Pluralform ist für die Abkürzungen fast immer wie im Englischen das angehängte „s“ gewählt worden.

Inhaltsverzeichnis

- VORWORT III**

- 1 EINFÜHRUNG 1**

 - 1.1 Datenbanksysteme, Organisationsstrukturen und wirtschaftlicher Wandel2**
 - 1.1.1 Die Entstehung heutiger Datenbanksysteme.....2
 - 1.1.2 Wirtschaftlicher Wandel und Verteilung.....4

 - 1.2 Grundlegende Begriffe6**
 - 1.2.1 Verteilte Datenbanken6
 - 1.2.2 Weitere Realisierungsarten9
 - 1.2.3 Netzeigenschaften13

- 2 VERTEILUNG..... 17**

 - 2.1 Anforderungswünsche an verteilte Datenbanken17**

 - 2.2 Relationale Datenbanksysteme19**

 - 2.3 Datenbankverteilung21**
 - 2.3.1 Fragmentierung23
 - 2.3.2 Allokation25
 - 2.3.3 Anfragebearbeitung, Synchronisationsverfahren und Kopienhandhabung.....27

- 3 MATHEMATISCHE MODELLE 33**

 - 3.1 Modelle von Ceri und Pelagatti35**
 - 3.1.1 Allokation bei horizontaler Fragmentierung.....36
 - 3.1.2 Vertikale Fragmentierung38
 - 3.1.3 Anwendungsbereich.....40

 - 3.2 Das Modell von Dadam40**
 - 3.2.1 Berechnungsmodell41
 - 3.2.2 Anwendungsbereich.....43

 - 3.3 Die Allokationsbetrachtungen von Özsu und Valduriez.....45**
 - 3.3.1 Kritik am „All-Beneficial-Sites“-Ansatz und Dadams Lösungsweg.....45
 - 3.3.2 Der Alternativvorschlag.....47
 - 3.3.3 Anwendungsbereich.....50

 - 3.4 Der graphbasierte Modellansatz von Apers51**
 - 3.4.1 Erstes Vorgehensmodell52
 - 3.4.2 Zweites Vorgehensmodell56
 - 3.4.3 Anwendungsbereich.....57

 - 3.5 Ein Vergleich der Modelle57**

4 DER AUFBAU EINER NEUEN DATENBANK ZUR ARBEITSVERMITTLUNG IN DER BUNDESANSTALT FÜR ARBEIT	60
4.1 Organisationsstrukturen und Umstrukturierungen innerhalb der Bundesanstalt für Arbeit	60
4.2 EDV-Organisation und EDV-Ausstattungsmerkmale in der Bundesanstalt für Arbeit	64
4.3 Software zur Unterstützung der Arbeitsvermittlung	67
4.4 Die Studie von Siemens Nixdorf zur Allokationsplanung.....	70
4.4.1 Die Voraussetzungen	72
4.4.2 Das mathematische Modell der Studie	78
4.4.2.1 Die stellenorientierte Bewerbersuche	78
4.4.2.2 Die bewerberorientierte Stellensuche	92
4.4.2.3 Tageszeitabhängigkeiten und Netzbelastung	108
4.4.3 Lösungsvorschlag der Studie	115
4.5 Grenzen des SNI-Modells.....	117
5 ZUSAMMENFASSUNG UND AUSBLICK.....	122
A VERZEICHNISSE.....	126
Abkürzungsverzeichnis	126
Abbildungsverzeichnis	128
Tabellenverzeichnis	129
Literaturverzeichnis	130
Bücher	130
Tagungsbände	131
Artikel	131
Erlasse, Schriften der Bundesanstalt für Arbeit sowie Gesprächspartner	131
B ANLAGEN	133
1 SNI-Studie	134
2 KURZFASSUNG ARBEITSAMT 2000	135
4 Diskette mit Kalkulationsbeispielen zur Studie	142

1 Einführung

Information ist heute zu einem Wirtschaftsgut geworden. Eine schnelle Beschaffung von Informationen kann für die Konkurrenzfähigkeit eines Unternehmens von entscheidender Bedeutung sein. Für viele Firmen ist es daher heute selbstverständlich, leistungsfähige computerunterstützte Systeme für diese Informationsbeschaffung einzusetzen. Kern solcher Informationssysteme sind Datenbanken. Während zunächst räumlich zentralisierte Datenbanken vorherrschten, werden heute auch auf verschiedene und ggf. auch weit voneinander entfernte Rechner verteilte - semantisch zusammengehörende - Datenbestände für das Informationsmanagement genutzt. Dies kann Einsparungen mit sich bringen oder aber Leistungsvorteile bedeuten.

Mathematische Modelle zur Allokationsplanung in verteilten Datenbanken sollen helfen, eine möglichst optimale Verteilung für eine solche Situation zu finden. Dabei berücksichtigen sie bestimmte Leistungs- und Kostenaspekte, als einige wenige Beispiele seien die Anzahl vorhandener Rechner, deren Speicherkapazität und Übertragungskosten zwischen den Rechnern genannt. Unter Berücksichtigung bekannter oder geschätzter Anforderungsparameter wird versucht, die kostengünstigste anforderungsgerechte Lösung zu finden.

Ziel dieser Arbeit ist es, eine konkrete Allokationsplanung per mathematischem Modell vorzustellen und dieses in den Kontext einer Reihe verschiedener mathematischer Modelle zu stellen.

Dazu werden in diesem einleitenden Kapitel zunächst allgemeine Grundlagen aufgezeigt. Neben der Klärung von Begriffen wird auf mögliche Anforderungen verteilter Datenbanken eingegangen. Ebenso dient das Vorstellen verschiedener Konzepte und Entstehungsvarianten von Rechnerverbunden zur Datenhaltung einer besseren Verständnisgrundlage. Da ferner der konkrete Anwendungsfall eng mit einer allgemeinen organisatorischen Umstrukturierung der betrachteten Behörde einhergeht, werden auch einige Sichtweisen von Zusammenhängen allgemeiner Organisations- und EDV-Strukturen vorgestellt. Eine umfassende Behandlung dieses Aspektes kann jedoch im Rahmen des auf mathematischen Modellen basierenden Themenschwerpunktes nicht geleistet werden.

Im zweiten Kapitel werden solche für die Verteilung wesentlichen Aspekte herausgearbeitet, die Grundlage der diversen, im Kapitel 3 vorgestellten, mathematischen Modelle sind oder deren bessere Beurteilung unterstützen.

Die Behandlung des realen Anwendungsbeispiels folgt schließlich in Kapitel 4. Ergänzend zu den hier betrachteten Planungsüberlegungen für eine neue Soft- und Hardwaregeneration wird zunächst auf die ursprüngliche Lösung und den gegenwärtigen Planungskontext eingegangen.

Kapitel 5 bildet mit einer Zusammenfassung und einem Ausblick den Abschluß dieser Arbeit.

1.1 Datenbanksysteme, Organisationsstrukturen und wirtschaftlicher Wandel

1.1.1 Die Entstehung heutiger Datenbanksysteme

Den gegenwärtigen Lösungsansätzen zur Verwaltung umfassender Datenbestände sind einige Entwicklungsstufen vorausgegangen (vgl. [HS95], S. 4 - 5).

So war Anfang der 60er Jahre noch eine enge Kopplung der Daten an ihre Programme üblich. Die Programme selbst übernahmen die Aufgabe, die von ihnen benötigten Massendaten zu typisieren und in Dateien abzulegen. Sofern es Datenüberdeckungen verschiedener Anwendungsprogramme gab, führte dies automatisch zu Redundanz und damit zu Konsistenzproblemen, da jedes Programm mit seinen eigenen Dateien arbeitete. Da die Programme selbst die Peripheriegeräte ansprechen mußten, bestand hierbei auch eine hohe Geräteabhängigkeit. Dieser Nachteil wurde beseitigt, als sich **Dateiverwaltungssysteme** durchsetzten, die den Anwendungen das Ansprechen der Datenträger abnahmen. Die Daten-Programmabhängigkeit und die Redundanzprobleme aber blieben bestehen (vgl. ebenda).

Ende der 60er Jahre gab es einen größeren Wandel. Die Dateiverwaltungssysteme wurden erstmals von **Datenbanksystemen** abgelöst. Waren zuvor die Daten den Anwendungen entsprechend strukturiert, so wurden sie nun unabhängig von den auf ihnen gewünschten Anwendungen zunächst abstrakt modelliert. Die Betrachtungsgrundlage waren nun alle relevanten im Unternehmen existierenden Daten. Diese neue Betrachtungsweise bei der Datenmodellierung führte allein schon zu längerfristig gültigen, integrierten, d. h. alle Daten des Unternehmens umfassenden, Datenbeständen mit weniger nachträglichen Änderungen. Außerdem vermied man quasi automatisch unerwünschte Redundanz. Die Aufgabe, interne Datenabhängigkeiten korrekt zu halten, wurde nun ebenso wie die Koordination und Realisierung von Mehrbenutzerzugriffen und die Lösung von Problemsituationen (z. B. Komplettausfall eines Speichers) einer eigenen Softwarekomponente, dem **Datenbankmanagementsystem** (DBMS) übertragen. Ein Zugriff auf die Daten (die sogenannte Datenbank) ist nun für alle Anwendungen nur noch über diese zentrale Komponente möglich. Das Datenbanksystem (DBS) wird aus dem DBMS und der Datenbank gemeinsam gebildet (vgl. ebenda, S. 4 - 8).

„In dieser Zeit war die Hardware noch eine teure und knappe Ressource, außerdem gab es nur wenig ausgebildetes Personal, das diese Systeme bedienen bzw. Anwendungen dafür entwickeln konnte. Die Unternehmen waren daher gezwungen, diese teuren Hardware-Ressourcen möglichst effizient zu nutzen. Dies führte schließlich zur Einführung von Zentralrechnern bzw. zentralen Rechenzentren und, als Folge davon, zu gravierenden Änderungen in den betrieblichen Abläufen. Die vormals häufig eher dezentral organisierten Unternehmensstrukturen und Arbeitsabläufe wurden zentralisiert und „EDV-gerecht“ umgestaltet. Informationsflüsse, die vorher direkt zwischen Abteilungen stattfanden, gingen von nun an über das zentrale Rechenzentrum ...“ ([Da96], S. 2).

Die Verbreitung von DBMS auf diesen Zentralrechnern wurde Anfang der 80er Jahre durch die Abwendung von im Stapel abgearbeiteten Programmen hin zur Dialogorientierung beschleunigt. Im **Dialogverfahren** müssen eingegebene Anfragen in akzeptabler Zeit eine Reaktion auf dem Bildschirm liefern. Veränderungen an den Daten müssen außerdem sofort wirksam werden. Diese wesentlich höheren Ansprüche an das Zeitverhalten der Anwendungen wurden zum einen erst durch deutliche Leistungssteigerungen der Hardware möglich. Zum anderen konnten sie durch ein DBMS viel effizienter bewältigt werden, als durch die älteren Dateiverwaltungssysteme (vgl. ebenda, S. 5).

Natürlich hat es bis heute weitere Entwicklungen im Bereich der Verwaltung umfangreicher Datenbestände gegeben. Unter den Ansätzen, die sich mit dem Einsatz mehrerer vernetzter Rechner beschäftigen, gibt es zum einen solche, die die Daten des Datenbestandes auf den Rechnern des Netzes verteilen, dennoch aber eine gemeinsame Verwaltung vorsehen. Ein anderer Ansatz verteilt statt dessen die Funktionalität, die auf den Daten erfolgt, während weiterhin ein Rechner mit zentralem Datenbestand beibehalten wird.

Erste Verteilungen von Daten auf verschiedene Rechner ergaben sich bereits in der Zeit des Zentralrechner-Einsatzes. Wenn ein solcher Rechner für die Bereitstellung aller Daten nicht leistungsfähig genug war oder ein leistungsfähigerer den finanziellen Rahmen gesprengt hätte, so übergab man diese Aufgabe mehreren Rechnern (vgl. ebenda, S. 4 - 5).

Später waren eher der Preisverfall der Hardware und die hiermit verbundenen neuen Möglichkeiten Anlaß dafür, über neue Lösungen der Datenverteilung nachzudenken. So konnte man nun wieder die Unternehmens-Informationsflüsse den Geschäftsprozessen entsprechend auf Rechnern organisieren. Eine Datenhaltung an der Stelle des häufigsten Gebrauchs, auf mehreren vernetzten Computern, kam beispielsweise wieder in Betracht. Dennoch mußte man der Anforderung, daß der gesamte Datenbestand wie eine einzige Datenbank gepflegt werden konnte, gerecht werden. Dieser Ansatz führt zu den eng integrierten, homogenen verteilten Datenbanken. Der Wunsch, verteilte Daten zu verwalten, konnte aber ebenso durch das Zusammenwachsen getrennt entstandener, möglicherweise auch völlig verschiedener Datenbanken aufkommen. Typische Ansätze aus dieser Entwicklung werden später unter den Begriffen „heterogene Datenbanken“ und „föderierte verteilte Datenbanken“ beschrieben.

Der andere, Funktionalität verteilende Ansatz, führte zu Client-Server-Datenbanken. Er wurde stark forciert durch die sich in den 80er Jahren rapide entwickelnden Personal Computer (PCs). Nicht nur deren Leistungsfähigkeit und Preis sondern auch die erstmals von Apple⁶-Computer 1984 geprägte bedienerfreundliche grafische Oberfläche dieser Geräte spielte zunehmend eine Rolle. Derartige Oberflächen konnte ein zentraler Hintergrundrechner für etliche Terminals unmöglich bereitstellen. Um dennoch auch in Großrechnerumgebungen derartige Oberflächen zu ermöglichen, ersetzte man die Terminals durch PCs und überließ diesen die Aufgabe der Oberflächengestaltung. Der Client-Server Ansatz geht allerdings hierüber noch hinaus. Bei einer Client-Server-Datenbank etwa liegt die Datenbank in einem lokalen Netz zentral auf einem Rechner, dem Datenbankserver⁷. Die mit diesem Rechner vernetzten Computer, in der Tat häufig PCs, übernehmen über die Gestaltung der Dialogoberfläche hinaus auch Funktionen der Anwendung. Das auf dem Server arbeitende DBMS reicht über eine Schnittstelle die Daten an die vernetzten Computer, die Clients, und erhält sie hierüber auch zurück. Die Clients können nun die erhaltenen Daten eigenständig manipulieren.

Weitere Entwicklungen betreffen das Verwalten von Nicht-Standard-Daten sowie den Aspekt der flexiblen Erweiterbarkeit in Bezug auf das Datenmodell. Andere Forschungen beschäftigen sich mit Wissensbanken und deduktiven Systemen. Auch den objektorientierten Datenbanksystemen (OODBS) wird viel Aufmerksamkeit gewidmet.

Im Rahmen dieser Arbeit spielen diejenigen Ansätze eine Rolle, die sich mit **physisch verteilten Daten** beschäftigen. Deshalb wird in späteren Abschnitten und Kapiteln auch nur hierauf eingegan-

⁶ Firmenname

⁷ Die Möglichkeit, daß der Datenbankserver auch realisiert sein kann als ein “(in der Regel dann homogenes, eng integriertes) verteiltes DBMS ... , das gegenüber den Anwendungen jedoch (in der Regel) wie ein DBMS auftritt” ([DA96], S. 22), spielt für unsere Betrachtungen keine Rolle.

gen. Ferner sind die Auswahl und der Grad der Ausführlichkeit der Betrachtungen von der Relevanz für die später vorgestellten mathematischen Modelle sowie für das Anwendungsbeispiel abhängig gemacht worden.

1.1.2 Wirtschaftlicher Wandel und Verteilung

Parallel zu der eben aufgezeigten Entwicklung kann man den Wandel der allgemeinen Organisationsstrukturen in den Unternehmen betrachten. Hierbei lassen sich erstaunliche Beziehungen zu den Veränderungen in der EDV-Organisation erkennen. Wurden die 60er Jahre im EDV-Sektor noch durch die Leistungsgrenzen und Kosten der Rechner bestimmt und forderten ihren Tribut in an diesen Vorgaben angepaßten Organisationsstrukturen der Unternehmen, so sind die 90er Jahre dafür von einer Suche nach wirtschaftlich optimalen allgemeinen Organisationsformen bestimmt, die zudem bestmöglich durch EDV unterstützt werden sollen.

Den in den Unternehmen weichenden hierarchischen Organisationsstrukturen entsprechend, wandeln sich die alten hierarchisch aufgebauten Host-Terminal-Systeme zu Rechnernetzen, die - wie bereits oben angeklungen - Dezentralisierungsstreben oder auch das Zusammenwachsen verschiedener Firmen oder vormals eigenständiger Insellösungen widerspiegeln können.

Als Ursache für die heute stark zu beobachtenden Veränderungen innerhalb der Organisation und der Informationstechnik in den Unternehmen und Behörden kann man die Reaktion auf den durch die Computertechnologie ausgelösten gegenwärtigen Strukturwandel ansehen (vgl. [KM94], S. 5). Die zunehmenden Kommunikationsmöglichkeiten der Unternehmen und Konsumenten führen zu einer immer stärkeren Globalisierung der Märkte und zu einem immer mehr wachsenden Konkurrenzdruck. Karer und Müller formulieren dies so:

„Aufgrund des gerade stattfindenden Strukturwandels sind heute viele Unternehmen gezwungen, sich neu zu orientieren und auszurichten. Schlagworte wie Lean Production und Lean Management, Down- und Rightsizing, Outsourcing und Design to Cost sind Schlaglichter für neue Organisationsformen in den Unternehmen. Wie auch immer: Mehr Flexibilität in den Abläufen, in der Reaktion auf die Anforderungen des Marktes, kürzere Reaktionszeiten in den Entscheidungsgremien - all das hat Rückwirkungen und wird beeinflußt von der Informatik im Unternehmen.“ (ebenda, S. 6)

Aber nicht nur in den Hardwarestrukturen werden Parallelen zu den Organisationsstrukturen der Unternehmen gesehen. Im Softwarebereich hat sich ein Wandel von der imperativen Programmierung hin zur Objektorientierung sowie von zentralen hin zu verteilten Anwendungen ergeben. Bei beiden Entwicklungen kann man Ähnlichkeiten mit selbständig agierenden Geschäftseinheiten, die untereinander kommunizieren, entdecken.

Wenn aber die treibende Kraft für Umstrukturierungen die Wettbewerbsfähigkeit der Unternehmen ist, so müssen die neuen Organisationsformen eindeutige Vorteile mit sich bringen. Neben eventuell möglichen finanziellen Einspareffekten wird insbesondere die Möglichkeit, schnellere Anpassungsleistungen zu erzielen, genannt. Diese Eigenschaft wird übrigens auch den Hard- und Softwareentsprechungen zugeschrieben. Außerdem sollen die neuen Strukturen durch die Rücknahme einer zu strikten Zerlegung von Abläufen und Zuständigkeiten die Verantwortung der Mitarbeiter und ihre Identifikation mit der geleisteten Arbeit und damit letztendlich auch ihre Motivation stärken.

Genau diese Aspekte werden auch als grundlegende Organisationsziele der Umstrukturierungen in der Bundesanstalt für Arbeit genannt. Hier haben sowohl die organisatorischen Veränderungen als auch die Vorhaben im Bereich der Informationstechnik (IT) folgende grundlegende Organisationsziele zu beachten (vgl. [See97], S.102, 105):

- Dienstleistungen sind am Kunden zu orientieren, (Kundenorientierung),
- Dienstleistungen sind wirksam zu erbringen (Effektivität),
- Dienstleistungen sind wirtschaftlich zu erbringen (Effizienz),
- Zufriedenheit und Arbeitserfolg der Mitarbeiter sind zu fördern (Mitarbeiterorientierung).

Man kann also zusammenfassend sagen, daß den gegenwärtigen Entwicklungen hin zu neuen Strukturen in der Wirtschaft und in Unternehmen strukturelle Entsprechungen in der Informationstechnologie, sowohl im Hard- wie auch im Softwarebereich gegenüberstehen. Dabei erhofft man sich von diesen neuen Lösungen Vorteile in unterschiedlichsten Bereichen (Qualität, Effektivität, Effizienz).

Inwieweit diese Hoffnungen berechtigt sind, bleibt abzuwarten, da nicht vergessen werden darf, daß diese Lösungsansätze neben gewünschten Vorteilen auch ganz eigene Probleme mit sich bringen. So sollen hier beispielsweise für die Entwicklungen im IT-Bereich das kompliziertere Transaktionsmanagement und auch die Probleme der genügend schnellen Datenübertragung in verteilten Datenbanken genannt werden (vgl. Kapitel 2). Ebenso sind auch die nach den neuen Strukturen gestalteten, objektorientierten Datenbanken keineswegs generell ihren Vorgängern überlegen. So sind in traditionellen Anwendungen häufig relationale Datenbanken von Vorteil.

Andere Bedenken werden am Ende des folgenden Zitats aus einem Beitrag zur Modernisierung des Staates durch Informationstechnik von Hans Brinckmann in ([Fr95], S. 85) geäußert: „Die Potentiale der Vernetzungstechnik können Interdependenzen unterstützen und zu neuer Integration von Verwaltungsfunktionen auf allen Ebenen innerhalb und zwischen Organisationen führen. ... Ist aber mehr Koordinationsleistung verfügbar, dann werden auch die gegenwärtig diskutierten Innovationen unserer öffentlichen Verwaltung greifbarer. Der Übergang von der starren Kopplung von Verwaltungseinheiten durch Hierarchie, Budget- und Personalrecht hin zu einem System von eher losen, über Ziele, Vorgaben und Controlling verkoppelten Einheiten mit mehr Eigenverantwortung, Wettbewerb und operationalen Zielvorgaben. Ob allerdings die Koordinationsvorleistung der Technik so schnell wächst wie der Koordinationsbedarf und ob insbesondere auch die erforderliche Qualifikation mit wächst, ist derzeit noch offen.“

Zugleich tritt in diesem Zitat auch stärker der Aspekt der sich gegenseitig unterstützenden oder behindernden Strukturen auf den Ebenen der IT und Organisation hervor. Und hierzu stellt Brinckmann am Ende seines Beitrags noch eine interessante Behauptung auf, die, wenn man ihr zustimmt, die gleichzeitigen Umstrukturierungsmaßnahmen und die Entwicklung neuer IT-Lösungen in der BA als einzig sinnvolle Vorgehensweise erscheinen läßt. Hier heißt es mit Bezug auf die Möglichkeit, mit Hilfe der Informations- und Kommunikationstechnik zugleich eine Steigerung der Komplexität, Anpassungsfähigkeit, Größe und Beweglichkeit von Organisationen zu erzielen: „Diese Chance kann allerdings nur genutzt werden, wenn an die Stelle der hierarchischen, rigiden Steuerung eine offene zielorientierte Führung tritt, die nicht auf einem Maximum an Kontrolle, sondern auf einem Optimum an Autonomie und Delegation beruht.“ (ebenda) Im Kapitel 4, in dem das Anwendungsbeispiel in den Vordergrund rückt, wird daher auch kurz auf die gegenwärtigen Organisationsstrukturen und die angestrebten Veränderungen eingegangen.

Es sei ebenfalls noch erwähnt, daß die von Karer und Müller vertretene Betrachtungsweise einer immer weiter zunehmenden Dezentralisierung keineswegs die einzig mögliche ist. So betont Bode in seinem Beitrag [Bo98] zu Management-Veränderungen durch Informationstechnik wechselnde Phasen von Dezentralisierung und Zentralisierung in der IT und sieht gegenwärtig sogar eher wieder Zentralisierungstendenzen: „Die Dezentralisierung war notwendig geworden, weil die Hierarchien nicht mehr schnell genug auf Veränderungen reagieren konnten. Doch jetzt werden, angetrieben durch die IT die Dinge wieder zentralisiert.“ (S. 16) Er untermauert diese Behauptung durch Beispiele für gegenwärtige Zentralisierungsmaßnahmen durch IT (ebenda, S. 18): „

- In digitalen Archiven, Universal Servern, und im Intranet werden alle Unternehmensdaten und Vorgänge zentral abgespeichert.
- Im Data Warehouse werden die Daten zentral für Auswertungen zusammengeführt.
- Mit Workflow-Systemen werden die betrieblichen Abläufe über alle Abteilungsgrenzen hinweg strukturiert.
- Die Kommunikation wird über Workgroup-Systeme unternehmensweit und sogar unternehmensübergreifend organisiert.
- Intra-, Inter- und Extranet schaffen einheitliche Plattformen für die Koordination.“

Die hier vorgestellten Auffassungen werden bei der Behandlung des Anwendungsbeispiels im Kapitel 4 wieder aufgegriffen. Im folgenden Abschnitt und in Kapitel 2 werden zunächst einmal Voraussetzungen für das grundsätzliche Verständnis der in Kapitel 3 vorgestellten Allokationsmodelle gelegt.

1.2 Grundlegende Begriffe

Wie bereits zuvor beschrieben wurde, können unterschiedliche historische Entwicklungen bei der Bildung von Datenbanken zu unterschiedlichen Formen der Verwaltung verteilter Datenbestände führen. Auch unterschiedliche Anforderungen können Ursache für divergierende Lösungsformen sein. In den folgenden beiden Abschnitten werden daher zunächst die hier üblichen Begrifflichkeiten eingeführt. Dabei wird zunächst etwas ausführlicher auf die verteilten Datenbanken eingegangen. Schon bei deren Behandlung wird deutlich werden, daß viele Aspekte bei der Organisation verteilter Datenbestände eine Rolle spielen und damit auch begriffsbestimmend sein können. Deshalb geht deren Betrachtung die Beschreibung weiterer Realisierungsarten voraus. Die in den folgenden Abschnitten vorgestellten Klassifizierungsüberlegungen werden benötigt, um später zu den behandelten mathematischen Modellen deren Wirkungsradius besser angeben zu können. Außerdem weichen in der Praxis vorkommende Realisierungen im allgemeinen von den abstrakten, idealisierten Definitionen theoretischer Ausführungen mehr oder weniger ab, so daß diese Abschnitte auch dazu dienen, die in Verbindung mit dem Praxisbeispiel angedachten Realisierungsvarianten besser einordnen zu können.

1.2.1 Verteilte Datenbanken

Künftig soll immer dann von **verteilten Datenbeständen** oder **verteilten Informationssystemen** gesprochen werden, wenn in einem Netz aus eigenständigen Rechnern auf Daten zugegriffen wird, die auf mehrere Rechner (echt größer 1) physisch verteilt sind. Dabei soll bei diesen Begriffen belanglos sein, wie der Zugriff auf diese Daten gelöst ist. Bei ersterem Begriff kommt mehr die Verteilung der Daten zur Geltung, bei letzterem wird der Aspekt deutlicher, daß diese natürlich auch

verwaltet werden müssen. Beide Begriffe stellen Oberbegriffe für die folgenden Betrachtungen zu Realisierungsformen dar, umfassen aber keine zentralen Datenbanksysteme.

Die Begriffsfestlegung für **verteilte Datenbanken** (vDB), **verteilte Datenbankmanagementsysteme** (vDBMS) sowie für die aus diesen beiden Komponenten gebildeten **verteilten Datenbanksysteme** (vDBS) wird in der Literatur nicht einheitlich vorgenommen.

Heuer und Sake beispielsweise heben die globale Sicht hervor. Zunächst grenzen sie in ihrer Definition verteilte Datenbanken gegenüber Client-Server-Datenbanken ab, indem sie betonen, daß dort „der *Datenbestand* physisch auf mehrere Knoten verteilt“ ist und nicht die Verarbeitungskapazität. Darüber hinaus fordern sie aber auch die folgende Eigenschaft: „Während die Datenbestände verteilt sind, erfolgt der Zugriff konzeptionell weiterhin über eine global verfügbare Schnittstelle. Die Verteilung ist für die Anwender transparent (also nicht sichtbar), d. h. die Anfragen werden so formuliert, als gäbe es einen globalen zentralen Datenbestand.“ ([HS95], S. 444)

Schwinn sieht dies in seiner Beschreibung verteilter Datenbanken als den Idealfall an und räumt damit auch Abweichungen ein. Außerdem hebt er auch die lokalen Komponenten hervor und macht Angaben über die Distanz zwischen den Rechnern und zu den wichtigsten Aufgaben (hierzu finden sich Erläuterungen in Kapitel 2) des DBMS ([Schw92], S. 229):

„Die Informationsverarbeitung in Unternehmen, die geographisch über mehrere Gebiete verstreut sind, verlangt nach einem Konzept für eine gemeinsame Organisation der Datenbestände. Im Unterschied zu den bisherigen Annahmen wird davon ausgegangen, daß mehrere Rechner (Stellen) in ein (lokales oder geographisch verteiltes) Kommunikations-Netz eingebunden sind. Dort wird eine autonome Datenverwaltung durch lokale DBMS vorausgesetzt.

Ein verteiltes Datenbanksystem integriert mehrere zentrale Datenbanksysteme zu einer Organisationseinheit, mit der Anwendungen im Idealfall genau so umgehen wie mit einem zentralen Datenbanksystem. Das zugrunde liegende Software-System einschließlich der lokalen DBMS wird als *verteiltetes Datenbank Managementsystem* (engl.: distributed database management system; DDBMS) bezeichnet. Eine seiner wichtigsten Aufgaben besteht darin, global formulierte Transaktionen in Subtransaktionen aufzugliedern und diese an die betreffenden Stellen zur Bearbeitung weiterzuleiten. Gleichzeitig muß es in Zusammenarbeit mit den lokalen DBMS die globale Serialisierbarkeit der verteilten und lokalen Transaktionen gewährleisten und das Recovery im verteilten System ermöglichen.“

Weske schließlich berücksichtigt globale Aspekte überhaupt nicht in seiner Definition: „Ein verteiltes Datenbanksystem (VDBS) besteht aus einer Menge von lokalen Datenbanksystemen, die durch ein Kommunikationsnetzwerk miteinander verbunden sind. Jedes lokale Datenbanksystem besteht aus einer lokalen Datenbank (LDB) und einem lokalen Datenbank-Management-System (LDBMS).“ ([We95], S. 214)

Nicht einmal die physische Verteilung wird von allen Autoren gefordert. So weisen Holloway und Warren in ihrem Beitrag in [Wi92] auf Seite 14 darauf hin, daß die physische Verteilung der Daten lediglich eine hinreichende, nicht aber eine notwendige Bedingung für verteilte Datenbanken sei: „A distributed database does not require geographic dispersement of the data. However, geographic dispersement of data is a sufficient condition to state that the database is distributed, since no one DBM could execute in multiple disjoint processors. However, geographic dispersion is not a necessary condition. It is possible for a database, existing totally in one site, to be partitioned and managed by multiple DBMs at that same site.“

Diese Beispiele unterschiedlicher Definitionsansätze dürften deutlich gemacht haben, daß es in Hinblick auf lokale und globale Anforderungen sowie auf die physische Verteilung der Daten unterschiedliche Forderungen geben kann. Bezüglich der lokalen Teilsysteme kann es nicht nur Unterschiede in Hinblick auf den Grad der lokalen Autonomie geben sondern auch in Hinblick auf den Aspekt, ob für diese ausschließlich Homogenität oder aber auch Heterogenität zugelassen ist.

Ergänzen möchte ich diesen Abschnitt nun noch um eine weitere Definition aus einer frühen Darstellung zu verteilten Datenbanken von Ceri und Pelagatti, die ihr einleitendes Kapitel mit einer Abgrenzung der Fälle, die sie betrachten möchten, beginnen.

“A distributed database is a collection of data which are distributed over different computers of a computer network. Each site of the network has autonomous processing capability and can perform local applications. Each site also participates in the execution of at least one global application, which requires accessing data at several sites using a communication subsystem.“ ([CP84], S. 6)

Mit dieser Formulierung werden sowohl die physische Verteilung und der Aspekt der logischen Zusammengehörigkeit des Datenbestandes (in lokalen oder geographisch verteilten Netzen, bei heterogenen oder homogenen Teilsystemen) umfaßt, andererseits aber auch explizit lokale und globale Anwendungen betont. Damit wollen die Autoren, wie sie es in ihrer Einführung auch deutlich machen, einerseits lediglich vernetzte, eigenständige Datenbanken eines gemeinsamen Unternehmens (ohne globale Komponente), wie sie die Definition von Weske durchaus noch umfassen würde, ausschließen. Andererseits sollen auch Multiprozessorsysteme, die keine lokalen Anwendungen unterstützen, von ihren Betrachtungen ausgenommen werden. Und schließlich fallen so auch die Client-Server-Datenbanken (vgl. Abschnitt 1.1.1) aus den Betrachtungen, da diese auf einen zentralen, integrierten Datenbestand zugreifen.

Trotz dieser Ausschließungen läßt diese Auffassung von verteilten Datenbanken mehr Möglichkeiten zu als etwa die von Heuer und Sake. Gerade deshalb möchte ich sie aber für diese Arbeit wählen, da hiermit in der Praxis vorkommenden Realisierungsformen besser eingefangen werden.

Der folgende Abschnitt wird weitere praxisrelevante Organisationsformen aufzeigen. Sie unterscheiden sich in Hinblick auf lokale Autonomie, Homogenität oder Heterogenität der Teilsysteme und die zur Verfügung stehenden globalen Möglichkeiten. Die vorgestellten Formen werden anschließend zur Einordnung des später betrachteten Praxisbeispiels verwandt.

1.2.2 Weitere Realisierungsarten

Neben den verteilten Datenbanken gibt es zahlreiche weitere Möglichkeiten, verteilte Datenbestände zu managen. Auch die hierzu in der Literatur vorkommenden Klassifizierungen weisen mehr oder weniger große Unterschiede auf. Daher kann man auf Abweichungen stoßen, wenn man zu den folgend vorgestellten Begriffen andere Quellen heranzieht. Da aber schon im vorausgegangenen Kapitel durch den dort erfolgten Vergleich unterschiedlicher Definitionsansätze die wichtigsten, verschiedenen Betrachtungsaspekte herausgearbeitet wurden, wird in diesem Abschnitt auf ähnlich umfangreiche, weitere vergleichende Betrachtungen verzichtet. Um aber den Kontext zum vorangegangenen Abschnitt herzustellen, wird zusätzlich darauf eingegangen, inwieweit die betrachtete Form unter die Definition von Ceri und Pelagatti fallen würde. Abschließend folgt dann die oben angekündigte Einordnung des Praxisbeispiels.

Auf **konventionelle Weise** greifen Anwendungen dann auf verteilte Informationen zu, wenn sie selbst deren Aufruf durch Kommunikation mit Anwendungen anderer Rechner lösen. Dies bedeutet, daß die Anwendungen wissen müssen, wo sich benötigte Daten befinden und unter welchem Programm mit welchen Datenformaten sie abzufragen sind. Außerdem muß geklärt sein, welche Operationen hierauf zulässig sind und wie die Daten zurückgereicht werden dürfen. Auch für die Fehlerbehandlung ist die Anwendung zuständig (vgl. [Da96], S. 9). Somit liegt hier zwar ein verteilter Datenbestand vor, aber es sind weder Verteilungstransparenz noch ein globales Transaktionskonzept verwirklicht. Dabei bedeutet „Verteilungstransparenz“, daß die Anwendungen, die auf die Datenbestände zugreifen, lediglich über eine globale Schnittstelle mit der Datenbank kommunizieren und über die Verteilung der Daten auf die verschiedenen Knoten keine Kenntnisse benötigen. Bei Vorliegen eines globalen Transaktionsmanagements kann man auch bei verteilten Datenbeständen von Konsistenz ausgehen, die dann das vDBMS gewährleistet. Bei der hier beschriebenen Variante sind dafür die Anwendungen selbst verantwortlich. Dies ist natürlich nicht unproblematisch. Da es keine gemeinsame knotenübergreifende globale Anwendung gibt - tatsächlich existiert weder eine Beschreibung des gesamten Datenbestandes noch eine Komponente, die sich einer globalen Anwendung widmen würde - fällt diese Variante nicht unter die von Ceri und Pelagatti gegebene Definition verteilter Datenbanken.

Auch in **offenen Multidatenbanksystemen** regeln die Anwendungen globale Aufgaben. Solche Systeme stellen einen sehr losen Verbund eigenständiger Datenbanksysteme dar. Im Gegensatz zum konventionellen Ansatz sind hier an jedem Teilknoten lokal Datenbanksystemmanagementssysteme vorhanden. Diese erbringen aber anders als die später vorgestellten föderierten, verteilten Datenbanksysteme keine (oder kaum) Partizipationsleistung. Durch Kommunikation mit den lokalen DBMS muß in offenen Multidatenbanksystemen die Anwendung die Integrationsleistung oberhalb der lokalen DBMS erbringen. Es gibt kein vDBMS, das für Verteilungstransparenz sorgt oder ein globales Transaktionsmanagement realisiert. Da die lokalen Systeme ein solches im allgemeinen auch nicht unterstützen, spielen in offenen Multidatenbanksystemen Kompensationsoperationen (Stornobuchungen) eine wichtige Rolle. Auch kann es sein, daß global nur lesender Zugriff erlaubt ist. Beispielsvorkommen dieser Realisationsform findet man etwa bei Literatursuchsystemen, die auf diverse eigenständige Bibliotheksdatenbanken zurückgreifen oder bei Buchungssystemen für Hotels oder Flüge. Derartig lose Zusammenschlüsse können nicht mehr als verteilte Datenbanken bezeichnet werden.

Alle folgenden Formen verteilter Datenbestände fallen unter den von Rahm ([Ra94], S. 27 ff.) gewählten Begriff der **Mehrrechnerdatenbanken**. Diese zeichnen sich lediglich durch eine Kooperation von Prozessoren bzw. DBMS aus und umfassen damit auch die verteilten Datenbanken. Über

eine Unterteilung in Mehrrechnerdatenbanken mit „Funktionaler Gleichstellung“ bzw. „Spezialisierung“, wobei er „funktional spezialisierte DBS“ auch „vertikal verteilte Mehrrechner-DBS“ nennt, werden (u. a.) die Client-Server-Architekturen abgegrenzt.

Es verbleiben die **horizontal verteilten Mehrrechner-DBS**, die Rahm als nächstes in solche **mit und ohne gemeinsame Externspeicherzuordnung** einteilt. Unter erstere fallen die von Ceri und Pelagatti ausgenommenen, in der Praxis sehr verbreiteten und den zentralen DBS noch sehr ähnlichen Multiprozessorsysteme, die Dadam ([Da96], S. 10) in von Rahm abweichender Terminologie zu den eng integrierten, homogenen verteilten Datenbanken zählt.

Typisch für **Multiprozessorsysteme** ist, daß jeder Prozessor Zugriff auf alle externen Speicher und je nach Variante auch auf einen gemeinsamen Hauptspeicher besitzt. Die beteiligten Rechner sind zudem nicht geographisch verteilt. Ziel dieser Konstellationen ist es, durch parallele Anfragebearbeitung einen erhöhten Datenbankdurchsatz zu erzielen. Das vDBMS erscheint nach außen wie ein einziges zentrales DBS. Die lokalen Schemata, die eine Computer-verständliche Beschreibung der lokalen Datenbanken⁸ darstellen, sind homogen und ihre Vereinigung liefert das globale Schema. Nur hierüber sind Datenbankabfragen möglich. Transaktionen werden auf den verteilten Datenbeständen wie in zentralen Datenbanken ganz oder gar nicht abgewickelt. Lokale Autonomie existiert dagegen nicht. So sind Schemaänderungen oder die Vergabe von Zugriffsrechten global zu vereinbaren. Derartige Formen werden normalerweise entsprechend vorliegender Gegebenheiten speziell entworfen.

Während Dadam mit den Multiprozessorsystemen zusammen auch fehlertolerante Systeme und die aus Dezentralisierung vormals zentraler Datenbanken entstandenen, homogenen verteilten Datenbanken zusammen als **eng integrierte, homogene verteilte Datenbanken**⁹ betrachtet, hat Rahm für letztere (für die aus Dezentralisierung entstandenen, homogenen verteilten Datenbanken) eine eigene Einteilung. Er nennt sie „**geographisch verteilte, integrierte Shared-Nothing-Mehrrechner-Datenbanksysteme**“. Nach der hier gewählten Auffassung ist diese Klasse noch zu den verteilten Datenbanken zu rechnen, auch wenn dabei die lokale Autonomie sehr eingeschränkt ist. Andererseits fallen diejenigen Formen, bei denen eine gemeinsame Peripherie-Nutzung und ggf. auch eine gemeinsame Hauptspeichernutzung vorliegt, aus der Begriffsdefinition von Ceri und Pelagatti heraus.

Die Realisierungsform der **geographisch verteilten, integrierten Shared-Nothing-Mehrrechner-Datenbanksysteme**, in der normalerweise dezentralisierte Datenbanken gestaltet werden, zeichnet sich durch Verteilungstransparenz aus. Hier erscheint idealerweise das vDBMS wie ein einziges zentrales DBMS (zDBMS). Die lokalen Schemata, die eine Computer-verständliche Beschreibung der lokalen Datenbanken darstellen, sind homogen und ihre Vereinigung liefert das globale Schema. Transaktionen werden auf den verteilten Datenbeständen ganz oder gar nicht abgewickelt. Dies entspricht dem Verhalten einer zentralen Datenbank. Dafür ist bei dieser Variante „die Autonomie der einzelnen Rechner/DBVS stark eingeschränkt, da Schemaänderungen, Vergabe von Zugriffsrechten etc. global koordiniert werden müssen.“ ([Ra94], S. 36) Nur über das vDBMS und somit das globale Schema sind Anfragen möglich. Diese können natürlich auf Datenbestände an einem oder mehreren Teilknoten begrenzt sein und werden dann an dem oder den entsprechenden Knoten abgewickelt.

⁸ Dies gilt auf allen Architekturebenen einer Datenbank: der konzeptionellen, der externen und der internen Ebene (vgl. [Schw92], S. 14, 15)

⁹ Die von Dadam gewählte Verwendung des Adjektivs “eng” weicht von der von Rahm gewählten deutlich ab! Bei Rahm kennzeichnet die Bezeichnung “eng” den gemeinsamen Zugriff auf Haupt- und Externspeicher. (Vgl. [Ra94], S. 37) Dieser liegt aber bei dezentralisierten Datenbanken nicht vor.

Aus dem Zusammenwachsen ehemals eigenständiger Datenbanken kommt man zu einer weiteren Realisierungsform, den **föderierten, verteilten Datenbanken** oder (in der Nomenklatur von Rahm) zu den **föderativen Shared-Nothing-Mehrrechner-Datenbanken**. Auch hier ist die Grundvoraussetzung eines verteilten Datenbestandes gegeben. Die Teilsysteme „bringen in der Regel nicht alle lokalen Daten in den globalen Verbund ein, sondern nur diejenigen, an denen ein 'globales Interesse' besteht“ ([Da96], S. 15). Für diese Daten gilt auch, daß Schemaveränderungen nur in gegenseitiger Absprache stattfinden dürfen, während die Teilsysteme für diejenigen Daten, die nur lokale Bedeutung besitzen, eigenmächtig Änderungen oder Erweiterungen vornehmen können. Damit bleibt eine gewisse lokale Autonomie gewahrt. Es gibt auch weiterhin rein lokale Anwendungen, die ausschließlich über die lokalen Datenbankmanagementsysteme abgewickelt werden. Für globale Anfragen ist das vDBMS zuständig. Es sorgt auch dafür, daß der in den globalen Verbund eingebrachte Datenbestand nach außen wie eine zentrale Datenbank erscheint (Transaktionsmanagement und Verteilungstransparenz). Für den gesamten Datenbestand gilt dies jedoch nicht!

Sehr verbreitet ist bei dieser Variante der Zusammenschluß **heterogener Systeme**. Liegen solche vor, werden zumindest Schematransformationen erforderlich. Ein sehr einfaches Beispiel hierfür ist das Anpassen verschiedener Bezeichner auf einen gemeinsamen globalen. Im allgemeinen werden auch Typanpassungen erforderlich. Darüber hinaus kann es sein, daß die Teilsysteme sogar nach verschiedenen Datenmodellen arbeiten (z. B. nach dem Netzwerk-Datenmodell, dem relationalen Datenmodell oder nach dem hierarchischen Datenmodell). Dann können sogar Modelltransformationen erforderlich werden um zu einem gemeinsamen globalen Schema zu gelangen. Wenngleich diese Transformationen wichtige Forschungsbereiche darstellen, werden sie in dieser Arbeit nicht berücksichtigt, auch wenn sich hieraus einige weitere Allokationsüberlegungen ergeben. So muß man bei heterogenen Systemen entscheiden, wo die Informationen für die erforderlichen Transformationen gehalten werden sollen und wie bei Änderungen andere Stellen informiert werden. Da sich gute konzeptuelle Schemata durch hohe Langlebigkeit auszeichnen, wird es meist sinnvoll sein, diese Regeln lokal repliziert zu speichern und nur bei Schema-Veränderungen Datenübertragungen und Updates in Kauf zu nehmen. Bei der Berechnung von Systemgrößen erhöhen sich natürlich durch die Transformationen die Antwortzeiten von Datenbank Anfragen. Das kann dafür sprechen, die Daten repliziert, sowohl in der gemeinsamen Form als auch in der jeweilig spezifischen Form, zu halten. Der Speicherplatz für die Transformationsregeln dürfte eine unwesentliche Rolle spielen.

Unter Vernachlässigung der Aspekte, die in Verbindung mit Transformationen stehen, lassen sich die später angestellten Allokationsüberlegungen aber auch auf die föderativen Datenbanken anwenden. Für das betrachtete Praxisbeispiel hat das Ausklammern der Transformationsaspekte keine Nachteile, da es sich hier um die Entwicklung einer komplett neuen Datenbankanwendung handelt, also eine homogene Situation gegeben ist. Dennoch ist in dem Projekt auch eine Stufe enthalten, in dem der Fall einer föderierten, heterogenen Datenbank vorliegt. Bei der Softwareumstellung auf die neue Generation, deren Dauer vermutlich mindestens ein Jahr beträgt, wird durch das Nebeneinander von Geschäftsstellen, die noch mit dem alten Verfahren arbeiten, und solchen, die bereits die neue Datenbank nutzen, genau der Fall des Zusammenarbeitens zweier heterogener Datenbanken desselben Datenmodells vorliegen. Diese vorübergehende Migrationsphase ist jedoch für die Allokationsüberlegungen der zu entwickelnden neuen Datenbank irrelevant.

Die folgende Tabelle zeigt eine zusammenfassende Übersicht über die vorgestellten Realisierungsarten.

Realisierungsart	Charakteristika	Vorkommen (Beispiele)	Definition vDBS dieser Arbeit erfüllt
konventionell	Anwendungsprogramme sind für Datenzugriffe, Fehlerbehandlung und Konsistenzprobleme verantwortlich	alte, aufgabenorientierte EDV-Lösungen in Unternehmen mit verteilten Daten	nein
Offene Multidatenbanksysteme	ein sehr loser Verbund von normalerweise heterogenen, lokal autonomen Datenbanken	Auskunfts- und Buchungssysteme	nein
Shared Everything oder Shared Disk-Mehrrechner DBS	von außen von einem zDBS nicht zu unterscheiden, keine lokale Autonomie an den beteiligten Rechnerknoten, lokale Anordnung der Rechner, d. h. keine geographische Verteilung der Rechner	Multiprozessorsysteme, sehr große Datenbanken mit strengen Antwortzeitanforderungen	nein
geographisch verteiltes, integriertes, Shared-Nothing-Mehrrechner-DBS	von außen von einem zentralen DBS nicht zu unterscheiden, keine lokale Autonomie an den beteiligten, geographisch verteilten Rechnerknoten	Dezentralisierte Datenbanken	ja
Föderierte, vDBS (Föderative, homogene und heterogene Shared-Nothing-Mehrrechner-DBS)	ein engerer Zusammenschluß heterogener oder homogener Datenbanken, Einigung über einen gemeinsamen globalen Datenbestand, lokale Autonomie über den rein lokalen Datenbeständen	Zusammenschluß ehemals eigenständiger Datenbanken	ja

Tabelle 1: Realisierungsarten verteilter Systeme

Abgeschlossen wird der Abschnitt nun mit der Einordnung der heutigen und zukünftig geplanten Datenbanklösung des Anwendungsbeispiels. Hierbei handelt es sich um die heute in der Arbeitsvermittlung eingesetzte Datenbanklösung der Bundesanstalt für Arbeit und deren Neuentwicklung, zu der später die Allokationsplanungen im Vordergrund stehen werden. Aufgabe dieser Software ist es, Arbeit suchende Personen mit ihren beruflichen Wünschen und Profilen ebenso wie Stellenangebote und Daten anbietender Firmen zu erfassen sowie Suchen auf diesen Daten zu ermöglichen, die zur Vermittlung von Stellen führen. (Genauere Ausführungen folgen in Kapitel 4.)

In der gegenwärtigen Lösung existieren in fast allen Arbeitsämtern eigenständige, **zentrale Datenbanken**, auf denen im jeweiligen Amt und ggf. auch in ausgelagerten, dem Amt zugeordneten Geschäftsstellen, über Terminals dialogorientiert gearbeitet wird. Diese Datenbanken betreiben asynchron (nachts) einen **konventionellen Datenaustausch** mit einem Zentralrechner über den in geringem Umfang Datenaustausch in andere Arbeitsämter stattfinden kann. Auch dieser Datenaustausch wird in Kapitel 4 noch näher beschrieben.

Im Rahmen der vorgestellten Allokationsstudie für die neue Softwaregeneration sind - bis auf einen zentralen Ansatz - alle angedachten Lösungsvarianten am ehesten den **geographisch verteilten, integrierten Shared-Nothing-Mehrrechner-DBS** zuzuordnen. In der ganz aktuellen Planung (Stand Juni 98) wird außerdem über **Multiprozessorsysteme** nachgedacht.

1.2.3 Netzeigenschaften

Bei der Behandlung der für die Allokationsplanung verteilter Datenbestände grundlegenden Begriffe fehlen in den unter 1.2 vorgenommenen Betrachtungen noch wichtige Einflußgrößen des Netzes. So wurden bisher weder der Aspekt der geographischen Ausdehnung der beteiligten Rechnerknoten, noch Auswirkungen unterschiedlicher Übertragungsformen angesprochen. Sie sollen in diesem Abschnitt behandelt werden. Hinzu kommt, daß viele der hier vorgestellten Netzeigenschaften in den später beschriebenen mathematischen Modellen nicht oder nur sehr pauschal berücksichtigt werden. Schon um auf diese Problematik hinweisen zu können, sind die hier eingeführten Begriffe unerläßlich. Und schließlich sind diese Aspekte auch für die gegenwärtig stattfindende Erwägung eines Multiprozessorsystems zur Lösung des Anwendungsbeispiels relevant.

Zunächst gibt es Informationssysteme, bei denen die Computer, die für die Datenhaltung eingesetzt werden, durch ein **lokales Netz** (Abk.: LAN für engl.: Lokal Area Network) miteinander verbunden sind. Ein lokales Netz zeichnet sich nach der Definition des Institute of Electrical and Electronics Engineers (IEEE) dadurch aus, daß

- es mehreren unabhängigen Rechnern die Möglichkeit gibt, miteinander zu kommunizieren,
- die Kommunikation auf ein eng begrenztes geographisches Gebiet mit einem Radius von höchstens einigen wenigen Kilometern begrenzt bleibt,
- ein Kommunikationsmedium mit mittlerer bis hoher Datenübertragungsgeschwindigkeit und geringer Fehlerrate eingesetzt wird,
- es der Kontrolle des Anwenders unterliegt und auf dessen Grundstück begrenzt bleibt, z. B. innerhalb eines Betriebsgeländes

(vgl. [Ku92], S. 117).

Ein Beispiel für ein Netz mit einer größeren, hier bundesweiten Ausdehnung, zeigt die Abbildung 1. Es verbindet die mit Knotenrechnern ausgestatteten Dienststellen der Bundesanstalt für Arbeit und trägt die Bezeichnung **IBAN**, eine Abkürzung für „Integriertes **BA**-Netz“. Sind Computer über solche, LAN-Ausdehnungen überschreitende Entfernungen miteinander verbunden, so spricht man von **geographisch verteilten Informationssystemen**. In den sie verbindenden **Wide Area Networks** (WANs) kommunizieren die Rechner im allgemeinen langsamer als in den oben beschriebenen LANs. Hinzu kommt, daß die Übertragungszeiten nur dann effektiv ausgeschöpft werden können, wenn lange Datenströme gesandt werden, da ansonsten das Verhältnis von Datenmenge zu Protokollinformationen zu schlecht ist. In Frage kommen Übertragungen über Telefon- und Datenleitungen oder per Satellit. Betrieben werden sie von öffentlichen oder privaten Netzbetreibern, die für die Nutzung der von ihnen bereitgestellten und gewarteten Kommunikationswege Gebühren verlangen. Um eine Verbindung von einem Rechner zu einem anderen in einem WAN herzustellen, gibt es meist mehrere Wege über unterschiedliche Verbindungsrechner. Dennoch ist bei WANs die „ständige Verfügbarkeit der Kommunikationsverbindungen zwischen den Teilsystemen nicht immer gesichert.“ ([Da96], S. 8) Deshalb ist eine Forderung nach möglichst hoher lokaler Autonomie der im Netz zusammengeschlossenen Rechner sinnvoll, damit zumindest lokale Anfragen und solche globalen, die unabhängig von den ausgefallenen oder nicht erreichbaren Knoten sind, weiter ausgeführt werden können. Aus diesem Grund kann auch eine vermehrte Haltung von Replikaten und Datenwegen (mit dem dazugehörigen zusätzlichen Verwaltungs- und Kommunikationsaufwand!) gewünscht sein.

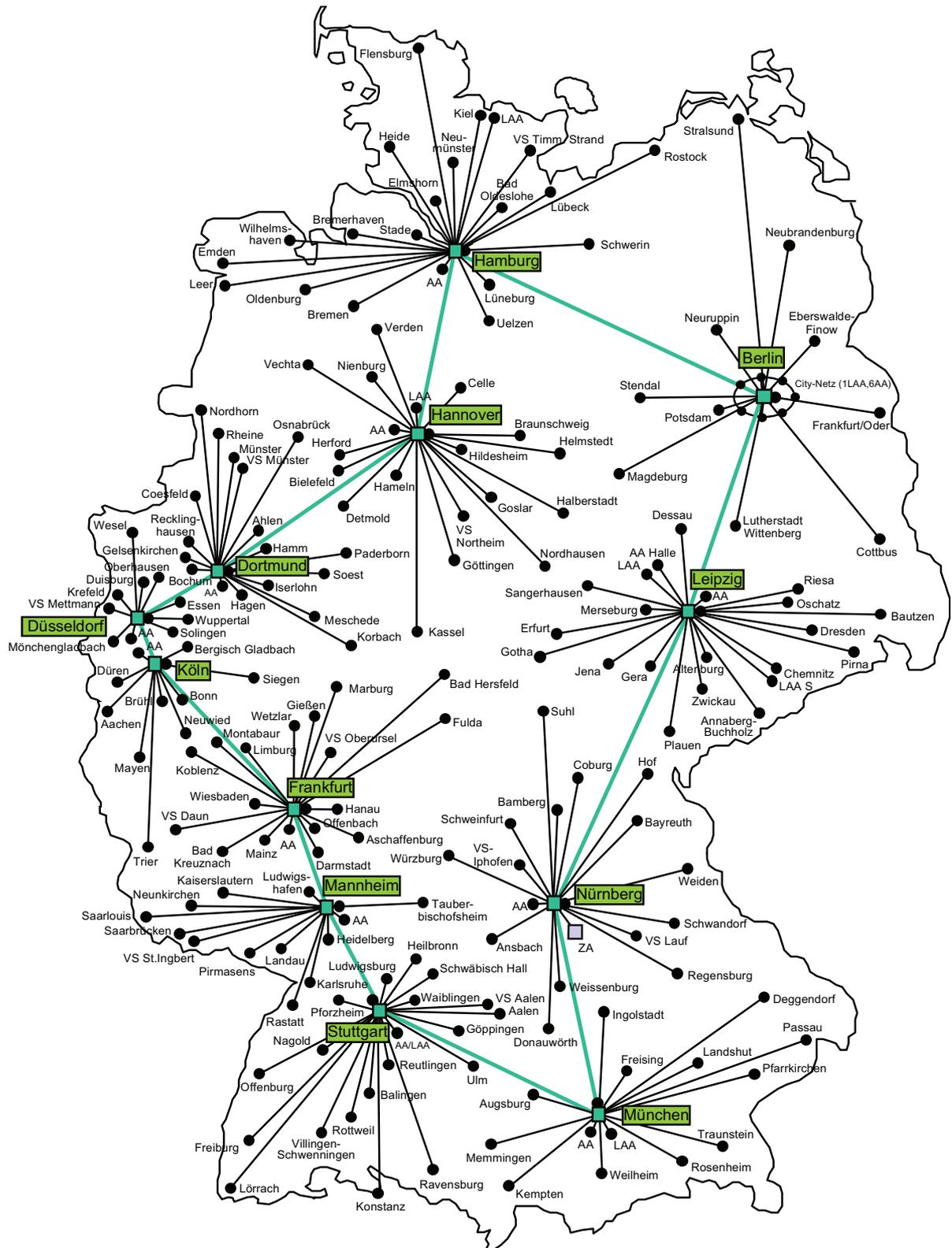


Abbildung 1: Das integrierte BA-Netz (BA-Dokument)

Damit sind bereits einige wesentliche Unterschiede zwischen WAN und LAN-Verbindungen genannt, die etwa bei der Entscheidung, eine zentrale Datenbank zu dezentralisieren, wichtig sein können. Durch die höhere Ausfallsicherheit der LANs und durch die hier möglichen kostenlosen Übertragungen, sind vDBS in LAN-Umgebungen sehr viel selbstverständlicher anzutreffen als in WANs.

Neben der auch vom Übertragungsmedium abhängigen, (im allgemeinen lokal grundsätzlich wesentlich höheren) möglichen Übertragungsgeschwindigkeit des vorhandenen oder zu gestaltenden Netzes ist auch die Zahl der physischen Verbindungen, die zwischen den Rechnern des Netzes bestehen oder bestehen sollen, von Bedeutung. Sie wirkt sich auf die Ausfallsicherheit des Systems und die Möglichkeiten paralleler Bearbeitung aus. Darüber hinaus haben auch der tatsächlich gewählte Übertragungsweg, die Netzkommunikationsprotokolle, die sich ebenfalls für LANs und WANs unterscheiden, sowie eingesetzte Komprimierungsalgorithmen Einfluß auf die Übertragungsgeschwindigkeit und die Fehlersicherheit.

Um dies etwas zu verdeutlichen, soll in ganz knapper Form auf einige Grundlagen von Rechnernetzen eingegangen werden. Die sogenannten Knotenrechner eines Rechnernetzes sind nicht unmittelbar miteinander verbunden, „sondern verfügen über spezielle Kommunikationsrechner (IMP) (engl.: Intermediate Message Processors), die für die gesamte Kommunikation mit anderen Knotenrechnern des Rechnernetzes zuständig sind.“ ([WE95], S. 350)

Sie verfügen über Speicherbereiche, in denen sie zu versendende Nachrichten zwischenspeichern. IMPs können auch lediglich spezielle Hardwarekomponenten sein, die in die eigentlichen Knotenrechner des Netzes eingebaut sind (vgl. Abbildung 2). In dem in Abbildung 1 abgebildeten IBAN sind die IMPs an den mit Rechtecken umrandeten Standorten vorzufinden.

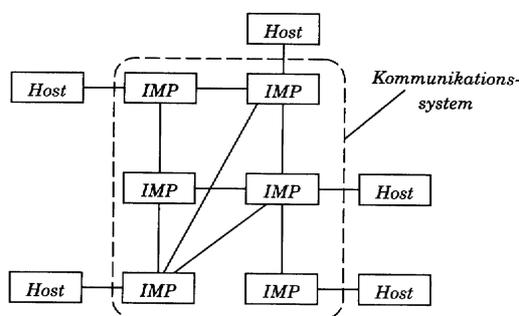


Abbildung 2: Aufbau eines Rechnernetzes
(entnommen aus [WE95], S. 349)

Man unterscheidet Broadcast-Netzwerke und Punkt-zu-Punkt-Netzwerke. Bei der ersten Variante kann eine Nachricht stets von jedem Knoten des Netzes empfangen werden, und sie ist typisch für LAN-Netze. Bei den Punkt-zu-Punkt-Netzwerken wird eine Nachricht stets von einem Knoten zu genau einem anderen gesendet. Die anderen Knoten können diese Nachricht nicht empfangen.

Da bei den Punkt-zu-Punkt-Netzwerken Nachrichten nicht unmittelbar vom Absender zum Empfänger gesandt werden können, wenn es sich nicht zufällig um Nachrichten zwischen benachbarten Rechnern handelt, kommt in diesen Netzen der Wegfindung, dem sogenannten Routing, eine besondere Bedeutung zu. Dabei werden größere Datenmengen in Datenpaketen übertragen, wobei es von der Art der Datenvermittlung abhängt, ob diese alle denselben Weg zum Zielknoten nehmen. Bei sogenannten virtuellen Verbindungen wird die Verbindung vor der eigentlichen Datenübertragung hergestellt. Dann nehmen alle Pakete denselben Weg. Eine andere Datenvermittlungsart, das Datagrammverfahren, gibt jedem Paket die Zieladresse mit auf den Weg, der hierbei nicht zuvor ermittelt wird. Jedes Paket kann einen anderen Weg nehmen. Bei sehr kurzen Nachrichten kann gegenüber der virtuellen Wegfindung das Verhältnis von Nachricht zu Adresse sehr ungünstig sein. Bei Paketen mit virtuellen Adressen, die kürzer als echte Adressen sind, ist es günstiger. Zur Weiterleitung der Pakete liegen in jedem IMP Tabellen vor, aus denen anhand der virtuellen Adresse der richtige Folgerechner ermittelt wird. Trotz des bei kürzeren Nachrichten schlechteren Verhältnisses der übertragenen Dateneinheiten zur Adressinformation, sind Datagramme für kurze,

allerdings nicht interaktive Anwendungen vorteilhafter, da hier der Aufwand, zunächst erst einmal eine Verbindung zu ermitteln, entfällt. Bei interaktiven Vorgängen hingegen lohnt es sich, einen Weg aufzubauen und für die Dauer der Interaktion beizubehalten. Den einzelnen, möglicherweise sehr kurzen Nachrichten kommt dann noch zu gute, daß wenig zusätzliche Adreßinformation mitgeschickt wird.

Weitere Unterschiede der beiden Verfahren liegen in der Fehlertoleranz (bei Datagrammen höher) und in der Tatsache, daß bei Datagrammen der IMP des Zielrechners die zusätzliche Aufgabe hat, die möglicherweise in veränderter Reihenfolge eingegangenen Pakete wieder in die Ausgangsreihenfolge zu bringen.

Neben diesen unterschiedlichen Möglichkeiten zur Mitgabe der Zieladresse hat außerdem das gewählte Routing-Verfahren größeren Einfluß auf die Netzübertragung. Man unterscheidet statische und adaptive Algorithmen. Bei ersteren Lösungsansätzen ist der Weg von Paketen unabhängig von aktuellen Netzbelastungen. Anhand eines Netzgraphen wird zu Beginn des Netzstartes für alle aus Netzanforderungen denkbaren Wege-Verbindungen ein Weg mit minimalem Aufwand ermittelt. Dabei können über Wertigkeiten entlang der Kanten des Graphen durchaus Aspekte wie Leitungsbreite und Belastung in die Berechnung mit eingehen. Minimal bezüglich der betrachteten Aspekte ist dann ein Weg vom Start zum Ziel durch den Graphen mit minimaler Kantensumme. Adaptive Verfahren hingegen reagieren bei jeder Anforderung neu auf die aktuelle Netzsituation. Bei diesen Verfahren unterscheidet man, je nachdem, wo das Verfahren hierzu allokiert ist, in zentrale und verteilte Algorithmen.

Wenn später Datenübertragungen betrachtet werden, so sind eigentlich auch die aus solchen zur Netzorganisation erforderlichen Datenströme zu betrachten. Solche Transfers entstehen z. B. bei Netzbelastungs- oder Zustandsabfragen, bei der Anforderung von Pufferspeicherplatz oder für das Entdecken, Verhindern oder Auflösen von Verklemmungen, die aus konkurrierendem Zugriff aus diese Speicher auftreten können.

Das IBAN ist ein Punkt-zu-Punkt-Netzwerk, in dem adaptive Routing-Algorithmen mit virtueller Datenvermittlung eingesetzt werden. An zwölf Standorten (vgl. Abbildung 1, S. 14) sind sogenannte Points of Present (POPs) eingerichtet, an denen zum Zwecke des Datentransfers Zwischenspeicherungen stattfinden.

2 Verteilung

Sowohl bei den integrierten wie auch bei den föderativen Shared-Nothing-Datenbanksystemen stellt die physische Verteilung des Datenbestandes einen wesentlichen Aspekt für die mögliche Leistungsfähigkeit des Systems dar. Sie ist bestimmend für den Kommunikationsaufwand zwischen den Knoten. Außerdem beeinflußt sie die Lastbalancierung, „da die einem Rechner zugeordneten Daten die von ihm zu verarbeitenden Operationen bestimmen. Durch die Möglichkeit, Teile der Datenbank repliziert zu speichern, wirkt sich die Datenverteilung auch auf die Verfügbarkeit aus.“ ([Ra94], S. 59)

In Abhängigkeit von den Wünschen an die Datenbank sind daher die Allokationsüberlegungen ein wesentlicher Bestandteil bei der Gestaltung verteilter Datenbanken.

Deshalb sollen in diesem Kapitel zum einen Beweggründe für die Verteilung, zum anderen Grundlagen zur Beschreibung verteilter Datenbestände und des Verteilungsvorganges gelegt werden. Im letzten Abschnitt wird zudem auch auf Aspekte der Umsetzung von Anwenderwünschen einschließlich möglicher Probleme eingegangen. Auswahl und Umfang der Betrachtungen sind durch die mathematischen Modelle, die in Kapitel 3 vorgestellt werden, bestimmt. So soll der Zugang zu diesen sowie ihre Beurteilung erleichtert werden.

Für die Beschreibung von Daten und ihrer Zerlegung wird üblicherweise ein bestimmtes Modell zugrunde gelegt. Im Bereich der verteilten Datenbanken ist dies vornehmlich das mengenorientierte, relationale Datenmodell. Das hängt damit zusammen, daß die Fragmentierung einer nicht relationalen Datenbank nur sehr eingeschränkt möglich ist (vgl. [Ra94], S. 11). Rahm weist an gleicher Stelle auch auf den Vorteil relational aufgebauter Datenbanken hin, ein weit größeres Anfrage-Optimierungspotential zu besitzen als satzorientierte, navigierende Anfragesprachen¹⁰. Den Abschnitten zur Verteilung, Fragmentierung und Allokation und zu Aspekten des Managements verteilter Datenbanken geht daher ein Abschnitt über relationale Datenbanksysteme voraus.

2.1 Anforderungswünsche an verteilte Datenbanken

Bereits zentrale DBMS erfüllen, wie Kapitel 1 gezeigt hat, eine Reihe wesentlicher Wünsche, die an das Verwalten umfangreicher Datenbestände gestellt werden, insbesondere die Unabhängigkeit der Anwendungen von internen Änderungen und die Vermeidung von Redundanz. In verteilten Datenbanken sind typische Wünsche für die globalen Anfragen die Ortstransparenz und ein Single-System-Image. Man möchte also weitestgehend, je nach zusätzlich gewünschter erforderlicher lokaler Autonomie, weiterhin sowohl den Eindruck einer zentralen Datenbank als auch die den gesamten Datenbestand umfassende Verwaltung erhalten.

Dazu gehört auch, daß das von zentralen Datenbanken bekannte Transaktionskonzept ebenso für verteilte Datenbanken umgesetzt wird. Das heißt, auch Transaktionen, die Daten an mehreren Kno-

¹⁰ Dennoch können, wie wir im vorangegangenen Kapitel gesehen haben, bei verteilten Datenbanken durch das Zusammenwachsen heterogener Datenbanken verschiedene Modelle aufeinandertreffen, somit also insbesondere auch andere Modelle als das relationale vorkommen. Da dieser Aspekt weder in den hier behandelten mathematischen Modellen noch in für den (auf dem relationalen Modell beruhenden) Anwendungsfall innerhalb der BA mit ihrer zentralen EDV-Planungsstruktur Relevanz besitzt (vgl. Abschnitt 1.2.2, S. 11), möchte ich auf nicht-relationale Modelle und mögliche Transformationsprobleme, die aus dem Zusammenschluß verschiedener Datenbanken mit unterschiedlichen zugrunde liegenden Modellen entstehen, in dieser Arbeit nicht eingehen.

ten betreffen, müssen stets ganz oder gar nicht ausgeführt werden (**Atomizität**), führen die Datenbank von einem konsistenten Zustand in einen weiteren (**Konsistenz**), sind in ihrer Wirkung unabhängig von anderen Transaktionen (**Isolation**) und besitzen **Dauerhaftigkeit**. Eigens für verteilte Datenbanken entwickelte Synchronisationsverfahren¹¹, die in Abschnitt 2.3.3 noch genauer betrachtet werden, setzen diese Forderungen um. Die gerade beschriebenen Eigenschaften müssen zudem bei Systemausfällen einzelner oder mehrerer Knoten und deren Wiederinbetriebnahme sichergestellt werden. Hierzu gibt es für verteilte Datenbanken entwickelte Recovery-Lösungen.

Für die Abkehr von einer zentralen Datenbank und einem zentralen DBMS und die Verteilung von Daten sprechen andere Gründe. Am häufigsten genannt (insbesondere für lokale Netze) werden **die Erhöhung der Ausfallsicherheit (Verfügbarkeit)** und die **Verbesserung des Durchsatzes**. Aber auch eine **allgemeine Kostenreduzierung**, z. B. in geographisch verteilten Netzen durch das Einsparen von Kommunikationskosten oder bei der Hardware durch den Einsatz mehrerer kostengünstiger PCs gegenüber einem Großrechner, kann Grund für die Realisierung verteilter Datenbanken sein.

Problematisch an diesen Forderungen und damit auch an der Bestimmung einer optimalen Datenverteilung ist, daß sie sich zum Teil gegenseitig behindern und zudem einer Vielzahl von Abhängigkeiten unterliegen. Diese sind noch dazu oft nur annäherungsweise zu bestimmen. Gegenätzliche Forderungen liegen etwa in dem Wunsch nach Verfügbarkeit und der Beschleunigung von Änderungsoperationen. So kann die Verfügbarkeit durch Replikate an verschiedenen Knoten ebenso wie die Geschwindigkeit von Leseoperationen stark verbessert werden. Der Aufwand bei Änderungen steigt demgegenüber jedoch stark an. Ebenso stehen sich Parallelisierung und lokale Anfragebearbeitung gegenüber. Um Kommunikationsvorgänge zu reduzieren, ist es sinnvoll, häufig gemeinsam benötigte Daten an einem Rechner zu halten; für die Parallelisierung verteilt man gerade solche Daten. Allerdings kann hier die Kenntnis der anfallenden Datenmenge eine Hilfe sein, denn Parallelisierung lohnt sich nur bei großen Datenmengen. Zudem können auch lokale Replikate in Betracht gezogen werden. Und schließlich kann Lokalisierung aus geographischen Gründen abweichen von einer gleichmäßigen Lastverteilung, also einer ausgewogenen Rechnerauslastung (vgl. [Ra94], S. 70, 71). Ferner werden Verfügbarkeit und Durchsatz auch durch die zu erwartenden Anfragen, den Umfang an Lese- und Änderungsaufträgen, die zu erwartende Ausfallsicherheit der Hardware und die eingesetzten Verfahren zur Anfragebearbeitung, zur Kommunikation, zur Transaktionsgestaltung und zur Fehlerbehebung beeinflußt.

Als weitere wünschenswerte Eigenschaft verteilter Datenbanken wird deren modulare Wachstums- bzw. **Anpassungsfähigkeit** angesehen. Auch die **Unterstützung dezentraler Organisationsstrukturen** kann für Verteilungsentscheidungen eine Rolle spielen. Zentrale Lösungen können z. B. auf lokale Akzeptanzprobleme stoßen. Dies kann bis hin zu zusätzlichen, vor Ort entwickelten, eigenständigen Lösungen führen, die dann unkontrollierte Redundanz mit sich bringen! Aber auch die **Berücksichtigung heterogener Datenbanken** innerhalb eines Unternehmens in einem globalen Konzept kann unter diesen Gesichtspunkt fallen.

¹¹ In diesem Kontext spielt auch die in Abschnitt 1.2.1 in dem Zitat von Schwinn erwähnte Serialisierbarkeit globaler Transaktionen eine Rolle (s. S. 7). Serialisierbarkeit ist ein für die genannten Eigenschaften wesentliches, überprüfbares Kriterium, das für jede Menge parallel bzw. verzahnt auszuführender Transaktionen gelten muß und besagt, daß Ausgaben und der durch den Ablauf der Teiloperationen der Transaktionen entstandene Datenbankzustand auch durch eine (beliebige) seriell entstandene Ausführungsreihenfolge der beteiligten Transaktionen hätten erzielt werden können.

Und schließlich soll bei all den gewünschten Eigenschaften dennoch die Systemadministration, die das kompliziertere Synchronisations- und das Recovery-Management umfaßt, so einfach wie nur möglich sein!

2.2 Relationale Datenbanksysteme

Relationale Datenbanksysteme basieren auf dem relationalen Datenmodell, dessen umfassende Behandlung allerdings den Rahmen dieser Arbeit sprengen würde. Hier sollen lediglich Voraussetzungen abgesichert und Schreibweisen festgelegt werden. Ansonsten verweise ich auf die vorhandene Standardliteratur zu Datenbanken und verteilten Datenbanken. Die hier nicht dargestellten Inhalte lassen sich in knapper Form gut in [Ra94] nachlesen, in etwas ausführlicherer und mathematischerer Darstellung findet man sie bei Heuer und Saake ([HS95]).

„In relationalen Datenbanken werden die Daten in Tabellen oder Relationen gespeichert. Jede Relation besteht aus einer bestimmten Anzahl von Spalten oder Attributen sowie einer Menge von Zeilen oder Tupeln (Sätzen). ... Jedem Attribut ist ein sogenannter Domain (Definitionsbereich) zugeordnet, der die zulässigen Werte festlegt. Die Beschreibung einer Relation umfaßt den Namen der Relation und deren Attribute sowie die zugehörigen Domains. Die Ausprägung der Relation ist durch die Menge der Tupel gegeben und entspricht einer Teilmenge des kartesischen Produktes über den Attribut-Domains. Die Mengeneigenschaft von Relationen bedeutet, daß kein Tupel mehrfach vorkommen darf und daß keine vorgegebene Reihenfolge unter den Tupeln besteht.“ ([Ra94], S. 12)¹²

Das Relationenmodell fordert die Einhaltung von zwei Integritätsbedingungen:

1. Jedes Tupel einer jeden Relation muß eindeutig über einen sogenannten Primärschlüssel der betreffenden Relation identifizierbar sein. Dieser kann aus einem oder einer Kombination aus mehreren (ggf. allen) Attributen der Relation bestehen. Es kann auch durchaus mehrere, verschiedene Schlüssel geben, z. B. eine Personalnummer und alternativ dazu die Kombination aus Name, Vorname und Anschrift. Alle Schlüssel haben jedoch die Eigenschaft, daß sie durch Entfernen bereits eines Attributs ihre Eigenschaft, jedes Element eindeutig zu identifizieren verlieren; man sagt hierzu, daß Schlüssel minimal sein müssen. Aus den möglichen Schlüsseln einer Relation ist der Primärschlüssel ein willkürlich ausgezeichneter.
2. Ferner muß für Relationen, in denen ein Fremdschlüssel auftaucht, d. h. ein Schlüssel einer anderen Relation, folgendes gelten: Das Element, auf das durch den Fremdschlüsselwert verwiesen wird, muß existieren.

¹² Heuer und Saake legen in ihrer Definition zudem fest, daß die Reihenfolge der Attribute in den Tabellen keine Rolle spielen soll, indem sie zur Definition nicht das kartesische Produkt heranziehen. Damit sind Relationen, die sich lediglich in der Reihenfolge ihrer Attribute unterscheiden, automatisch gleich. Bei Rahm wären derartige Relationen nur äquivalent. Auch wenn hier Rahm zitiert wurde, sollen dennoch äquivalente Relationen als gleich angesehen werden.

Eine Menge von Relationenschemata bildet dann ein Datenbankschema; eine Datenbank wird durch eine Menge von Relationen gebildet. Außerdem sollen für die Tupel der Datenbank die beiden gerade genannten Integritätsbedingungen gelten. Doch selbst wenn diese Minimalforderungen erfüllt sind, bleiben weitere Probleme, z. B. Anomalien¹³ oder nicht verwaltete Redundanzen oder Abhängigkeiten¹⁴. Deshalb wird für viele Datenbanken neben den oben angegebenen zwei Integritätsregeln das Vorliegen einer Normalform¹⁵ verlangt. Da unsere späteren Allokationsüberlegungen auf bereits gebildeten Relationen aufsetzen, wird hier nicht weiter auf diese Theorien eingegangen.

Eine verbreitete Beschreibungsform für relationale Datenbanksysteme liegt in der Drei-Schichten-Architektur. Hierbei wird die Datenbank auf drei Ebenen beschrieben, der konzeptuellen, der internen und der externen Ebene. Dabei legt die konzeptuelle Ebene den logischen Aufbau der Datenbank fest. Im konzeptuellen Schema finden sich beim relationalen Datenbankmodell dann u. a. die Definition sämtlicher Relationenschemata und Domains sowie alle Integritätsbedingungen wieder. Die externen Schemata beschreiben die Sichten der Benutzer und Anwendungsprogramme, die auf die Datenbank zugreifen. Sie sind normalerweise genau auf deren Bedürfnisse zugeschnitten und umfassen nicht alle Objekte der konzeptionellen Ebene. Das interne Schema schließlich liefert die Beschreibung für die physische Speicherung und Indexe. Auf allen Beschreibungsebenen und natürlich auch bei der Formulierung der Transformationsregeln spielt aber das relationale Modell zur Identifizierung der betrachteten Datenmenge eine wesentliche Rolle.

Teilbestände von relationalen Datenbanken stellen eventuell durch Mengenoperationen verknüpfte Teilmengen der Relationen der Datenbank dar. Ihre Beschreibung spielt nicht nur im Rahmen der Fragmentierung von Datenbanken, sondern auch zur Spezifizierung eines Anfrageergebnisses oder eines zu ändernden Datenbestandes eine wesentliche Rolle. Dabei lassen sie sich in Relationenalgebra, im Relationenkalkül oder in hierfür entwickelten Anfragesprachen formulieren. Unter letzteren ist SQL (Structured Query Language) mit allerdings zahlreichen Versionen besonders verbreitet. Für unsere Betrachtungen spielt die Möglichkeit, die Zerlegung von Datenbanken übersichtlich beschreiben zu können, eine besondere Rolle. Deshalb ist für uns die Relationenalgebra mit ihrer kompakten Darstellungsform und Mengenorientierung am vorteilhaftesten.

In der folgenden Tabelle sind die relevanten Operationen und ihre für diese Arbeit gewählte, jeweilige Schreibweise dargestellt. Erläuterungen und Beispiele finden sich in der oben angegebenen Basisliteratur. Die Großbuchstaben R und S stehen für Relationenschemata, die Großbuchstaben A und B für Attribute. Mit Hilfe der dort aufgeführten Operationen (tatsächlich würden weniger reichen, da diese kein minimales System bilden) kann man beliebig verschachtelte Ausdrücke bilden. Diese werden immer - allerdings unter Berücksichtigung von Klammern - von links nach rechts

¹³ Anomalien können bei ungeschickt ausgewählten Relationen auftreten, z. B. eine Einfügeanomalie, wenn für die Eingabe eines Tupels zwingend ein bestimmter Attributwert erforderlich ist und eine Situation eintritt, für die dieser Wert fehlt.

¹⁴ Redundanz kann durch Mehrfachspeicherung vorliegen oder aber auch in der Abhängigkeit verschiedener Attribute liegen (vgl. auch Fußnote 15). So hängt z. B. vom Attribut "Familienstand" ab, welche Werte bei "Steuerklasse" zulässig sind. Über eine Kontrolle durch das Datenbanksystem kann die Integrität der Datenbank sichergestellt bzw. unterstützt werden (fehlerhafte, aber plausible Daten aufgrund von Fehleingaben sind unvermeidbar). So müssen bei Änderung eines Wertes auch zugleich vorhandene Replikate angepaßt bzw. Plausibilitätsprüfungen abhängiger anderer Werte und ggf. Rückfragen ausgelöst werden.

¹⁵ "Relationen müssen adäquat strukturiert sein, damit der Datenbank-Betrieb nicht durch unerwünschte Effekte gestört ist. Vor allem ist der Gefahr mehrfach gespeicherter - redundanter - Daten zu begegnen. Mit der Theorie der Normalformen werden Relationen systematisch auf Schwachstellen hin analysiert und ggf. umstrukturiert." ([Schw92], S. 66) Damit helfen Normalformen, geeignete Relationen für gegebene Modellierungssituationen zu finden. Eine Darstellung der ersten bis fünften Normalform sowie der Boyce-Codd Normalform und umfangreiche Hinweise zu weiterer Literatur findet man in der angegebenen Quelle auf den Seiten 66 - 81.

abgearbeitet. Bei der Ermittlung des Ergebnisses sind doppelte Tupel zu eliminieren, damit die Mengeneigenschaft der Relationen gewahrt bleibt.

Operation	Schreibweise	Anmerkung
Vereinigung	$R \cup S,$ $\cup_{i \in F} R_i$	R, S bzw. die endliche Familie R_i sind Relationen und müssen dasselbe Relationenschema besitzen, F ist Indexmenge.
Durchschnitt	$R \cap S,$ $\cap_{i \in F} R_i$	R, S, R_i, F wie oben.
Differenz	$R - S$	R, S wie oben.
Kartesisches Produkt	$R \times S$	
Selektion	$\sigma_{\text{Selektionsbedingung}} R$	Die Selektionsbedingung ist ein überprüfbarer boolescher Ausdruck auf den Tupeln von R.
Projektion	$\pi_{\{\text{Attributmenge}\}} R$	Die Attributmenge stellt diejenige Teilmenge der Attribute von R dar, auf die projiziert wird.
Verbund (Join)	$R [A\theta B] S$	A Attribut von R, B von S mit demselben Wertebereich, θ Vergleichsoperator; das Ergebnis ergibt sich durch Auswertung des Operators und Konkatenation über alle Attribute.
Gleichverbund (Equi-Join)	$R [A=B] S$	Verbund, bei dem θ der Gleichheitsoperator ist.
Natürlicher Verbund (Natural Join)	$R \bowtie S$	Verknüpfung per Gleichheit über alle gemeinsamen Attribute; dabei werden gleiche Spalten nur einmalig in das Ergebnis aufgenommen.
Semi-Verbund (Semi-Join)	$R \ltimes_{[A\theta B]} S$	Nach dem Verbund von S und R wird auf die Attribute von R projiziert.
Umbenennung	$\beta_{\text{Attributbezeichner} \leftarrow \text{Attributbezeichner}} R$	Die Relation bleibt gleich, nur das Schema ändert sich.

Tabelle 2: Relationenalgebra - Operationen

2.3 Datenbankverteilung

Für die Zuordnung der Daten zu den beteiligten Knoten des Netzes sind zwei Ansätze typisch, erstens das **Top-Down-Verfahren**, zweitens die **Bottom-Up-Methode**. Dabei ist der Top-Down-Ansatz gefragt, wenn bereits ein globales konzeptuelles Schema (GKS) vorliegt und es gilt, festzulegen, „wie die Objekte (Relationen) des GKS auf die einzelnen Rechner und deren LKS verteilt werden.“ ([Ra94], S. 59) Das Bottom-Up-Vorgehen hingegen ist bei der Bildung föderierter Datenbanken von Bedeutung, bei denen ausgehend von den lokalen Datenbanken zunächst erst einmal ein gemeinsames GKS gebildet werden muß.

Das Top-Down-Vorgehen erfolgt in einem zweistufigen Prozeß. Zunächst wird eine geeignete Fragmentierung¹⁶ für das globale Schema gesucht, dann wird die Allokation festgelegt. Dabei versteht man unter **Fragmentierung** eine Zerlegung der Relationen des globalen konzeptuellen Schemas, aus der die ursprünglichen Relationen rekonstruiert werden können. Nur für das Sicherstellen der Rekonstruktion ist es erlaubt, nicht disjunkte Mengen zu bilden (vgl. [KS91], S. 479). Die **Allokation** hingegen beschreibt die Zuordnung der Fragmente zu den Rechnern (Ortszuweisung). Dabei handelt es sich nicht um eine funktionale Zuordnung, da Replikate möglich sind.

Die Unterscheidung in Fragmentierung und Allokation entspricht einer Unterscheidung der Verteilung auf logischer und physischer Ebene. Sie ist nicht zwingend. Da wir aber für die Fragmentierung die maximal mögliche Redundanzfreiheit gefordert haben, wird mit der Beschreibung der Allokation dann sofort offensichtlich, wo Redundanzen vorliegen.

Ein Beispiel für die Gestaltung von Fragmentierung und Allokation zeigt Abbildung 3.

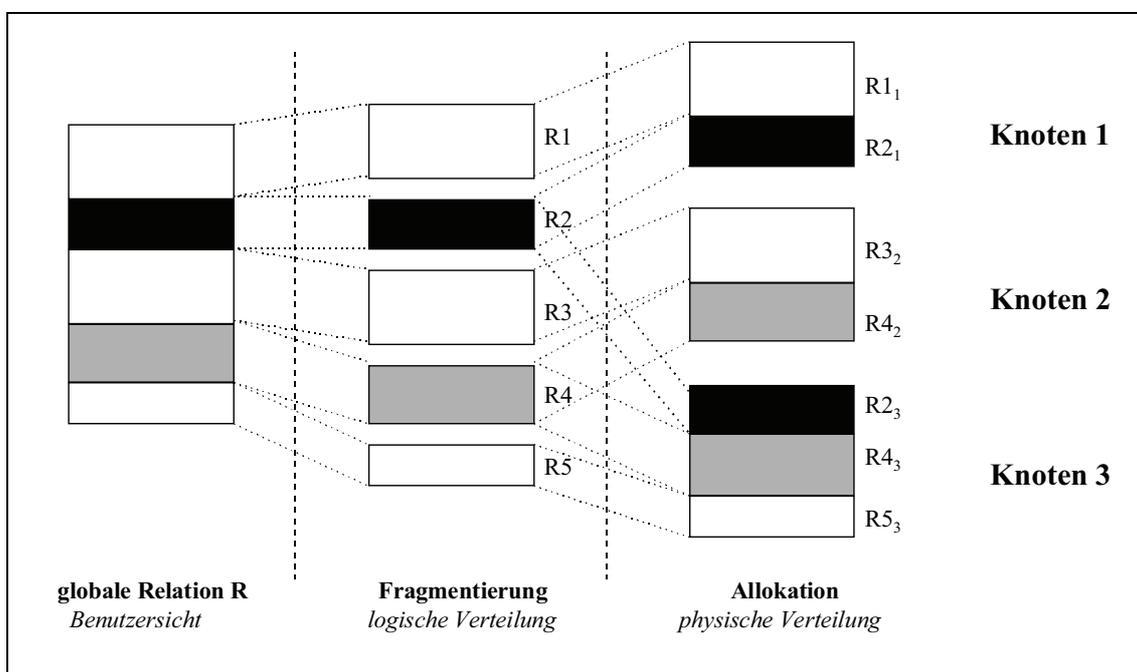


Abbildung 3: Fragmentierung und Allokation (nach [CP84], S. 41)

Die lokalen Relationen der beteiligten Rechner ergeben sich aus den dem jeweiligen Rechner zugeordneten Fragmenten, die sich wiederum über die oben behandelten Operationen in Relationenalgebra beschreiben lassen.

Bei föderierten Datenbanken ist im Planungsprozeß hingegen zu überlegen, welche Daten in das globale Schema eingebracht werden sollen, welche Transformationen hierfür von Seiten der lokalen Datenbanken erforderlich sind und ob zugunsten gewünschter Anforderungen zusätzliche Replikate an anderen Rechnern aufgenommen werden sollen. Für die bereits bestehenden Datenbestände fallen hingegen normalerweise keine neuen Fragmentierungs- und Allokationsüberlegungen an! Angenommen, das globale Schema wurde relational modelliert. Dann kann man wiederum (diesmal allerdings nicht als Planungsprozeß) Fragmente und Allokation der globalen Relationen beschrei-

¹⁶ In der deutschsprachigen Literatur wird auch der Begriff Partitionierung verwandt. In der englischen Sprache ist "fragmentation" üblich. Auch wenn hierdurch häufigere Wortwiederholungen auftreten, wird innerhalb dieser Arbeit zum leichteren Verständnis nur der Begriff Fragmentierung verwandt..

ben. Für globale und lokale Zugriffsmöglichkeiten sorgen Transformationsregeln. Da Probleme der Heterogenität von Modellen und Schemata in dieser Arbeit ausgeklammert sind (s. o., S. 11), können wir unsere Betrachtungen auf Fragmentierung und Allokation und das Arbeiten auf verteilten Datenbeständen beschränken. Letzteres umfaßt Ausführungen zur Anfragebearbeitung, Synchronisation und Kopienhandhabung.

2.3.1 Fragmentierung

Da eine relationale Datenbank aus einer Menge von Relationen besteht, ist die einfachste Form der Fragmentierung das Einteilen in komplette Relationen. Diese werden dann im Rahmen der Allokation diversen Rechnern zugeordnet. Vorteilhaft hieran sind, durch die vergleichsweise geringe Zahl an Verteilungseinheiten, die einfachen Allokationsüberlegungen sowie geringe Kommunikationskosten bei häufigen, auf einzelne Relationen begrenzten Operationen. Ebenso sind Integritätsprüfungen innerhalb einer Relation, z. B. die Schlüsselbedingung oder interne Abhängigkeiten leicht durchzuführen (vgl. [Ra94], S. 60).

Andererseits kann man durch feinere Fragmentierung häufig eine bessere Lastbalancierung, also eine gleichmäßigere Auslastung aller beteiligten Rechner erreichen. Außerdem können nur hierdurch geographische Gesichtspunkte sinnvoll berücksichtigt werden. So könnte man bei obigem Ansatz in einem über mehrere Orte mit je einem Rechnerknoten ausgestatteten verstreuten Unternehmen die Kundendaten nur komplett an jedem beliebigem Knoten des Netzes, bei gewünschten Replikaten auch an mehreren, halten. Dies wäre mit hohem Kommunikationsaufwand zur Verwaltung der Replikate oder für den entfernten Datenzugriff verbunden. Eine feinere Fragmentierung hingegen ermöglicht die Aufspaltung der Relation „Kunde“ nach Filialen und die dortige Allokation. Wenn die Fragmentierung zudem so gewählt ist, daß sie eine häufig vorkommende Selektionsbedingung darstellt, dann kann sie auch für eine Reduzierung des Verarbeitungsumfangs genutzt werden. Und schließlich kann durch feinere Fragmentierung auch die Parallelverarbeitung gefördert werden. So wäre es vorstellbar, daß bei einer bundesweiten Stellenangebotssuche in einer nach Regionen fragmentierten und gespeicherten Datenbank zeitgleich, regional verstreute Suchen stattfinden und trotz des Zeitaufwandes für die Kommunikation zu einer geringeren Bearbeitungszeit führen als im nicht oder anders fragmentierten (z. B. nach der Art der Stelle) Falle.

Eine Relation kann horizontal oder vertikal fragmentiert werden. Die horizontale Fragmentierung entspricht einer zeilenweisen disjunkten Aufteilung. Sie läßt sich im Relationenkalkül durch Selektionsbedingungen (ggf. verknüpften) beschreiben (vgl. Abbildung 4). Die Vereinigung der disjunkten Fragmente muß die ursprüngliche Relation ergeben. Die horizontale Fragmentierung ist insbesondere zur Unterstützung paralleler Anfragen geeignet (s. o.). Darüber hinaus existiert noch die abgeleitete Fragmentierung. Hierbei wird eine Relation R horizontal fragmentiert, die Einteilung hängt aber von Werten einer anderen Relation S ab. Dies entspricht in Relationenalgebra einem Semi-Join $R \text{ SJ}_{[A\theta B]} S$, wobei A und B Attribute darstellen und θ einen Vergleichsoperator. Die vertikale Fragmentierung basiert auf einer spaltenweisen Zerlegung der Relation und wird daher im Relationenkalkül durch Projektionen auf Attribute beschrieben (vgl. Abbildung 4). Hier muß jedoch, um die ursprüngliche Relation über einen natürlichen Verbund wieder herstellen zu können, jedem Fragment ein Schlüssel, normalerweise wird der Primärschlüssel gewählt¹⁷,

¹⁷ Die Wahl des Primärschlüssels ist nicht zwingend. Besser wäre es zu fordern, daß ein eindeutiger, möglichst kurzer Schlüssel zu wählen ist, denn diese beiden Eigenschaften sind wesentlich. So ist ein kurzer Schlüssel vorteilhaft, da er mehrfach abgelegt werden muß. Denkbar wäre also etwa eine durch das System vergebene Nummer. Außerdem

mitgegeben werden. Die Disjunktheit der Zerlegung bezieht sich in diesem Falle daher um der Rekonstruktion willen nur auf solche Attribute, die nicht Teil des Schlüssels sind, der zur Bildung des Natürlichen Verbundes in mehreren Fragmenten vorkommt. Ferner sei noch darauf hingewiesen, daß auch Mischformen aus den vorgestellten Fragmentierungsarten auftreten können, z. B. wenn man, wie auch in Abbildung 4 erfolgt, Fragmente erneut auf abweichende Weise zergliedert.

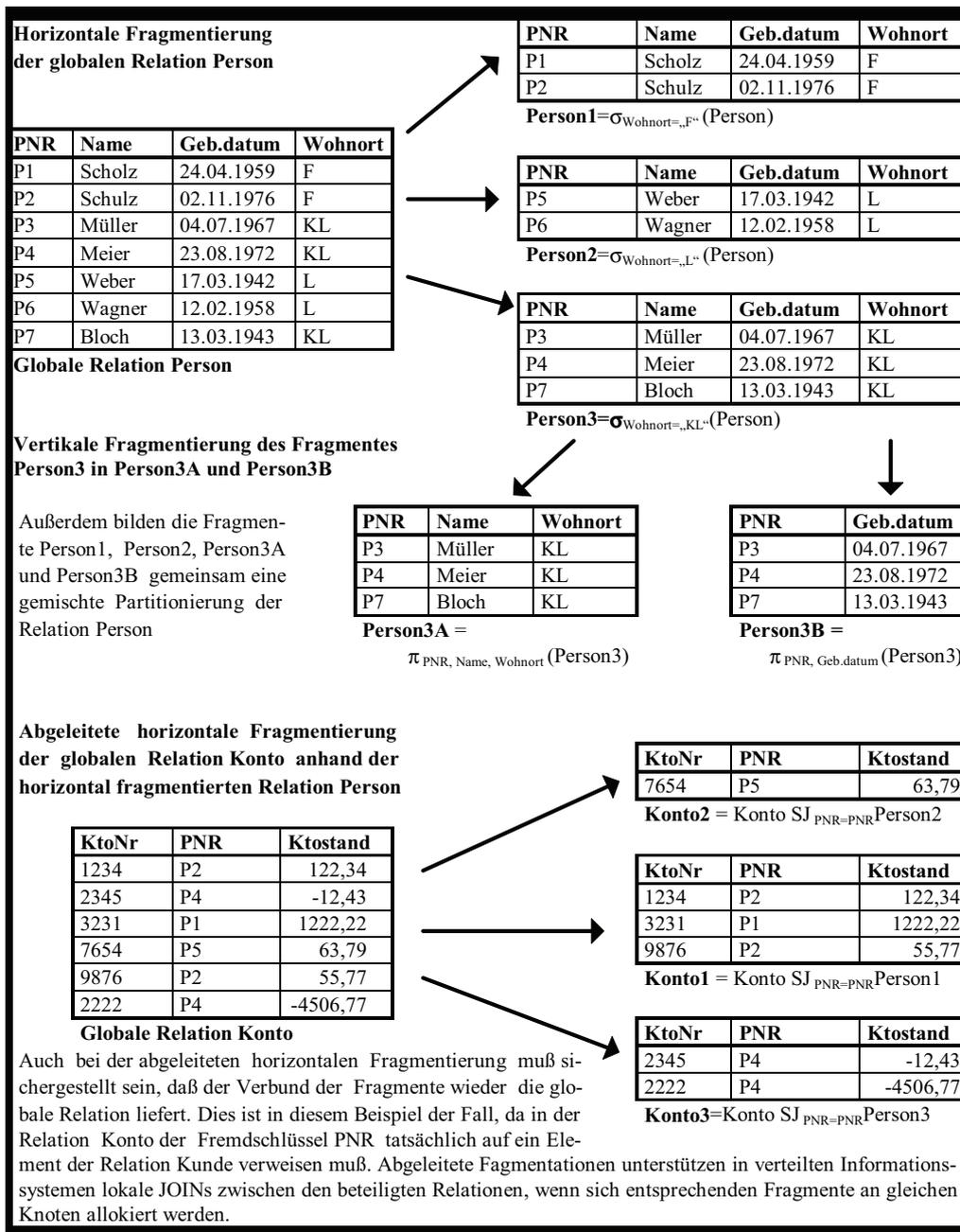


Abbildung 4: Fragmentierungsarten (vgl. Rahm S. 63, 64)

Kennzeichnend für die gerade beschriebenen, sogenannten **statischen Verteilungsformen** ist, daß sie einmalig für die Datenbank entworfen werden und eine Vorstufe zur Allokationsplanung darstellen. Anders hingegen arbeiten dynamische Verteilungsstrategien.

ist die Eindeutigkeit wesentlich, damit keine weiteren Zuordnungen anfallen. Nur dieser Aspekt ist bei der Festlegung auf den Primärschlüssel automatisch (per Definition) gewährleistet.

„Eine dynamische Verteilung wird durch Zugriffsstatistiken gesteuert:

Jedes Datum (oder zumindest jeder vordefinierte Teil der Datenbasis) wandert an den Knoten, von dem aus es zuletzt verlangt wurde. Wenn ein ausgeprägtes Lokalitätsverhalten vorausgesetzt wird, finden im eingeschwungenen Zustand kaum noch Datenwanderungen statt. Kopien von Daten werden vom System automatisch erzeugt, wenn die Zugriffshäufigkeit eine vorgegebene Schwelle übersteigt; sie können ebenso wieder aus dem Verkehr gezogen werden.

Als größter Nachteil der dynamischen Verteilung erweist sich aber, daß auf keinem Knoten der aktuelle Zustand des gesamten Systems bekannt ist; Daten, die nicht lokal verfügbar sind, müssen im ganzen Netz gesucht werden. Ein solcher Suchvorgang kann aber zu unververtretbaren Verzögerungszeiten führen.“ ([OB90], S. 59)

Somit umfaßt die dynamische Verteilung zugleich Fragmentierungs- und Allokationsüberlegungen.

2.3.2 Allokation

Bei statischen Verteilungsformen setzen die Überlegungen zur Allokation im allgemeinen auf den Fragmentierungsentscheidungen auf. Es wird anhand der vordringlich gewünschten Eigenschaften entschieden, von welchen Fragmenten Replikate gebildet und wo welche Fragmente und Replikate gehalten werden sollen. Dem Nutzer soll aber das Vorhandensein von Replikaten verborgen bleiben (Replikattransparenz). Insbesondere muß das vDBMS sicherstellen, daß eine Information, die zu einem repliziert gehaltenen Datum herausgegeben wird, unabhängig von dem gewählten Replikat ist (Konsistenzbedingung).

Für das Anlegen von Replikaten sind vielfältige Formen denkbar. Sie bewegen sich zwischen den Extremen einer **voll replizierten** Datenbank (die Datenbank liegt an allen Knoten in kompletter Kopie vor) und einer **partitionierten** Datenbank (jedes Fragment ist an genau einem Knoten gespeichert). Damit verbundene Eigenschaften sind in der folgenden, aus [ÖV91] entnommenen Tabelle dargestellt (S. 104).

	Full replication	Partial replication	Partitioning
QUERY PROCESSING	Easy	← Same difficulty →	
DIRECTORY MANAGEMENT	Easy or nonexistent	← Same difficulty →	
CONCURRENCY CONTROL	Moderate	Difficult	Easy
RELIABILITY	Very high	high	Low
REALITY	Possible application	Realistic	Possible application

Tabelle 3: Allokationsmöglichkeiten und ihre Handhabung, (aus: [ÖV91], S. 104)

Die Bestimmung einer Datenallokation in verteilten Systemen ist ein intensiv untersuchtes Optimierungsproblem. Rahm führt hierzu aus: „Das Allokationsproblem wurde meist für die Zuordnung ganzer Dateien betrachtet (file allocation problem, FAP), die entwickelten Ansätze lassen sich oft jedoch auch zur Allokation von Fragmenten heranziehen. Das Optimierungsziel ist i. a. die Minimierung einer globalen Kostenfunktion unter Berücksichtigung bestimmter Randbedingungen. Bei der Ausgestaltung von Kostenfunktion sowie den betrachteten Randbedingungen bestehen erhebliche Unterschiede zwischen den einzelnen Ansätzen. Im einfachsten Fall wird lediglich eine Minimierung von Kommunikationsvorgängen durch Unterstützung von Lokalität angestrebt; komplexere Modelle berücksichtigen daneben auch die Speicherkosten sowie lokale Verarbeitungskosten. Mögliche Randbedingungen sind die Beachtung knotenspezifischer Obergrenzen hinsichtlich der Speicher- oder Verarbeitungskapazität, um z. B. eine Lastbalancierung zu unterstützen.“ ([Ra96], S. 72)

Kapitel 3 wird sich mit einigen derartigen Lösungsansätzen befassen; darüber hinaus auch mit einem, von den restlichen etwas abweichenden Lösungsvorschlag, bei dem der Begriff der Schedule eine wesentliche Rolle spielt. Deshalb und weil sie im Kontext anfallender Datenübertragungen in verteilten Datenbanken Bedeutung besitzen, soll im folgenden noch ein etwas genauerer Blick auf Anfragebearbeitung, Synchronisationsverfahren und Kopienhandhabung geworfen werden.

2.3.3 Anfragebearbeitung, Synchronisationsverfahren und Kopienhandhabung

Durch Datenbankmanagementsysteme ist eine bequeme Nutzung von Datenbeständen möglich geworden, denn sie übernehmen neben der Verwaltung der Daten die Ausführung typischer Anwenderwünsche, wie etwa eine flexible, zielgerichtete und angemessen schnelle Auswertung oder Anpassung (Update) der Daten und dies zudem im allgemeinen in Mehrbenutzerumgebungen. Derartige Anwenderwünsche nennt man auch **Datenbankanfragen**. Dabei umfaßt der Begriff „Anfrage“ durchaus lesende und schreibende Datenzugriffswünsche. Zur Formulierung stellen moderne DBMS den Anwendern spezielle Datenbankanfragesprachen oder Oberflächen zur Verfügung (vgl. Abschnitt 2.2). Das DBMS wandelt die Datenbankanfragen in Folgen von Schreib- und Leseoperationen auf dem Datenbestand um; die genau gewünschten Daten sind über die oben beschriebenen Mengenoperationen identifizierbar. Zur Formulierung der jeweiligen Terme sind die tatsächlich gespeicherten Relationen und ihre Allokation zu berücksichtigen. Doch ist eine solche Transformation nicht eindeutig festgelegt. Ein Grund hierfür ist, daß sich Mengen auf unterschiedliche aber dennoch äquivalente Weise identifizieren lassen. Das Abarbeiten äquivalenter Terme kann aber mit einem sehr unterschiedlichen Aufwand verbunden sein. So kann etwa das Entfernen redundanter Teilterme deutliche Vorteile mit sich bringen. Auch das Vorziehen von Selektionen oder anderen das Ergebnis verkleinernden Operationen kann die Anfragebearbeitung verbessern. Ebenso kann es sein, daß mit einer bestimmten Ausführungsreihenfolge vorhandene Indexe besonders gut genutzt werden können.

Im Idealfall sollte sich der Anwender weder um die gespeicherten Relationen noch um deren Allokation noch um Optimierungsaspekte Gedanken machen müssen, d. h. die Art und Weise, wie er seine Frage in der Datenbankanfragesprache formuliert hat, sollte für die Ausführung der Anfrage irrelevant sein. Dies setzt voraus, daß das DBMS in der Lage ist, die gestellte Anfrage anhand der gerade genannten Kriterien umzuwandeln, d. h. eine sogenannte **automatische Anfragebearbeitung** und **Anfrageoptimierung** durchzuführen. In verteilten Datenbanken sind bei der Anfrageoptimierung mehr Aspekte einzubeziehen als in zentralen Datenbanken. Unter Berücksichtigung der vorliegenden Fragmente und ihrer Allokation sind z. B. bei Lesevorgängen in partiell oder total replizierten Datenbanken die Replikate zu ermitteln, die am günstigsten zu erreichen sind. Ganz allgemein ist in verteilten Systemen unter Berücksichtigung der vordringlichsten Ziele herauszufinden, bei welcher Anfragetransformation die geringsten Datentransferkosten oder geringsten Übertragungszeiten erzielt werden können. Bei Anfragen, die Informationen aus globalen Relationen, die auf mehrere Knoten verteilt sind, für ihre Beantwortung benötigen, kann es große Unterschiede geben, je nachdem, ob man zunächst lokale Lösungen ermittelt und diese dann zusammenführt oder aber zunächst eine lokale Zusammenführung der benötigten Daten herstellt und dann weiter selektiert. Ebenso gibt es verschiedene Möglichkeiten, Verbunde zwischen Daten verschiedener Rechnerknoten herzustellen, die je nach Art der beteiligten Fragmente Vor- oder Nachteile besitzen (vgl. [Da96], Kapitel 6).

Unter dem Begriff der **Synchronisation** faßt man schließlich die notwendigen Schritte zusammen, die das korrekte Ausführen der Anwenderwünsche nach den ACID-Eigenschaften (Atomarität, Konsistenz, Isolation, Dauerhaftigkeit, vgl. S. 17) sicherstellen. Dabei kann es in verteilten Datenbanken Transaktionen geben, die rein lokaler Natur sind, d. h. zu deren Ausführung lediglich Operationen an einem Rechner des Netzes erforderlich sind. Andere, sogenannte globale Transaktionen, erfordern hingegen Zugriffe auf mehr als einen Rechner. Das kann sowohl damit zu tun haben, daß die beteiligten Daten nur durch Zugriff auf mehrere Rechnerknoten vorteilhaft oder überhaupt erst

zu erreichen sind oder aber, daß die lokal möglichen Operationen Auswirkungen auch auf Kopien an anderen Rechnern haben. Synchrone Updates stellen ein Hauptproblem verteilter Datenbanken dar, denn wenn der Anwender ein Single-System-Image wünscht, so muß sichergestellt werden, daß man auf die Datenbank stets zu einem Zeitpunkt dieselbe Sicht hat, egal, an welchem Knoten und auf welches Replikat man auf sie zugreift¹⁸. Es gibt aber auch für den verteilten Fall Lösungen für das Sicherstellen der ACID-Eigenschaften, sie sind allerdings etwas komplizierter als im zentralen Fall. Zum einen erfordern sie Koordinationsvorgänge zwischen den Knoten, zum anderen müssen sie allgemeinerer Natur sein, um ein Weiterarbeiten an der Datenbank auch bei lokalen Ausfällen von Rechnern oder Datenübertragungswegen auf den noch funktionsfähigen bzw. erreichbaren Rechnern sicherzustellen. Und schließlich muß auch im Falle der Wiederinbetriebnahme eines ausgefallen bzw. eines zeitweise nicht mehr erreichten Rechners für das Nachziehen der dort vorliegenden, in der Zwischenzeit von Transaktionen betroffenen Daten gesorgt sein.

Im Kontext der in Kapitel 3 vorgestellten Allokationsmodelle soll im folgenden auf Lösungsansätze zur Synchronisation und Kopienanpassung, die sogenannten Replikationsverfahren, eingegangen werden, denn hier gibt es eine Vielzahl von Anlässen, die einen Datenaustausch zwischen den verschiedenen Rechnern des Netzes erforderlich machen und die zu zusätzlichen Netzbelastungen und zur Erhöhung der Antwortzeit sowie zur Speicherung zusätzlicher Informationen führen können. Da diese Themen gänzlich eigene Schwerpunkte in der Literatur verteilter Datenbanken darstellen, kann hier nur eine knappe Behandlung erfolgen. Diese basiert vor allem auf Darstellungen in [Da96]¹⁹. Nur in Verbindung mit einer Berücksichtigung dieser Vorgänge kann man die Lösungsvorschläge von Apers verstehen, der von einer gegenseitigen Abhängigkeit von Allokation und Anfragebearbeitung ausgeht und bei dem der Begriff der Schedule, der unten erläutert wird, eine zentrale Rolle spielt. Zum anderen dienen diese Ausführungen dem Verständnis von vorhandenen Grenzen der später vorgestellten mathematischen Modellbetrachtungen.

Synchronisation

Das Bilden optimaler Folgen von auf dem Datenbestand ausführbaren Schreib- und Leseoperationen zu jeder Anfrage allein reicht noch nicht aus, um ein Korrektheit erhaltendes Arbeiten auf dem Datenbestand sicherzustellen. Um die ACID-Eigenschaften zu gewährleisten, müssen diese Folgen zudem noch in Transaktionen gegliedert werden, die die Datenbank, vollständig ausgeführt, jeweils von einem korrekten Zustand in den nächsten überführen. Gemeinsam an einem Knoten abgewickelte Vorgänge müssen in einer gemeinsamen Abarbeitungsfolge, der sogenannten **Schedule** Platz finden. Hierbei dürfen durchaus Operationen unterschiedlicher Transaktionen miteinander verzahnt vorkommen. Ein stets sequentielles Abarbeiten der erforderlichen Transaktionen würde im allgemeinen zu langsam sein. Durch die Verzahnung wird es aber erforderlich, Zugriffe auf Daten, die von noch nicht vollständig abgewickelten Transaktionen verändert wurden, zu verhindern. Ansonsten würde ein eventuell zur Konsistenzerhaltung erforderliches, isoliertes Zurücksetzen unmöglich. Eine korrekte Ausführung der Transaktionen kann immer dann erreicht werden, wenn die ausgeführte Schedule **serialisierbar** ist. D. h. es muß eine zu dem Ergebnis der tatsächlich ausgeführten Schedule äquivalente, serielle Abarbeitungsreihenfolge der beteiligten Transaktionen geben.

Im Unterschied zu zentralen Datenbanken muß man bei verteilten Datenbanken für die globalen Transaktionen, an deren Ausführung mehrere Knoten beteiligt sind, noch ein weiteres Se-

¹⁸ In manchen Entwürfen wird dem Datendurchsatz ein so hoher Vorrang eingeräumt, daß vorübergehende, sich in Grenzen haltende, Inkonsistenzen in Kauf genommen werden (vgl. [DA96], S. 185).

¹⁹ Hier findet man ausführlichere Darstellungen, zahlreiche Beispiele und weiterführende Literaturangaben.

rialisierbarkeitskriterium verlangen. Globale Transaktionen können in einem verteilten System an einem beliebigen Knoten, ihrem sogenannten Primärknoten starten. Am Primärknoten einer Transaktion werden sowohl die lokal ausführbaren Operationen gebildet und in die lokale Schedule eingearbeitet als auch die Vorgänge an den weiter erforderlichen Knoten, nämlich das Starten der sogenannten Subtransaktionen, ausgelöst sowie die Teilergebnisse oder Rückmeldungen dieser Subtransaktionen koordiniert²⁰. Darüber hinaus wird am Primärknoten eine eventuell erforderliche Fehlerbehandlung durchgeführt und die globale Transaktion protokolliert. Trotz serialisierbarer lokaler Schedules kann es bei derartigen Transaktionen zu knotenübergreifenden Konflikten kommen, so z. B. wenn zwei Transaktionen an unterschiedlichen Knoten Objekte blockieren und gegenseitig auf die Freigabe dieser Objekte warten. Derartige Situationen kann man mit der Forderung verhindern, daß es möglich sein muß, aus der Menge der zu den globalen Transaktionen gehörenden äquivalenten, seriellen Schedules eine widerspruchsfreie serielle Ausführungsreihenfolge für die globalen Transaktionen abzuleiten (vgl. ebenda, S. 198).

Für das Erreichen sowohl der lokalen als auch der globalen Serialisierbarkeitsbedingungen gibt es unterschiedliche Verfahren, auf die im kommenden Absatz noch eingegangen wird. Allen gemeinsam ist jedoch, daß neben den eigentlichen Schreib- und Leseoperationen auch Vorgänge zur Festlegung des Beginns und des Endes einer Transaktion oder einer Subtransaktion, je nach Verfahren auch zum Sperren und Freigeben von Datenobjekten für andere Transaktionen oder zum Zurücksetzen von Transaktionen erforderlich sind.

Ein gegenüber zentralen Verfahren erweitertes Transaktionskonzept, das sogenannte **Zwei-Phasen-Commit-Protokoll**, unterstützt in vielen verteilten Datenbanken die Einhaltung der oben geschilderten Forderungen. Es beinhaltet Regeln für den Kommunikationsaustausch zwischen den Knoten, die den Beginn, den Ausführungsstand und das Ende von Transaktionen und Subtransaktionen betreffen und somit das vollständige Ausführen von Transaktionen, ggf. ihre isolierte Rücksetzung oder aber auch ihr Nachziehen bei Rechnerausfällen oder Erreichbarkeitsproblemen von Rechnern sicherstellen. Dazu wird den beteiligten Rechnern jeweils der Beginn einer neuen Transaktion mitgeteilt (Start-Botschaft). Ferner fordert der Primärknoten nach der Initiierung aller erforderlichen Subtransaktionen dazu auf, das Commit vorzubereiten („Prepare-to-Commit“-Botschaft), d. h. die gemachten Änderungen zu sichern. Erst mit Vorliegen einer „Ready-to-Commit“-Botschaft von allen beteiligten Knoten sendet der Primärknoten wiederum an alle Knoten das „COMMIT“. Falls nur eine „Ready-to-Commit“-Botschaft nicht vorliegt, wird allen Knoten der Abbruch der Transaktion mitgeteilt und lokal sorgen die betroffenen Knoten für ein isoliertes Zurücksetzen der Transaktion und erneutes Freigeben der betroffenen Daten. Mit dem Protokoll alleine, von dem es natürlich auch Varianten gibt, ist aber weder festgelegt, wie die einzelnen Knoten die Operationen einer Schedule ausführen und deren Serialisierbarkeit sicherstellen, noch wie globale Serialisierbarkeit sichergestellt werden kann.

Eine Möglichkeit hierfür besteht in der Anwendung sogenannter **Optimistischer Synchronisationsverfahren**. Diese lesen ihre benötigten Objekte aus der Datenbank und halten Veränderungen zunächst auf privaten Kopien fest. Vor Abschluß einer (Sub-)Transaktion wird geprüft, ob konkurrierende Zugriffe stattgefunden haben (Validationsphase). In solchen Fehlerfällen sind die betroffenen Transaktionen zurückzunehmen, anderenfalls folgt die Schreibphase. „Im *verteilten* Fall muß die Freigabe der lokalen Änderungen, also der Eintritt in die Schreibphase mit dem Zwei-Phasen-Commit koordiniert werden, sonst besteht die Gefahr nicht-serialisierbarer globaler Schedules. ... Es muß sichergestellt sein, daß alle in Konflikt stehenden globalen Transaktionen lokal jeweils in exakt derselben Reihenfolge (erfolgreich) validieren.“ (ebenda, S. 220) Weil Konflikte erst

²⁰ Dies ist zumindest der Regelfall. Andere Organisationsformen sind aber auch möglich.

in der Validationsphase erkennbar sind, hat dies die Konsequenz, daß der Eintritt in die lokalen Validationsphasen über ein globales Zeitstempel- oder ähnliches Verfahren geregelt werden muß (vgl. ebenda). Auf Zeitstempelverfahren wird bei der Beschreibung der Kopienbehandlung noch eingegangen. Zur Ermittlung des Synchronisationsaufwandes in mit optimistischen Verfahren arbeitenden verteilten Systemen ist daher ein Blick auch auf den Aufwand dieser Kopienbehandlungsmethoden zu werfen.

Während die optimistischen Verfahren davon ausgehen, daß relativ selten Konfliktfälle auftreten und diese deshalb erst beheben, wenn sie aufgetreten sind, verhindern **pessimistische Verfahren** gemeinsame Datenzugriffe verschiedener Transaktionen, sofern diese Konflikte bergen²¹, bereits im voraus. Dazu werden für die Transaktionen Sperren auf die benötigten Daten gesetzt. Auch wenn hierdurch bei Einhaltung des **Zwei-Phasen-Sperrprotokolls**²² Serialisierbarkeit erreicht werden kann, können neue Probleme auftreten: Bereits lokal - aber auch über verschiedene Knoten hinweg - kann es hierbei zu Verklemmungen kommen, wenn zwei Transaktionen gegenseitig benötigte Daten gesperrt haben und beide auf die Freigabe der durch die andere Transaktion gesetzten Sperre warten. Ein solcher Zustand muß dem System zunächst erst einmal auffallen. Hierfür gibt es verschiedene Lösungsstrategien, zum einen Verfahren zur expliziten Analyse auf Verklemmungen, zum anderen die Beobachtung der Ausführungsdauer aktiver Transaktionen. Eine explizite Analyse auf Verklemmungen kann z. B. durch die Suche nach Zyklen in einem mitgeführten und bei Sperrverweigerung analysierten Transaktionsgraphen erfolgen. Bei der Beobachtung der Ausführungsdauer aktiver Transaktionen sorgt das Überschreiten von Zeitschranken dafür, daß das System von einer Verklemmung ausgeht. Deren Auflösung erfolgt bei allen Ansätzen durch das Zurücknehmen einer beteiligten Transaktion und ihrem erneuten Start.

Um genau auf diese Weise auch globale Verklemmungen zu vermeiden, müßten die beim Zwei-Phasen-Sperrprotokoll angefertigten Sperrtabellen zentral geführt werden. Dieses Vorgehen brächte aber gravierende Nachteile mit sich. Zum einen würde eine starke Abhängigkeit von dem betreffenden Knoten entstehen, zum anderen entstünden sehr umfangreiche zusätzliche Datenübertragungen, denn alle Transaktionen müßten für jeden neuen Datenzugriff Einträge in die Sperrtabelle vornehmen. Darüber hinaus ist auch das Festlegen von Zeitschranken problematisch, wenn Datenübertragungszeiten mit zu berücksichtigen sind. Damit können das Zwei-Phasen-Sperrprotokoll und die Verfahren zur Aufhebung von Verklemmungen in der beschriebenen Form nur lokal eingesetzt werden.

Globale Verklemmungen müssen anders verhindert werden. Die globale Serialisierbarkeit kann allerdings schon durch lokalen Einsatz des Zwei-Phasen-Sperrprotokolls an allen Rechnerknoten sichergestellt werden. Hier muß lediglich das Zwei-Phasen-Commit-Protokoll für alle Teiltransaktionen zur Erreichung des „Ready-to-Commit“-Zustandes einen gesperrten Zustand sämtlicher von der Teiltransaktion benötigter Objekte verlangen. Somit bleibt die Aufgabe, ohne eine zentrale Sperrtabelle globale Verklemmungen zu entdecken und aufzulösen. Eine Lösung besteht aus einem zentralen Analyserechner. Dieser kann mit weniger Informationen von den Netzknoten auskommen, als zur Aufstellung der Sperrtabelle erforderlich sind: er erhält von lokalen Auswerterroutinen nur Informationen, wenn knotenübergreifende Wartesituationen existieren. Die Abhängigkeit von einem Knoten wäre aber auch bei dieser Lösung ein Nachteil. Eine andere Möglichkeit ist es, globale Verklemmungen durch ein verteiltes Verfahren zu entdecken. Hierbei werden zunächst

²¹ Verhindert werden müssen gemeinsame schreibende sowie schreibend-lesende Zugriffe von Transaktionen. Ausschließlich gemeinsames Lesen ist unproblematisch.

²² Das Zwei-Phasen-Sperrprotokoll legt fest, daß eine Transaktion vor jedem Datenzugriff auf die benötigten Daten eine Sperre anfordern und setzen muß. Nach dem ersten Lösen einer Sperre darf keine Sperre mehr gesetzt werden. Dieses Vorgehen sichert die Serialisierbarkeit.

lokal Auswertungen der Transaktionen vorgenommen und für jede gefundene, mögliche globale Verklemmung wird jeweils eine Information zu einem der verwickelten Knoten übertragen. Durch Numerierung der Transaktionen wird zudem die Zahl der Übertragungen halbiert sowie die Gefahr, zu viele Transaktionen zurückzunehmen, reduziert, da nur die Verklemmungen weitergemeldet werden, bei denen eine Transaktion mit höherer Nummer auf eine mit niedrigerer Nummer wartet. Dennoch kann es durch das eigenständige Entscheiden der Knoten, welche Transaktion im Verklemmungsfalle zurückgesetzt wird, zu überflüssigen Rücksetzungen kommen.

Kopienbehandlung

Die einfachste Möglichkeit, alle Kopien eines Elementes zu behandeln, ist ihr automatisches, gleichzeitiges Update innerhalb der jeweiligen Transaktion. Dies hat den Vorteil, daß alle Kopien grundsätzlich auf aktuellem Stand und zum Lesen geeignet sind. Von Nachteil ist, daß die Transaktion nur erfolgreich abgeschlossen werden kann, wenn alle Kopien zum Update gesperrt sind. Schon der Ausfall oder das Nichterreichen eines Rechners führt dazu, daß die Transaktion scheitert, denn dieser kann den „Ready-to-Commit“-Zustand nicht beisteuern. Weil aber Kopien zumeist wegen der erhöhten Verfügbarkeit des Systems gehalten werden und dieses Vorgehen die Verfügbarkeit herabsetzt, wird dieses Verfahren so nicht eingesetzt.

Es gibt aber eine Vielzahl von Alternativen. Alle begnügen sich damit, die Transaktion bereits dann auszuführen, wenn eine Teilmenge sämtlicher Replikate synchron aktualisierbar ist. Die restlichen Updates müssen anschließend nachgezogen werden.

Sofern es sich bei der Teilmenge um eine **Primärkopie** handelt und diese in eines der oben beschriebenen Synchronisationsverfahren einbezogen ist, so ist die Serialisierbarkeit gesichert. Die Primärkopie ist anschließend für die Aktualisierung der restlichen Kopien verantwortlich. Für lesende Zugriffe wird hierbei im allgemeinen in Kauf genommen, daß bei einem Zugriff auf eine Kopie, die noch nicht nachgezogen wurde, ein leicht veralteter Zustand gesehen wird. Die Dauer der Transaktion hängt hier also ausschließlich vom Zugriff auf jeweils eine Kopie ab. Insbesondere, wenn hierdurch Übertragungszeiten entfallen, kann dies deutliche Vorteile mit sich bringen. Die Übertragungsmenge im Rahmen eines kompletten Updates bleibt jedoch unverändert.

Ohne Primärkopie ist die Serialisierbarkeit schwerer sicherzustellen. So muß beispielsweise gewährleistet werden, daß eine aktuelle Kopie als Grundlage herangezogen wird (man denke etwa an eine Operation, die einen Objektwert verdoppeln soll). Außerdem muß sichergestellt werden, daß Updates an allen Knoten in der richtigen Reihenfolge eingehen. Hierzu werden **Zeitstempelverfahren** eingesetzt. Jeder Transaktion wird ein global eindeutiger Zeitstempel mitgegeben, Datenobjekte werden über zwei Einträge sowohl für lesende als auch schreibende Zugriffe durch diese gekennzeichnet²³ und zwar so, daß sich bei Zugriff auf das Objekt immer der maximale Zeitstempel aller bisher auf das Objekt zugegriffenen Lese- bzw. Schreiboperationen durchsetzt. Ein Update ist nur dann sinnvoll, wenn der Zeitstempel der Transaktion größer ist als alle Objektzeitstempel der von ihr benötigten Objekte, sonst würde die Transaktion einen veralteten Wert setzen. Auch ein „Hellsehen“ von Objekten mit größerem Zeitstempel als ihn die Transaktion aufweist, kann hierüber verhindert werden. In Verbindung mit einer Phaseneinteilung in der zunächst gelesen, dann für ein Schreiben gesperrt und schließlich geschrieben wird, kann man so geeignet synchronisieren.

²³ Dies ist eine mögliche Variante.

Wie bereits zu Beginn dieses Abschnitts auf Seite 28 angeklungen, müssen sowohl bei den Synchronisationsverfahren als auch bei den Kopienbehandlungen Lösungen für Fehlerfälle, die durch Knotenausfall oder Netzausfall entstehen können, vorgesehen sein. Fehlerfälle sind jedoch nicht Grundlage der Kostenberechnungen in Allokationsmodellen. Es wird auf sie daher nicht weiter eingegangen.

3 Mathematische Modelle

Dieses Kapitel soll einen kleinen, allgemeinen Einblick in mathematische Modelle geben, die dem Zweck dienen, eine möglichst „günstige“ (statische²⁴) Verteilung eines Datenbestandes auf diverse Knoten eines Rechnernetzes bei zugelassener Redundanz zu finden. Dabei ist eine Allokation im mathematischen Sinne günstig, wenn eine aufgestellte Zielfunktion einen absoluten Optimalwert annimmt. Eine Zielfunktion beispielsweise, die die anfallenden Übertragungskosten verschiedener Szenarien beschreibt, wird nach einem Minimum hin untersucht. Ferner können über Nebenbedingungen weitere Aspekte in die Berechnung eingehen, z. B. die jeweils mögliche Speicherkapazität an jedem Knoten. Bei allen Modellen gehen Informationen über das Netz sowie über die stattfindenden Operationen auf dem Datenbestand ein. Diese müssen daher im voraus bekannt oder zumindest abschätzbar sein. Bei den Verfahren unterscheidet man solche, die durch mehr oder weniger vollständiges Ausprobieren sämtlicher Möglichkeiten zur optimalen Lösung finden als auch sogenannte heuristische Ansätze, die dieses Vorgehen aus Aufwandsgründen vermeiden. Ihr Ziel ist es, zumindest zu einer möglichst optimalen Lösung zu führen. Ein von Ceri und Pelagatti vorgeschlagenes Vorgehen, das dieser Klasse angehört, geht beispielsweise von der besten replikatfreien Allokation aus und fügt schrittweise Kopien hinzu bis keine Verbesserung mehr erzielt werden kann.

Die Modellbetrachtungen der ersten drei herangezogenen Literaturgrundlagen ähneln einander, sind aber von zunehmender Komplexität. Ausgangspunkt ist bei ihnen jeweils eine vorliegende Fragmentierung und natürlich eine Menge von Knoten, auf die die Verteilung erfolgen soll. Ebenso erfordern sie Wissen über die Anwendungen, die auf den Daten stattfinden sollen. Einen Sonderfall nimmt der Ansatz zur vertikalen Fragmentierung von Ceri und Pelagatti ein. Hier ist der Fragmentierungsvorgang noch nicht abgeschlossen, sondern wird ggf. durch vertikale Fragmentierung noch verbessert. Abschließend werden zwei auf einem graphischen Modell basierende Lösungsansätze von Apers vorgestellt, die sich noch deutlicher von den vorangegangenen unterscheiden. Aufgrund des Graphenmodells wird dieser Ansatz in der Literatur auch als „graphbasierter Lösungsansatz“ bezeichnet. Die graphische Betrachtungsweise unterstützt sicherlich das Verständnis für seinen Ansatz, stellt aber aus meiner Sicht nicht den wesentlichen Unterschied zu den vorangegangenen Modellen dar²⁵. Die eigentliche Abgrenzung liegt in anderen Aspekten: So geht Apers überhaupt nicht von einer vorliegenden Fragmentierungsentscheidung aus, sondern ermittelt horizontale und vertikale Fragmentierung und Allokation gemeinsam. Charakteristisch für sein Vorgehen ist, daß er sich für die Ermittlung der optimalen Allokation stufenweise von einem Anfangszustand auf die Lösung hin bewegt. Dabei berücksichtigt sein Verfahren die Vorgehensweise des DBMS bei der Umwandlung der Anforderungen in die zugehörigen Transaktions- und schließlich in die Operationsfolgen. In einem seiner Lösungsansätze wird zudem die gegenseitige Abhängigkeit der Anfragebearbeitung durch das DBMS und der Allokationsentscheidungen mit einbezogen.

Zur Erleichterung eines Vergleichs der verschiedenen Modelle sollen vor deren Beschreibung gemeinsam benötigte Variablen einheitlich bezeichnet werden. Um wiederholtes Rückblättern überflüssig zu machen, befindet sich zudem unter den Anlagen eine herausnehmbare Übersicht aller in diesem Kapitel verwendeten Bezeichner!

²⁴ Grund für diese Beschränkung ist, daß auch für das Anwendungsbeispiel ausschließlich statische Betrachtungen eine Rolle spielen.

²⁵ Dies ist auch der Grund dafür, daß sein Modellansatz in der Gliederung keine besondere Position einnimmt.

K_i	Knoten i	$\#K$	Anzahl Knoten
		i, j	Knotenindexvariablen
P_n	Fragment (Partition) n	$\#P$	Anzahl von zu allozierenden Fragmenten (disjunkt) der globalen Relationen, entspricht der Kardinalität von PS
PS	Menge der Fragmente	n	Indexvariable für Fragmente
t	Indexvariable für den Typ einer Datenbankanforderung	$\#T$	Anzahl an Typen von Datenbankanforderungen

- M_i maximale Speicherkapazität in Dateneinheiten am Knoten K_i ($i = 1, \dots, \#K$)
- S_i Speicherkosten pro Dateneinheit am Knoten K_i ($i = 1, \dots, \#K$)
- $U_{\#K, \#K}$ Matrix der Übertragungskosten; jedes Element u_{ij} der Matrix $U_{\#K, \#K}$ liefert die Übertragungskosten je Übertragungseinheit von Knoten K_i nach Knoten K_j
- $V_{\#P, \#K}$ Indikatormatrix zur Verteilung der Fragmente auf die Knoten

Für alle Elemente v_{ni} der Matrix $V_{\#P, \#K}$ gilt: $v_{ni} = \begin{cases} 1, & \text{falls } P_n \text{ an } K_i \text{ alloziert ist} \\ 0 & \text{sonst} \end{cases}$

Eine weitere Erleichterung beim Lesen soll durch eine einheitliche Darstellung der Summenbildung bewirkt werden. An allen Stellen, an denen dies sinnvoll möglich war, ist hier durch die Wahl der Variante A deutlich gemacht worden, welche Werte die betreffende Variable innerhalb der natürlichen Zahlen durchläuft. Einige Autoren haben statt dessen lediglich die Laufvariable unter das Summenzeichen gesetzt (Variante C), oder die Schreibweise aus Variante B bevorzugt.

$\sum_{i=1}^{\#K} M_i$	$\sum_{i \in I} M_i$	$\sum_i M_i$
A	B	C

Abbildung 5:
Summenschreibweisen

Man beachte jedoch, daß diese Wahl keine Festlegung in Hinsicht auf die algorithmische Abarbeitung der Formeln darstellen soll. Da nur endliche Summen betrachtet werden, sind Umordnungen der Summanden ohne Beeinträchtigung des Ergebnisses möglich. Diesem Aspekt hätte man an sich durch die Darstellungsweise B mit exakter Angabe der jeweiligen Menge, im Beispiel durch $I = \{1, \dots, \#K\}$ am besten Rechnung getragen. Dagegen sprachen Platzgründe.

3.1 Modelle von Ceri und Pelagatti

Auf die von Ceri und Pelagatti 1984 veröffentlichten Allokationsmodelle ([CP84], S. 82 ff.) wird in der entsprechenden Literatur oft verwiesen. Auch bei den Veröffentlichungen von Dadam sowie Özsü und Valduries, die anschließend vorgestellt werden, haben sie Berücksichtigung gefunden. Das hauptsächliche Anliegen der Autoren ist es, für diverse Situationen Formeln aufzustellen, die einen möglicherweise vorhandenen Vorteil für eine angedachte Allokation widerspiegeln können. In diesem Bemühen unterscheiden sie Formeln für horizontale und vertikale Fragmentierung. Für die horizontale geben sie zudem für den nicht redundanten Fall eine gesonderte, besonders einfache Formel an, mit der in diesem Kapitel begonnen werden soll. Danach folgen Betrachtungen für redundante Datenhaltung.

Sowohl die für den partitionierten Fall als auch die erste für den redundanten Fall entwickelte Formel sollen genutzt werden, um alle möglichen Allokationen durchzurechnen und basierend auf diesen Ergebnissen eine Entscheidung zu finden. Im partitionierten Fall wird für jedes Fragment der Knoten ermittelt, an dem am meisten Zugriffsaufwand entsteht und dieser für die Allokation ausgewählt. Dieses Vorgehen wird dann auf den Fall zugelassener Redundanz übertragen. Hier wird zu jedem Fragment jeder Knoten ermittelt, bei dem der Zugriff auf eine Kopie an diesem Knoten mehr Vor- als Nachteile mit sich bringt. Das Fragment ist dann an allen diesen Knoten zu speichern. Allerdings weisen die Autoren auf die Problematik der Quantifizierung der Vor- und Nachteile für jede einzelne Fragment-Allokation hin, da hier gegenseitige Effekte, die durch das Allokieren verschiedener Kopien desselben Fragmentes auftreten können, unberücksichtigt bleiben. Die dritte Formel schließlich dient der Anwendung in einem heuristischen Ansatz. Neu berücksichtigt wird hier, daß das Vorhandensein von ein bis zwei Kopien eines Fragmentes einen weit größeren Vorteil mit sich bringt, z. B. in Hinblick auf Ausfallsicherheitsaspekte, als die Hinzunahme noch weiterer Kopien. Ausgehend von der optimalen partitionierten Lösung soll so zu einer insgesamt möglichst optimalen Lösung gefunden werden.

Die Ansätze zur horizontalen Fragmentierung setzen voraus, daß die Fragmentierungsentscheidung bereits vorliegt und nicht mehr verändert wird. Diese Eigenschaft teilen sie mit den Modellen der Abschnitte 3.2 und 3.3. Demgegenüber wird bei der von Ceri und Pelagatti zur vertikalen Fragmentierung vorgestellten Formel die Fragmentierung der Ausgangssituation ggf. noch verfeinert.

Um die von Ceri und Pelagatti definierten Formeln vorstellen zu können, müssen noch eine Reihe weiterer Variablenbezeichner eingeführt werden:

$\#B_{ni}$ Anzahl lokaler Zugriffe auf das n-te Fragment P_n am Knoten K_i

$\#F_{ti}$ Absolute Häufigkeit einer Anforderung vom Typ t am Knoten K_i

$\#\Phi_{r_{tn}}$ Anzahl der Lesezugriffe der Anforderung vom Typ t auf das Fragment n

$\#\Phi_{w_{tn}}$ Anzahl der schreibenden Zugriffe der Anforderung vom Typ t auf das Fragment n

$\#Z_{tn}$ Anzahl der Zugriffe der Anforderung vom Typ t auf das Fragment n

B_{ni} Maß zur Beurteilung des Vorteils durch eine Allokation von Fragment P_n am Knoten K_i (B für Benefit)

$B_{nr \rightarrow sv}$ Maß zur Beurteilung des Vorteils durch vertikale Aufspaltung von Fragment P_n von Knoten K_r an K_s und K_v ; falls entweder $v = r$ oder $s = r$ gilt, wird dies im Index vermerkt

$\#P_n$ Redundanzgrad des Fragmentes P_n , der angibt, an wie vielen Knoten P_n allokiert ist

Bv_n Größe für den Vorteil durch volle Replikation des Fragmentes P_n an allen Knoten K_i

3.1.1 Allokation bei horizontaler Fragmentierung

Die erste von Ceri und Pelagatti vorgestellte Formel gilt für eine **partitionierte Datenbank** mit horizontaler Fragmentierung. Hier wird die Anzahl der lokalen Zugriffe B_{ni} auf das n -te Fragment P_n am Knoten K_i über das Auftreten aller Typen von Anforderungen an dem betreffenden Knoten wie folgt ermittelt:

$$\#B_{ni} = \sum_{t=1}^{\#T} \#F_{ti} \times (\#\Phi_{r_{tn}} + \#\Phi_{w_{tn}}) = \sum_{t=1}^{\#T} F_{ti} \#Z_{tn} \quad (\text{Best-Fit-Ansatz})$$

Man wird nun das Fragment P_n an dem Knoten K_{i^*} allokiert, für den $\#B_{ni^*}$ maximal ist, d. h. an dem Knoten, an dem die Zahl lesender und schreibender Zugriffe auf das Fragment am größten ist. In diesem Fall gilt daher, daß die Anzahl der lokalen Zugriffe auf das Fragment P_n am Knoten K_i selbst schon als Maß zur Beurteilung des Vorteils einer Allokation von Fragment P_n am Knoten K_i geeignet ist.

Sofern **Redundanz zugelassen** ist, wird das Fragment P_n an allen Knoten K_i gespeichert, an denen die von Anforderungen verursachten Kosten für lesende Zugriffe diejenigen übersteigen, die insgesamt durch alle an anderen Knoten auftretenden, schreibenden Anforderungen gegen dieses Fragment entstehen. Dabei ist für diese Betrachtung (und die Formel) der Speicherort des Fragments zunächst noch völlig unbedeutend. Wesentlich in der Betrachtung ist, von welchem Knoten aus eine Anforderung gegen dieses Fragment initiiert wird. Es gilt:

$$B_{ni} = \sum_{t=1}^{\#T} \#F_{ti} \#\Phi_{r_{tn}} - C \times \sum_{t=1}^{\#T} \sum_{\substack{i'=1 \\ i' \neq i}}^{\#K} \#F_{ti'} \#\Phi_{w_{tn}} \quad C \geq 1 \quad (\text{All-Beneficial-Sites-Ansatz})$$

Die erste Summe liefert die Zahl aller lesenden Zugriffe gegen das Fragment P_n am Knoten K_i . Die zweite Summe über t betrachtet über alle Anforderungen hinweg alle schreibenden Zugriffe an allen vom Knoten K_i verschiedenen Knoten gegen das Fragment P_n . C ist eine Konstante, die das Verhältnis zwischen Schreib- und Lesezugriffskosten ausdrückt. Man kann davon ausgehen, daß C mindestens den Wert 1 besitzt, denn typischerweise sind die Kosten für einen schreibenden Zugriff höher als für einen lesenden, da sie mehr Steuerinformationen²⁶ und Angaben zu lokalen Operationen beinhalten. Ferner reicht für Lesevorgänge häufig ein Cache-Zugriff, für persistente Schreibvorgänge hingegen nicht!

Das Fragment P_n wird nun an allen Knoten K_{i^*} allokiert, für die die Differenz B_{ni^*} positiv ist. Ist für keinen Knotenindex i der Wert der Differenz positiv, so wählt man das i^* aus, für das B_{ni^*} zumindest maximal ist. Replikate entstehen also an Knoten, an denen der Aufwand für lesende Zugriffe

²⁶ Z. B. zum Setzen von Schreibsperrern auf Original und Kopien (vgl. Abschnitt 2.3.3).

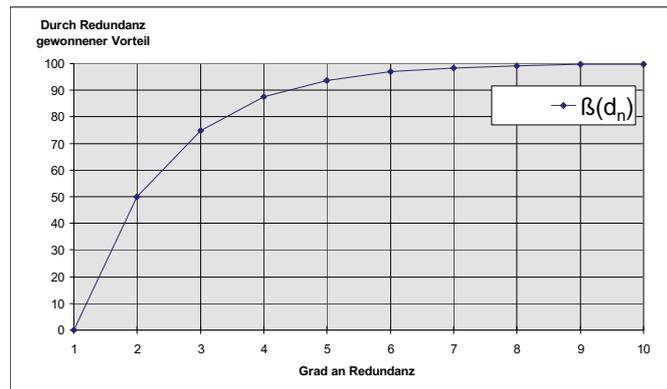
den für schreibende Zugriffe an allen anderen Knoten übersteigt. Somit berücksichtigt dieses Modell den Aspekt, daß Kopien im Falle schreibender Zugriffe nachgezogen werden müssen.

Die dritte Formel schließlich soll in Verbindung mit einer heuristischen Strategie, ausgehend von der optimalen partitionierten Lösung zum Einsatz kommen. Der Reihe nach soll für jedes Fragment geprüft werden, ob das Anlegen weiterer Kopien vorteilhaft ist. In die Formel, die diesen Vorteil ermittelt, gehen zwei zusätzliche Größen ein. Mit $\#P_n$ wird der Grad an Redundanz des Fragmentes P_n erfaßt. $\#P_n$ gibt an, an wie vielen Knoten P_n allokiert ist. B_{V_n} soll den Vorteil beziffern, der durch volle Replikation des Fragmentes P_n an allen Knoten K_i entsteht. Er muß als Eingabedatum in das Verfahren eingehen. Wesentlich für das Verständnis der in der Methode eingesetzten Formel ist der funktionale Zusammenhang dieser Größen, den die Autoren mit der Funktion β definieren.

Sie wächst für steigende Werte von $\#P_n$ immer schwächer:

$$\beta(\#P_n) = (1 - 2^{1-d_n})B_{V_n}.$$

Dabei gilt: $\beta(1) = 0,$
 $\beta(2) = \frac{1}{2} B_{V_n}$
 $\beta(3) = \frac{3}{4} B_{V_n}, \dots$



Mit Hilfe dieser Funktion läßt sich nun auch eine Formel aufstellen, bei der der zunehmend schwächere Vorteil durch das Hinzufügen weiterer Replikate berücksichtigt wird:

$$B_{ni} = \sum_{t=1}^{\#T} \#F_{ti} \# \Phi_{r_{tn}} - C \times \sum_{t=1}^{\#T} \sum_{\substack{i'=1 \\ i' \neq i}}^{\#K} \#F_{ti'} \# \Phi_{w_{tn}} + \beta(d_n) \quad C \geq 1$$

(Additional-Replication-Ansatz)

Verglichen mit der Formel vom „All-Beneficial-Sites“-Ansatz kommt hier noch der zuvor eingeführte Funktionswert für die gerade betrachtete Allokation als nicht negativer Summand hinzu. Dieser Summand hat mit zunehmendem Grad an Redundanz im Vergleich zum Zugriffsaufwand immer geringere Bedeutung.

3.1.2 Vertikale Fragmentierung

Ceri und Pelagatti befassen sich außerdem mit der Frage, welcher Vorteil entsteht, wenn man ein Fragment P_n , das am Knoten K_r gehalten wird, durch vertikale Fragmentierung in die zwei Fragmente P_s bzw. P_v an den Knoten K_s und K_v (von K_r verschieden) aufteilt. (Am Knoten K_r verschwindet P_n dann.)

Für das Aufstellen einer Formel zur Bewertung einer solchen Aufspaltung werden die Anforderungen in folgende Fälle unterschieden:

1. A_s und A_v umfassen die Anforderungen, die an den Knoten K_s bzw. K_v ausgeführt werden und die nach der vertikalen Fragmentierung nur noch lediglich Attribute von Fragmenten des betreffenden Knoten benötigen. Diese Anforderungen sparen einen entfernten Zugriff.
2. Die Menge A_1 umfaßt die Anforderungen, die zuvor nur lokal auf das Fragment am Knoten K_r zugegriffen haben und nun die betreffenden Attribute ausschließlich entweder von P_s oder von P_v benötigen. Durch diese Anforderungen entsteht jeweils ein zusätzlicher entfernter Zugriff zu dem betreffenden Knoten.
3. Die Menge A_2 umfaßt diejenigen Anforderungen, die zuvor nur lokal auf das Fragment am Knoten K_r zugegriffen haben und nun Attribute sowohl von P_s als auch von P_v benötigen. Durch diese Anforderungen entstehen nun jeweils zwei zusätzliche entfernte Zugriffe zu den beiden Knoten K_s und K_v .
4. Die Menge A_3 schließlich bestehe aus denjenigen Anforderungen, die an Knoten ausgeführt werden, die von K_r , K_s und K_v verschieden sind und die Attribute sowohl von P_s als auch von P_v benötigen. Während zuvor nur entfernt auf den Knoten K_r zugegriffen wurde, muß nun auf zwei Knoten entfernt zugegriffen werden. Daher entsteht hier jeweils ein zusätzlicher entfernter Zugriff.

Bevor die Formel zur Berechnung des möglichen Vorteils durch vertikale Fragmentierung aufgestellt wird, soll kurz überlegt werden, daß mit diesen von Ceri und Pelagatti aufgeführten Fällen tatsächlich alle Fälle, bei denen sich der Aufwand verändert, erfaßt sind.

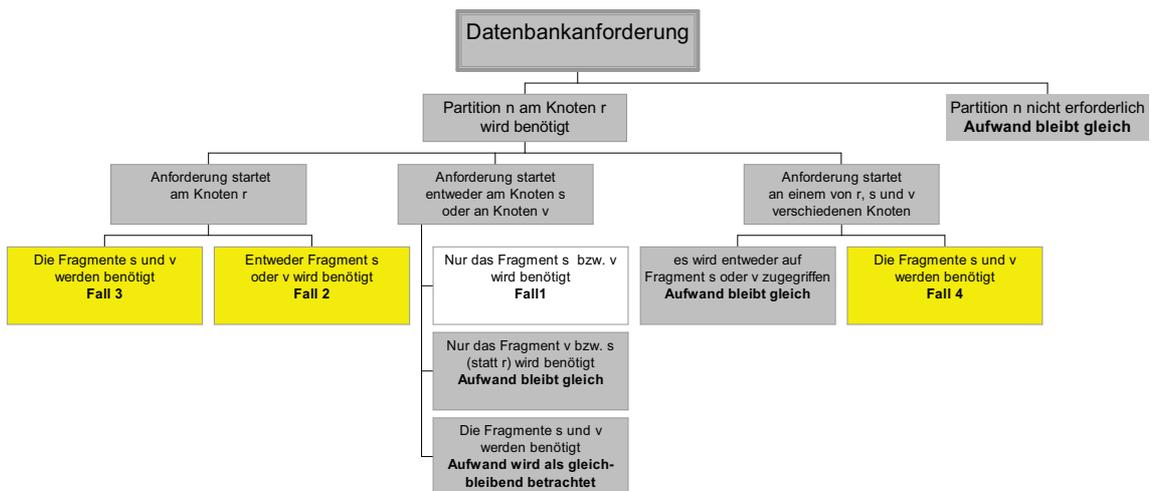


Abbildung 6: Fälle zu B_{nr-sv} bei vertikaler Fragmentierung nach Ceri und Pelagatti

Wie an der obigen Abbildung deutlich wird, verändert sich nur der Aufwand von Anforderungen, die auch Fragment P_n benötigen. Anforderungen, die von Knoten K_r ausgehen, führen zu den oben aufgeführten Fällen 2 und 3. Startet die Anfrage am Knoten K_s (analog für K_v), so ist sofort offensichtlich, daß sofern nur noch Attribute des betreffenden Knotens benötigt werden, der Aufwand um einen entfernten Zugriff abnimmt. Falls genau das andere Fragment benötigt wird, bleibt der Aufwand identisch, da lediglich der Knoten, auf den entfernt zugegriffen wird, wechselt. Werden beide Fragmente benötigt, so bleibt für die Anfrage ein entfernter Zugriff erforderlich, auch wenn er weniger Attribute erfaßt. Die Autoren behandeln diesen (je nach Attributmenge möglicherweise geringen) Vorteil allerdings nicht als solchen. Er ist daher in obigem Diagramm mit gleichbleibendem Aufwand eingetragen und schlägt sich auch in der folgenden Formel nicht nieder. Falls die Anforderung von einem von den Knoten K_r , K_s und K_v abweichendem Rechner kommt, so sind neben dem aufgeführten vierten Fall auch zwei Situationen denkbar, bei denen auch hier der Aufwand identisch bleibt, nämlich dann, wenn statt auf Knoten K_r nun lediglich entweder auf K_s oder K_t zugegriffen wird.

Nach diesen Betrachtungen folgt nun die Formel für den Vorteil (Benefit) der vertikalen Fragmentierung bei Aufspaltung in zwei Fragmente an die vom bisherigen Knoten K_r abweichenden Rechnerknoten K_s und K_v :

$$B_{nr-sv} = \sum_{t \in A_s} \#F_{ts} \#Z_{tn} + \sum_{t \in A_v} \#F_{tv} \#Z_{tn} - \sum_{t \in A_1} \#F_{tr} \#Z_{tn} - \sum_{t \in A_2} 2 \times F_{tr} \#Z_{tn} - \sum_{t \in A_3} \sum_{\substack{j=1 \\ j \notin \{r, s, v\}}}^{\#K} \#F_{tj} \#Z_{tn}$$

Für den Fall, daß nur ein Fragment an einem anderen Knoten allokiert werden soll, betrachten wir erneut eine Fallaufteilung. Ohne Beschränkung der Allgemeinheit wird davon ausgegangen, daß das zweite Fragment am Knoten K_s Platz finden soll, also Knoten K_r mit Knoten K_v zusammenfällt.

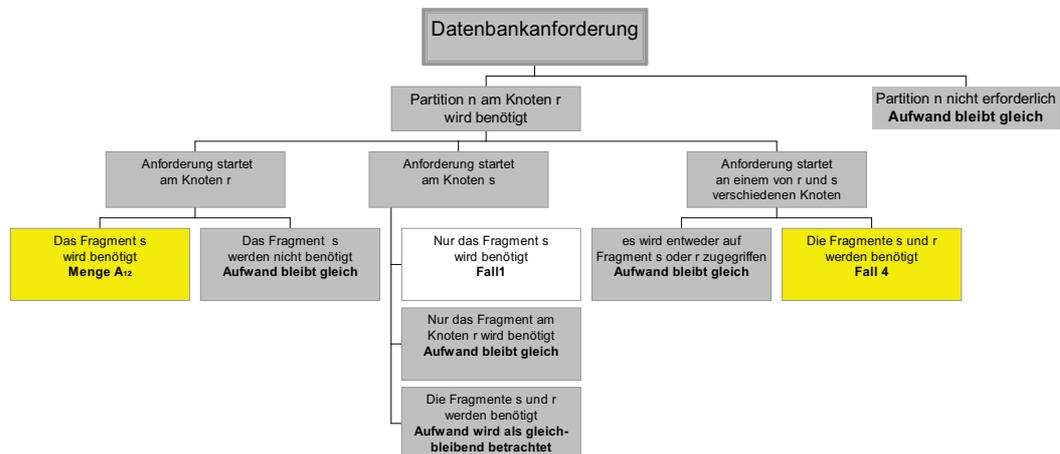


Abbildung 7: Fälle zu $B_{nr-sv=r}$ bei vertikaler Fragmentierung nach Ceri und Pelagatti

Geht man erneut alle Fälle durch, so stellt man fest, daß sich nur die Situation für Anforderungen, die von Knoten K_r ausgehen, verändert. Hier ist eine andere Menge, die Menge A_{12} , mit je einem zusätzlichen, entfernten Zugriff verbunden. Diese umfaßt Teile der vormaligen Menge A_1 und alle Fälle der vorherigen Menge A_2 . Entsprechend lautet die Formel dann für diesen Fall:

#K

$$B_{nr-sv=r} = \sum_{t \in A_s} \#F_{ts} \#Z_{tn} - \sum_{t \in A_{12}} \#F_{tr} \#Z_{tn} - \sum_{t \in A_3} \sum_{\substack{j=1, \\ j \notin \{r, s\}}} \#F_{tj} \#Z_{tn}$$

Der Einfachheit halber berücksichtigen diese Formeln nur die Zahl der Zugriffe und unterscheiden nicht zwischen lesendem und schreibendem Zugriff. Will man diesen Aspekt berücksichtigen, so ersetzt man $\#Z_{tn}$ durch $(\#\Phi_{rtn} + C \times \#\Phi_{wtn})$.

Mit Hilfe dieser Formeln kann man nun ein Verfahren gestalten, das aufsetzend auf einer bereits getroffenen Allokationsentscheidung prüft, ob eine angedachte vertikale Fragmentierung zu weiteren positiven Effekten führt. Hierzu wird man für jedes Fragment der Reihe nach alle möglichen Allokationen der vertikalen Fragmente durchgehen. Eine Aufteilung wird vorgenommen, sofern es mindestens ein positives B_{nr-sv} gibt. Unter den gefundenen Paaren wird das ausgesucht, für das B_{nr-sv} bzw. $B_{nr-s=r}$ oder $B_{nr-sv=r}$ maximal wird.

3.1.3 Anwendungsbereich

Die Modelle von Ceri und Pelagatti berücksichtigen diverse Aspekte der Anforderungen, die an den Knoten des Netzes mit unterschiedlich häufigem Auftreten entstehen, z. B. muß zu jedem Anforderungstyp ermittelt werden, von wo aus und wie dieser schreibend und lesend auf die Daten der einzelnen Fragmente zugreift. Diese Daten müssen bekannt sein und feste Größen darstellen. Über die Konstante C können Unterschiede im Aufwand schreibender und lesender Zugriffe in pauschalisierter Weise eingehen. Alle Verfahren setzen auf einer abgeschlossenen horizontalen Fragmentierung auf. Beim Verfahren zur vertikalen Fragmentierung wird diese jedoch noch weiter entwickelt, folgen somit nicht nur Allokationsentscheidungen sondern zugleich auch weitere Fragmentierungsentscheidungen.

Hauptsächliches Kriterium zur Bewertung des Aufwandes ist die Zahl der entfernten Zugriffe. **Die transportierte oder zu speichernde Datenmenge spielt in den Ansätzen jedoch keine Rolle.** Ebenso bleiben **Zugriffe, die durch die Verwaltung des Datenbestandes auftreten**, in allen Ansätzen unberücksichtigt. Nur im „Additional-replication“-Ansatz läßt sich dieser Aspekt in die Gewichtung, die einer zunehmende Kopienzahl zugeordnet wird, die aber auch Sicherheitsaspekte widerspiegelt, einbringen.

3.2 Das Modell von Dadam

Das von Dadam vorgestellte Berechnungsmodell ([DA96], S. 75 ff.) hat Ähnlichkeiten mit dem „All-Beneficial-Sites“-Ansatz von Ceri und Pelagatti, berücksichtigt aber zudem noch weitere Aspekte. Wie auch das oben vorgestellte Verfahren geht es von der Annahme aus, daß Datenbankanforderungen von jedem Knoten des Netzes aus initiiert werden können. An dem Ausgangsknoten findet die ggf. erforderliche Zerlegung dieser Anforderung in Teiloperationen zur Weiterleitung an die diversen beteiligten, entfernten Knoten zur dortigen Bearbeitung statt. Der initiiierende Knoten erhält von allen beteiligten Knoten eventuelle Resultate zurück. Bei Dadam werden aber die Daten in Speichereinheiten gemessen. Auf diese Weise kann man leicht unterschiedlichen Platzbedarf verschiedener Daten im Berechnungsmodell sowie den Übertragungsaufwand berücksichtigen. Mit diesen Annahmen berechnet Dadam zu allen möglichen Allokationen die sich aus Speicher- und

Übertragungskosten ergebenden Kosten und wählt dann diejenige Allokation aus, die am günstigsten ist.

Auch Dadam stellt den nicht redundanten Fall gesondert dar. Hier lassen sich einfachere Formeln bilden als für den redundanten Fall. Da aber der redundante Fall letztendlich den nicht redundanten umfaßt, werde ich hier nur den redundanten Fall wiedergeben. Zur Angabe der Formeln werden zunächst noch einige Variablenbezeichner eingeführt. Wer eine algorithmische Beschreibung des Verfahrens sucht, findet diese in der angegebenen Literatur.

S Speicherkosten (insgesamt)

U Übertragungskosten (insgesamt)

G_n Größe (gemessen in Dateneinheiten) von Fragment n ($n = 1, \dots, \#P$)

R_{tn} Größe (gemessen in Dateneinheiten) des Resultats einer Anforderung vom Typ t gegen Fragment P_n ($t = 1, \dots, \#T, n = 1, \dots, \#P$)

A_{tn} Größe (gemessen in Dateneinheiten) einer Anforderung (d. h. des „Anfragestrings“) vom Typ t gegen Fragment P_n ($t = 1, \dots, \#T, n = 1, \dots, \#P$)

3.2.1 Berechnungsmodell

Da bei Leseoperationen typischerweise die in Dateneinheiten gemessene Größe des Resultats R_{tn} einer Operation vom Typ t gegen Fragment P_n deutlich größer als der Anfragestring A_{tn} zu dieser Operation ist, kann A_{tn} unter Umständen vernachlässigt werden, umgekehrt verhält es sich bei Änderungsoperationen. Hier kann A_{tn} relativ groß sein, während R_{tn} lediglich aus der Bestätigung des Änderungsvorganges besteht.

Die **Speicherkosten S** ergeben sich aus der Formel:

$$S = \sum_{n=1}^{\#P} \sum_{i=1}^{\#K} G_n v_{ni} S_i$$

Dabei liefert ein einzelner Summand die Kosten eines Fragments am Knoten K_i . Der Wert v_{ni} der Indikatormatrix $V_{\#P, \#K}$ zeigt an, ob Fragment P_n an dem betrachteten Knoten K_i vorliegt, S_i gibt an, wie hoch die Kosten je Dateneinheit an diesem Knoten sind, und G_n liefert die Größe des gerade betrachteten Fragments.

Für den Fall, daß ein Fragment redundant gespeichert wird, bringt es mehrere von Null verschiedene Summanden in die Formel mit ein. Da jedes Fragment zumindest einmalig allokiert sein muß, ist somit zu fordern:

$$\sum_{i=1}^{\#K} v_{ni} \geq 1 \quad \text{für alle } n \in \{1, \dots, \#P\}.$$

Im Falle, daß keine Redundanz zulässig ist, also eine partitionierte Datenbank gewünscht ist, muß die Nebenbedingung etwas strenger sein:

$$\text{Hier mu\ss gelte: } \sum_{i=1}^{\#K} v_{ni} = 1 \quad \text{f\ur alle } n \in \{1, \dots, \#P\}.$$

Damit die maximalen Speicherkapazit\aten nicht \uberschritten werden, darf an keinem Knoten K_i die maximale Speicherkapazit\at M_i \uberschritten werden. Dabei ermittelt sich der Speicherbedarf an einem Knoten \uuber die Summe des Speicherbedarfs aller an diesem Knoten vorliegenden Partitionen.

$$\sum_{n=1}^{\#P} G_n v_{ni} \leq M_i \quad \text{f\ur alle } i \in \{1, \dots, \#K\}$$

Die **\uUbertragungskosten** U lassen sich mit folgender Formel, die nachfolgend noch erl\autert wird, beschreiben:

$$U = \sum_{i=1}^{\#K} \sum_{t=1}^{\#T} \sum_{n=1}^{\#P} \#F_{ti} \Gamma_t \left(A_{tn} U_{ij} + R_{tn} U_{ji} \right)$$

F\ur jedes m\ogliche Tripel (i, t, n) wird betrachtet, wie h\aufig $(\#F_{ti})$ die zugeh\orige Operation t am Knoten K_i auftritt. Entsprechend oft gehen \uUbertragungskosten zu dem gerade betrachteten Fragment P_n ein, f\ur das aus A_{tn} und R_{tn} hervorgeht, ob die Operation t \uuberhaupt darauf zugreift. Dabei wird zwischen lesenden und schreibenden Zugriffen unterschieden:

- Da eine vom Knoten K_i initiierte Leseoperation, f\ur die P_n ben\otigt wird, an den Knoten K_j gesandt wird, f\ur den eine Kopie mit geringsten \uUbertragungskosten erreicht werden kann, werden in der Formel in der rechten Klammer zun\achst die \uUbertragungskosten zu allen Knoten, an denen P_n lagert (f\ur diese gilt: $v_{nj=1}$), ermittelt und unter ihnen dann der niedrigste Wert herausgesucht. Dementsprechend ist Γ_t , falls t eine Leseoperation ist, ein Minimum-Operator. Falls das Fragment am Knoten K_i selbst lagert, also $v_{ni=1}$ gilt, ergibt sich als Minimum der Wert 0, da in diesem Fall $U_{ii} = 0$ gelten mu\ss.
- Hingegen werden \uAnderungsanforderungen, die P_n betreffen, von K_i aus an alle Knoten, an denen P_n allokiert ist, verschickt. Somit mu\ss hier auf die rechte Klammer, in der f\ur jeden Knoten, an dem P_n allokiert ist, die \uUbertragungskosten ermittelt werden, der Summenoperator angewandt werden, d. h. Γ_t liefert hier die Summe \uuber s\amtlliche \uUbertragungskosten.

F\ur die entstehenden \uUbertragungskosten zu einem Knoten K_j , an dem P_n allokiert ist, wird der Aufwand f\ur Anfrage- und Antwort\uUbertragungen ber\ucksichtigt. Deren Gr\o\ss e in Dateneinheiten wird durch A_{tn} und R_{tn} beschrieben. Dabei wird die anfallende Datenmenge auf dem Hinweg von Knoten K_i nach Knoten K_j mit den dort anfallenden \uUbertragungskosten je Dateneinheit (U_{ij}) multipliziert, entsprechend wird mit R_{tn} auf dem R\uckweg verfahren. Hier geht als Faktor U_{ji} ein.

Diese Formel zur Berechnung der \uUbertragungskosten gilt auch f\ur eine partitionierte Datenbank. Hier l\asst sie sich jedoch vereinfachen, da im Anfragefall eine Minimumbildung \uuberfl\ussig bleibt und im Antwortfall die Summe stets nur einen

Summand besitzt. Dies läßt sich unter Einbeziehung der Indikatorvariablen v auch so ausdrücken:

nicht redundanter Fall:

$$U = \sum_{i=1}^{\#K} \sum_{t=1}^{\#T} \sum_{n=1}^{\#P} \sum_{j=1}^{\#K} \#F_{it} v_{nj} \times (A_{tn} U_{ij} + R_{tn} U_{ji})$$

Abschließend muß nun noch die Summe aus Speicher- und Übertragungskosten unter den genannten Nebenbedingungen minimiert werden! Sowohl das Berechnen der Formeln als auch die Lösung der Zielfunktion läßt sich mit dem eingangs erwähnten Algorithmus durchführen.

3.2.2 Anwendungsbereich

Dadam setzt grundsätzlich voraus, daß sämtliche Fragmente (ob vertikal oder horizontal) bereits fest vorliegen. Hiervon ausgehend berücksichtigt er in seinem Berechnungsmodell ([DA96], S. 75 ff.) sowohl die Häufigkeit als auch die Art von Datenbankanforderungen, zählt aber nicht nur die entfernten Zugriffe, sondern bezieht hierbei auch die Größe der übertragenen Daten ein. Dabei berücksichtigt sein Modell sogar die genaue Datenmenge des „Anfragestrings“ und der Antwort. Bei beschreibenden Zugriffen geht außerdem ein, daß sämtliche Replikate angepaßt werden müssen. Darüber hinaus erfaßt sein Modell auch die zur Verfügung stehende Speicherkapazität der Knoten als begrenzende Größe sowie den Aspekt der Speicherkosten. Der höhere Grad der Berücksichtigung von Teilaspekten gegenüber dem zuvor dargestellten Best-Fit-Ansatz erfordert natürlich entsprechend umfangreichere Informationen über die Systemanforderungen und Gegebenheiten.

Da die Zeitkomplexität für das Durchrechnen sämtlicher Fälle, zumindest wenn man von einer großzügig ausgelegten Speicherkapazität der einzelnen Knoten ausgeht, sowohl im partitionierten Falle als auch bei zugelassener Redundanz exponentiell von der Anzahl der Fragmente abhängt, dürfte das Verfahren allerdings nur bei einer relativ geringen Zahl an Fragmenten oder bei sehr strengen Speicherplatzvorgaben, die für die Allokation wenig Spielraum lassen, vom Aufwand her akzeptabel sein. Dies sollen die folgenden, diesen Abschnitt abschließenden Betrachtungen verdeutlichen.

Da man, wenn man von der Nebenbedingung des Speicherplatzangebotes an den Knotenrechnern einmal absieht, die Verteilung der Fragmente auf die Knoten als unabhängige Ereignisse betrachten kann, ergeben sich für den partitionierten Fall $\#K^{\#P}$ mögliche Anordnungen, denn jedes Fragment läßt sich genau einem der $\#K$ Rechner zuordnen.

#K/#P	1	2	3	4	5	10	15	20
1	1	1	1	1	1	1	1	1
2		4	8	16	32	1.024	32.768	1.048.576
3			27	81	243	59.049	1,435E+07	3,487E+09
4				256	1024	1.048.576	1,074E+09	1,100E+12
5					3125	9.765.625	3,052E+10	9,537E+13
6						6,047E+07	4,702E+11	3,656E+15
7						2,825E+08	4,748E+12	7,979E+16
8						1,074E+09	3,518E+13	1,153E+18
9						3,487E+09	2,059E+14	1,216E+19
10						1,000E+10	1,000E+15	1,000E+20

Tabelle 4: Anzahl möglicher Allokationen im partitionierten Fall ($\#K^{\#P}$)

(hellgrau schattierte Bereiche sind uninteressante Fälle)

Bei zugelassener Redundanz hat man sogar $(2^{\#K}-1)^{\#P}$ Möglichkeiten. Zu berücksichtigen ist hier, daß an keinem Knoten ein Fragment doppelt allokiert werden darf und jedes Fragment mindestens einmal allokiert werden muß. Die Anzahl an Möglichkeiten, ein einziges Fragment zu allokiieren, läßt sich leicht angeben, wenn man zunächst den Fall, daß dieses gar nicht allokiert wird, mit berücksichtigt: dann gibt es genau $2^{\#K}$ Möglichkeiten, da an jedem Knoten unabhängig vom vorherigen eine ja/nein-Entscheidung für oder gegen die Allokation fallen kann. Da nur genau ein einziger Fall existiert, bei dem das Fragment überhaupt nicht allokiert wird (an allen Knoten wird „Nein“ gewählt), gibt es also genau $2^{\#K}-1$ Möglichkeiten. Somit ergibt sich bei angenommener Unabhängigkeit zwischen den Verteilungen der verschiedenen Fragmente die oben angegebenen Zahl von $(2^{\#K}-1)^{\#P}$ Möglichkeiten.

#K/#P	1	2	3	4	5	10	15	20
1	1	1	1	1	1	1	1	1
2		9	27	81	243	59.049	1,435E+07	3,487E+09
3			343	2.401	16.807	2,825E+08	4,748E+12	7,979E+16
4				50.625	759.375	5,767E+11	4,379E+17	3,325E+23
5					2,863E+07	8,196E+14	2,347E+22	6,718E+29
6						9,849E+17	9,775E+26	9,701E+35
7						1,092E+21	3,606E+31	1,191E+42
8						1,163E+24	1,253E+36	1,351E+48
9						1,214E+27	4,230E+40	1,474E+54
10						1,255E+30	1,406E+45	1,576E+60

Tabelle 5: Anzahl möglicher Allokationen, falls Replikate zugelassen sind $((2^{\#K}-1)^{\#P})$

(hellgrau schattierte Bereiche sind uninteressante Fälle)

3.3 Die Allokationsbetrachtungen von Özsu und Valduriez

An dem Beitrag von Özsu und Valduriez zur Allokationsproblematik in [ÖV91] sind für diese Arbeit zwei Aspekte besonders interessant. Einerseits führen sie eine Reihe von Kritikpunkten zu einem Modell an, das dem „All-Beneficial-Sites“-Ansatz von Ceri- und Pelagatti und dem von Dadam sehr ähnelt und die daher bei dem dieses Kapitel abschließenden Vergleich der Modelle hilfreich sein können, zum anderen gehen die Autoren darauf ein, welche Aspekte ein geeigneteres, alternatives Modell umfassen müßte, auch wenn sie hierzu keine algorithmische Lösung vorstellen. Beiden Aspekten soll in den folgenden Abschnitten nachgegangen werden.

3.3.1 Kritik am „All-Beneficial-Sites“-Ansatz und Dadams Lösungsweg

Der Schwerpunkt dieses Abschnitts wird in den Kritikpunkten liegen, die Özsu und Valduriez zu einem von ihnen vorgestellten, ersten Modell auflisten, das etwas genauer als der „All-Beneficial-Sites“-Ansatz von Ceri und Pelagatti und etwas begrenzter als der von Dadam ist. Das Modell selbst soll hier nur kurz informell, vergleichend mit den früher behandelten Modellen, beschrieben werden, da es nichts wesentlich neues beinhaltet. Durch den Vergleich sollen die Zusammenhänge zu den zwei vorangegangenen Modellen so weit hergestellt werden, daß verständlich wird, daß die vorgebrachten Kritikpunkte ebenso auf diese Modelle zutreffen.

Özsu und Valduriez benötigen für ihr zuerst vorgestelltes Modell wie die bisher beschriebenen Modelle Kenntnis über die aus den Anforderungen hervorgehende Zahl der Lesezugriffe und Schreib- anforderungen gegen ein jedes Fragment. Allerdings wird in den Formeln nur noch die Summe aller Schreibzugriffe sowie die Summe aller Lesezugriffe an den jeweiligen Knoten betrachtet. Aus welchen Anforderungen diese ursprünglich entstanden sind, ist irrelevant. Darüber hinaus geht die Größe der Fragmente ein. Mit Hilfe dieser Daten minimiert das Modell Speicher- und Übertragungskosten, wobei bei letzteren auch die transportierte Speichermenge berücksichtigt wird. Dies entspricht dem Vorgehen von Dadam. Damit ist es wie dieses etwas genauer als der „All-Beneficial-Sites“-Ansatz von Ceri und Pelagatti, bei dem lediglich die Zahl der entfernten Schreib- und Lesezugriffe betrachtet wurde. Unterschiede zwischen schreibenden und lesenden Zugriffen, die bei Ceri und Pelagatti nur über die gemeinsame Konstante C berücksichtigt wurden, werden hier wie bei Dadam einmal durch Suche nach dem preiswertesten Weg, zum anderen durch Summenbildung über alle Wege modelliert. Zwischen Anfrage- und Antwortdatenfolgen wird nicht unterschieden. Zudem wird vereinfachend angenommen, daß es keine Kapazitätsbedingungen für die Knoten des Netzes und die Übertragungswege gibt. Aus diesen Gründen halte ich meine oben getroffene Einordnung des Modells „zwischen“ dem All-Beneficial-Sites-Ansatz und Dadams Vorgehen für gerechtfertigt. Nur in zwei, aus meiner Sicht eher unwesentlichen Aspekten, geht es über die Modellierung bei Dadam hinaus:

- Zum einen wird ermöglicht, daß die Kosten im Netz von einem zum anderen Knoten nicht symmetrisch sein müssen, d. h., die Übertragungskosten von K_1 nach K_2 können von denen von K_2 nach K_1 abweichen.
- Die Speicherkosten für die Fragmente an den Netzknoten werden durch eine Menge von Werten für jedes Fragment beschrieben. Ein solcher Wert beschreibt für ein Paar aus einem Fragment und einem Netzknoten die Kosten der Speicherung dieses Fragmentes am zugeordneten Knoten. Dies beinhaltet ein bißchen mehr Spielraum als die Vorgehensweise von Dadam, bei dem sich die Speicherkosten als Produkt der Fragmentgrößen und der spezifischen Speicherkosten eines

jeden Knotens ergaben. Die Sichtweise von Özsu und Valduriez läßt hingegen offen, wie die Kosten ermittelt werden. Auf diese Weise könnte man z. B. auch mit zunehmender Fragmentgröße nicht linear ansteigende Kosten abbilden, die der Vorgehensweise eines DBMS entsprechen können, sofern dieses ab einer bestimmten Fragmentgröße zusätzliche Indexe anlegt. Ich vermute allerdings nicht, daß dieser Aspekt bei der Beschreibung eine Rolle gespielt hat, zumindest wird er nicht erwähnt.

Die aufgestellte Formel zur Lösung des Minimierungsproblems ist bis auf die genannten Unterschiede nicht von der von Dadam zu unterscheiden. Genau wie bei ihm wird für jede denkbare, Redundanz zulassende Allokation eines jeden Fragmentes ermittelt, welche Kosten durch Allokation dieses Fragmentes an den Knoten, an denen es in der gerade betrachteten Allokationsmöglichkeit vorliegt, durch schreibende sowie lesende Zugriffe auf das Fragment entstehen. Beim Schreiben werden Übertragungskosten zu allen Kopien berücksichtigt, beim Lesen nur die niedrigsten Kosten zu einer Kopie. Insgesamt werden sämtliche, von allen Knoten ausgehende Anforderungen durchgegangen. Zu diesen Kosten kommen die Speicherkosten noch hinzu. Gesucht wird die Allokation, bei der die Gesamtsumme minimal wird.

In ihrer Kritik an diesem Modell weisen die Autoren zum einen auf dessen hohen Aufwand hin, wie er hier ja auch bereits für das Modell von Dadam beispielhaft aufgezeigt wurde. Dabei geben sie Literatur an, in der die NP-Vollständigkeit des Problems bewiesen wurde. Damit zählt es zu einer Menge von Problemen in der Informatik, für die für größere Eingabewerte - hier wären es die Anzahl der Knoten und die Anzahl der Fragmente - gegenwärtig keine in angemessener Zeit ablaufenden Lösungsalgorithmen angegeben werden können und von denen man weiß, daß die Suche nach einer solchen Lösung entweder hoffnungslos²⁷ oder zumindest ausgesprochen schwierig sein dürfte. Fände man wider Erwarten doch eine in angemessener Zeit ausführbare Lösung, so ließe sich diese auf eine ganze Klasse von Problemen übertragen (vgl. [Se91]), S. 718 - 721). Kritikpunkte anderer Art sind die folgenden:

- Die isolierte Allokationsentscheidung für jedes einzelne Fragment ist unrealistisch, da die Änderung der Allokation eines Fragments die Kosten für Zugriffe auf andere Fragmente, die für einen Vorgang, z. B. eine Join-Operation, gemeinsam benötigt werden, ebenfalls verändern kann.
- Das Zugriffsmodell ist zu einfach. Es geht lediglich davon aus, daß an jedem beliebigen Knoten Daten des Datenbestandes angefordert werden und dann aus den betroffenen Fragmenten zu diesem Knoten transportiert werden müssen. Daß es aber einen Zusammenhang zwischen der betrachteten Allokation und der Bildung abzuarbeitenden Schedules durch den Anfragebearbeitungsalgorithmus gibt, bleibt unberücksichtigt.
- Kosten, die zur Aufrechterhaltung der Integritätsbedingungen erforderlich sind, kommen im Modell überhaupt nicht vor. Falls aber Integritätsbedingungen Daten an verschiedenen Knoten betreffen, kann dieser Aspekt durchaus relevant sein.
- Kosten, die durch die Organisation des Mehrbenutzerzugriffs entstehen, bleiben ebenfalls außen vor.

²⁷ Dies hängt davon ab, ob $NP \neq P$ oder $NP = P$ gilt. Dabei ist P „umgangssprachlich“ beschrieben die Menge aller Probleme, die mit Hilfe eines deterministischen Algorithmus in polynomialer Zeit gelöst werden können (der Zeitbedarf ist proportional zu einem Polynom über der Anzahl der Bits der Eingabedaten). Ebenso kann man NP vereinfachend beschreiben als die Menge aller Probleme, die mit Hilfe eines nichtdeterministischen Algorithmus in polynomialer Zeit gelöst werden können. Um nachzuweisen, daß ein Problem in NP liegt, muß man einen in polynomialer Zeit ablaufenden Algorithmus finden, der zu einer gegebenen Lösung nachprüft, ob sie richtig ist (vgl. [Se91]), S. 718 - 719).

Özsu und Valduriez stellen daher neben das gerade betrachtete Modell einen anderen, umfassenderen Ansatz, zu dem sie allerdings lediglich einige grundsätzlich mögliche heuristische Lösungsstrategien angeben, keine davon jedoch ausführen.

3.3.2 Der Alternativvorschlag

Das alternative Modell von Özsu und Valduriez berücksichtigt insbesondere mehr Aspekte als die bisher vorgestellten Modelle. Es erfordert etwas detailliertere Angaben über die Datenbank und die darauf ablaufenden Anwendungen sowie über die Leistungsfähigkeit der Rechner im Netz. In der grundsätzlichen Vorgehensweise bleibt es jedoch den bisherigen ähnlich.

Auch zur Darstellung dieses Ansatzes habe ich soweit möglich die bereits verwandten Bezeichner gewählt. Dennoch kommen noch einige wenige hinzu: Das Zugriffsverhalten auf die Daten, die in der verteilten Datenbank gespeichert werden sollen, beschreiben die Autoren über die **Selektivität eines Fragmentes**. Diese Eigenschaft eines Fragmentes ermittelt man, indem man für jede Anforderung²⁸ t ($t = 1, \dots, \#T$) die Zahl der Tupel im Fragment, auf die bei einer solchen Anforderung zugegriffen wird, bestimmt. Man bezeichnet diese Zahl als $sel_{t,n}$. Ebenfalls neu ist die Verwendung zweier Indikatormatrizen $(I\Phi_r)_{t,n}$ und $(I\Phi_w)_{t,n}$ mit Elementen aus $\{0,1\}$, die lediglich anzeigen, ob die Anforderung t auf das Fragment P_n lesend bzw. schreibend zugreift. $O(t)$, $t \in \{1, \dots, \#T\}$ liefert den Index von dem Knoten, an dem die Anforderung t startet. Ferner soll für jede Anforderung t ihre **maximal zulässige Antwortzeit** erfaßt werden. Diese Zeiten können im Modell in Form von Nebenbedingungen einfließen. Hierauf gehen die Autoren allerdings nicht genauer ein. Ferner möchten die Autoren zusätzlich zu dem bei Dadam betrachteten Aspekt der Speicherkapazität der Knoten (S_i) auch die **Verarbeitungsleistung** der Rechner über Schätz- oder Meßwerte berücksichtigen. Mit L_i werden die **Kosten für das Abarbeiten einer Verarbeitungseinheit** am Knoten K_i erfaßt. Eine solche Einheit soll dem Aufwand eines Lese- bzw. Schreibvorganges entsprechen.

Als bereits bekannter Parameter taucht erneut die Größe G_n eines Fragments auf. Allerdings führen Özsu und Valduriez sie auf die Anzahl der Tupel des betreffenden Fragmentes (die Kardinalität von P_n) und die Länge (in Bytes) eines Tupels zurück. Dies ist aber für alle weiteren Betrachtungen unwesentlich und wird hier ignoriert. Auch die Zahl der Lese- und Schreibzugriffe einer Anforderung t auf ein Fragment P_n wird im Modell gebraucht. Sie wurden schon bei Ceri und Pelagatti unter $\Phi_{r,n}$ und $\Phi_{w,n}$ benötigt²⁹. Bezüglich der Netzeigenschaften werden nur die Kosten der Übertragungswege durch die eine Matrix U_{ij} widergespiegelt. Sie erfaßt für jede Knotenverbindung einen Wert für die Kosten je Dateneinheit auf dieser Verbindung. Für die Berücksichtigung weiterer Aspekte, wie z. B. Kanalkapazitäten, Abstand zwischen Knoten oder Protokollaufwand verweisen die Autoren auf Literatur zu Netzwerkmodellen. Die Elemente v_{ni} der von früher bekannten Matrix $V_{\#P,\#K}$ sind dafür vorgesehen, die getroffenen Allokationsentscheidungen festzuhalten. Im folgenden sind alle neu hinzugekommenen Bezeichner nochmals aufgelistet (Gesamtübersicht siehe Anlage 3).

t	Hier: leicht abweichend als Variable für eine Anforderung verwandt
$\#T$	Hier: Anzahl aller auftretender Anforderungen

²⁸ Dadam hatte Typen von Anforderungen betrachtet. Dennoch wird hier dieselbe Variable verwendet.

²⁹ Özsu und Valduriez weisen zusätzlich darauf hin, daß hierfür z. B. die Zahl der Block-Zugriffe auf dem Speichermedium gewählt werden kann. Auch dies ist für unsere Betrachtungen unwesentlich.

sel_{tn}	Selektivität des Fragments P_n unter der Anforderung t
$(I\Phi_r)_{\#T,\#P}$	Indikatormatrix, hat an allen Stellen $i\phi_{r_{tn}}$ den Wert 1, für die die Anforderung t auf Fragment P_n lesend zugreift, sonst 0
$(I\Phi_w)_{\#T,\#P}$	Indikatormatrix, hat an allen Stellen $i\phi_{w_{tn}}$ den Wert 1, für die die Anforderung t auf Fragment n schreibend zugreift, ansonsten 0
$o(t)$	Gibt den Index des Knotens an, an dem die Anforderung t initiiert wurde
L_i	Verarbeitungskosten je Lese- bzw. Schreibzugriff am Knoten K_i

Die Formeln, die die beiden Autoren nun auf Grundlage der hiermit bezeichneten Eingabewerte entwickeln, dienen der Minimierung der gesamten Verarbeitungs- und Speicherkosten, während gleichzeitig versucht wird, die vorgegebenen Begrenzungen hinsichtlich der Antwortzeiten zu den diversen Anfragen sowie der Speicherkapazität und Prozessorleistung der Knoten des Netzes einzuhalten.

In die Gesamtkosten fließen als Summanden zum einen die zu allen Anfragen entstehenden Kosten für die Anfragebearbeitung, zum anderen die Kosten für die Speicherung der Allokationen an den Knoten des Netzes ein.

$$\begin{aligned}
 \text{Gesamtkosten} &= \text{Anfragebearbeitungskosten} + \text{Speicherkosten} \\
 \text{Gesamtkosten} &= \sum_{t=1}^{\#T} \text{Verarbeitungskosten}(t) + \text{Übertragungskosten}(t) + \sum_{i=1}^{\#K} \sum_{n=1}^{\#P} S_i G_n v_{ni}
 \end{aligned}$$

Die Formel für die Speicherkosten stimmt bis auf die Reihenfolge der Argumente exakt mit der von Dadam überein. (Speicherkosten fallen nur an, falls ein Fragment P_n an einem Knoten K_i allokiert ist. Dies zeigt v_{ni} an. Das Produkt aus den spezifischen Speicherkosten je Speichereinheit des gerade betrachteten Knotens K_i und aus der Größe des Fragmentes P_n liefert den gewünschten Wert. Das Durchlaufen aller Indizes führt schließlich zu den Gesamtkosten.)

Auch die Ermittlung der Übertragungskosten ist ähnlich wie bei Dadam modelliert. Allerdings wird kein gemeinsamer Operator verwendet, sondern auf die eingeführten Indikatormatrizen zurückgegriffen. Außerdem wird die zuvor eingeführte Selektivität der Anforderungen in Hinblick auf die Fragmente, auf die zugegriffen wird, berücksichtigt.

$$\begin{aligned}
 \text{Übertragungskosten}(t) &= \text{Ü}_r(t) + \text{Ü}_w(t) \\
 &\quad \swarrow \quad \searrow \\
 &\text{(Übertragungskosten für Leseanforderungen}(t)) \quad \text{Übertragungskosten für Schreibanforderungen}(t)
 \end{aligned}$$

$\ddot{U}_r(t)$ und $\ddot{U}_w(t)$ schlüsseln sich hierbei wie folgt auf:

$$\ddot{U}_r(t) = \sum_{n=1}^{\#P} \text{Min}_{i \in \{1, \dots, \#K\}} \left(i\phi_{r_{tn}} v_{ni} U_{o(t)i} + i\phi_{r_{tn}} v_{ni} U_{i_{o(t)}} (\text{sel}_{tn} / G_n) \right)$$

$$\ddot{U}_w(t) = \sum_{n=1}^{\#P} \sum_{i=1}^{\#K} \left(i\phi_{w_{tn}} v_{ni} U_{o(t)i} \right) + \sum_{n=1}^{\#P} \sum_{i=1}^{\#K} \left(i\phi_{w_{tn}} v_{ni} U_{i_{o(t)}} \right)$$

Dabei ergeben sich die Übertragungskosten für Leseanforderungen einer Anforderung t aus der Summe der jeweils günstigsten Kosten zum Lesen der erforderlichen Fragmente. Für alle Fragmente wird ermittelt, für welchen Knoten die Summe aus den Übertragungskosten für den Transfer der Anfrage (1. Summand in der Klammer) und den Kosten zur Übersendung des Anfrageergebnisses (2. Summand in der Klammer) von und zum Ausgangsknoten $o(t)$ am geringsten sind. Anhand der Formel läßt sich ferner erkennen, daß für die Anfrage selbst mit einer Einheit ($i\phi_{r_{tn}}$), für die Antwort mit der genauen Größe der Ergebnismenge ($i\phi_{r_{tn}} \times (\text{sel}_{tn} / G_n)$) gerechnet wird. Hier hatte Dadam selbst für die Anfrage die genaue Größe eingehen lassen.

Da bei Schreibzugriffen jedes Fragment bearbeitet werden muß, wird hier kein Minimum ermittelt. Durch die doppelten Summen werden jeweils alle Knoten und alle Fragmente durchlaufen, also jede Allokation eines jeden Fragmentes auch berücksichtigt. Die einzelnen Kosten ergeben sich aus denen zur Übertragung der Schreibanforderung ausgehend vom Startknoten (erste Klammer) und dem Zurücksenden einer anschließenden Bestätigung zu diesem Knoten $o(t)$. Da auch hier lediglich auf die Indikatorwerte zurückgegriffen wird, geht jede einzelne solche Übertragung mit je einer Größeneinheit ein. Dadam hatte hier für Hin- und Rückweg die exakten Werte berücksichtigt.

Verarbeitungskosten wurden bei den bisherigen Modellen nicht einbezogen. Özsu und Valduriez setzen diese Kosten zusammen

- aus den Zugriffskosten $Zk(t)$,
- aus den Kosten für den Aufwand aus Integritätsbedingungen $Ik(t)$ und
- aus konkurrierenden Anforderungen $Kk(t)$.

Verarbeitungskosten $(t) = Zk(t) + Ik(t) + Kk(t)$, also insgesamt:

$$\text{Gesamtkosten} = \sum_{t=1}^{\#T} \left(Zk(t) + Ik(t) + Kk(t) + \text{Übertragungskosten}(t) \right) + \text{Speicherkosten}$$

Dabei verweisen die Autoren darauf, daß die detaillierte Spezifikation der Verarbeitungskosten von den zur Lösung eingesetzten Algorithmen abhängt und führen nur die Zugriffskosten beispielhaft weiter aus:

$$Zk(t) = \sum_{i=1}^{\#K} \sum_{n=1}^{\#P} \left(\underbrace{i\phi_{w_{tn}}}_{\text{Schreibzugriffe}} \underbrace{\# \Phi_{w_{tn}}}_{\text{Lesezugriffe}} + \underbrace{i\phi_{r_{tn}}}_{\text{Lesezugriffe}} \underbrace{\# \Phi_{r_{tn}}}_{\text{Kosten je Zugriff auf Fragment } P_n \text{ am Knoten } K_i} \right) \times v_{ni} L_i$$

Diese werden zu jeder Anforderung t gesondert bestimmt. Zu jedem Knoten K_i und zu jedem Fragment P_n wird bestimmt, wie oft die Anwendung t auf P_n lesend bzw. schreibend zugreift. Falls P_n an dem betreffenden Knoten allokiert ist, so wird die Zahl der Zugriffe mit den dort anfallenden Verarbeitungskosten je Zugriff multipliziert. Dabei weisen die Autoren selbst darauf hin, daß sie die Verarbeitungskosten nicht nach lesendem und schreibendem Zugriff differenzieren. Wollte man dies, so benötigte man Werte Lr_i und Lw_i für jeden Knoten K_i . Etwas seltsam mutet es aber an, daß im Gegensatz zu der Gestaltung der Übertragungskosten hier die Lesezugriffe auf alle Kopien eines Fragmentes eingerechnet werden, obwohl bei den Übertragungskosten nur die günstigste Kopie berücksichtigt wird. Da wäre es doch eigentlich sinnvoll, auch hier nur die dort entstehenden Lesezugriffskosten eingehen zu lassen oder noch besser, für Leseanforderungen zunächst den Index herauszufinden, für den die Verarbeitungs- und Übertragungskosten zusammen minimal werden und dann diese Kosten zu berücksichtigen.

Selbstkritik üben die Autoren in Hinblick auf die Tatsache, daß auch diese Formel leider nur eine sehr einfache Abarbeitung von Anfragen modelliert. Damit bleiben auch für dieses Alternativmodell der Autoren einige der selbst aufgestellten Kritikpunkte von oben (isolierte Allokationsentscheidungen, Art der Join-Ausführung, zu einfaches Zugriffsmodell) bestehen.

Die möglichen Nebenbedingungen skizzieren die Autoren lediglich. Für das Antwortzeitverhalten sollte für jede Anfrage t eine Ungleichung aufgestellt werden, die einfordert, daß diese Zeit unter der maximal akzeptablen Antwortzeit der Anfrage bleibt. Über eine Ungleichung lassen sich ebenso Nebenbedingungen zur Speicherkapazität an den diversen Knoten des Netzes formulieren. Auf diese Weise wurde der Aspekt auch bei dem von Dadam vorgestellten Modell berücksichtigt. Und schließlich fordern die Autoren noch als weitere Bedingung, daß die Summe der Verarbeitungskosten, die durch die Bearbeitung der Anfrage t am Knoten K_i entsteht, die zulässigen Verarbeitungskosten an dem betreffenden Knoten nicht überschreiten darf.

Einen Lösungsalgorithmus, der diese mehr oder weniger detailliert ausgearbeiteten Formeln berücksichtigen soll, geben die Autoren nicht an, er liegt jedoch auf der Hand, sofern man von Problemen der Komplexität absieht und davon ausgeht, daß die nicht aufgeführten Kosten zu ähnlich „einfachen“ Formeln führen, wie es die Autoren auch im Text andeuten. Dann sind für alle vorkommenden Anfragen für alle Knoten alle Fragmente zu durchlaufen sowie jeweils die Nebenbedingungen zu prüfen. Özsu und Valduriez weisen aufgrund der Komplexitätsproblematik darauf hin, daß einem heuristischen Lösungsansatz mit geringerer Komplexität der Vorzug zu geben ist. Dabei ist natürlich wesentlich, wie gut ein solcher Ansatz an die optimale Lösung heranreicht. Hierzu verweisen sie auf generelle heuristische Lösungsansätze für verwandte Probleme. Zudem weisen sie darauf hin, daß Lösungsansätze durch Simulationstechniken für das Allokationsproblem bisher noch wenig erforscht sind.

3.3.3 Anwendungsbereich

Wie bereits oben angeklungen, liegt der hauptsächliche Gewinn in diesem Alternativvorschlag darin, daß mehr Aspekte als in den zuvor vorgestellten berücksichtigt werden. Damit werden die realen Gegebenheiten genauer modelliert, wobei immer noch vereinfachte Sichtweisen in Hinblick auf die Anfragebearbeitung bestehen bleiben. Allerdings muß auch hier die generelle Frage gestellt werden, ob die erforderlichen, zahlreichen Eingabedaten im vorhinein hinreichend genau geschätzt werden können oder gar feststehen. Geht man davon aus, daß dies nicht der Fall ist, etwa, wenn die Datenbank innerhalb des Betriebes starken Schwankungen unterliegt, so vermag die zusätzliche Genauigkeit bedeutungslos werden. Ebenso bleibt der von den Autoren selbst genannte Problempunkt, daß

aufgrund der Komplexität der Problemsituation nicht gleichzeitig Kosten- und Durchsatz- sowie Sicherheitsgesichtspunkte optimiert werden können.

3.4 Der graphbasierte Modellansatz von Apers

Apers stellt in seinem Beitrag [Ap88] zwei Lösungsansätze vor, die es möglich machen, die Kosten möglicher Allokationen in Abhängigkeiten zu auftretenden Transaktionsfolgen zu ermitteln und hiervon ausgehend die beste Allokation oder, jeweils mit einem etwas weniger aufwendigem heuristischen Ansatz eine verhältnismäßig gute Allokation zu erreichen. Apers betrachtet im Gegensatz zu den zuvor vorgestellten Modellen auch die Programme, die aus den zu erwartenden Transaktionen die abzuarbeitenden Operationsfolgen, die Schedules³⁰ bilden. Da er eine zirkuläre Abhängigkeit zwischen Allokationsüberlegungen und Schedule-Bildung sieht (vgl. ebenda, S. 269), berücksichtigt er diese in einem seiner Lösungsansätze, den er dynamisch nennt. Diese Bezeichnung hat aber nichts mit dynamischer Allokationsplanung zu tun, sondern soll den Aspekt betonen, daß während der Ermittlung der günstigsten Allokation immer wieder auch die Anfrageoperationsfolgen überarbeitet und angepaßt werden. Den anderen Lösungsansatz nennt er im Gegensatz dazu selbst „statisch“. Zur Vermeidung dieser beiden Begriffe, die nicht der bisherigen Verwendung in dieser Arbeit entsprechen, wird hier sein „statischer Lösungsvorschlag“ als „erstes“, sein „dynamischer“ als „zweites Vorgehensmodell“ vorgestellt. Ein weiterer Unterschied seiner Ansätze zu den bisherigen besteht darin, daß er die Fragmente nicht gleich auf denkbare Weisen den geplanten Netzknoten zuordnet, sondern zunächst auf virtuelle Knoten verteilt, für die sich dann im Laufe der Verfahren herausstellt, an welchem realen Netzknoten sie am besten anzusiedeln sind.

Erstes und zweites Vorgehensmodell sind graphbasiert; beide verwenden gerichtete Graphen zur Repräsentation der möglichen Allokationsorte und der abzuarbeitenden Schedules. Wie bereits erwähnt, halte ich diese Visualisierung jedoch im Vergleich zu den oben dargestellten Charakteristika seiner Vorgehensweise zur Abgrenzung gegenüber den bisher betrachteten Modellen für sekundär. Eine weitere Gemeinsamkeit des ersten und zweiten Vorgehensmodells besteht darin, daß Kopien zugelassen sind und als Minimierungsfunktion die Übertragungskosten gewählt wurden. Update-Anforderungen werden auf Leseanfragen zurückgeführt.

Die folgenden beiden Abschnitte versuchen, Apers Lösungsansätze so vorzustellen, daß ihre Funktionsweise prinzipiell verständlich wird. Grundlegende Erläuterungen zu Bezeichnungen und seinem graphischen Modell werden bei der Behandlung des ersten Vorgehensmodells gegeben. Deswegen setzt das Lesen des zweiten Verfahrens das des ersten Modells voraus. Für Details ist die Quelle heranzuziehen.

³⁰ Dieser Begriff wurde in Abschnitt 2.3.3 eingeführt.

3.4.1 Erstes Vorgehensmodell

Apers setzt Angaben zu der Art der Anfragen und Updates, ihrer Frequenz und dem Ort, zu dem das Ergebnis gesandt werden soll, voraus. Unter Verwendung der bisherigen Variablen benötigen wir daher wieder die Variable t , die hier die diversen Transaktionen bezeichnen wird, die im Rahmen der Anforderungen an die Datenbank auftreten. Den Bezeichner F_{ti} als Maß für die Häufigkeit der Transaktion t am Knoten K_i hatten wir bereits bei Ceri und Pelagatti benötigt; leicht abgewandelt taucht er hier wieder auf.

Die zur Ausführung jeder Transaktion erforderlichen Operationen auf dem Datenbestand verteilt Apers neben den Daten und deren möglichen Kopien auf die Knoten des Netzes. Außerdem nimmt er **Synchronisations-** und **Verzweigungsprozesse** in sein Modell auf, die jeder Operation auf dem Datenbestand sowie den Übertragungsvorgängen vor- und nachgeschaltet werden. Dabei läßt ein Synchronisationsprozeß eine Operation warten, bis alle oder zumindest ein Teil der benötigten Eingabedaten vorliegen; ein Verzweigungsprozeß ermöglicht dem Ergebnis einer Operation oder einer Übertragung ein Eingabewert einer weiteren Operation oder auch gleich von mehreren zu werden. Letzterer Fall umfaßt das Erzeugen von Kopien. Verzweigungsprozesse werden im Graphenmodell graphisch dargestellt.

Weitere grundlegende Begriffe in Apers Modell sind „nucleus site“, „physical site“ und „virtuell site“, die ich im folgenden als **Kernpunkte**, **physikalische** und **virtuelle Punkte** bezeichne. Dabei ist „Kernpunkt“ der Oberbegriff für entweder physikalische oder virtuelle Punkte und definiert als ein Paar (PS, OS) um, wie oben beschrieben, sowohl eine Menge von Fragmenten (PS) als auch eine Menge von Operationen (OS) aufnehmen zu können. Im Falle eines physikalischen Punktes umfaßt PS die Fragmente, die dort allokiert werden sollen, im Falle eines virtuellen Punktes wird deutlich, daß die in PS liegenden Fragmente gemeinsam allokiert werden sollen, der genaue Standort ist jedoch noch offen. Virtuelle Punkte können (maximal) einem physikalischen Punkt zugeordnet werden. Einem physikalischen Punkt können durchaus mehrere virtuelle Punkte zugeordnet sein. Eine **Zuordnung** ist reversibel und für das Durchgehen denkbarer Lösungen sinnvoll. Demgegenüber stellt eine **Vereinigung** von einem virtuellen und einem Kernpunkt eine endgültige Zusammenlegung (zweier virtueller Punkte oder eines virtuellen und eines physikalischen Punktes) dar. Bei ihr werden auch formal die Mengen PS und OS der jeweiligen Kernpunkte vereinigt; zugeordnete Punkte eines physikalischen Punktes bleiben. Solange in den Modellbetrachtungen noch virtuelle Punkte auftreten, liegt lediglich eine **partiell spezifizierte Allokation** vor. Gibt es nur noch physikalische Punkte, so hat man eine **komplett spezifizierte Allokation** erreicht. Innerhalb seines graphischen Modells werden Punkte als **Knoten** bezeichnet.

Die in OS aufgeführten **Operationen** werden von Apers durch Tripel (t, F, x) gekennzeichnet, wobei t die Transaktion angibt, zu der die betreffende Operation gehört, F die Häufigkeit, mit der die Transaktion t auftritt und x die Ausführungszeit liefert. Während beim ersten Modellansatz von Ceri und Pelagatti mit F_{ti} die Häufigkeit einer Anforderung vom Typ t am Knoten i beschrieben wurde, ist hier durch das Tripel klar, zu welcher Transaktion F gehört. An welchem Knoten die Operation erfolgt, kann man erschließen, indem man verfolgt, welchem Punkt die Menge OS, der die Operation zugeordnet ist, angehört. Neben den Datenmengen, Operationen und Verzweigungsprozessen gehen in das Graphenmodell auch explizit die Übertragungsvorgänge in Form von Tripeln (t, F, D) ein. Aus ihnen geht neben der Transaktion, die die Übertragung erforderlich macht, sowie deren Frequenz die Datenmenge D hervor, die transportiert wird. Die aus diesen Tripel ab-

leitbaren Produkte $F \times D$ sind Grundlage für die Berechnung der Kostenfunktion, die die totalen Übertragungskosten als Summe aller durch die Anfragen entstehenden Datenübertragungen liefert.³¹

Im folgenden sind die gerade genannten Bezeichner nochmals kurz aufgelistet:

- t Transaktion vom Typ t
- PS Menge von Fragmenten
- OS Menge von Operationen
- $K_i = (PS, OS)_i$ physikalischer Punkt, im Graphenmodell physikalischer Knoten i,
- $N_i = (PS, OS)_i$ Kernpunkt i, im Graphenmodell Kernknoten i
- $V_i = (PS, OS)_i$ Virtueller Punkt, im Graphenmodell virtueller Knoten i
- (t, F, x) Operation der Transaktion t, wobei t mit Frequenz F auftritt und die Operation die Ausführungszeit x benötigt
- (t, F, D) Angabe zu einer Datenübertragung der Größe D (in Dateneinheiten), die durch Transaktion t entsteht, wobei t mit Frequenz F auftritt.

Ausgangssituation in Apers erstem Modell ist eine partiell spezifizierte Allokation (s. o.). Hierzu muß Apers eine anfängliche Fragmentierung der Daten vornehmen. Diese basiert auf den Anforderungen an die Datenbank. Dabei ist es für ihn irrelevant, ob lesend oder schreibend auf die Daten zugegriffen wird, vielmehr interessieren ihn die relationalen Operatoren, die zur Auswahl der benötigten Datenmengen dienen, wobei er nur **Selektion**, **Projektion** und bei mehrere Relationen betreffenden Situationen auch den **Join** zuläßt. Abhängig von den Selektionsprädikaten (vgl. Abbildung 8, hier sind es $\sigma_1 R$ mit $\sigma_1 = A \text{ op } a$ und $\sigma_2 R$ mit $\sigma_2 = B \text{ op}' b$) wird die horizontale Fragmentierung vorgenommen, die zu 2^n Fragmenten führt, von denen aber eines uninteressant ist, weil dessen Bedingungen in den vorliegenden Selektionen nicht vorkommt. Hiervon ausgehend werden weitere vertikale Fragmentierungen vorgenommen. Diese sind von den Attributen der Relationen abhängig: Ein Fragment ergibt sich jeweils durch Projektion auf den Primärschlüssel, weitere durch Projektion auf alle durch den Primärschlüssel und ein weiteres Attribut entstehende Attributmengen. Dies sind $\#Att - \#PrAtt + 1$ Fälle, wobei $\#Att$ und $\#PrAtt$ Variablen für die Anzahl der Attribute bzw. der Attribute des Primärschlüssels sind.

Bei einer Relation mit 4 Attributen und einem Primärschlüssel von 2 Attributen ergeben sich auf diese Weise bei Vorgabe zweier Selektionsbedingungen durch jeweils 3 vertikale Fragmentierungen insgesamt 9 Fragmente (je 3 Fragmente für die Fälle 1 - 3).

Falls mehrere Relationen betroffen sind, so lassen sich ebenso die vom Join betroffenen Fragmente durch Selektionen und Projektionen bezüglich der beteiligten Relationen beschreiben

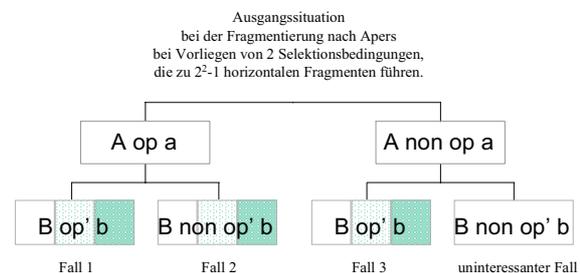


Abbildung 8: Horizontale Anfangsfragmentierung bei Apers

³¹ Apers verweist allerdings darauf, daß sein Vorgehen auch auf andere Kostenfunktionen übertragen werden könne. Außerdem sei es möglich, Nebenbedingungen, wie z. B. die Prozessor-Auslastung, mit zu berücksichtigen. Im Falle einer Überschreitung muß der Kostenfunktion dann der Wert „unendlich“ zugeordnet werden.

und das Vorgehen bleibt anwendbar³². Lediglich das Auftreten nicht wünschenswerter, sehr kleiner Fragmente, muß vermieden werden. Dies wäre über Grenzwerte für die Größe von Fragmenten möglich.

Als nächstes muß der Ausgangsgraph für die zu suchende Allokation aufgestellt werden. Apers nennt seine **Graphen** „Processing-schedules-graph“, ich werde im folgenden vom **PSG** sprechen. Der PSG erhält für jeden physikalischen und virtuellen Punkt (auch wenn PS und OS nur leere Mengen sind) einen Knoten. Diese werden durch Quadrate mit einem Punkt in der Mitte dargestellt, die physikalischen werden häufig größer wiedergegeben (vgl. Abbildung 9). Auf den Seiten der Quadrate der physikalischen Knoten werden im Laufe des Verfahrens zugeordnete virtuelle Knoten eingetragen. Einer Datenübertragung entspricht eine gerichtete Kante. Die Kanten werden bei der Ermittlung der jeweiligen Operations-Ausführungsfolge durch einen üblichen Anfragebearbeitungsalgorithmus zu den einzelnen Transaktionen gebildet. Bögen an abgehenden Kanten weisen auf Verzweigungsprozesse hin.

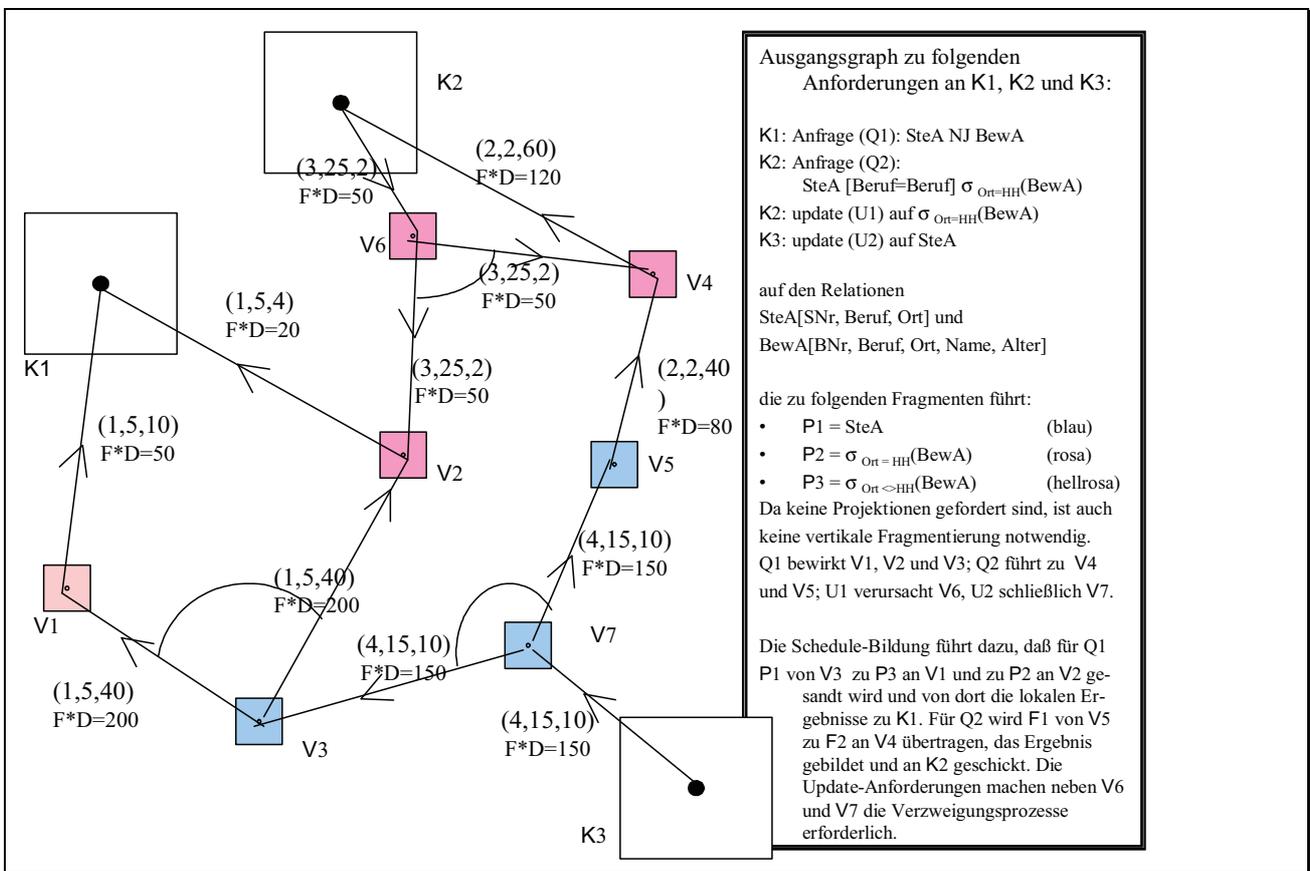


Abbildung 9: PSG (in Anlehnung an Figur 10, S. 287 aus [Ap88])

³² Apers beweist zudem, daß bei dem Ziel, die Übertragungskosten minimieren zu wollen, eine weitere Fragmentierung keine Vorteile mehr erbringt, vorausgesetzt, daß man tatsächlich alle Kenntnisse über die Datenbankanforderungen zur Bildung der Anfangsfragmente verwandt hat. Hierfür geht er von einer bestimmten Abarbeitung der relationalen Operationen aus, die er nicht nur von der Operation selbst, sondern auch von der Folgeoperation abhängig macht. Andererseits weist er auch darauf hin, daß bei einem anderen Minimierungsziel wie z. B. dem Erreichen einer möglichst geringen Antwortzeit eine weitere horizontale Fragmentierung durch die Erhöhung des Parallelisierungsgrades zu Vorteilen führen kann.

Der **Ausgangsgraph** erhält für alle physikalischen Punkte einen physikalischen Knoten κ_i und für jede Anfrage- bzw. jede Update-Anforderung für jedes benötigte Fragment einen virtuellen Knoten. Die Fragmente, die für diese Anforderung erforderlich sind, werden an diesem virtuellen Knoten allokiert. Kein virtueller Knoten ν_j wird bereits einem physikalischen Knoten κ_i zugeordnet. Zu dieser Situation werden der Anfragebearbeitungsalgorithmus durchlaufen und die Einträge in den PSG vorgenommen. Ferner werden virtuelle Knoten, die verschiedene Kopien desselben Fragmentes aufweisen und bei denen das Fragment in einen Update-Vorgang involviert ist, per Verzweigungsprozeß miteinander durch die graphische Darstellung eines Verzweigungsprozesses verbunden.

Die Ermittlung einer optimalen Datenallokation erfolgt dann durch systematisches, schrittweises Zuordnen und Vereinigen von Knoten sowie dem Entfernen von Kanten und Verzweigungsoperationen des Ausgangsgraphen hin zu einer komplett spezifizierten, optimalen, bzw. möglichst optimalen Lösung. Dabei vermeidet Apers in seinem vorgestellten Verfahren das Durchprobieren aller möglichen Veränderungen des Ausgangsgraphen, indem er Entscheidungsbäume verwendet und in diesen Abschätzungen vornimmt. Jeder Knoten eines solchen Baumes, der während des Verfahrens mitgeführt wird, enthält zunächst Angaben zu der gerade getroffenen Vereinigungs- oder Zuordnungsentscheidung, sowie die für diese Teillösung schon feststehenden, anfallenden Übertragungskosten. Hinzu kommt dann noch ein Schätzwert für die Kosten der Übertragungen, die „für die weitere Entwicklung“ noch zu erwarten sind. Dieser schätzt unter allen von diesem Knoten ausgehenden und zu komplett spezifizierten Lösungen führenden Pfaden die Kosten für einen günstigsten Pfad ab. Somit liefert er einen Schätzwert für die Kosten der preiswertesten komplett spezifizierten Lösung, die von diesem Knoten aus erreichbar wäre. Beachtenswert hieran ist, daß Apers diesen Schätzwert nicht durch Bildung der zugehörigen Graphen zu all diesen Pfaden ermittelt, sondern anhand eines Verfahrens auf dem Graphen der bisher erreichten partiell spezifizierten Lösung!

Auf jeder Stufe des Entscheidungsbaumes wird dann nur derjenige Knoten weiter verfolgt, bei dem dieser Schätzwert zusammen mit den bisher entstandenen Kosten minimal ist. Für diesen Fall werden alle denkbaren Vereinigungen eines jeden ν_i mit allen κ_i durchlaufen.

Apers beweist in seinem Beitrag, daß man ohne ein Anpassen der Schedules während der Anpassungsschritte ausgehend von der Startsituation zu keinem besseren Ergebnis gelangen kann, also das optimale Ergebnis erreicht. Durch Einbeziehung einiger Grundüberlegungen, wie etwa, daß es sinnvoll ist, nach der ersten Stufe zunächst virtuelle Knoten untereinander zu vereinigen und Kanten mit großen Datentransfers vor solchen mit weniger großen zu betrachten, entwickelt er zudem einen **heuristischen Ansatz** für denselben Transfervorgang und zeigt, daß diese Lösung für zumindest eine bestimmte Klasse von Ausgangsgraphen mit der optimalen Lösung zusammenfällt.

Apers hat Zeitvergleiche zwischen beiden Ansätzen durchgeführt. In Abhängigkeit von der Zahl der Transaktionen (≤ 4), der Anzahl der Fragmente (≤ 3) und einem Komplexitätsparameter für die Gestalt zufällig produzierter Graphen untersucht er vornehmlich die Güte des heuristischen Ansatzes und erhält als Nebenresultat, daß bei solch kleinen Werten die Ermittlung der optimalen Lösung das fünffache der Zeit kostet, die beim heuristischen Ansatz erforderlich ist. Er weist aber darauf hin, daß bei größeren Werten von einem rapide ansteigenden Unterschied auszugehen ist. Schon bei diesen kleinen Werten traten bei der Ermittlung der optimalen Lösung Speicherüberläufe auf. Konnten beide Verfahren zu Ende kommen, betrug die Güteabweichung im Durchschnitt etwas mehr als 3% der Übertragungskosten.

3.4.2 Zweites Vorgehensmodell

In seinem zweiten Vorgehensmodell berücksichtigt Apers bei jedem Entwicklungsschritt die sich hierdurch verändernden Schedules. Dies hat zwei Vorteile. Zum einen beinhaltet dann das Ergebnis Schedules, wie sie auch während des Betriebs der Datenbank auftreten, zum anderen sind auf diese Weise die Übertragungskosten der endgültigen Allokation meist günstiger, als bei dem durch das erste Vorgehensmodell erzielten Ergebnis. Hauptnachteil ist, daß der Berechnungsaufwand noch höher ist. Dies liegt nicht nur an dem bei jedem Veränderungsschritt erneut erforderlichen Aufruf des Algorithmus zur Schedule-Bildung, sondern insbesondere daran, daß eine Abschätzung auf einem partiell spezifizierten Graphen nun nicht mehr möglich ist. Dies würde zunächst bedeuten, daß von jedem Entscheidungsknoten aus alle denkbaren Lösungen berücksichtigt werden müßten. Damit würde bereits eine einzige Abschätzung mehr als polynomiale Zeit - bezogen auf die Anzahl der Kernknoten³³ - in Anspruch nehmen. Hierdurch müßte man nicht nur deutlich mehr Fälle als im vorangegangenen Beispiel betrachten, es sind auch wesentlich mehr als in allen vorigen Modellen, da Apers ja durch die Einbeziehung der virtuellen Knoten bei der realistischen Annahme, daß mehr Fragmente als Knoten vorliegen, insgesamt von viel mehr Knoten ausgeht. Nur durch Berücksichtigung der realistischen Annahme, daß Anfragen bzw. Updates im Normalfall nur geringe Bereiche des Graphen tangieren und eine vollständige Überprüfung von Anpassungsvorgängen in hiervon betroffenen Teilgraphen ausreichend ist, hat Apers auch für sein zweites Vorgehensmodell ein verkürztes Verfahren mit einem Abschätzvorgang gefunden. Auch dieses kommt aus, ohne tatsächlich alle Fälle durchzugehen.

Darüber hinaus hat er auch einen zweiten heuristischen Ansatz gefunden. Hierzu ergänzt er den PSG durch Kanten zwischen je zwei Knoten und Angaben an diesen Kanten, die besagen, um wieviel sich die Übertragungskosten verändern, wenn die beiden an der Kante anliegenden Knoten miteinander vereinigt oder einander zugeordnet würden. Um diese Werte zu berechnen wird der Anfragebearbeitungsalgorithmus für jede Kante zweimal durchlaufen. Einmal ohne ihr Zusammenfallen und einmal mit. Die Differenz wird eingetragen. In Verbindung mit diesen zusätzlichen Einträgen im Ausgangsgraphen, ist das Vorgehen ähnlich dem heuristischen Vorgehen im ersten Falle, nur daß dabei Schedules neu zu berechnen sind:

“First the virtual sites are individually assigned to physical sites such that the total transmission cost is minimized. Then pairs of virtual sites are considered for uniting in descending order of the labels of the edges between them. The cost of removing the two virtual sites from the physical sites to which they are assigned is the sum of the labels of the edges between the two virtual sites and the virtual sites that have already been assigned to the physical sites involved. Uniting them will decrease the cost function by an amount denoted by the label of the edge between the virtual sites. However, the decrease in the cost function when the union is assigned to a physical site is not yet known. Therefore, the schedules of the queries involved have to be recomputed and an assignment of the union to each physical site must be considered.“ ([Ap88], S. 298)

Wenn unter dieser Zuordnung zum physikalischen Knoten eine Verringerung oder Beibehaltung der Übertragungskosten erreicht wird, so werden die beiden virtuellen Knoten vereinigt und dem betreffenden Knoten zugeordnet. Dann nimmt man sich das nächste Paar virtueller Knoten vor, bis

³³ Sei $\#N$ die Anzahl der Kernknoten. Dann ist der Zeitaufwand zur Berechnung der Lösung auf keinem realistischen Maschinenmodell durch ein Polynom in Abhängigkeit von $\#N$ nach oben hin begrenzt. Funktionen, die diese Eigenschaft besitzen, sind für größere Eingabewerte praktisch nicht mehr berechenbar. Falls daher Probleme untersucht werden, bei denen größere Eingabewerte realistisch sind, wie dies hier der Fall ist, sind grundsätzlich zeitlich polynomial begrenzte Lösungen anzustreben.

schließlich keine Verbesserung der totalen Übertragungskosten mehr erzielt werden kann. Abschließend werden die Zuordnungen zu den physikalischen Knoten in Vereinigungen umgewandelt (vgl. ebenda).

Durch das Berücksichtigen der Schedules und hier insbesondere des später tatsächlich eingesetzten Anfragebearbeitungsalgorithmus während des Allokationsprozesses können effizientere Lösungen gefunden werden. Der Lösungsaufwand hierfür kann aber möglicherweise unvertretbar hoch sein.

3.4.3 Anwendungsbereich

Bei Apers werden, wie in den vorangegangenen Modellen, Kenntnisse zu den anfallenden Anforderungen (Anfragen sowie Updates) vorausgesetzt (s. o.). Als Minimierungsfunktion wählt er die Übertragungskosten, wobei nicht nur wie bei Ceri und Pelagatti die Zahl der entfernten Zugriffe, sondern auch die Datenmenge wie bei Dadam berücksichtigt wird. Eine Hinzunahme von Nebenbedingungen wäre möglich, ebenso die Anpassung an andere Minimierungsfunktionen.

Durch die Verzweigungsprozesse, die er in der Schedule-Bildung berücksichtigt, werden schreibbedingte Zugriffe auf Kopien exakt mit berücksichtigt. Datenübertragungen, die durch die Meldung von Anforderungen auftreten, vernachlässigt er im Gegensatz zu Dadam. In seinem zweiten Ansatz berücksichtigt er zudem als einziger die Tatsache, daß verschiedenen Allokationen unterschiedliche Schedules entsprechen können. Dabei ist vorteilhaft, daß der tatsächlich zum Einsatz kommende Algorithmus zur Anfragebearbeitung in seinem Modell berücksichtigt werden kann.

Auch wenn in keinem seiner Ansätze alle denkbaren Möglichkeiten verfolgt werden müssen, so scheint der Aufwand für sein Vorgehen sehr hoch. Allerdings umfaßt es ja auch zusätzlich die Entscheidung über die Kopienwahl, die in den anderen Fällen bereits festliegt.

3.5 Ein Vergleich der Modelle

Für dieses Kapitel wurden gezielt Modelle zur statischen Allokationsplanung ausgewählt. Zudem sollten die Modelle Redundanz zulassen. Ausnahmen in Hinblick auf den letztgenannten Aspekt wurden nur im Kontext darauf aufbauender Modelle mit Kopienhaltung betrachtet.

Eine statische Allokationsplanung setzt Kenntnisse über die Anforderungen an die Datenbank voraus. Je detaillierter hierbei der Planungsansatz ist, desto mehr und exaktere Angaben müssen vorliegen. Mehr, weil einfach eine größere Zahl an Eingabeparametern eingeht, exaktere, damit sich der zusätzliche Aufwand der durch die genaueren Berechnungen entsteht, nicht durch Ungenauigkeiten der Datenerhebungen wieder aufhebt.

Weitere Ungenauigkeiten, die allen Modellen inhärent sind, liegen in der fehlenden Modellierung exakter Synchronisations- und Kopienbehandlung (vgl. Abschnitt 2.3.3), wie es ja auch bereits von Özsü und Valduriez bemängelt wurde, auch wenn ihr alternatives Modell über die Verarbeitungskosten einen Schritt in diese Richtung weist. Ebenso bleiben in allen Fällen die meisten Faktoren des Netzes, wie Kommunikationskosten zur Netzbetriebung, Kosten durch Routing, Veränderungen der Datenmenge durch Komprimierungs- und Verschlüsselungsalgorithmen sowie Auslastung und Datenübertragungsgeschwindigkeit (vgl. Abschnitt 1.2.3, insbesondere S. 16) unberücksichtigt. Nur

die Kosten pro Dateneinheit für Datenübertragungen auf angenommenen Verbindungen zwischen je zwei Rechnern im Netz gehen, wenn überhaupt, in die Modelle ein. Dies erscheint vor allem in denjenigen Modellen, die die Menge der übertragenen Daten berücksichtigen (Dadam, Özsu, Apers), problematisch. So könnten Dateneinheiten, die aufgrund des Netzmanagements transferiert werden müssen (Protokollinformationen, Adressen etc.) kleine Datenmengen, die aufgrund der Datenbankanforderungen übertragen werden müssen, deutlich verzerren. Dementsprechend erscheint mir der Einsatz dieser Modelle ohne eine entsprechende Berücksichtigung von Netzgrößen nur bei einem durchschnittlich hohem Datenaufkommen pro Abfrage sinnvoll. Wird hingegen, wie bei Ceri und Pelagatti lediglich die Zahl der entfernten Zugriffe betrachtet, so taucht statt dessen als Problem auf, daß alle Anfragen, egal wie umfangreiche Datentransfers sie verursachen, gleich gewichtet werden.

Daß eine Einbeziehung der konkret vor Ort eingesetzten Anfragebearbeitungsalgorithmen in die Allokationslösung möglich ist, hat zumindest Apers aufgezeigt.

Die abweichende Abarbeitung schreibender und lesender Anforderungen wird unterschiedlich gut nachgebildet. Dabei berücksichtigen alle Modelle, daß bei schreibenden Zugriffen alle Kopien angepaßt werden müssen. Differenzen bestehen aber darin, inwieweit nur die entfernten Zugriffe (Ceri und Pelagatti) oder die Datenübertragungseinheiten (alle anderen) gezählt werden. Erstere berücksichtigen zudem über eine Konstante, daß schreibende Zugriffe aufwendiger sind als lesende. Özsu und Valduriez können diesen Aspekt im Rahmen der von ihnen modellierten Verarbeitungskosten noch genauer, spezifisch für jeden Knoten und nach der Größe der Datenmenge, auf die zugegriffen wird, abbilden.

Problematisch erscheint das Einbeziehen völlig unterschiedlicher Kosten in eine einzige Formel. Lassen sich wirklich Speicher- und Datenübertragungskosten, ggf. auch noch Verarbeitungskosten einfach aufsummieren oder sind hier nicht eher vergleichende Betrachtungen anzustellen? Eine Speichererweiterung dürfte leichter nachzuziehen sein als eine Netzanpassung. Gewichtungen, die über Konstanten einfließen können, so wie bei Ceri und Pelagatti zur Abwägung von schreibenden und lesenden Zugriffen, sind möglicherweise schwer festzulegen und können zu einer scheinbaren Objektivierung des Ergebnisses führen.

Darüber hinaus wurde bereits mehrfach die hohe Komplexität der Fragestellung und dementsprechend auch der Verfahren angeführt. Sowohl die Verfahren mit vollständigem Ausprobieren aller denkbaren Fälle als auch noch die heuristischen Ansätze, die dieses Vorgehen aus Aufwandsgründen vermeiden, sind von hoher Komplexität. Auch unter diesem Aspekt erscheint das Einbeziehen begrenzender Parameter sinnvoll. Zum einen reduzieren sie die Zahl der möglichen Allokationen zum anderen bilden sie wesentliche Aspekte der real geplanten Situation ab. In Verbindung mit den auf diesen Parametern aufsetzenden Nebenbedingungen können neben den Speicher- und Übertragungskosten auch Lastbalancierungsgesichtspunkte einfließen (vgl. oben, S. 26, Zitat von Rahm) und somit das Vermeiden der Anschaffung von Über- bzw. Unterkapazitäten möglich machen.

Eine Schwierigkeit bleibt, wie auch in [ÖV91] hervorgehoben, daß es leider keinen gemeinsamen Lösungsansatz für die zwei existierenden Optimierungswünsche

- Minimierung der Kosten (Speicherkosten, Bearbeitungskosten, Übertragungskosten) und
- Verbesserung des Datendurchsatzes und der Antwortzeiten

gibt (vgl. S. 137). Da sich die hier vorgestellten Modelle das erste Minimierungsziel gesetzt haben, ist es nicht verwunderlich, daß Aspekte der Parallelisierung von keinem der Autoren herausgearbeitet wurden. Allenfalls im „Additive-Replication“-Ansatz von Ceri und Pelagatti könnte der zweite

Gesichtspunkt über die Gewichtung der Kopienzahl grob modelliert werden. Bei Özsü und Valdúriez ließen sich über Grenzwerte für die einzelnen Knoten zumindest Obergrenzen erzwingen.

Offen bleibt in den herangezogenen Textgrundlagen auch, wie derartige Modelle in realen Situationen Anwendung finden. Nur in Apers Beitrag fanden sich Angaben zu simulierten Planungsvorgängen, diese waren aber von denkbaren realen Situationen weit entfernt. Veröffentlichungen aus dem Bereich der Wirtschaftswissenschaften belegen aber, daß zumindest Formen, wie sie von Dadam oder Özsü und Valdúriez vorgestellt wurden, Verwendung finden (vgl. z. B. [Pa85], [Wei83]). Zudem wird das nächste Kapitel aufzeigen, daß mathematische Modelle auch heute noch als Entscheidungshilfe bei der Allokationsplanung eingesetzt werden.

4 Der Aufbau einer neuen Datenbank zur Arbeitsvermittlung in der Bundesanstalt für Arbeit

Bei der Neugestaltung einer verteilten Datenbank zur Stellen- und Bewerbervermittlung in der Bundesanstalt für Arbeit (BA) wurden mathematische Modelle als Entscheidungshilfe in die Allokationsplanung mit einbezogen. Hierzu wurde von der Firma Siemens-Nixdorf eine Studie angefertigt, deren Vorstellung den eigentlichen Schwerpunkt dieses Kapitels und dieser Arbeit darstellt. In ihr werden mit Hilfe eines mathematischen Modells mehrere mögliche Allokationsformen verglichen. Zum besseren Verständnis der hierbei angestellten Überlegungen soll zunächst auf den Kontext, in dem diese Datenbankentwicklung steht, eingegangen werden.

Dazu wird in zwei einführenden Abschnitten kurz auf Organisationsstrukturen der BA eingegangen. Rechtliche Aspekte wurden hierbei ausgeklammert. Wer sich hierfür interessiert, findet diese in [HR94] kurz und übersichtlich dargestellt. Ein weiterer, vorausgehender dritter Abschnitt stellt schließlich knapp die heutige Softwarelösung und das Projekt, in dem die neue Datenbank im Rahmen einer neuen Softwaregeneration entwickelt wird, vor.

4.1 Organisationsstrukturen und Umstrukturierungen innerhalb der Bundesanstalt für Arbeit

Die Hauptstelle der Bundesanstalt für Arbeit befindet sich in Nürnberg. Ihr sind über das gesamte Bundesgebiet verteilt gegenwärtig 10 Landesarbeitsämter (Singular: LAA, Plural: LAÄ), 181 Arbeitsämter (AA, AÄ) mit weiteren 660 Geschäftsstellen und außerdem 16 besondere Dienststellen untergeordnet (vgl. auch Abbildung 10 und Abbildung 11).

„Die Hauptstelle stellt durch grundsätzliche Weisungen sicher, daß die fachlichen Aufgaben der BA im gesamten Bundesgebiet sachdienlich und einheitlich erfüllt werden.“ ([BA97], S. 12)

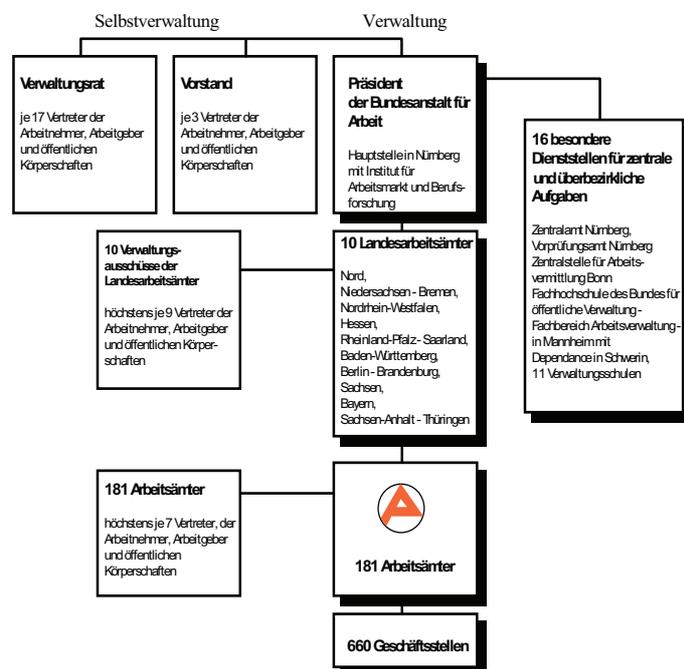


Abbildung 10: Organisation der BA
(nach [IT98], Teil A 1, S. 2)

Landesarbeitsämter und Arbeitsämter

Zeichenerklärung

- ▲ Sitz der Hauptstelle der BA (zugleich Sitz des Landesarbeitsamtes Nordbayern und des Arbeitsamtes Nürnberg)
- Sitz der Landesarbeitsämter (zugleich Sitz eines Arbeitsamtes)
- Sitz der Arbeitsämter

Niedersachsen - Bremen

- | | |
|--------------|---------------|
| Hannover | Nienburg |
| Braunschweig | Nordhorn |
| Bremen | Oldenburg |
| Bremerhaven | Osnabrück |
| Celle | Stade |
| Emden | Uelzen |
| Goslar | Verden |
| Göttingen | Wilhelmshaven |
| Hamel | |
| Helmstedt | |
| Hildesheim | |
| Leer | |
| Lüneburg | |

Nordrhein-Westfalen

- | | |
|-------------------|----------------|
| Düsseldorf | Recklinghausen |
| Aachen | Rheina |
| Ahlen | Siegen |
| Bergisch Gladbach | Soest |
| Bielefeld | Solingen |
| Bochum | Wesel |
| Bonn | Wuppertal |
| Brühl | |
| Coesfeld | |
| Detmold | |
| Dortmund | |
| Düren | |
| Duisburg | |
| Essen | |
| Gelsenkirchen | |
| Hegen | |
| Hamm | |
| Herford | |
| Iserlohn | |
| Köln | |
| Krefeld | |
| Meschede | |
| Mönchengladbach | |
| Münster | |
| Oberhausen | |
| Paderborn | |

Rheinland-Pfalz - Saarland

- | | |
|----------------|-------------|
| Saarbrücken | Mayen |
| Bad Kreuznach | Montabaur |
| Kaiserslautern | Neunkirchen |
| Koblenz | Neuwied |
| Landau | Pirmasens |
| Ludwigshafen | Saarlouis |
| Malnz | Trier |

Hessen

- | | |
|--------------|-----------|
| Frankfurt | Korbach |
| Bad Hersfeld | Limburg |
| Darmstadt | Marburg |
| Fulda | Offenbach |
| Gießen | Wetzlar |
| Hanau | Wiesbaden |
| Kassel | |

Baden-Württemberg

- | | |
|-------------|------------------------|
| Stuttgart | Pforzheim |
| Aalen | Rastatt |
| Balingen | Ravensburg |
| Freiburg | Reutlingen |
| Göppingen | Rottweil |
| Heidelberg | Schwäbisch Hall |
| Heilbronn | Tauberbischofsheim |
| Karlsruhe | Ulm (Donau) |
| Konstanz | Villingen-Schwenningen |
| Lörrach | Waiblingen |
| Ludwigsburg | |
| Mannheim | |
| Nagold | |
| Offenburg | |

Bayern

- | | |
|---------------|--------------|
| Nürnberg | München |
| Ansbach | Augsburg |
| Aschaffenburg | Deggendorf |
| Bamberg | Donauwörth |
| Bayreuth | Freising |
| Coburg | Ingolstadt |
| Hof | Kempton |
| Regensburg | Landshut |
| Schwandorf | Memmingen |
| Schweinfurt | Passau |
| Weiden | Pfarrkirchen |
| Weißenburg | Rosenheim |
| Würzburg | Traunstein |
| | Weilheim |

Nord

- Kiel
- Bad Oldesloe
- Elmshorn
- Flensburg
- Hamburg
- Helde
- Lübeck
- Neubrandenburg
- Neumünster
- Rostock
- Schwerin
- Stralsund

Berlin-Brandenburg

- Berlin
- I Berlin
- II Berlin
- III Berlin
- IV Berlin
- V Berlin
- VI Berlin
- VII Berlin
- VIII Berlin
- IX Berlin
- Cottbus
- Eberswalde
- Frankfurt/Oder
- Neuruppin
- Potsdam

Sachsen-Anhalt - Thüringen

- Halle/Saale
- Altenburg
- Dessau
- Erfurt
- Gera
- Gotha
- Halberstadt
- Jena
- Magdeburg
- Merseburg
- Nordhausen
- Sangerhausen
- Stendal
- Suhl
- Wittenberg

Sachsen

- Chemnitz
- Annaberg
- Bautzen
- Dresden
- Leipzig
- Oschatz
- Pirna
- Plauen
- Riesa
- Zwickau

Abbildung 11: Räumliche Gliederung der LAÄ und AÄ (vgl. S. 13 in [BA97])

Zu diesen Aufgaben zählen unter anderem die Arbeitsvermittlung, die Berufsberatung einschließlich der Vermittlung beruflicher Ausbildungsstellen sowie eine Vielzahl von mit dem Arbeitsmarkt in Verbindung stehenden finanziellen Leistungsgewährungen an Arbeitnehmer, Arbeitgeber oder Träger: Zum ersten Bereich der Leistungsgewährungen zählen z. B. Entgelt-

ersatzleistungen, darunter Arbeitslosengeld, Arbeitslosenhilfe, Kurzarbeitergeld und Konkursausfallgeld, aber auch Leistungen wie Überbrückungsgeld, Wintergeld oder Hilfen zur Aufnahme einer Beschäftigung. Beispiele für Leistungen an Arbeitgeber sind Beschäftigungshilfen für Langzeitarbeitslose, Eingliederungszuschüsse sowie Hilfen zur beruflichen Eingliederung Behinderter. Auch Träger von Aus- oder Weiterbildungsmaßnahmen können Hilfen, z. B. im Rahmen der Förderung von Arbeitsbeschaffungsmaßnahmen oder von Strukturanpassungsmaßnahmen bekommen. Zudem ist die BA für die Zahlung von Kindergeld sowie für die Arbeitsmarkt- und Berufsforschung zuständig.

Die der Hauptstelle zur Aufgabenerfüllung untergeordneten LAÄ wiederum koordinieren die Arbeitsämter mit ihren Geschäftsstellen in ihren fachlichen Aufgaben. Die besonderen Dienststellen schließlich übernehmen zentrale oder überbezirkliche Aufgaben.

Die gewachsenen hierarchischen Strukturen wurden von Hoffjan [Ho97] mit Bezug auf ([Ki81], S.121) und ([HLP92], S. 290) im Rahmen einer zusammenfassenden Betrachtung Controlling-relevanter Kontextvariablen kritisch untersucht: „Behindert wird die zielorientierte Führung durch den normengesteuerten Aufgabenvollzug in der Bundesanstalt. In Gestalt gesetzlicher und anordnungsrechtlicher Regelungen wird der Aufgabenvollzug zentral gelenkt. Die Durchsetzung der Normen wird institutionell durch **zentralistische Aufbauorganisation** abgesichert. Die internen Handlungs- und Entscheidungsstrukturen werden durch das Überwiegen hierarchischer Elemente und durch ein kontrollintensives Arbeitsklima zwischen Landesarbeitsämtern und Arbeitsämtern geprägt. Der hierarchischen Struktur der Weisungs- und Entscheidungsbefugnisse entspricht die Informationspolitik innerhalb der BA.“ Bringen wir diese Beschreibung mit dem Zitat von Brinckmann aus Abschnitt 1.1.2, Seite 5, in Verbindung, in dem darauf hingewiesen wird, daß gerade die oben angeführten Eigenschaften einem wirkungsvollen Einsatz von IT-Maßnahmen entgegenstehen, dann wird deutlich, daß den 1994 mit großer Intensität begonnenen Umstrukturierungsmaßnahmen innerhalb der BA auch aus EDV-Sicht eine sehr wichtige Rolle zukommt.

Diese Maßnahmen haben die Neuorganisation der Geschäftsprozesse in allen Dienststellen (vgl. [RdErl 24/94]) zur Aufgabe und laufen unter dem Stichwort „AA 2000“. Ihre Ausgangslage wird in einem internen Bericht (s. [AA96] / Anlage 2, S. 2), der Kurzfassung zur Grobbeschreibung des Organisationskonzeptes Arbeitsamt 2000, wie folgt beschrieben:

“Die gegenwärtige organisatorische Gliederung der Dienststellen der BA in Abteilungen (Sparten) richtet sich an einer fachlich-funktionalen Aufgliederung der übertragenen Aufgabenkomplexe aus. So sind die Arbeitsvermittlung und Arbeitsberatung, die Berufsberatung und die Gewährung von Leistungen in verschiedenen Abteilungen angesiedelt. Zwar ist es mit dieser Organisationsform gelungen, die Komplexität der Aufgaben und auch besondere Herausforderungen zu beherrschen, wie zum Beispiel Aufbau und Integration der Arbeitsverwaltung in den neuen Bundesländern zeigen. Sie hat aber zwangsläufig eine differenzierte arbeitsteilige Gliederung der Arbeitsprozesse und damit eine tayloristisch geprägte Struktur zur Folge - häufig mit mehreren Ansprechpartnern in verschiedenen Abteilungen für zusammenhängende Kundenanliegen. Bei der Suche nach verbesserten Organisationsformen ist verstärkt zu berücksichtigen, daß die Bürger eine möglichst **ganzheitliche** Erledigung ihrer Anliegen erwarten, die von der bisherigen Organisationsform so nicht geleistet werden kann.“

Hieran wird ersichtlich, daß die Bundesanstalt einen Handlungsbedarf durchaus erkannt hat. Um die in Abschnitt 1.1.2 bereits erwähnten Organisationsziele Kunden- und Mitarbeiterorientierung sowie Wirksamkeit und Wirtschaftlichkeit zu erreichen, wurde dem Handlungsfeld

„Einrichtung kundenorientierter Mitarbeiterteams“ eine Schlüsselrolle zugewiesen. Diese Teams übernehmen nun für den Kunden im Arbeitsamt anstelle früher völlig getrennter Abteilungen die weitestgehend ganzheitliche Erledigung seines Anliegen. Weitere Kernpunkte des Lösungsansatzes sind „**Dezentralisierung** und **Integration** der Aufgaben“ sowie „die Möglichkeiten zur **Selbstinformation** für die Kunden zu verbessern“ (ebenda).

Am 01.07.97 begannen zu diesen Umstrukturierungsmaßnahmen die ersten Erprobungen in vier Arbeitsämtern³⁴. Mit der anderen Aufgabenerledigung sind auch neue Verantwortlichkeiten eingeführt worden. Die Neuorganisation von Arbeitsprozessen und Verantwortlichkeiten umfaßt aber nicht nur die Strukturen in AA und Geschäftsstellen, sondern auch die von LAÄ und Hauptstelle.

Weitere Handlungsfelder, zu denen auch der Ausbau der Informationstechnik gehört, zeigt die folgende Abbildung.

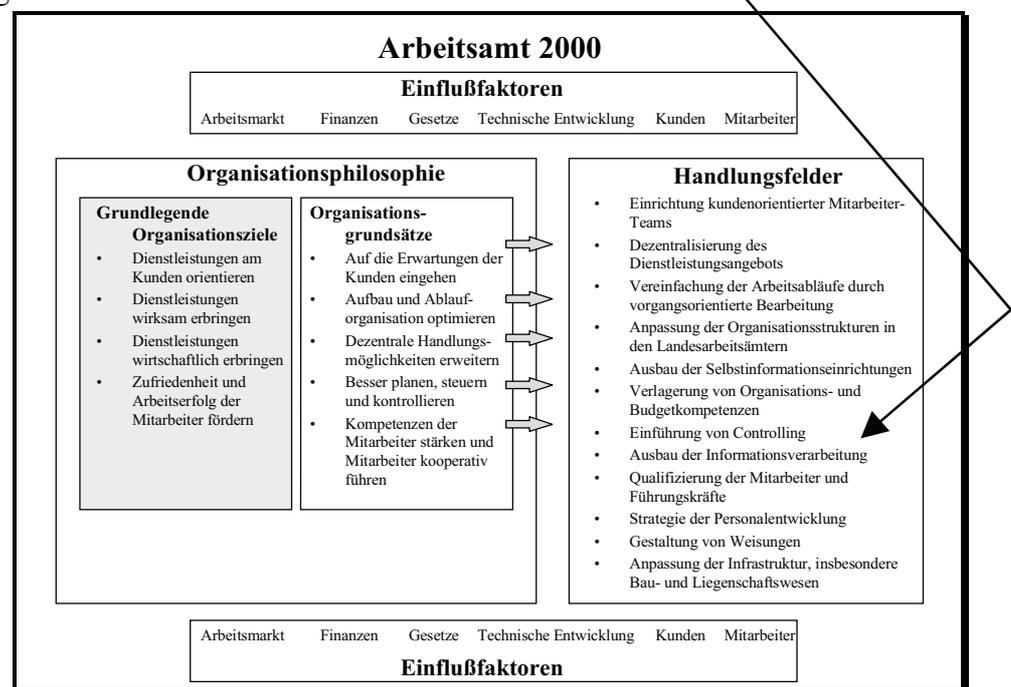


Abbildung 12:
Arbeitsamt 2000
- Einflussfaktoren, nach ([AA2000] / Anlage 2; dort Anhang II, Abb. 1)

Im Bereich der Informationstechnik hat sich die BA das Ziel gesetzt, EDV-Ausstattung und Verfahren dem heutigem Standard und natürlich auch den veränderten Organisationsstrukturen anzupassen. In ([AA96] / Anlage 2, S. 10) heißt es hierzu weiter:

„Die Umsetzung des Konzepts „Arbeitsamt 2000“, das eine ganzheitliche objekt- bzw. prozessorientierte Vorgangsbearbeitung vorsieht, erfordert **die Weiterentwicklung der Informations-Technik (IT)**. Lösungsansätze hierzu wurden aufgrund der Komplexität der Materie und des Erfordernisses, verstärkt informationstechnischen Sachverstand einzusetzen, nicht im Rahmen der Projektorganisation „Arbeitsamt 2000“ entwickelt. Die Weiterentwicklung der Informations-Technik ist jedoch mit dem Konzept „Arbeitsamt 2000“ abzustimmen und im Grundsatz „organisationsneutral“ zu gestalten.

Grundlegende **Defizite** der derzeitigen Anwendungssoftware - insbesondere uneinheitliche Anwendungs- und Datenstrukturen - können nur im Wege einer **grundsätzlichen Neugestaltung** behoben werden. Dies soll mit dem „IT-Konzept 2000“ erreicht werden. Eine ge-

³⁴ Dortmund, Halberstadt, Heilbronn und Saarbrücken

meinsame BA-Datenbasis mit einheitlichen Benutzer- (Bediener-) Oberflächen ist dabei ein wesentliches Ziel. Unter anderem soll ein BA-Netz alle wesentlichen Dienste der BA unterstützen und Anwender sowie Daten miteinander verbinden. Zur Umsetzung ist ein Stufenkonzept vorgesehen.“³⁵

Diese IT-spezifischen Aspekte werden in den folgenden Abschnitten im Mittelpunkt stehen.

4.2 EDV-Organisation und EDV-Ausstattungsmerkmale in der Bundesanstalt für Arbeit

Eine hierarchische Gliederung kennzeichnet auch die EDV-Organisation der Bundesanstalt. In der Hauptstelle findet die Planung und Bereitstellung sowohl der Infrastruktur der Informations- und Kommunikationstechniken als auch die der notwendigen Anwendungssoftware statt. Darüber hinaus wird von hier aus auch die Aus- und Fortbildung im informationstechnologischen Bereich geplant, gesteuert und koordiniert.

Das Zentralamt, eine der besonderen Dienststellen der BA, betreibt ein zentrales Rechenzentrum und ist zudem für die Umsetzungen der Infrastrukturvorgaben und für die systemtechnische Betreuung der dezentralen Rechenzentren zuständig. Darüber hinaus organisiert das Zentralamt die Aus- und Fortbildung der Fachkräfte und Anwender aus dem Bereich der Informationsverarbeitung (IV) und ist für die Entwicklung von IV-Schulungsmaterial zuständig. Ferner werden zentrale Schulungen angeboten.

Weitere Ausbildungsorte im IT-Bereich sind die Fachhochschule in Mannheim und die zugehörige *Dépendance* Schwerin sowie die Verwaltungsschulen. Sie sind organisatorisch direkt der Hauptstelle zugeordnet (vgl. Abbildung 10). Darüber hinaus finden aber auch zahlreiche Fortbildungen in den Ämtern vor Ort statt.

Die LAÄ führen die Aufsicht über die Infrastruktur ihres Bezirkes und stellen den Betrieb der Systeme für regionale und die eigenen Aufgaben sicher. Daneben erheben sie aus den einzelnen Meldungen der AÄ den Fortbildungsbedarf im IT-Bereich ihres Bezirkes.

Die Arbeitsämter schließlich betreiben eigene Rechenzentren, deren Rechner seit 1996 über ein bereits für die Zukunft ausgerichtetes Netz, das IBAN (Integriertes BA-Netz, vgl. Kapitel 1.2.3, Abbildung 1) miteinander und mit der Hauptstelle verbunden sind. Daneben sind sie für die Betreuung der Einzelplatzrechner und eigenständige Teilnetze vor Ort zuständig. Sie ermitteln zudem den IV-Fortbildungsbedarf vor Ort und leiten ihn an das LAA weiter. Da dem Netz im Rahmen der Allokationsüberlegungen eine besondere Bedeutung zukommt, soll es nun ein wenig ausführlicher beschrieben werden.

Das IBAN ist ein Weitverkehrsnetz, das alle lokalen Netze und hierarchischen Systeme in den Liegenschaften der BA untereinander verbindet. Jedes der über 60.000 Endgeräte (Terminals oder PCs) kann über eine eindeutige Adresse erreicht werden. Jedes Amt - und im Rahmen eines noch anstehenden weiteren Projektes auch jede Geschäftsstelle - kann hierüber mit einer Übertragungsgeschwindigkeit von 64 Kbit/sec bis 2 Mbit/sec, manche Standorte sogar mit bis zu 34 Mbit/sec versorgt werden. Es wird universell genutzt, d. h. für alle in der BA zum Tra-

³⁵ Dabei soll mit dem Begriff „organisationsneutral“ ausgedrückt werden, daß die neu entstehende Hard- und Softwarelandschaft so flexibel sein muß, daß sie sich vergleichsweise leicht an veränderte organisatorische Vorgaben anpassen kann.

gen kommenden Anwendungen. Dazu zählen interne Email, Dateitransfer, Datenaustausch, wie er im Rahmen der heute eingesetzten Verfahren erforderlich ist (vgl. Abschnitt 4.3), sowie zukünftig die Verbindung zu einem Firewall-Rechner, über den die Mitarbeiter Zugang zum Internet bekommen sollen. Gegenwärtig stehen hierfür noch Einzelplatzrechner zur Verfügung. Als Netzwerk- und Transportprotokoll wird das TCP/IP-Protokoll verwandt. Verschlüsselung und Komprimierung obliegen den Anwendungen selbst und erfolgen nicht im Rahmen der Übertragungsalgorithmen. Die Netzauslastung des WAN, das von einem externen Anbieter³⁶ betrieben wird, wird ca. alle 15 Minuten in allen Liegenschaften gemessen und zentral ausgewertet. Übersteigen oder unterschreiten die Daten gewisse Grenzwerte, so können Vertragsänderungen mit dem Netzbetreiber veranlaßt werden. Das Netz verfügt über zwölf sogenannte Points of Present (POPs), an denen die über das Netz gesendeten Nachrichten zwischengepuffert werden. Von Seiten der BA wurde aus Verfügbarkeitsanforderungen heraus ein adaptives Routing-Verfahren verlangt. Nach der Beschreibung des Netzverhaltens ist davon auszugehen, daß der Netzbetreiber (die Telekom), den Datenfluß mit virtuellen Adressen steuert.

Gegenwärtig findet ein Austausch der Großrechner-Terminal-Systeme in den Ämtern durch lokale Windows-NT-Netze statt. Die ehemaligen Terminals werden hierbei durch PCs (NT-Clients) ersetzt, vorhandene „Stand-alone-PCs“ ebenfalls weitestgehend als NT-Clients integriert. Die Netzverwaltung wird von NT-Domain-Servern übernommen. Eingebundene UNIX-Maschinen dienen der Datenhaltung für die alten Verfahren, die zunächst noch weiter mit Hilfe von Emulationsprogrammen eingesetzt werden. Als Netzbandbreite wurde in Hinblick auf die neue IT-Generation für die LANs 1 Mbit/s pro Endgerät definiert (für das WAN bis zu maximal 10Kbit/s pro Endgerät) und bei Neuverkabelungen und Anpassungen der Übertragungseinheiten berücksichtigt (vgl. [IT98], Teil A 4.2, Kap. 2 und 4). Bezüglich der weiteren IT-Infrastrukturgestaltung in Zusammenhang mit der neuen Softwaregeneration existieren zwar eine Reihe von Planungsstudien, eine Hardwareentscheidung ist jedoch noch nicht gefallen. Insbesondere ist noch offen, auf wie vielen und welchen Rechnern und Peripheriegeräten und in wie vielen Rechenzentren die neue, gemeinsame Datenbank installiert werden wird.

Die gegenwärtige Software-Situation wird in einem Artikel der BA-internen Zeitschrift DIALOG vom September/Oktober '95 wie folgt beschrieben:

„Die heutigen, über viele Jahre gewachsenen IT-Anwendungen sind vorwiegend auf die Belange der einzelnen Fachabteilungen zugeschnitten. Jede Anwendung hat ihre eigene Datenbasis. Die Nutzung mehrerer Anwendungen durch einen Mitarbeiter wird durch unterschiedliche Bedienerführungen erschwert. Der Arbeits- und Informationsfluß wird durch Medienbrüche beeinträchtigt.“³⁷ ([DIALOG], Jg. 2, Nr. 5, „Schöne neue Computerwelt“, Titelseite)

Dazu sei ergänzt, daß unterschiedliche Eingabefolgen für analoge Vorgänge leider auch innerhalb einzelner Verfahren vorkommen. Abkürzungen für Feldbezeichner sind aufgrund der nicht grafikfähigen Programme mit ihren geringen Auflösungsmöglichkeiten verbreitet; ein Abrufen der Langform ist jedoch im Normalfall nicht möglich. Hilfetexte sind ebenfalls nur in den neueren Verfahren in genügendem Umfang abrufbar. Inflexible interne Programmstrukturen haben dazu geführt, daß Anpassungen nur mit sehr hohem Aufwand erfolgen können. Folgen dieser Situation sind neben Redundanzen in den Daten und Programmen und den damit verbundenen Mehrfacheingaben und komplizierteren und aufwendigeren Wartungsarbeiten

³⁶ Gegenwärtig von der Telekom

³⁷ Fachabteilungen sind Arbeitsvermittlung, Leistungsgewährung, Berufsberatung, Verwaltung

auch in einem unverhältnismäßig hohen Schulungsbedarf zu sehen. Eine fachübergreifende Arbeitserledigung kann mit diesen Verfahren nicht sinnvoll unterstützt werden.

Deshalb wurde 1995 ein Projekt PlaIT200 (Planung der Informationstechnik 2000) ins Leben gerufen. Um der erforderlichen, Ressort-übergreifenden Softwareentwicklung auch organisatorisch gerecht zu werden, wurde hierfür zunächst die Projektform und darüber hinaus ein ausgelagerter Entwicklungsstandort gewählt. Inzwischen wurden sogar gänzlich neue Organisationsstrukturen für den IT-Bereich eingeführt.

Ziel des 1997 in die Realisierungsphase übergewechselten Planungsprojektes ist es, die durch die Neuorganisation in den AÄ veränderten Geschäftsprozesse zukünftig durch eine ihnen gerecht werdende EDV zu unterstützen. Praktische Erprobungen des nun unter dem Projektnamen IT2000 laufenden Entwicklungsvorhabens werden in einem Modellamt in Ingolstadt seit September '97 durchgeführt und sollen Ende '98 abgeschlossen sein. Die hier in den verschiedenen Umsetzungsstufen gemachten Erfahrungen fließen in die Weiterentwicklung der neuen Software sowie in die Planung der Gesamteinführung ein.

In allen Dienststellen werden dann auf den die ehemaligen Terminals ersetzenden NT-Clients Standard-Office-Programme verfügbar sein. Außerdem wird von jedem Arbeitsplatz aus die prinzipielle Möglichkeit bestehen, über das oben beschriebene, interne BA-Netz (IBAN) mit jedem am Netz angeschlossenen Rechner der BA beliebig Daten auszutauschen. Dabei wird es aber von der Software abhängen, in welchem Umfang dies tatsächlich unterstützt wird. Dieser Aspekt wird in den nächsten Abschnitten im Rahmen des dort betrachteten Teilproblems der Vermittlung noch deutlicher werden. Zudem gelten für die Übertragung sensibler Daten, wie z. B. personenbezogener Daten oder für Paßwörter Verschlüsselungsvorschriften. Außerdem wird es möglich sein, mittels des IBAN über einen Firewall-Rechner auf das Internet zuzugreifen. Der hierfür verwandte Browser wird auch im BA-Intranet verwandt werden, so daß hier eine Homogenität der Oberfläche gegeben ist.

4.3 Software zur Unterstützung der Arbeitsvermittlung

Innerhalb der BA gibt es eine Vielzahl von zentralen³⁸ und dezentralen³⁹ Verfahren, die die unterschiedlichen Aufgaben der BA unterstützen. Die Arbeitsvermittlung wird heute im wesentlichen durch ein dezentrales Verfahren auf den in den AÄ stehenden Großrechnern⁴⁰ unterstützt. Bei diesem Programm, das den Namen coArb (**computerunterstützte Arbeitsvermittlung**) trägt, handelt es sich um ein menüorientiertes Dialogverfahren. Hiermit werden für die Arbeitsvermittlung relevante Bewerber- und Stellendaten erfaßt, sowie statistische Daten ermittelt. Wichtigste Aufgabe dieses Programms ist es, Suchen auf den Datenbeständen zu unterstützen. Dabei unterscheidet man **bewerberorientierte Suchen**, die von einem konkreten Bewerberangebot (BewA) ausgehend passende Stellen suchen, **stellenangebotsorientierte Suchen**, bei denen analog von einem konkreten Stellenangebot (SteA) ausgegangen wird und hierzu passende Bewerber ermittelt werden. (Auf den folgenden Seiten sind einige Bildschirmansichten des Verfahrens coArb hierzu abgebildet.) Ebenso existiert die Möglichkeit, unabhängige Suchen zu starten, die z. B. zur Einschätzung des Arbeitsmarktes sinnvoll sein können. Eine Suche wird heute grundsätzlich durch einen Mitarbeiter initiiert und **erfolgt ausschließlich auf dem Datenbestand des Zentralrechners des zugehörigen Arbeitsamtes**. Fernübertragungen sind nur erforderlich, sofern das Terminal in einer ausgelagerten Geschäftsstelle steht. Über das Programm werden weiterhin Vermittlungs- und Beratungsvermerke aufgenommen sowie Standardschreiben und Termingestaltung unterstützt. Gewisse Daten, die leistungsrelevant sind, werden automatisch an das Verfahren zur Leistungsgewährung weitergeleitet. Hierzu zählt z. B. die Wiederaufnahme von Arbeit eines Bewerbers, der Arbeitslosengeld bezieht.

Jeder in der Vermittlung tätige Mitarbeiter hat über einen Terminal (per Identifizierung und Authentifizierung) Zugang zu diesem Verfahren. Ein Zugriff auf Vermittlungsdaten, die nicht auf dem zugehörigen AA-Rechner liegen, ist nicht möglich. Deshalb ist für die Suchergebnisse heute **entscheidend, welcher Datenbestand sich auf dem AA-Rechner befindet**. Das sind heute **die dort erfaßten Bewerber- und Stellenangebote sowie über einen zentralen automatisierten Austausch alle Stellen spezieller Berufe⁴¹ und des Tagespendelbereiches**. Letzterer ist für jedes Amt über eine Menge von Dienststellennummern regional benachbarter Arbeitsämter definiert. Arbeitsorte der zu diesen Ämtern gehörenden Gebiete können im allgemeinen durch tägliches Pendeln vom Wohnort aus erreicht werden. Gehört beispielsweise das AA 1 zum Tagespendelbereich der Ämter AA 2, AA 4 und AA 5, so werden die täglich in AA 1 neu erfaßten Stellen über einen nächtlichen Datentransfer via Zentralamt den Rechnern der anderen Ämter zur Verfügung gestellt. Unabhängig vom Pendelbereich läßt das Verfahren zu, ein Stellenangebot in bis zu drei⁴² weiteren Arbeitsämtern mitzuführen. Auch dieses Replizieren erfolgt nachts über das zentrale Rechenzentrum in Nürnberg. Da die Datenübertragungen nur nachts erfolgen, sind die hierdurch ausgetauschten Daten nie tagesaktuell. Dassel-

³⁸ In der BA werden diejenigen Programme als „zentrale Verfahren“ bezeichnet, die im Zentralamt im Nürnberg, ggf. unter Einbeziehung von Daten aus den Ämtern, ausgeführt werden.

³⁹ Hierunter werden die in den Ämtern lokal eingesetzten Programme verstanden. In vielen Fällen übertragen diese Programme zudem Daten an zentrale Verfahren.

⁴⁰ Gegebenenfalls erfolgt die Lösung auch über mehrere eng gekoppelte Großrechner (in Shared Everything oder Shared Disk Architektur). Dies spielt für unsere Betrachtungen aber keine Rolle.

⁴¹ Hierbei handelt es sich um solche Berufe, für die überregionale Nachfragen aufgrund hoher Qualifikationsanforderungen oder aufgrund der beruflichen Anforderungen typisch sind.

⁴² Die Oberfläche läßt nicht mehr Eingaben zu! Inwieweit diese Restriktion ursprünglich einmal aufgrund der zu erwartenden Datenübertragungsbegrenzungen oder schlicht aus Gestaltungsaspekten der Oberfläche heraus gewählt wurde oder welche sonstigen Entscheidungen hierzu führten, ist mir unbekannt.

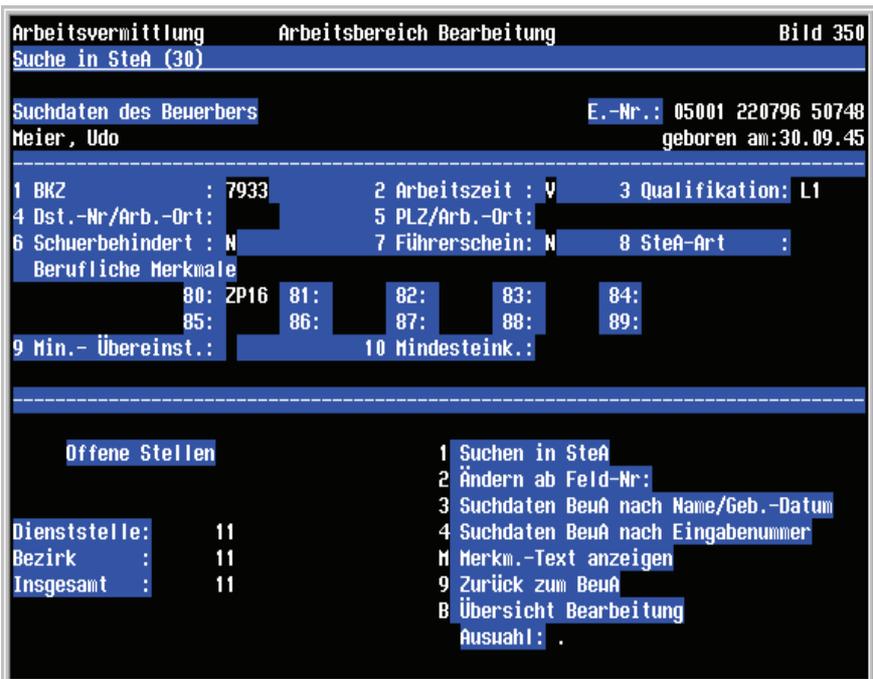
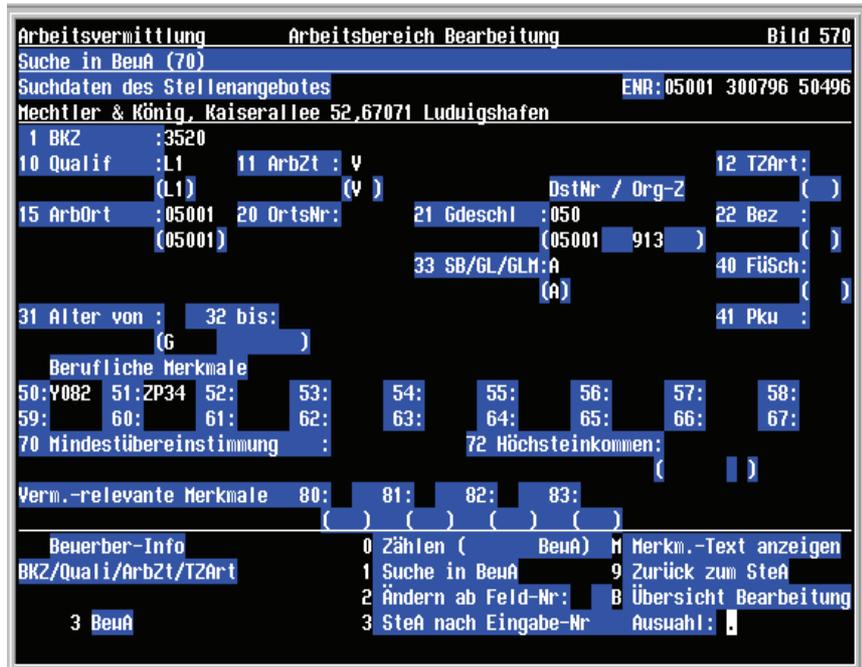


Abbildung 14:
„Suchbilder“ in coArb
a) Suche in SteA

Ruft man in den in Abbildung 13 dargestellten Kurzformen den Menüpunkt 1 auf, so werden die dem BewA bzw. SteA entsprechenden Daten automatisch in die hier abgebildeten Suchbilder eingetragen. Diese Suchen nennt man bewerberangebots- bzw. stellenangebotsorientiert.

b) Suche in BewA

Der Arbeitsvermittler hat in diesen Bildern noch Veränderungsmöglichkeiten, um das Suchergebnis weiter einschränken oder möglicherweise vergrößern zu können. Außerdem kann er sie nicht nur über ein BewA oder SteA, sondern auch direkt aufrufen und selbstständig ausfüllen. So können unabhängige Suchen gestaltet werden.



Neben dem Verfahren coArb gibt es in der BA weitere Programme zur Arbeitsvermittlung, die sogenannten Selbstinformationssysteme. Dies sind Verfahren, die es der Öffentlichkeit ermöglichen, eigene Recherchen vorzunehmen.

In den Kundenzonen der Arbeitsämter haben Arbeitsuchende über das Verfahren SIS (Stelleninformationssystem) die Möglichkeit, eigene Suchen auf dem Stellenbestand des jeweiligen Amtes durchzuführen und Ergebnisse auszudrucken. Allerdings werden nur solche Stellen in den Suchen berücksichtigt, für die eine hierzu erforderliche Zustimmung vorliegt. Die Abfragen können die Interessenten über grafikfähige Bildschirmarbeitsplätze mit einer aus wenigen Steuerungstasten bestehenden Tastatur gestalten. Die Bedienung ist einfach. Es bestehen

allerdings auch nicht derartig vielfältige Suchmöglichkeiten, wie sie den Mitarbeitern offen stehen. Nur regionale und berufsfachliche Suchwege werden unterstützt. Wendet sich ein Arbeitssuchender nach einer eigenen Recherche im Amt an einen Arbeitsvermittler, so kann dieser die Suche über coArb nachvollziehen, da beiden Verfahren derselbe Datenbestand zugrunde liegt.

Die BA bietet der Öffentlichkeit darüber hinaus weitere eigene Recherchemöglichkeiten zu Stellen- und Bewerberangeboten an. Diese sind zugänglich über „T-Online *Arbeitsamt#“ oder alternativ im Internet unter der Adresse „<http://www.arbeitsamt.de>“ abrufbar. Auch hier müssen Arbeitgeber bzw. Arbeitssuchende der Veröffentlichung ausdrücklich zustimmen. Das Verfahren zur Stellensuche heißt ebenfalls SIS und ist der dezentralen, d. h. der in den Ämtern zum Einsatz kommenden Lösung, sehr ähnlich. Als Arbeitgeberinformationssystem, kurz AIS wird das Verfahren zur Bewerbersuche, die ebenfalls nach regionalen oder berufsfachlichen Kriterien erfolgt, bezeichnet.⁴³

Die für SIS und AIS benötigten Daten werden auf einem Rechner in Nürnberg bereit gehalten. Ihre Erfassung erfolgt über das Verfahren coArb in den Ämtern, wobei zur Veröffentlichung vorgesehene Angebote speziell gekennzeichnet werden. Dies gilt auch dann, wenn Arbeitgeber die Möglichkeit nutzen, ein Stellenangebot per Internet an das regional zuständige Arbeitsamt zu schicken. Nachts findet ein Transfer der zur Veröffentlichung gekennzeichneten, neu erfaßten oder veränderten SteA-Daten statt. BewA-Daten werden per Datenträger geschickt! Da über das Internet bundesweit Angebote zusammenfließen - zumindest die zur Veröffentlichung freigegebenen - können gegenwärtig über Internet Suchergebnisse erzielt werden, die an den diversen Standorten mit coArb nicht aufzurufen wären! Auf diesen Gesichtspunkt wird bei der Beurteilung möglicher Lösungen am Ende dieses Kapitels noch einmal zurückzukommen sein. Die Veröffentlichung der Angebote erfordert zudem eine besondere Sorgfalt der Mitarbeiter der BA in Hinblick auf die Aktualität der Angebote. Regelmäßige und häufige Kontakte zu Bewerbern und Stellenanbietern sind notwendig, um nicht mehr vakante Angebote auch zügig aus dem System zu entfernen, denn nur ein aktueller Datenbestand wird auf Dauer gut angenommen werden und sich etablieren können. Entsprechend werden andere Übertragungswege dringend erforderlich.

Die BA hat die Notwendigkeit der Entwicklung einer gemeinsamen, neuen Datenbank für die Vermittlung erkannt. An ihrer Modellierung wird gegenwärtig gearbeitet. Neben einer Beschreibung der Anforderungen, die auf diesem Datenbestand zu erfolgen haben, wird auch, wie wir in Kapitel 2 gesehen haben, die Allokation der Daten dieser Datenbank einen leistungs- und kostenbestimmenden Faktor darstellen. Deshalb wurden bereits im Planungsprojekt PLaIT2000 hierzu Anfragen an diverse Softwarefirmen gestellt. Eine in diesem Kontext entstandene mathematische Studie zu denkbaren Allokationen für eine verteilte Datenbank für die Vermittlung soll im folgenden Abschnitt vorgestellt werden.

4.4 Die Studie von Siemens Nixdorf zur Allokationsplanung

Im Rahmen der Entwicklung der neuen Softwaregeneration wurde bereits in der Planungsphase des IT-Projekts 2000 von einem der möglichen Lösungsanbieter, der Firma Siemens

⁴³ Darüber hinaus gibt es Speziallösungen zur Vermittlung von Künstlern sowie zur Unterstützung einer Stellensuche im Ausland, z. B. durch Adressen von Partnerorganisationen aber auch durch Vermittlungsangebote.

Nixdorf, in einem Vortrag aufgezeigt, daß methodische Allokationsplanungen möglich sind und für eine Planung der neuen, verteilten Datenbank herangezogen werden könnten. Mit diesem Anspruch (- nicht mit dem Anspruch sämtliche mögliche Lösungen für eine neue Datenbank vorzustellen -) entstand die Studie zur Allokationsplanung ohne Auftrag durch die BA. Von Seiten der Firma SNI wollte man zeigen, daß man einer systematischen Planung gewachsen sei und bei Wunsch von Seiten der BA auch eine umfangreichere Allokationsplanung würde ausführen können (laut Auskunft von [Kr99]).

Die Unterlagen zu diesem Vortrag sind in den Literaturangaben unter [SNI96] angeführt und zudem im Anhang als erste Anlage zu finden. Innerhalb des Kapitels 4 wird bei einem Bezug auf diese Quelle daher nur noch auf Anlage 1 verwiesen. Dazu wird dann die Originalseitenzahl der Quelle angegeben. Da die Unterlagen aus zwei Teilen, zum einen einer textuellen Beschreibung, zum anderen aus Abbildungen, Tabellen und Modellrechnungen bestehen und beide Teile mit einer Numerierung von 1 beginnen, wird bei einem Verweis auf den zweiten Teil ergänzend „Teil 2“ angegeben.

Die Studie soll hier exemplarisch als Beispiel für den Einsatz mathematischer Modelle in realen Softwareentwicklungsprozessen vorgestellt werden. **Die Studie untersucht die in der Vermittlung täglich anfallenden Transaktionen auf dem Datenbestand sowie die hieraus resultierenden Netzbelastungen für insgesamt sieben mögliche Allokationsszenarien.** Diese berücksichtigen die gegenwärtig vorhandenen Rechenzentren in den Arbeitsämtern, Landesarbeitsämtern und im Zentralamt.

In der Studie heißt es:

„Es gibt nun mehrere Möglichkeiten die Daten auf die verschiedenen Rechner zu verteilen. Die Art der Verteilung bestimmt das Zusammenwirken der einzelnen Rechner und u. a. auch

- die Verfahren der Speicherung der Daten
- die Methoden des Suchens in den Daten
- die Treffermengen
- die Anforderungen an die Verfügbarkeit der Verbindungen zwischen den Rechnern
- die Performance des Gesamtsystems“

(Anlage 1, S. 13).

Auffallend ist, daß behauptet wird, die Treffermengen seien von der Allokation der Daten abhängig. Warum dies so ist und an welchen Stellen vereinfachende Annahmen, wie sie bei Modellbildungen allgemein üblich sind, gemacht wurden, soll im folgenden Abschnitt herausgearbeitet werden, der damit die angenommenen konzeptuellen Voraussetzungen der Studie wiedergibt. Nach einem weiteren Abschnitt, in dem die mathematische Modellbildung der Studie beschrieben wird und der somit der formalen Untersuchung der diversen Szenarien gewidmet ist, wird schließlich vorgestellt, welche Variante von der Studie favorisiert wird.

Abgeschlossen wird das Kapitel mit einer Beschreibung der Grenzen des SNI-Modells. Hierbei wird eine Einordnung der Studie in die in Kapitel 3 vorgestellten mathematischen Modelle vorgenommen, so daß die dort herausgestellten Resultate auf die Studie übertragen werden können. Darüber hinaus werden zum einen Anmerkungen zu Formeln, die des besseren Verständnisses wegen schon während der Darstellung der Studie getroffen werden, wieder aufgegriffen, zum anderen aber auch generelle Überlegungen zum konzeptionellen Ansatz angestrengt.

4.4.1 Die Voraussetzungen

Im Modellansatz der Studie werden sowohl stellenorientierte Suchen im BewA-Datenbestand als auch bewerberorientierte Stellensuchen betrachtet. Diese Suchen besitzen gegenüber den oben ebenfalls bereits erwähnten freien Suchen bei der täglichen Arbeit deutliche Priorität.

Um die Vorgänge in den Ämtern zu modellieren, sind einige konzeptuelle Annahmen zu treffen. So werden eine Reihe von Objekten benötigt, zwischen denen zudem bestimmte Beziehungen angenommen werden (vgl. Tabelle 6). Es wird davon ausgegangen, daß Stellen grundsätzlich eindeutig einer anbietenden Firma oder einer deren Niederlassungen zugeordnet werden können. Die Betriebsinformationen werden unabhängig von angebotenen Stellen benötigt⁴⁴. Jede Firma und jede Niederlassung wiederum ist eindeutig dem Zuständigkeitsbereich eines Arbeitsamtes zugeordnet. Dies gilt auch für Arbeitsuchende. Ebenso läßt sich jede Stelle und jeder Arbeitsuchende⁴⁵ einem Berufsbereich zuordnen. Über diese Zuordnungen ist geregelt, welches Team aus Sachbearbeitern für die Bearbeitung der Suchvorgänge zuständig ist. Diese Annahmen entsprechen in etwa den heute existierenden Gegebenheiten.

n : 1		1 : 1	
• Stelle - Firma	bzw. Stelle - Niederlassung	• Berufsbereich - Team aus Sachbearbeitern	
• Firma - AA	bzw. Niederlassung - AA	• Team - Teamdatei ⁴⁶ mit zu bearbeitenden Vorgängen	
• Arbeitsuchender - AA			
• Stelle - Berufsbereich			
• Arbeitsuchender - Berufsbereich			

Tabelle 6: Grundlegende Objekte der SNI-Studie und ihre Beziehungen

Jedes Team kann sowohl stellenorientierte als auch bewerberorientierte Vermittlungen durchführen. Die Suche erfolgt über einen Abgleich der beruflichen und regionalen Kriterien. Die folgende Annahme macht schließlich deutlich, daß das Suchergebnis in der Tat von der regionalen Zuordnung der Daten und damit je nach Verteilung auch von der Lokalität der Daten abhängen kann:

„Grundsätzlich wird bei der Suche so vorgegangen, daß

- zunächst versucht wird, die Suche im eigenen Arbeitsamt zu befriedigen
- und nur dann, wenn dies nicht möglich ist, andere Arbeitsämter oder ein zentrales System einzubeziehen.“ (Anlage 1, S. 6).

Nur wenn sich im Zuständigkeitsbereich des eigenen Amtes nicht genug Lösungen finden lassen (vgl. folgenden Absatz), können also auch Lösungen aus Datenbeständen von entfernten

⁴⁴ Die von Stellenangeboten unabhängige Datenhaltung der Betriebsinformationen von Betrieben mit sozialversicherungspflichtig Beschäftigten ist durch das Sozialgesetzbuch (SGB4, 2. Datenerfassungsverordnung und 2. Datenübermittlungsverordnung vom 29.05.89) sowie durch den Einigungsvertrag für die neuen Bundesländer gesetzlich geregelt.

⁴⁵ Dies ist eine vereinfachte Sichtweise, sofern ein Arbeitsuchender mehrere verschiedene berufliche Qualifikationen aus unterschiedlichen Berufsbereichen aufweisen sollte. Aber auch in diesem Falle ließe sich eine eindeutige Zuordnung zu dem Berufsbereich des Hauptwunsches gestalten.

⁴⁶ Warum das Modell hierfür Dateien betrachtet und diese Informationen nicht als Bestandteil in die Datenbank einbezieht, geht aus der Studie nicht hervor.

Rechnern angezeigt werden. Der Wechsel des zuständigen Arbeitsamtes wäre daher im Normalfall (bei Vorhandensein lokaler Lösungen) mit anderen Ergebnissen verbunden⁴⁷.

Für alle Suchen ist die Anzahl der Suchergebnisse durch einen Maximalwert nach oben begrenzt. Bei Stellenangeboten kann die maximal gewünschte Bewerberzahl auch durch den Betrieb vorgegeben werden. Diese Grenze darf nicht überschritten werden. Suchen für Bewerber werden beim Erstkontakt so mit ihm abgestimmt, daß die erzielte Ergebnismenge dem Bewerber so gerecht wie möglich wird. Als Besonderheit wird deshalb diese mit dem Bewerber abgestimmte Suchabfrage für spätere Suchen gespeichert.

Den Suchen liegt ein räumliches Konzept zugrunde. Sowohl jedes Stellenangebot als auch jedes Bewerberangebot erhält einen Pendelbereich zugeordnet. Dies soll die räumliche Umgebung sein, aus der eine Firma in Frage kommende Arbeitsuchende wünscht oder die Umgebung eines Arbeitsuchenden, in der ihm zuzumuten⁴⁸ ist, eine Stelle anzunehmen. Dabei wird davon ausgegangen, daß die Pendelbereiche von Firmen im allgemeinen weit größer sind als die von Bewerbern. Pendelbereiche sind generell als Suchbedingung formulierbar. Über eine notwendige Teilmengenbeziehung zwischen Stellen- und Bewerberpendelbereichen wird keine Aussage getroffen, sie scheint aber nicht zwingend zu sein. Interessant in Hinblick auf die Suchen sind Schnittmengen. Bei der Ausgestaltung dieses Konzeptes sind unterschiedliche Feinheitsgrade denkbar. So könnten Pendelbereiche über Angabe von Arbeitsämtern bzw. Rechnern, deren Original-Datenbestand berücksichtigt werden soll oder aber über das Abbilden tatsächlich gewünschter Regionen (Städte, Stadtteile, Straßen, ...) gelöst werden. Im zweiten Falle sind deutlich höhere Anstrengungen erforderlich um schnelle Suchergebnisse sicherzustellen. Die Studie macht hierzu keine näheren Angaben. In der Beschreibung der Szenarien wird allgemein vom „Stellen-Pendelbereich“ und „Bewerber-Pendelbereich“ gesprochen, in den Formeln mit durchschnittlich anzutreffenden Werten für die Zahl der Arbeitsämter, die zu einem Stellen- oder Bewerber-Pendelbereich gehören, gearbeitet. Ein Beispiel denkbarer Pendelbereiche (anscheinend individuell festlegbar und feiner als Arbeitsamtszuständigkeitsbereiche) findet man im Anlage 1, Teil 2 auf der Abbildung Seite 1. Im Altverfahren coArb wird demgegenüber mit an das Amt gebundenen Pendelbereichen gearbeitet (vgl. S. 67). Auf den Aspekt, daß dieses Konzept zwar in vielen Fällen zutreffen mag, aber nicht alle denkbar möglichen Vermittlungswünsche von Firmen oder Bewerbern abdeckt, wird im Abschnitt 4 noch eingegangen.

Unter diesen Voraussetzungen kann in den meisten Szenarien eine Reihe von Suchen bereits lokal gelöst werden. Der Rest wird je nach Szenario und Suchrichtung (stellen- oder bewerberorientiert) unterschiedlich über entfernte Anfragen gelöst.

Das SNI-Modell berücksichtigt Bewerber- und Stellensuchen, die bei Neuaufnahme und in regelmäßigen Zeitabständen sowie durch vom Vermittler vorgegebene Zeiten (z. B. zu folgenden Beratungsterminen, bei Spontanbesuchen) initiiert werden⁴⁹. Außerdem gehen regel-

⁴⁷ Diese (letzte) Eigenschaft gilt auch für das gegenwärtig im Einsatz befindliche Verfahren coArb. Ob sie aus heutiger Sicht auch für eine neu zu entwickelnde Software sinnvoll ist, wird noch in Abschnitt **Fehler! Verweisquelle konnte nicht gefunden werden.** erörtert.

⁴⁸ Dieser Begriff ist geprägt von der Sicht, daß Leistungsempfänger nach Stellen suchen. Hier ist per Gesetz unter anderem eine regionale Zumutbarkeitsgrenze definiert. Die heute im Altverfahren coArb festgelegten Tagespendelbereiche (vgl. S. 67) der Arbeitsämter umfassen diese Region. Damit ist sichergestellt, daß einem Leistungsempfänger alle möglichen passenden täglich erreichbaren Stellen angeboten werden können. Über die regionalen Kriterien lassen sich (sowohl in coArb als auch vom Modellansatz der Studie her) aber genauso bei Nicht-Leistungsempfängern regionale Wünsche eingeben.

⁴⁹ Es besitzt daher Züge eines aktiven Informationssystems.

mäßige Auswertungen zu noch offenen und nicht mehr vakanten Stellen in das Modell ein, wenn auch nur vereinfacht über einen konstanten Faktor. Genauere Ausführungen hierzu folgen an entsprechender Stelle bei Beschreibung der Formeln.

Nach allen Suchen wird grundsätzlich noch manuell nachbereitet. Dies erfolgt durch das zuständige Team. Die Zuständigkeit für die Bearbeitung der gefundenen Suchergebnisse ist grundsätzlich über die Bewerber geregelt. Bei stellenorientierten Suchen und bei nicht selbst ausgelösten bewerberorientierten Suchen erfährt das Team durch die eindeutig zugeordnete Teamdatei (vgl. Fußnote 46) von vorliegenden Suchergebnissen. In diesen Fällen wird für jedes gefundene Paar aus Bewerber und Stellenangebot ein Eintrag mit einem Verweis auf beide Datenobjekte in diejenige Teamdatei geschrieben, die dem gefundenen Bewerber zugeordnet ist. (Dies ist eine abgeleitete 1:1-Beziehung). Bei bewerberorientierten Suchen, die das Team unmittelbar selbst auslöst, ist ein solcher Eintrag nicht erforderlich. Zur Nachbereitung gehört bei entsprechender Entscheidung die Benachrichtigung der Bewerber (schriftlich oder telefonisch) sowie das Eingeben der Daten über diesen Vermittlungsvorschlag (Zähler im Stellenangebot erhöhen, Bewerber- bzw. Stellenvorschläge eintragen). Ebenso werden in den Bewerberdaten manuell verworfene Lösungen vermerkt, damit diese nicht wiederholt auftreten können.

Nach einer manuell ausgewerteten Suche legen die Sachbearbeiter Folgeberatungen fest. Hierzu können sie gezielt terminierte Suchen veranlassen, die ihrer Teamdatei zugeordnet werden. Besetzte Stellen werden für die Suche gesperrt. Das Modell geht davon aus, daß sie u. a. für statistische Auswertungen noch im Bestand bleiben (vgl. Anlage 1, S. 10) und sieht daher nur die Löschung von Mehrfachspeicherungen vor.

Das mathematische Modell betrachtet sieben verschiedene Möglichkeiten der Datenallokation. Grundsätzlich wird berücksichtigt, an welchem Rechner eine Transaktion gestartet wurde. Das Speicherkonzept der folgend kurz vorgestellten Szenarien sieht in sechs Fällen ausschließlich eine lokale, einmalige Speicherung der Bewerberdaten vor. Nur im siebten Fall werden bei den Bewerberangeboten auch Replikatе gehalten, dies jedoch nur von den vermittlungsrelevanten Daten. Für die Stellenangebote differieren die Varianten bezüglich der Replikatе stärker. Grundsätzlich wird aber immer davon ausgegangen, daß ein Server Hauptstandort des Stellenangebotes ist. Hier ist immer die komplette Information zum Stellenangebot zu finden. Für die stellenangebotsorientierte Suche wird in allen Verfahren ausschließlich auf diesem Rechner, in zwei Szenarien auch auf den Rechnern des Stellen-Pendelbereiches gesucht.

1. Zentrales Verfahren (SZ 1)

Dieses Szenario nimmt eine Sonderrolle ein. Es dient vor allem der Lastberechnung ohne Datenverteilung. Es wird davon ausgegangen, daß alle Daten, auch die Teamdateien zentral gehalten werden, die PCs in den Ämtern sind direkt mit dem hierfür bereitgestellten leistungsfähigen Zentralrechner verbunden, die vermittlerischen Aktivitäten finden auf dem Zentralrechner (unter Berücksichtigung des Pendelbereiches) statt.

2. Zentrale Stellenverwaltung, dezentrale Bewerberverwaltung (SZ 2)

Alle Stellendaten werden zusätzlich zentral gehalten, ansonsten werden Stellen- und Bewerberdaten auf genau einem lokalen Rechner im Amt gehalten. Notwendige Einträge in die Stellendaten erfolgen zentral und dezentral. Stellenorientierte Suchen unter den Bewerberangebo-

ten finden nur lokal statt, bewerberorientierte Stellensuchen finden zunächst lokal, bei zu geringer Ergebnismenge auch zentral (unter Berücksichtigung des Pendelbereichs) statt.

3. Teilweise zentrale Stellenverwaltung, dezentrale Bewerberverwaltung (SZ 3)

Der Ansatz ist hier wie beim zweiten Szenario, aber $X\%$ der Stellen - bestimmte Berufsgruppen, bei den Modellrechnungen um 50% - werden nur lokal geführt. Das bedeutet, daß hier auch die diese Berufsgruppe betreffenden bewerberorientierten Suchen nur lokal stattfinden.

4. Suche im Pendelbereich (SZ 4)

Alle Bewerberdaten werden auf genau einem lokalen, dezentralen Arbeitsamtsrechner gespeichert. Dies gilt ebenso für fast alle Stellenangebote. Ein geringer Teil an Stellenangeboten bestimmter Berufsgruppen - ca. 5% - wird allerdings auch zusätzlich zentral verwaltet. Dieser wird in der Studie mit $(100 - X^*)\%$ gekennzeichnet. Die PCs sind mit dem AA-Server, dieser wiederum mit dem zentralen Rechner verbunden. Alle dezentralen Server haben Zugriff auf die Daten der umliegenden Arbeitsämter des Pendelbereichs. Die Informationen über die Zuordnung von Stellen und Bewerbern werden lokal und nur für bestimmte Berufsgruppen auch zentral geführt.

5. Speicherung der Stellen im Pendelbereich (SZ 5)

Alle Bewerberdaten werden auf genau einem lokalen dezentralen Arbeitsamtsrechner gespeichert. Wieder werden $(100 - X^*)\%$ der Stellenangebote nach bestimmten Berufsgruppen wie im vorigen Szenario im zugeordneten Amt und zudem zentral verwaltet. Zu allen anderen Stellenangeboten werden bei der Ersterfassung Replikate auf allen Servern der Arbeitsämter des Pendelbereichs allokiert. Die Zuordnungsinformation zwischen Stellenangeboten und Bewerberangeboten erfolgt immer nur auf dem jeweils zuständigen Amtsrechner.

Hierdurch sind alle bewerberorientierten Suchabfragen im Pendelbereich lokal, Zuordnungen erfordern Netzübertragungen. Damit ein Stellenangebot bei erreichter Maximalzahl an Vorschlägen nicht berücksichtigt wird, wird es, wenn dieser Fall eintritt, in allen Replikaten gekennzeichnet. Eine erneute Aktivierung ist möglich. Diese Transaktionen sind im Modell nicht berücksichtigt, da es davon ausgeht, daß sie asynchron erfolgen.

6. Speicherung der Stellen im LAA (SZ 6)

Hier wird zwischen die Rechner der AA und dem zentralen Rechner noch eine weitere Ebene von LAA-Rechnern eingeführt. Die Bewerberdaten werden nur lokal (disjunkt) gehalten. Nur auf diesem Bestand erfolgt die stellenorientierte Bewerbersuche. Die Stellendaten werden lokal und im Rechner des zuständigen LAA gehalten. Stellen, die auch bei Szenario 4 und 5 zentral waren, werden auch hier zusätzlich zentral gespeichert. Die Zuordnung wird lokal, in den LAA-Daten und auch zentral geführt. Bei der bewerberorientierten Suche wird je nach Suchergebnis im lokalen Datenbestand, in den LAA-Daten oder zentral gesucht.

7. Speicherung der Stellen- und Bewerberdaten im Pendelbereich (SZ 7)

Hier werden diejenigen Bewerberdaten, die für die Vermittlung unabdingbar sind, sowie Stellendaten repliziert im Pendelbereich allokiert. Hierbei können Pendelbereiche von Bewerbern und Stellenangeboten divergieren. Durch diese Replikate entfallen nun auch Netzbelastungen

bei stellenorientierten Bewerbersuchen. Die Zuordnungsinformationen erfolgen nur an einem lokalen Rechner. Das Schreiben in die Teamdatei und das Update der Stellendaten erfolgt über das Netz.

Auf der folgenden Seite sind alle Szenarien noch einmal in einer tabellarischen Übersicht wiedergegeben.

Speicherkonzept und Suchstrategie in den diversen Szenarien

SZ	Ebenen	Bewerberangebote (BewA)	ZI	Stellenangebote (SteA)	ZI	TD	Suchstrategie SteA->BewA	Abh.	Suchstrategie BewA->SteA	Abh.
1	1	ZR	ja	alle Originaldaten	ja	ja	zentral	nein	zentral	nein
2	2	ZR		100% (Gesamtbestand) als Replikate	ja	-	nur lokal auf dem AA-R.	ja	erst lokal auf dem AA-R., ggf. zentral	nein
		AA	ja	Originale (Zerschlagung)	ja	ja				
3	2	ZR		100 - X% (Teilbestand nach Berufsgr.) als Replikate, z. B. X = 50%	ja	-	nur lokal auf dem AA-R.	ja	wie bei 2 für 100-X% X% nur lokal auf dem AA-R.	ja ja
		AA	ja	Originale (Zerschlagung)	ja	ja				
4	2	ZR		100 - X*% (Teilbestand nach Berufsgr.) als Replikate, X* ≈ 95	ja	-	auf den AA.-R. des PB	ja	erst lokal auf dem AA-R., ggf. im PB oder zentral	ja
		AA	ja	Originale (Zerschlagung)	ja	ja				
5	2	ZR		100 - X*% (Teilbestand nach Berufsgr.) als Replikate, X* ≈ 95	zum Teil	-	auf den AA-R. des PB	ja	erst lokal auf dem AA-R., ggf. zentral	ja
		AA	ja	Originale (Zerschlagung)	ja	ja				
				Replikate im PB (ohne zentrale Berufe)	nein	-				
6	3	ZR		100 - X*% (Teilbestand nach Berufsgr.) als Replikate, X* ≈ 95	ja	-	nur lokal auf dem AA-R.	ja	erst lokal auf dem AA-R., ggf. auf dem LAA-R., ggf. zentral	ja
		LAA		Replikate aller SteA der zug. AA-R.	ja	-				
		AA	ja	Originale (Zerschlagung)	ja	ja				
7	2	ZR		100 - X*% (Teilbestand nach Berufsgr.) als Replikate, X* ≈ 95	zum Teil	-	nur lokal auf dem AA-R.	ja	erst lokal auf dem AA-R., ggf. zentral	ja
		AA	ja	Originale und	ja	ja				
			nein	Replikate der VrD im PB	nein	-				

Erläuterung der Abkürzungen

AA: Arbeitsamt	LAA: Landesarbeitsamt	SteA: Stellenangebot	VrD: Vermittlungsrelevante Daten
Abh.: Abhängigkeit der Treffermengen von der Allokation der Daten	PB: Pendelbereich	SteA->BewA: stellenorientierte Bewerbersuche	ZR: Zentralrechner
Berufsgr.: Berufsgruppen	R.: Rechner	SZ: Szenario	ZI: Zuordnungsinformationen
BewA->SteA: bewerberorientierte Stellensuche		TD: Teamdatei	- : nein, (Schlußfolgerung)
zum Teil: es wird nur gespeichert, ob die gesuchte Zahl an Bewerbern bereits erreicht wurde oder noch nicht			

Tabelle 7: Szenarien-Übersicht

4.4.2 Das mathematische Modell der Studie

Das von der Firma Siemens Nixdorf aufgestellte mathematische Modell - im folgenden zur leichteren Abgrenzung gegenüber den in Kapitel 3 vorgestellten mathematischen Modellen als SNI-Modell bezeichnet - ermittelt zunächst einmal die erforderlichen Transaktionen auf den Daten und hieraus Aussagen über die Transaktionsrate, die jeder Rechner der ggf. verteilten Lösung zu erbringen hat. Außerdem werden die entfernten Zugriffe betrachtet, um den jeweiligen Umfang der Netzlast der diversen Szenarien beurteilen zu können.

Abweichend zu Kapitel 3 möchte ich an dieser Stelle nicht vorab die benötigten Variablenbezeichner einführen, da das SNI-Modell sehr viele verwendet. Sukzessive, wie erforderlich, werden sie erklärt. Zudem habe ich - um ein vergleichendes Nachschlagen in der Quelle im Anhang zu erleichtern - die Bezeichner wie im Original belassen und keine Anpassung an die in Kapitel 3 verwandten Bezeichner vorgenommen. **Fett** habe ich diejenigen Bezeichner dargestellt, die Ausgangsvariablen darstellen, d. h. für die gemessene oder zur Kalkulation mögliche denkbare Werte eingesetzt werden müssen. Alle anderen Variablen werden aus diesen Werten ermittelt. Der besseren Lesbarkeit dient auch wieder eine herausnehmbare Übersicht über alle Bezeichner als Anlage 3.

Im SNI-Modell werden Formeln für die Zahl der erforderlichen Transaktionen an den diversen Rechnerebenen (AA, LAA, zentral) sowohl für die Durchführung der stellenorientierten Bewerbersuche als auch für die bewerberorientierte Stellensuche aufgestellt. Hierzu wird außerdem eine Beispielberechnung durchgeführt. Diese berücksichtigt abschließend auch eine Aufteilung nach Tageszeitabschnitten und die Netzbelastung. Ich folge den in der Studie gemachten Ausführungen, fasse jedoch, sofern möglich, abweichend zum Original die Formeln der Szenarien 2 bis 7 zusammen. Wenn hierfür oder ansonsten an einigen Stellen weitere Variablen erforderlich wurden, so habe ich deren Bezeichner *kursiv* dargestellt; bei der Wahl des Namens habe ich mich am Original orientiert. Da das zentrale Szenario 1 stärker von den anderen abweicht und formelmäßig auch besonders einfach zu handhaben ist, werden die hierfür zutreffenden Formeln gesondert aufgeführt.

4.4.2.1 Die stellenorientierte Bewerbersuche

Die Formeln zur stellenorientierten Bewerbersuche berücksichtigen die Neuerfassung, Suchvorgänge, das Eintragen von Lösungen in die Teamdateien, die Entscheidungsfindung der Sachbearbeiter, Folgebearbeitungen und das Löschen von Stellen.

Neuerfassung

Bei der Neuerfassung einer Stelle wird davon ausgegangen, daß auf dem Amtsrechner die Betriebsdaten des Stellenanbieters bereits vorliegen (vgl. Abschnitt 4.4.1, S. 72). Damit werden für eine Neuerfassung Transaktionen nur zum Lesen der zugehörigen Betriebsdaten, auf die ein Verweis eingetragen werden muß sowie für die Speicherung der Stellendaten, zu denen dann natürlich auch der herausgefundene Verweis gehört, erforderlich. Möglicherweise kommen noch weitere Speicherungen für Replikate hinzu.

Pro Tag bedeutet dies für das Szenario 1, bei dem bundesweit alle vermittlungsrelevanten Daten auf einem Zentralrechner abgelegt werden, ein Transaktionsaufkommen entsprechend der durchschnittlichen täglichen Anzahl neu einzugebender Stellenangebote (TN) multipliziert mit dem durchschnittlichen Transaktionsaufwand zur Erfassung eines Stellenangebotes (transSpN),

der sich, wie gerade erwähnt, aus den durchschnittlichen Transaktionszahlen zum Lesen der Betriebsdaten (**transSuFi**) und zur Speicherung der Stellendaten (**transSpN***) ergibt.

$$\text{TransSpzenN} = \text{transSpN} \times \text{TN} = (\mathbf{transSuFi} + \mathbf{transSpN*}) \times \text{TN} \quad (\mathbf{SZ 1})$$

mit:

TransSpzenN	durchschnittliche Anzahl täglich anfallender Transaktionen zur zentralen Speicherung von SteAs
transSpN	durchschnittliche Anzahl der Transaktionen zur Erfassung eines SteAs
N	Anzahl der Stellen, die bundesweit jährlich neu besetzt werden sollen. Diese Stellen werden gespeichert.
AT	Anzahl der Tage, in denen ein Arbeitsamt im Jahr geöffnet ist
TN	durchschnittliche Anzahl der täglich im Bundesgebiet neu zu erfassenden SteAs, $\text{TN} = \text{N} / \text{AT}$
transSuFi	durchschnittliche Anzahl der Transaktionen zum Lesen der Betriebsdaten
transSpN*	durchschnittliche Anzahl der Transaktionen zur reinen Speicherung eines SteAs inklusive des Verweises auf die zugehörigen Betriebsdaten

Für die Szenarien 2 bis 7 wird zum einen die Zahl der lokal sich durchschnittlich in einem „mittleren Arbeitsamt“ ergebenden Transaktionen (**TransSpmN**) betrachtet, zum anderen die durchschnittliche Zahl der Transaktionen, die durch zusätzliche zentrale Speicherung entstehen. Von einem „mittleren Arbeitsamt“ wird gesprochen, da hier die Zahl **TmN** herangezogen wird. Diese liefert das arithmetische Mittel über die bundesweit durchschnittlich täglich neu zu erfassenden Stellen bezogen auf die Anzahl vorhandener Rechner auf Arbeitsamtsebene (**RAA**).

Folgende Bezeichner werden im folgenden benötigt:

RAA	Anzahl der Rechner auf Arbeitsamtsebene. Diese ist nicht identisch zur bundesweiten Anzahl an Arbeitsämtern, da manche Ämter kooperieren.
LAA	Anzahl der Rechner auf Landesarbeitsamtsebene in Modell 6. Diese ist identisch zur bundesweiten Anzahl an LAÄ.
TmN	durchschnittlich täglich neu zu erfassende Stellenangebote in einem mittleren Arbeitsamt, $\text{TmN} = \text{TN} / \mathbf{RAA} = \mathbf{N} / \mathbf{AT} / \mathbf{RAA}$
q	durchschnittliche Anzahl der Arbeitsämter im Pendelbereich von Stellen
I_{SZ}	I_{SZ} = 1 für alle Szenarien, deren Ordnungszahl in der Menge SZ, $\text{SZ} \subseteq \{2, \dots, 7\}$ aufgeführt ist, ansonsten 0. Verkürzt werden die Mengenklammern von SZ weglassen.
TransSpmN	durchschnittliche Anzahl täglicher Transaktionen zur Speicherung neuer Stellendaten in einem „mittleren“ Arbeitsamt
TransSpLAAN*	durchschnittliche Anzahl täglicher Transaktionen zur Speicherung von Replikaten neu erfaßter Stellenangebote auf einem LAA-Rechner (SZ 6)
X	prozentualer Anteil der in SZ 3 nur lokal angebotenen Stellen
X*	prozentualer Anteil in SZ 4 - 7 nur lokal angebotener Stellen
X	prozentualer Anteil in SZ 2 - 7 nur lokal angebotener Stellen, $X = 0$ für SZ 2, $X = \mathbf{X}$ für SZ 3, $X = \mathbf{X*}$ für SZ 4 - 7

Die durchschnittliche Anzahl täglich lokal anfallender Transaktionen zur Speicherung neuer Stellendaten setzt sich aus zwei Komponenten zusammen (s. Formel (F) weiter unten). Zunächst einmal ist der Aufwand zur Speicherung der Originaldaten zu berücksichtigen, wobei der Aufwand für das Lesen der Betriebsdaten zur Gestaltung der Zuordnungsinformation ein- geht.

Andererseits kann Aufwand für das Speichern von Stellen durch Anforderungen anderer Ämter aus dem Stellen-Pendelbereich entstehen.

Dies gilt für die Szenarien 5 und 7, für die in der folgenden Formel durch die Indikatorvariable I_{SZ} eine Anzahl von q solchen Anforderungen berücksichtigt werden, wobei q die durchschnittliche Zahl an Arbeitsämtern eines Stellen-Pendelbereichs wiedergibt. Da es sich bei q aber um einen Durchschnittswert handelt, werden regionale Unterschiede, wie sie in Flächenländern oder dicht besiedelten Gebieten auftreten können, sowie die individuell durch die Betriebe festlegbaren Pendelbereichsgrenzen verwischt. Die in der Studie angegebene Formel berücksichtigt nicht, daß im Pendelbereich bei den Szenarien 5 und 7 nach der textuellen Beschreibung nur der nicht zentral gehaltene Stellenbestand verteilt wird. Korrekter wäre, hier mit $(T_{mN} \times X\%)$ zu multiplizieren, wobei mit X der prozentuale Anteil der nicht zentral gespeicherten Stellen angegeben ist.

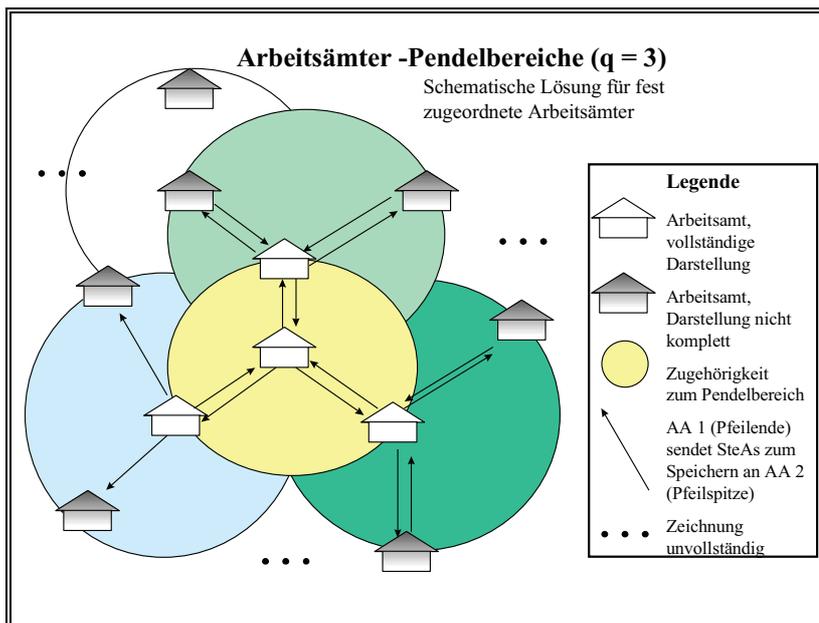


Abbildung 15: Arbeitsämter - Pendelbereiche

$$[(F^*) \text{ TransSpmN} = \text{transSpN} \times T_{mN} + I_{5,7} \times q \times \text{transSpN}^* \times X\% \times T_{mN}]$$

(F) für Originaldaten für Stellendaten der Ämter des Tagespendelbereich

$$\text{TransSpmN} = (\text{transSpN} + I_{5,7} \times q \times \text{transSpN}^*) \times T_{mN} \quad (SZ 2 - 7)$$

$$= (\text{transSuFi} + \text{transSpN}^* + I_{5,7} \times q \times \text{transSpN}^*) \times$$

SZ	X
2	0
3	= X, z. B. ≈ 50
4	= X* ≈ 95
5	= X* ≈ 95
6	= X* ≈ 95
7	= X* ≈ 95

T_{mN}

T_{mN}

$$= [\text{transSuFi} + \text{transSpN}^* \times (1 + I_{5,7} \times q)] \times$$

Die zusätzlich in den Szenarien 2 bis 7 anfallenden Speicherungen auf zentralen Rechnern lassen sich mit folgender Formel beschreiben, in der X den Prozentsatz der nicht zentral gespeicherten Stellendaten angibt.

$$\text{TransSpzenN}^* = \text{transSpN}^* \times TN \times (100 - X\%) \quad (SZ 2 - 7)$$

Man beachte, daß durch TN hier im Gegensatz zu obiger Formel wieder die im gesamten Bundesgebiet täglich neu eingehende Stellenanzahl berücksichtigt wird.

Im Szenario 6 sind zudem noch die Rechner in den LAÄ zu betrachten. Hierfür geht die Formel aus den zur Studie gehörigen Abbildungen (Anlage 1, Teil 2, S.11) hervor:

$$TransSpLAAN^* = \mathbf{transSpN^*} \times TN / \mathbf{LAA}$$

(SZ 6)

Da der gesamte Stellenbestand auf die LAÄ verteilt wird, geht die Formel von den bundesweit durchschnittlich täglich neu eingehenden Stellendaten aus. Regionale Spezifika bleiben unberücksichtigt.

Löschen

Da die Formeln für den Transaktionsaufwand zum Löschen denen zum Speichern sehr ähnlich sind, werden sie bereits an dieser Stelle aufgeführt. Es wird von dem Normalfall ausgegangen, daß die Betriebe den Ämtern melden, wenn eine Stelle besetzt wurde und daraufhin der Vermittler das Löschen der Stellendaten durchführt. Hierunter wird in der Studie verstanden, daß die Originaldaten auf den lokalen Rechnern bzw. auf dem Zentralrechner im Szenario 1 als vermittelt gekennzeichnet werden und sämtliche sonstigen Replikate tatsächlich gelöscht werden. Von den Originaldaten wird angenommen, daß sie noch für statistische Zwecke zur Verfügung stehen müssen. Ebenso kann es vorkommen, daß Stellenangebote wieder reaktiviert werden, etwa, wenn eine Probezeit zur Auflösung des getroffenen Arbeitsverhältnisses führt oder wenn es sich um saisonale Stellenangebote handelt. Hierdurch entstehende Auswirkungen auf die Neuerfassungen sowie Überlegungen zum hieraus resultierenden zusätzlichen Speicherbedarf werden in der Studie nicht angestrengt. In Hinblick auf den Speicherbedarf ist aber davon auszugehen, daß in der Realität endgültig vermittelte, in absehbarer Zeit nicht mehr benötigte Stellenangebote nach Eingang in die Statistik vollständig aus dem Bestand entfernt werden. Dies wäre asynchron möglich und stellt daher keinen sehr kritischen Vorgang dar.

Die folgenden Formeln ähneln denen zum Speichern sehr, aber die durchschnittliche Zahl an Transaktionen zum Speichern muß nicht mit der zum Löschen übereinstimmen. **TransLöN** ist der Aufwand, der für das Setzen des Kennzeichens „vermittelt“ erforderlich ist, **transLöN*** der Aufwand für das wirkliche Löschen eines Replikates.

Die Formeln zum Löschen lauten:

$$TransLözenN = \mathbf{transLöN} \times TN \quad (\mathbf{SZ 1})$$

$$TransLömN = (\mathbf{transLöN} + I_{5,7} \times q \times \mathbf{transLöN^*}) \times TmN \quad (\mathbf{lokal SZ 2 -7})$$

$$TransLözenN^* = \mathbf{transLöN^*} \times TN \times (100 - X)\% \quad (\mathbf{zentral SZ 2 -7})$$

$$TransLöLAAN^* = \mathbf{transLöN^*} \times TN / \mathbf{LAA} \quad (\mathbf{LAA SZ 6})$$

Als neue Variablenbezeichner wurden benötigt:

TransLözenN	durchschnittliche Anzahl täglich anfallender Transaktionen zum Löschen von zentral gelagerten Originalstellendaten in SZ 1
transLöN	durchschnittliche Anzahl der Transaktionen zur Kennzeichnung eines SteAs als „vermittelt“
transLöN*	durchschnittliche Anzahl der Transaktionen zum wirklichen Löschen eines SteAs (wird nur für Replikate vorgenommen)
TransLömN	durchschnittliche Anzahl täglicher Transaktionen zum Löschen von Stellendaten in einem „mittleren“ Arbeitsamt
TransLözenN*	durchschnittliche Anzahl täglicher Transaktionen zum Löschen zentral replizierter SteAs

*TransLöLAAN** durchschnittliche Anzahl täglicher Transaktionen auf einem LAA-Rechner zum Löschen replizierter SteAs (SZ 6)

Interessant ist, daß in den Formeln für das Löschen die Variable TN, also die täglich bundesweit anfallende Zahl neu zu erfassender Stellen, eingeht. Damit geht das Modell von einer ausgeglichenen Vermittlungssituation aus.

Auch bei diesen Fällen wurde (folgerichtig) nicht berücksichtigt, daß in den Szenarien 5 und 7 nur die nicht zentral gespeicherten Stellen auf den Rechnern des Pendelbereichs verteilt werden. Hat man bereits beim Speichern die alternative Formel gewählt (vgl. F*, S. 80), so ist ein analoges Vorgehen für den Löschvorgang natürlich ebenso erforderlich. In diesem Falle würde die Formel für den lokalen Löschaufwand in den Szenarien 2 bis 7 wie folgt lauten:

$$[(F^{**}) \quad \text{TransLöMn} = \text{transLöN} \times \text{TmN} + I_{5,7} \times q \times \text{transLöN}^* \times \text{TmN} \times X\%]$$

Suchvorgänge

Auf jede Ersterfassung eines Stellenangebotes erfolgt eine stellenorientierte Bewerbersuche. Für stellenorientierte Bewerbersuchen gibt es eine maximale Obergrenze **O** für die Anzahl der Suchergebnisse. Diese kann als Mittelwert über Grenzen, die für die unterschiedlichen Berufsgruppen sinnvoll sein mögen, sowie über unterschiedlichste Arbeitgeberwünsche angesehen werden. In einem bestimmten Prozentsatz (**r1**) der Suchen wird diese Grenze bereits auf dem lokalen Datenbestand erreicht, für einen verbleibenden prozentualen Anteil **r2** werden in den Szenarien 4 und 5 weitere Suchen auf entfernten Rechnern angestoßen, wobei **r1 + r2 = 100** gelte⁵⁰. Wenn man wie zuvor das Transaktionsaufkommen der diversen Rechner wissen möchte, so sind auf den Arbeitsamtsrechnern in den Szenarien 4 und 5 mit Anfragen im Pendelbereich (SZ 4, 5) auch die von ihnen ausgehenden **r2%** entfernten Anfragen sowie die von anderen Rechnern ausgelösten Suchen (durchschnittlich **q × r2%**) zu berücksichtigen.

Im zentralen Fall ist die Formel wieder am einfachsten. Hier entstehen Suchtransaktionen entsprechend der Anzahl täglich bundesweit neu aufkommender Stellenangebote:

$$\text{TransSuzenB} = \text{transSulokB} \times \text{TN} \quad (\text{SZ 1})$$

mit:

TransSuzenB	durchschnittliche Anzahl täglich anfallender Transaktionen für stellenorientierte Bewerbersuchen auf dem Zentralrechner in SZ 1
transSulokB	durchschnittliche Anzahl an Transaktionen für eine stellenorientierte Bewerbersuche

Dabei ist in **transSulokB** auch die Anzahl an Transaktionen berücksichtigt, die möglicherweise für Probeversuche zur Gestaltung der genauen Suchbedingungen erforderlich ist.

Für die Szenarien 2 bis 7 gilt demgegenüber: (SZ 2 - 7)

$$\text{TransSumB} = \text{TmN} \times (\text{transSulokB} + \underbrace{I_{4,5} \times (1 + q)}_{\text{an andere Rechner}} \times \underbrace{r2\%}_{\text{von anderen Rechnern}} \times \text{transSuremB})$$

⁵⁰ Dieser Formel könnte die Annahme zugrunde liegen, daß bei Stellenbeschreibungen enge regionale Vorgaben untypisch sind. Dann kommen Fälle, bei denen entfernte Suchen nicht mehr in Betracht zu ziehen sind, weil die regionalen Bedingungen dort nicht mehr erfüllbar sind, praktisch nicht vor.

mit:

TransSumB	durchschnittliche Anzahl täglicher Transaktionen für stellenorientierte Bewerbungen auf einem Rechner eines "mittleren" Arbeitsamtes
transSuremB	durchschnittliche Anzahl an Transaktionen, die für eine entfernte Suche oder durch eine Suche von einem entfernten Rechner im Pendelbereich im Rahmen der stellenorientierten Bewerbersuche erforderlich werden. Die Studie nimmt an: transSuremB = 1.
O	durchschnittliche Obergrenze für die Anzahl der zu suchenden Bewerber
r1	prozentualer Anteil der stellenorientierten Bewerbersuchen, die lokal vollständig bis zur Obergrenze gelöst werden können
r2	prozentualer Anteil der stellenorientierten Bewerbersuchen, die entfernt angestoßen werden müssen um (ggf.) die Obergrenze noch zu erreichen
	$r2 = 100 - r1$

Für die Rechner der LAÄ ist keine gesonderte Betrachtung erforderlich, da in Szenario 6 die stellenorientierte Bewerbersuche ausschließlich lokal erfolgt. Für die Variable **transSuremB** wird angenommen, daß sie normalerweise den Wert 1 annimmt, „da die Suchfrage bereits bei der lokalen Suche festgelegt wurde. Es wird davon ausgegangen, daß das Suchen in den anderen Rechnern aus der Sicht des lokalen Rechners **eine** verteilte Transaktion ist, es entstehen jedoch zusätzliche Transaktionen auf den anderen Rechnern. Dies bedeutet, daß am Anfang der Transaktion die Suchanfrage an alle Server im Pendelbereich geschickt wird und danach auf die Ergebnisse von den einzelnen Servern gewartet wird. Besteht zu einem der Rechner keine Verbindung, so wird dort nicht gesucht. Es wird davon ausgegangen, daß eine nicht vorhandene, abgehende Verbindung in Zukunft ein seltenes Ereignis ist.“ (Anlage 1, S. 31) Ebenso geht aus der Formel hervor, daß auf anderen Rechnern ausgelöste Suchen bezüglich der Transaktionszahl genauso behandelt werden wie eingehende Suchanfragen anderer Pendelbereichsrechner.

Schreiben in die Teamdatei

Bei der stellenorientierten Bewerbersuche wird für jeden gefundenen Bewerber ein Eintrag in die Teamdatei geschrieben, die zu dem Team gehört, dem der Bewerber zugeordnet ist. Jeder solche Eintrag zählt als eine Transaktion. Für die Modellierung wird die Variable L zur Beschreibung der durchschnittlichen Anzahl gefundener Bewerber pro Stelle benötigt. In der Studie wird aufgrund der gegenwärtigen Situation auf dem Arbeitsmarkt (berufsunspezifisch) davon ausgegangen, daß vornehmlich $L = O$ gilt, also die maximal zulässige Anzahl möglicher Bewerber gefunden wird (vgl. ebenda, S. 33). Aus den Formeln und der Definition der Variablen läßt sich schließen, daß L bereits auf dem Bestand der Daten des Pendelbereiches erreicht wird, denn für **L1** und **L2**, die für den jeweiligen prozentualen Anteil bereits lokal gefundener bzw. aus dem Zuständigkeitsbereich anderer Ämter gefundener Bewerber stehen, soll gelten: $L1 + L2 = 100$. Die Anzahl der Transaktionen für die Schreibvorgänge in die Teamdateien ergibt sich wie folgt:

$$\text{TransWrzenTD} = L \times \text{TN}$$

(SZ 1)

$$\begin{aligned}
\text{TransWrmTD} &= \mathbf{L1\%} \times L \times \text{TmN} \\
&\quad + \mathbf{I_{4,5,7}} \times \mathbf{L2\%} \times L \times \text{TmN} \\
&\quad \quad \quad + \mathbf{I_{4,5,7}} \times \mathbf{L2\%} \times L \times \text{TmN} \\
(\mathbf{F^{***}}) \quad &= (\mathbf{L1\%} + \mathbf{I_{4,5,7}} \times 2 \mathbf{L2\%}) \times L \times \text{TmN} \qquad \qquad \qquad (\mathbf{SZ\ 2 - 7})
\end{aligned}$$

}	= $\mathbf{L1\%} \times L \times \text{TmN}$	für:	(SZ 2, 3, 6)
	= $(1 + \mathbf{L2\%}) \times L \times \text{TmN}$	für:	(SZ 4, 5, 7)
	$\mathbf{L1\%} + \mathbf{L2\%} = 100$		in der Studie verwandte Formel

Im zentralen Fall wird täglich zu jedem gefundenen Bewerber aus den täglich initiierten Suchvorgängen zu allen Stellenersterfassungen ein Eintrag in die Teamdatei vorgenommen und es war angenommen worden, daß jeweils L gefunden werden. In der Formel für die anderen Szenarien gibt der erste Summand die Zahl der Transaktionen für die lokal gefundenen Bewerber wieder, der zweite Summand diejenigen, die für Schreiben in entfernte Teamdateien erforderlich sind und schließlich beziffert der dritte Summand, wieviel Transaktionen für Einträge auf dem Amtsrechner durch Suchen von Rechnern des Pendelbereiches entstehen. Diese Formel ist auch für Szenario 7 gültig, selbst wenn dort nur lokal gesucht wird, denn die Einträge in die Teamdateien müssen auf den Rechnern erfolgen, die für die Bewerberdaten zuständig sind.

Auf diese Weise finden in einem mittleren Arbeitsamt mit rein lokaler Suche durchschnittlich am Tage $\mathbf{L1\%} \times L \times \text{TmN}$ Einträge statt, in denjenigen, die in die Suche die Bewerber des Stellen-Pendelbereiches mit einbeziehen $L \times \text{TmN}$ Einträge und in SZ 1 sind es $L \times \text{TN}$ Einträge auf dem Zentralrechner. Da auf den LAÄ-Rechnern keine Teamdateien gespeichert werden, fallen hier auch keine Transaktionen an und es wird keine gesonderte Formel erforderlich.

In der Studie wird an einem Beispiel verdeutlicht, daß $\mathbf{L1}$ und $\mathbf{L2}$ im Normalfall nicht mit $\mathbf{r1}$ und $\mathbf{r2}$ übereinstimmen: Wenn in einem Amt 100 Stellen neu zu besetzen sind, maximal 3 Bewerber je Stelle gesucht werden, so würden sich bei $\mathbf{r1} = 50\%$ 150 Bewerber bereits durch die Suchen ergeben, die vollständig lokal abgewickelt werden. Verbleiben die anderen 50%. Angenommen bei diesen lokal nicht vollständig erfolgreichen Suchen wurden dennoch immerhin 2 Lösungen in 20% aller Fälle lokal gefunden, 1 Lösung ebenfalls in 20% der Suchen und gar keine Lösung nur in 10% der Fälle. Dann wären auf dem lokalen Rechner bereits $150 + 40 + 20 = 210$ Lösungen zusammen gekommen. Dies sind 70% aller 300 gesuchten Bewerber (vgl. Anlage 1, Teil 2, Abb. S. 2).

Eine wichtige Annahme für die Korrektheit der Formeln und die Vergleichbarkeit der Modelle liegt darin, daß davon ausgegangen werden muß, daß alle L Lösungen bereits im Pendelbereich erzielt werden. Ansonsten könnten in Szenario 1 mehr Lösungen auftreten als in allen anderen Szenarien und L wäre nicht mehr eindeutig über alle Szenarien.

Deshalb wäre es aus meiner Sicht sinnvoller gewesen, die Variablen etwas allgemeiner zu vereinbaren und die Bedingung $\mathbf{L1} + \mathbf{L2} = 100$ entfallen zu lassen (vgl. Fußnote 51). Damit könnte man zugleich auch einer sich wandelnden Arbeitsmarktlage gerechter werden. $\mathbf{L1}$ würde dann den Anteil lokal gefundener Lösungen, $\mathbf{L2}$ der Anteil an Lösungen, die aus dem Pendelbereich zustande kommen und \mathbf{L} die auf dem bundesweit vorhandenen Datenbestand gefundenen Lösungen beziffern, wobei nach wie vor von $\mathbf{L} \leq \mathbf{O}$ ausgegangen würde. In Szenario 1 erhielte man dann insgesamt \mathbf{L} Lösungen, in den Szenarien 2, 3 und 6 wären es $\mathbf{L1\%} \times \mathbf{L}$ und in den Szenarien 4, 5 und 7 jeweils $(\mathbf{L1} + \mathbf{L2})\% \times \mathbf{L}$. Die Annahme, daß fast immer die Obergrenze $\mathbf{O} = \mathbf{L}$ bei den Suchen erreicht wird, könnte man ebenfalls zurücknehmen. Zur Gestaltung der Modellrechnungen wäre es denkbar, $\mathbf{L1}$ und $\mathbf{L2}$ funktional abhängig von \mathbf{O} so anzusetzen, daß hohe Werte von \mathbf{O} niedrigere Werte von $\mathbf{L1}$ zur Folge haben, ggf. auch, daß

$(L1+L2)\% \times L$ deutlich kleiner als O sind. Niedrige Werte von O könnten zu $L1 = O$ und $L2 = 0$ führen. Die obigen Formeln könnten beibehalten werden, die Formel F*** (s. S. 84) dürfte nicht mehr weiter umgeformt werden. Auswirkungen auf künftige Formeln durch eine solche Setzung der Variablen $L1$, $L2$ und L werden künftig in Fußnoten oder Anmerkungen angegeben.

für das Schreiben in die Teamdatei benötigte Variablenbezeichner:

<i>TransWrzenTD</i>	durchschnittliche Anzahl täglicher Transaktionen für Einträge in die Teamdateien, die in SZ 1 aus stellenorientierten Bewerbersuchen nach Stellen-Ersterfassung entstehen
L	durchschnittliche Anzahl gefundener Bewerber je Stelle ⁵¹ . In den Modellrechnungen wird angenommen, daß stets gilt: $L = O$ (O Obergrenze)
$L1$	prozentualer Anteil lokal gefundener Bewerber von L (vgl. Fußnote 51)
$L2$	prozentualer Anteil der durch Anfragen im Pendelbereich in SZ 4 und SZ 5 gefundenen Bewerber von L (vgl. Fußnote 51), $L2 = 100 - L1$
<i>TransWrmTD</i>	durchschnittliche Anzahl an täglich in einem „mittleren“ Arbeitsamt entstehenden Transaktionen für Einträge in die Teamdateien, die aus stellenorientierten Bewerbersuchen nach Stellen-Ersterfassung resultieren

Nachbearbeitungen (Entscheidung der Sachbearbeiter und Folgebearbeitungen)

Durch das Abarbeiten der Teamdatei-Einträge, die in allen Szenarien mit Ausnahme von SZ 1 auf lokale Bewerberdaten verweisen, entstehen täglich durchschnittlich eine Anzahl weiterer Transaktionen auf den Rechnern der Ämter bzw. im Szenario 1 auf dem dortigen Zentralrechner. Diese sind in der Tabelle der folgenden Seite wiedergegeben. Dabei wurden die auf Seite 84 in dem markierten Absatz angegebenen Formeln zur Zahl der Einträge in die Teamdateien an den diversen Rechnern berücksichtigt. Der neu vorkommende Bezeichner I gibt an, wie hoch der Anteil an den gefundenen Bewerbern ist, der durchschnittlich von den Sachbearbeitern als geeignet akzeptiert wird.

I	durchschnittlicher prozentualer Anteil an den gefundenen Bewerbern L , der von den Sachbearbeitern als geeignet akzeptiert wird
-----	---

Aus der Tabelle gehen durch Zusammenfassung der einzelnen Transaktionsanforderungen Formeln für die diversen Szenarien hervor. Ausgehend von diesen Formeln werden auf den der Tabelle folgenden Seiten eine Reihe von Umformungen vorgenommen. Diese führen zu den jeweiligen Darstellungen, die in zur Studie in den Abbildungen der Seiten 9 bis 11 (Formelübersicht für die stellenorientierte Bewerbervermittlung) im Anlage 1 gewählt wurden.

⁵¹ alternativ:

L	durchschnittliche Anzahl gefundener Bewerber je Stelle auf zentralem Bewerberdatenbestand, Annahme $L \leq O$ (Obergrenze für Lösungen)
$L1$	durchschnittlicher prozentualer Anteil von L , der aus lokalen Arbeitsamtsrechnerdaten hervorgeht
$L2$	durchschnittlicher prozentualer Anteil von L , der aus den Daten von Rechnern des Pendelbereiches oder Replikaten dieser Daten hervorgeht, in SZ 2, 3, 6 gilt: $L2 = 0$, denn diese Daten sind dort bei bewerberorientierten Stellensuchen unsichtbar

Spalte 1 Vorgang Transaktionen	2 auf dem Zen- tralrechner in SZ 1	3 auf einem Rechner eines „mittleren“ AA in SZ 2, 3, 6	4 auf einem Rechner eines „mittleren“ AA in SZ 4	5 auf einem Rechner eines „mittleren“ AA in SZ 5 und 7	6 <i>alternativ für SZ 2 - 7, L, L1, L2, I_x wie in Fußnote 51</i>
zum Lesen der Einträge	L × TN	L1% × L × TmN	L × TmN	L × TmN	(L1+L2)% × L × TmN
zum Lesen der lokalen BewAs	L × TN			L × TmN (SZ 7)	(L1+L2)% × L × TmN
• aus lokal gestarteten Suchen		L1% × L × TmN	L1% × L × TmN	L1% × L × TmN (SZ 5)	
• aus entfernt gestarteten			L2% × L × TmN	L2% × L × TmN (SZ 5)	
zum Lesen der Stellendaten	L × TN				
• lokal		L1% × L × TmN	L1% × L × TmN	L × TmN.	(L1% + L2% × I _{5,7}) × L × TmN
• entfernt			L2% × L × TmN	liegen lokal als Replika- te vor	L2% × I ₄ × L × TmN
zum Lesen lokaler Stelleninformationen aufgrund von entfernten Anfragen			L2% × L × TmN	in SZ 5 nicht erforderlich, da diese Informationen auf den entfernten Rechnern vorliegen in SZ 7 nur lokale Suche	L2% × I ₄ × L × TmN
zum Update akzeptierter BewAs	I% × L × TN	I% × L1% × L × TmN	I% × L × TmN	I% × L × TmN	I% × (L1+L2)% × L × TmN
zum Update der Stellendaten	I% × L × TN				
• lokal		I% × L1% × L × TmN	I% × L1% × L × TmN	I% × L1% × L × TmN	I% × L1% × L × TmN
• entfernt			I% × L2% × L × TmN	I% × L2% × L × TmN	I% × L2% × L × TmN
Updateanforderungen auf lokalen Stellendaten von anderen Rechnern			I% × L2% × L × TmN	I% × L2% × L × TmN	I% × L2 × L × TmN

Tabelle 8: Durchschnittlich täglich hervorgerufene Transaktionen aufgrund von Sachbearbeiterentscheidungen nach stellerorientierten Bewerbersuchen

Anmerkungen:- dunkelgrau unterlegte Bereiche sind mit Datenfernübertragungen verbunden und sind daher auch gesondert ausgewiesen.

- In den Spalten 2 bis 5 (und nur dort) gilt $L1 + L2 = 100$.

- Das Modell berücksichtigt keine Transaktionen, die durch Updates der Zuordnungsinformationen der zentral gehaltenen Stelleninformationen entstehen.

Daraus ergibt sich für Szenario 1:

$$TransSachzenB = (3 + 2 I\%) \times L \times TN \quad (SZ 1)$$

und für die Rechner der AÄ:

a) für die Szenarien, bei denen ausschließlich auf dem lokalen BewA-Bestand gesucht wird

$$\begin{aligned} TransSachmB_{2,3,6} &= (3 L1\% + 2 I\% \times L1\%) \times L \times TmN \\ &= (3 + 2 I\%) \times L1\% \times L \times TmN \quad (SZ 2, 3, 6) \\ &= (1 + 2 + 2 I\%) \times L1\% \times L \times TmN \\ &= (1 + 2 \times (1 + I\%)) \times L1\% \times L \times TmN \end{aligned}$$

b) für Szenario 4 (Suche im Datenbestand des Pendelbereichs ohne Replikate) (SZ 4)

$$\begin{aligned} TransSachmB_4 &= [1 + 2 (L1\% + L2\%) + L2\% + I\% \times (1 + L1\% + 2 L2\%)] \times L \times TmN \\ &= [3 + L2\% + I\% \times (1 + (L1 + L2)\% + L2\%)] \times L \times TmN \\ &\quad L1+L2=100 \\ &= (3 + L2\% + 2 I\% + I\% \times L2\%) \times L \times TmN \\ &= (1 + 2 + L2\% + 2 I\% + I\% \times L2\%) \times L \times TmN \\ &= (1 + (1 + I\%) \times (2 + L2\%)) \times L \times TmN \end{aligned}$$

c) für die Szenarien 5 und 7 (Suche im Datenbestand des Pendelbereichs mit Replikaten)

$$\begin{aligned} TransSachmB_5 &= [3 + I\% \times (1 + L1\% + 2 \times L2\%)] \times L \times TmN \\ &= [3 + I\% \times (2 + L2\%)] \times L \times TmN \quad (SZ 5, 7) \\ &\quad L1+L2=100 \end{aligned}$$

Die alternative Formel würde zusammengefaßt lauten:

$$\begin{aligned} TransSachmB &= [3 L1 \% + 2 L2\% + L2\% \times I_{5,7} + 2 L2\% \times I_4 \\ &\quad + I\% \times (2 L1\% + 3 L2\%)] \times L \times TmN \quad (SZ 2 - 7) \end{aligned}$$

Dabei wurden die folgenden Bezeichner benötigt:

<i>TransSachzenB</i>	durchschnittliche Anzahl an täglich entstehenden Transaktionen, die durch das Abarbeiten der Teamdateien durch die Sachbearbeiter im zentralen Fall aufgrund der stellenorientierten Bewerbervermittlung entstehen
<i>TransSachmB</i>	durchschnittliche Anzahl an täglich in einem mittleren Arbeitsamt entstehenden Transaktionen, die durch das Abarbeiten der Teamdateien durch die Sachbearbeiter aufgrund der stellenorientierten Bewerbervermittlung entstehen. Angegebene Indizes weisen darauf hin, daß die Formel nur für das diesem Index entsprechende Szenario gilt.

Folgebearbeitung

Folgebearbeitungen der stellenorientierten Bewerbersuche nach Neuerfassung sind zum Beispiel weitere Suchen in festgelegten zeitlichen Intervallen und die hieraus resultierenden erneuten Einträge in Teamdateien sowie Auswirkungen auf die Sachbearbeitertätigkeiten. Sie sind im Modell pauschalisiert über einen wählbaren Faktor **fa** berücksichtigt. Er soll je nach Auslastung der Sachbearbeiter und in Abhängigkeit von den Wünschen der Betriebe gesetzt werden. Die Zahl der Transaktionen für Folgebearbeitungen ergibt sich dann als Produkt aus dem Faktor **fa** und den in den Szenarien für die Vorgänge Suchen, Einträge in die Teamdateien und Sachbearbeiterentscheidungen angesetzten Anzahlen an Transaktionen.

Die Formeln lauten im zentralen Fall:

$$TransFolzenN = \mathbf{fa} \times (TransSuzenB + TransWrzenTD + TransSachzenB) \quad (\mathbf{SZ\ 1})$$

und für die Rechner der Arbeitsamtsebene:

$$TransFolmN = \mathbf{fa} \times (TransSumB + TransWrmTD + TransSachmB). \quad (\mathbf{SZ\ 2 - 7})$$

Die Rechner der Landesarbeitsämter bleiben unberührt, weil auf diesen keine stellenorientierten Suchen erfolgen.

Da die Variablen *TransSuzenB* und *TransSumB* auf die Variable **TransSulokB** zurückgreifen und in dieser auch probeweise Suchen berücksichtigt sind, umfaßt die Formel der Folgebearbeitung entweder eine erneute Anpassung der Suchbedingungen oder aber arbeitet mit einer zu hohen Zahl für die zur Suche erforderlichen Transaktionen.

Weitere verwendete Bezeichner:

<i>TransFolzenN</i>	durchschnittliche Anzahl an täglich entstehenden Transaktionen im zentralen Fall aufgrund von Folgebearbeitungen aus stellenorientierten Bewerbervermittlungen
<i>TransFolmN</i>	durchschnittliche Anzahl an täglich in einem mittleren Arbeitsamt entstehenden Transaktionen aufgrund von Folgebearbeitungen aus stellenorientierten Bewerbervermittlungen
fa	pauschalisiert wählbarer Faktor zur Beschreibung des Aufwands aus stellenorientierten Bewerbervermittlungen, die nicht mit Neuerfassung zusammenhängen

Gesamtanforderungen

Die Gesamtanforderungen ergeben sich getrennt nach Rechnerebenen und Szenarien durch Summation der Transaktionsanforderungen. Eine übersichtliche Darstellung aller in der Studie entwickelten Formeln zur stellenorientierten Bewerbervermittlung findet sich auf den Seiten 9 bis 11 in Anlage 1, Teil 2. Daher zeigt Tabelle 9 nur die von mir zusammengefaßten Formeln.

		lokale Transaktionen SZ 2 - 7
Speichern Stellen (SpST)		$= (\text{transSpN} + I_{5,7} \times q \times \text{transSpN}^*) \times \text{TmN}$
Löschen Stellen (LöST)		$= (\text{transLöN} + I_{5,7} \times q \times \text{transLöN}^*) \times \text{TmN}$
Suchen (SuBw)		$= (\text{transSulokB} + I_{4,5} \times (1 + q) \times r2\% \times \text{transSuremB}) \times \text{TmN}$
Teamdatei-Einträge (STd)		$= (L1 + I_{4,5,7} \times 2 \times L2)\% \times L \times \text{TmN}$
Sachbearbeitertätigkeiten (Sach)		$= (3 \times L1\% + 2 \times L2\% + L2\% \times I_{5,7} + 2 \times L2\% \times I_4 + I\% \times (2 \times L1 + 3 \times L2)\%) \times L \times \text{TmN}$
Folgeaktivitäten (FoAkt)		$= fa \times (\text{SuBw} + \text{STd} + \text{Sach})$

	SZ1, zentral	SZ 6, LAA	SZ 2 - 7 zentral
SpST	$= \text{transSpN} \times \text{TN}$	$= \text{transSpN}^* \times \text{TN/LAA}$	$= \text{transSpN}^* \times \text{TN} \times (100 - X)\%$
LöST	$= \text{transLöN} \times \text{TN}$	$= \text{transLöN}^* \times \text{TN/LAA}$	$= \text{transLöN}^* \times \text{TN} \times (100 - X)\%$
SuBw	$= \text{transSulokB} \times \text{TN}$		
STd	$= L \times \text{TN}$		
Sach	$= (3 + 2 \times I\%) \times L \times \text{TN}$		
FoAkt	$= fa \times (\text{SuBw} + \text{STd} + \text{Sach})$		

Tabelle 9: Formeln für die stellenorientierte Bewerbersuche

Von Siemens Nixdorf wurde eine Beispielrechnung durchgeführt und dargestellt (s. Abbildung 16, S. 91, Tabelle A1, basierend auf: Anlage 1, Teil 2, S. 12 - 13). Auf sie soll an dieser Stelle kurz eingegangen werden.

Es wird von einigen generellen, durchschnittlichen Annahmen ausgegangen, wie etwa von einer Neuaufnahme von täglich 10.000 Stellen im Bundesgebiet, von 200 Arbeitsämtern und 5 Ämtern im Stellen-Pendelbereich. Weitere Ausgangsgrößen betreffen die Zahl der angenommenen Transaktionen für die diversen oben beschriebenen Aktionen. Bezüglich der Suche wird angesetzt, daß 50% hiervon vollständig lokal befriedigt werden können, 70% der Lösungen lokal und nur die verbleibenden 30% entfernt gefunden werden. Es wird nur nach 3 Lösungen gesucht, 40% der Lösungen werden verworfen. Als Faktor für die Folgebearbeitungen wurde 2 gewählt. Bei dem Prozentsatz der abgelehnten Lösungen erscheint es allerdings sinnvoll, von vornherein nach mehr Lösungen zu suchen. Vollzieht man die angegebene Lösung nach (und verwendet zudem für die Variable **transLöN** den Wert 2), so kann man derartige Überlegungen bequem durchkalkulieren⁵².

Ebenso lassen sich auch die Auswirkungen der von mir leicht modifizierten Formeln austesten (s. Abbildung 16). So wurden in Version B die Formeln **F*** und **F**** (s. S. 82 u. 84) berücksichtigt. Im Bereich der Beispieldaten spielen sie eine sehr geringe Rolle (grüne Felder in A1 und B1). Bei einem deutlichen Heraufsetzen der gewünschten Lösungen (z. B. auf $L = 20$) und der Annahme, daß im Pendelbereich nicht mehr alle Lösungen gefunden werden (nur noch 80%, davon 50% lokal), fallen allerdings Abweichungen auf (grüne Felder in A2 und B2). Sie bleiben aber vergleichsweise gering.

Ein Vergleich von A2 (Studie) und C2 an den hellgelben Feldern verdeutlicht zudem die für die hier gewählten Eingabedaten entstehenden Verzerrungen aufgrund der in der Studie erfolgten Annahme einer kompletten Ergebnisfindung im Pendelbereich, denn für C2 wurden die allgemeingültigeren Formeln aus Tabelle 9 verwandt.

⁵² Hierbei findet man an den in der Abbildung 16 (s. S. 91) an den 4 hellgelb markierten Felder andere Werte als in der Beispielrechnung der Studie. Dunkelgelb gekennzeichnet sind Felder, die sich in Folge hiervon ändern. Diese - vermutlich unbeabsichtigten - von den in der Studie entwickelten und auf den Seite 9 - 11 dargestellten Formeln abweichenden Werte habe ich entsprechend der Formeln der Studie abgeändert.

Neben einer Variation der gerade betrachteten Werte hat vor allem - wie auch in der Studie betont - der Faktor der Folgebearbeitungen gravierende Auswirkungen auf die Zahl der Transaktionen.

Möchte man auch Kennzeichnungen in den zentral gespeicherten Stellen-Replikaten der Bundesebene und der LAA-Ebene berücksichtigen, so entstehen auch stärkere Veränderungen bei dem Aufkommen zentraler Transaktionen (Variante C2, blaue Felder), die sich zudem auf die Folgebearbeitung auswirken (lila Felder). Hier wurde bei den lediglich lokal suchenden Szenarien die Formel

$$\text{zentraleKz} = \mathbf{I\%} \times \mathbf{L1\%} \times L \times \text{TN} \times (100 - X)\% \quad (\text{SZ 2, 3, 6})$$

verwandt, bei den anderen Szenarien die Formel:

$$\text{zentraleKz} = \mathbf{I\%} \times (\mathbf{L1\%} + \mathbf{L2\%}) \times L \times \text{TN} \times (100 - X)\% \quad (\text{SZ 4, 5, 7})$$

Zusammengefaßt gilt also:

$$\text{zentraleKz} = \mathbf{I\%} \times (\mathbf{L1\%} + \mathbf{L2\%} \times \mathbf{I_{4,5,7}}) \times L \times \text{TN} \times (100 - X)\% \quad (\text{SZ 2 - 7})$$

Für die LAA-Rechner gilt folgendes:

$$\text{zentraleKzLAA} = \mathbf{I\%} \times (\mathbf{L1\%}) \times L \times \text{TN} / \mathbf{LAA} \quad (\text{SZ 6 LAA-Ebene})$$

mit:

<i>zentraleKz</i>	durchschnittliche Anzahl an Transaktionen zur Kennzeichnung der zentral gehaltenen Stellenreplikate im Rahmen der stellenorientierten Bewerbersuche
<i>zentraleKzLAA</i>	durchschnittliche Anzahl an Transaktionen zur Kennzeichnung der in einem LAA gehaltenen Stellenreplikate im Rahmen der stellenorientierten Bewerbersuche

Dabei wurde davon ausgegangen, daß diese Kennzeichnungen lokal keine Veränderungen mehr bewirken, da sie im Rahmen der Updates der Originalstellendaten angestoßen werden können.

Weitere Aspekte sollen erst in einer abschließenden Gesamtbetrachtung nach Behandlung der bewerberorientierten Stellensuche angesprochen werden.

Beispielrechnungen zur stellenorientierten Bewerbervermittlung

Die Ausgangsdaten zu den Tabellen A1 und B1 wurden der Studie entnommen (Anlage 1, Teil 2, S. 13), sind hier aber zur besseren Übersicht ebenfalls angeführt. Die teilweise hinter diesen Werten stehenden Angaben in Klammern stellen die veränderten Annahmen von A2, B2 und C2 dar. Für die Tabellen A wurden die Formeln der Studie verwandt, für B an den grün markierten Stellen die auf S. 80 u. 82 alternativ entwickelten Formeln und in C gingen zusätzlich die zentrale Kennzeichnung und die allgemeineren Festlegungen von L1, L2 und L ein.

Die **Beschriftungen** in der linken Spalte bedeuten von oben nach unten: Speichern einer Stelle, Löschen einer Stelle, Suche nach Bewerbern, Schreiben in Teamdatei, Aktivitäten der Sachbearbeiter, Folgebearbeitungen.

Die Ausgangsdaten sind: TN: 10.000 TmN = 50 q = 5 transSuFi = 1
 transSpN* = 1 X = 50% X* = 95% transLöN* = 1 transSulokB = 5 transSuremB = 1
 l = 60% fa = 2 r1 = 50%, (30%) L1 = 70%, (50%) L2 = 30% L = 3, (20)

A1

	Szenario 1		Szenario 2		Szenario 3		Szenario 4		Szenario 5		Szenario 6			Szenario 7	
	Zentral	lokal	LAA	zentral	lokal	zentral									
SpSt	20.000	100	10.000	100	5.000	100	500	350	500	100	1.000	500	350	500	
LöSt	20.000	100	10.000	100	5.000	100	500	350	500	100	1.000	500	350	500	
SuBw	50.000	250		250		400		400		250			250		
Std	30.000	105		105		195		195		105			195		
Sach	126.000	441		441		702		657		441			657		
FoAkt	412.000	1.592		1.592		2.594		2.504		1.592			2.204		
Zusammen	658.000	2.588	20.000	2.588	10.000	4.091	1.000	4.456	1.000	2.588	2.000	1.000	4.006	1.000	

B1

	Zentral	lokal	LAA	zentral	lokal	zentral								
SpSt	20.000	100	10.000	100	5.000	100	500	338	500	100	1.000	500	338	500
LöSt	20.000	100	10.000	100	5.000	100	500	338	500	100	1.000	500	338	500
SuBw	50.000	250		250		400		400		250			250	
Std	30.000	105		105		195		195		105			195	
Sach	126.000	441		441		702		657		441			657	
FoAkt	412.000	1.592		1.592		2.594		2.504		1.592			2.204	
Zusammen	658.000	2.588	20.000	2.588	10.000	4.091	1.000	4.431	1.000	2.588	2.000	1.000	3.981	1.000

A2

	Zentral	lokal	zentral	lokal	zentral	lokal	zentral	lokal	zentral	lokal	LAA	zentral	lokal	zentral
SpSt	20.000	100	10.000	100	5.000	100	500	350	500	100	1.000	500	350	500
LöSt	20.000	100	10.000	100	5.000	100	500	350	500	100	1.000	500	350	500
SuBw	50.000	250		250		460		460		250			250	
Std	200.000	500		500		1.300		1.300		500			1.300	
Sach	840.000	2.100		2.100		4.680		4.380		2.100			4.380	
FoAkt	2.180.000	5.700		5.700		12.880		12.280		5.700			11.860	
Zusammen	3.310.000	8.750	20.000	8.750	10.000	19.520	1.000	19.120	1.000	8.750	2.000	1.000	18.490	1.000

B2

	Zentral	lokal	zentral	lokal	zentral	lokal	zentral	lokal	zentral	lokal	LAA	zentral	lokal	zentral
SpSt	20.000	100	10.000	100	5.000	100	500	338	500	100	1.000	500	338	500
LöSt	20.000	100	10.000	100	5.000	100	500	338	500	100	1.000	500	338	500
SuBw	50.000	250		250		460		460		250			250	
Std	200.000	500		500		1.300		1.300		500			1.300	
Sach	840.000	2.100		2.100		4.360		4.260		2.100			4.260	
FoAkt	2.180.000	5.700		5.700		12.240		12.040		5.700			11.620	
Zusammen	3.310.000	8.750	20.000	8.750	10.000	18.560	1.000	18.735	1.000	8.750	2.000	1.000	18.105	1.000

C2

	Zentral	lokal	zentral	lokal	zentral	lokal	zentral	lokal	zentral	lokal	LAA	zentral	lokal	zentral
SpSt	20.000	100	10.000	100	5.000	100	500	338	500	100	1.000	500	338	500
LöSt	20.000	100	10.000	100	5.000	100	500	338	500	100	1.000	500	338	500
SuBw	50.000	250		250		460		460		250			250	
Std	200.000	500		500		1.100		1.100		500			1.100	
Sach	840.000	2.100		2.100		3.840		3.540		2.100			3.540	
zentrale Kz			60.000		30.000		4.800		4.800		6.000	3.000		4.800
FoAkt	2.180.000	5.700	120.000	5.700	60.000	10.800	9.600	10.200	9.600	5.700	12.000	6.000	9.780	9.600
Zusammen	3.310.000	8.750	200.000	8.750	100.000	16.400	15.400	15.975	15.400	8.750	20.000	10.000	15.345	15.400

Abbildung 16: Stellenorientierte Bewerbervermittlung, Beispielrechnungen

4.4.2.2 Die bewerberorientierte Stellensuche

Die Formeln zur bewerberorientierten Stellensuche berücksichtigen Aktivitäten, die beim Erstkontakt mit einem Bewerber auftreten; dies sind die Erfassung der Personendaten, die Suche nach geeigneten Stellen und die sich hieraus ergebenden Updates der Personen- und Stellendaten. Darüber hinaus werden Folgeberatungen berücksichtigt.

Zur Bewertung der Ergebnisse der einzelnen Szenarien werden in einem späteren Abschnitt der Studie Transaktionsbelastungen in unterschiedlichen Tageszeitbereichen herangezogen. In diesem Kontext wird für den Vorgang des Erstkontaktes bereits an dieser Stelle eine Zeitbedarfsmodellierung vorgenommen.

Erstkontakt - Zeitbedarf

Die durchschnittliche tägliche Anzahl an Erstkontakten ergibt sich, sofern man von einer gleichmäßigen Verteilung dieser Kontakte über die Ämter ausgeht, für jeden einzelnen Sachbearbeiter durch die jährlich hinzukommenden Arbeitsuchenden (als Wert hierfür wäre z. B. der des Vorjahres möglich), die sich auf die Sachbearbeiter und deren Arbeitstage verteilen.

$$TE_{\text{ber}} = AS+ / \text{arbD} / \text{ATS}$$

Sofern ein Erstkontakt durchschnittlich **ZEber** an Zeit benötigt, ergibt sich die tägliche zeitliche Beanspruchung je Sachbearbeiter für diese Kontakte durch:

$$TZE_{\text{ber}} = TE_{\text{ber}} \times ZE_{\text{ber}}$$

Bei einer täglichen Öffnungszeitdauer für Publikum von **TZopen** verbleibt für Folgeberatungen mit Publikumsverkehr eine Zeit von

$$TZF_{\text{ber}} = TZ_{\text{open}} - TE_{\text{ber}}$$

mit:

TEber	durchschnittliche tägliche Anzahl an Erstkontakten je Sachbearbeiter
ZEber	durchschnittliche Zeitdauer für einen Erstkontakt
TZEber	durchschnittlicher Tageszeitbedarf für Erstkontakte eines Vermittlers
TZopen	durchschnittliche tägliche Öffnungszeitdauer mit Publikumsverkehr
TZFber	durchschnittliche Zeit für Folgeberatungen mit Publikumsverkehr
AS	Gesamtzahl der bundesweit registrierten und in Datenbanken gespeicherten Arbeit-suchenden
AS+	Gesamtanzahl Arbeitsuchender, die bundesweit jährlich neu registriert werden
arbD	Anzahl der PCs, an denen Arbeitsvermittlung durchgeführt wird. Sie entspricht der Anzahl der Sachbearbeiter für die Arbeitsvermittlung (bundesweit).
ATS	Anzahl der Arbeitstage eines Sachbearbeiters, an denen er seine Aufgabe durch-führt

Bei jährlich bundesweit 2,5 Millionen neuen Erstkontakten und 7500 Vermittlern mit je 200 Arbeitstagen sowie einer durchschnittlichen Dauer von 30 Minuten für einen Erstkontakt be-trägt die durchschnittlich täglich benötigte Zeit für die mit Erstkontakten verbundenen Ar-beiten 50 Minuten. Falls für den Publikumsverkehr täglich 4 Stunden vorgesehen sind, so bleibt für Folgeberatungen während des Publikumsverkehr ein Zeitraum von 190 Minuten übrig. Weitere 3,5 Stunden verbleiben bei einer täglichen Arbeitszeit von 7,5 Stunden dann für Tätigkeiten außerhalb der Besuchszeit (vgl. Anlage 1, Teil 2, S. 14).

Erstkontakt - Erfassen der Bewerberdaten

Bei der Neuerfassung einer Person geht die Studie von drei Summanden aus: den Transaktionen zur Suche der Person, zur Speicherung der Personendaten und zur Erfassung der Berufsdaten. Die Suche nach einer Person im Datenbestand macht Sinn um doppelte Erfassungen auszuschließen. So kann es durchaus sein, daß Personendaten aufgrund früherer Anliegen noch oder aufgrund anderer über das Arbeitsamt abgewickelter Aufgaben bereits im Datenbestand sind (vgl. Fußnote 54, S. 104). Die Transaktionen zur Neuerfassung fallen in den Szenarien 1 bis 6 jeweils genau einmal für jeden hinzukommenden Arbeitsuchenden an, in Szenario 1 über entfernten Zugriff, ansonsten lokal.

Im zentralen Fall des Szenarios 1 lautet die Formel für das tägliche Transaktionsaufkommen durch Neuerfassung von Bewerbern daher:

$$\text{TransErfzenB} = (\text{transSuAS} + \text{transSpPers} + \text{transSpBer}) \times \text{TAS} + \quad (\text{SZ 1})$$

In SZ 7 müssen die erfaßten Bewerberdaten (Personal- und Berufsdaten) zudem in den Bewerber-Pendelbereich übertragen werden, zu dem durchschnittlich (in allen Szenarien) **p** Ämter zählen. Für eine Übertragung werden im Durchschnitt **transErfBP** Transaktionen benötigt. Damit ergibt sich als Formel für die Rechner in den Ämtern der Szenarien 2 bis 7:

$$\text{transErfmB} = \text{transErfB} \times \text{TmAS} + I_7 \times \mathbf{p} \times \text{transErfBP}$$

mit:

TAS+	durchschnittlich täglich bundesweit hinzukommende Zahl an Arbeitsuchenden $\text{TAS} + = \text{AS} + / \text{AT}$
TmAS+	durchschnittlich täglich in einem „mittleren Arbeitsamt“ hinzukommende Zahl an Arbeitsuchenden $\text{TmAS} + = \text{AS} + / \text{AT} / \text{RAA}$
<i>TransErfzenB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zur Erfassung neuer Bewerberdaten im zentralen Szenario 1
<i>TransErfmB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zur Erfassung der Bewerberdaten in einem „mittleren“ AA
transErfB	durchschnittliche Anzahl an Transaktionen zur Ersterfassung von Bewerberdaten $\text{transErfB} = \text{transSuAS} + \text{transSpPers} + \text{transSpBer}$
p	durchschnittliche Anzahl an AA-Rechnern im Pendelbereich eines BewAs
transSuAS	durchschnittliche Anzahl an Transaktionen zum Suchen von Personendaten
transSpPers	durchschnittliche Anzahl an Transaktionen zum Speichern von Personendaten
transSpBer	durchschnittliche Anzahl an Transaktionen zum Speichern der Berufsdaten
transErfBP	durchschnittliche Anzahl an Transaktionen zum Übersenden der Bewerberdaten in den Pendelbereich

Erstkontakt - Suche nach geeigneten Stellen

Jedem Erstkontakt folgt eine mit dem Bewerber abgestimmte Suche im Stellenbestand. Die Suche erfolgt automatisch zunächst auf den lokalen Daten. Wenn hier noch nicht genügend Stellen gefunden werden, wird im Pendelbereich oder zentral weitergesucht.

Im folgenden werden eine Reihe von Variablen beschrieben, die für die Formulierung der Suchformeln verwandt werden. Sie sind denen, die für die stellenorientierte Bewerbersuche

benötigt wurden zwar vom gedanklichen Modell her sehr ähnlich, müssen aber bezüglich der Werte nicht mit ihnen übereinstimmen. Darauf wird zukünftig nicht mehr hingewiesen. Da in der Studie in solchen Fällen neue Bezeichner gewählt wurden, ist sowieso offensichtlich, wo Abweichungen vorliegen können.

Die Anzahl der Suchergebnisse der bewerberorientierten Stellensuche wird durch die Obergrenze O^* begrenzt. Mit **transSulokV** wird die durchschnittliche Anzahl der lokalen Suchtransaktionen angegeben. Sie umfaßt auch die Probeversuche, die durch die Abstimmungsprozesse mit dem Bewerber erfolgen. Die Variable **transSuremV** gibt die durchschnittliche Anzahl an Suchtransaktionen nach geeigneten Stellen in der Ferne wieder. „Eine solche Suchtransaktion ist lokal eine verteilte Transaktion erzeugt aber in jedem Rechner, der die Suchfrage erhält, jeweils eine Transaktion.“ (Anlage 1, S. 41) In der Studie wird daher lokal wie zentral mit dieser Variablen gearbeitet.

Darüber hinaus werden im Modell die drei Größen s_1 , s_2 und s_3 verwandt, die für die prozentualen Anteile von nur lokalen Suchen, von Suchen auf dem Pendelbereich und von den zentralen Suchen für besondere Berufsgruppen benötigt werden und deren Summe 100 ergeben soll.

Benötigt werden also:

transSulokV	durchschnittliche Anzahl der lokalen Suchtransaktionen bei einer mit dem Bewerber abgestimmten bewerberorientierten Stellensuche
transSuremV	durchschnittliche Anzahl an Suchtransaktionen nach geeigneten Stellen in der Ferne für eine bewerberorientierten Stellensuche
O^*	Obergrenze O^* , begrenzt die Anzahl der Stellen, die als Suchergebnis angezeigt werden
s_1	gibt den prozentualer Anteil der Fälle wieder, bei denen in der bewerberorientierten Stellensuche nur lokal gesucht wird
s_2	gibt den prozentualen Anteil der Fälle wieder, bei denen in der bewerberorientierten Stellensuche auch im Pendelbereich gesucht wird
s_3	gibt den prozentualer Anteil der Fälle wieder, bei denen in der bewerberorientierten Stellensuche für besondere Berufsgruppen auch zentral gesucht wird
<i>TransSuzenV</i>	durchschnittliche Anzahl an Transaktionen, die täglich zur bewerberorientierten Stellensuche zentral anfallen
<i>TransSumV</i>	durchschnittliche Anzahl an Transaktionen, die täglich zur bewerberorientierten Stellensuche auf einem Rechner eines „mittleren“ Arbeitsamtes anfallen
<i>TransSuLAAV</i>	durchschnittliche Anzahl an Transaktionen, die täglich zur bewerberorientierten Stellensuche in SZ6 auf einem LAA-Rechner anfallen

Unter Verwendung dieser Variablen und der Indikatorvariablen I_{SZ} und der bereits für die stellenorientierten Suchen eingeführten Variablen X lauten die Formeln in der Studie für die anfallenden durchschnittlichen Transaktionszahlen für tägliche Suchvorgänge nach einer BewA-Ersterfassung zusammengefaßt wie folgt:

$TransSuzenV = TAS+ \times transSulokV$	im zentralen Fall (SZ 1)
$TransSumV =$ $TmAS+ \times transSulokV +$ $TmAS+ \times transSuremV \times [(100 - s1)\% \times I_{2,6}$ $+ (100 - s1)\% \times (100 - X)\% \times I_3$ $+ s3\% \times I_{5,7}$ $+ ((1 + p) \times s2\% + s3\%) \times I_4]$	lokal an den Rechnern in den Ämtern
$TransSuLAAV = TAS+ / LAA \times (100 - s1)\% \times transSuremV$	auf den LAA-Rechnern in Szenario 6
$TransSuzenV =$ $TAS+ \times transSuremV \times [(100 - s1)\% \times (100 - X)\% \times I_{2,3}$ $+ s3\% \times transSuremV \times I_{4,5,6,7}$	zentral in SZ 2 - 7

Im zentralen Fall ist die Formel wieder am einfachsten. Hier entstehen Suchtransaktionen entsprechend der Anzahl täglich bundesweit neu aufkommender Stellenangebote.

An den Rechnern der Arbeitsämter findet die Suche zunächst lokal statt (1. Zeile). Über die Variable **transSulokV** werden hierbei Probesuchen berücksichtigt. Die lokal nicht abgeschlossenen Suchen $(100 - s1)\%$ sorgen für entfernte Zugriffe. In Szenario 2, 3 und 6 werden diese direkt an den Zentralrechner bzw. Rechner des LAA geschickt. Bei Szenario 2 ist zu berücksichtigen, daß eine Weiterleitung der Suche nur bei den zentral gespeicherten Berufsgruppen Sinn ergibt. Dies ist über den Faktor $(100 - X)$ modelliert worden. In den Fällen 5 und 7 wird auf Replikaten des Pendelbereiches gesucht. Hier wird nur in $s3\%$ aller Fälle die Suche zentral weiter verfolgt. Bei Szenario 4 wird auf den Rechnern des Pendelbereiches gesucht. Es fallen dort also $s2\%$ Suchen an und es werden von q Rechnern jeweils ebenso viele entgegengenommen. $s3\%$ Suchen werden direkt zur zentralen Recherche geschickt. Die Formel für die zentral entstehenden Transaktionen und das LAA ergeben sich auf diese Weise unmittelbar, wobei berücksichtigt werden muß, daß der LAA-Rechner wiederum $s3\%$ Anfragen weiterleitet, was sich in der 4. Formel niederschlägt, ihm selbst allerdings in Formel 3 nicht gesondert berechnet wird.

Auch wenn diese Modellierung auf den ersten Blick plausibel erscheint, so birgt sie doch gewisse Ungereimtheiten, die mit den Variablen $s1$, $s2$ und $s3$, der Speicherung nur bestimmter Berufsgruppen auf bestimmten Ebenen in den diversen Szenarien sowie der Forderung, daß deren Summe 100 ergeben muß, zusammenhängen. In Szenario 2 beispielsweise wird mit $100 - s1$ ein anderer Anteil an Suchen weitergeleitet als er durch $s2 + s3$ gegeben wäre, denn in $s3$ werden auch Suchen weitergeleitet, die im Pendelbereich über anderen als den 5%-Berufen der Szenarien 4 bis 7 noch nicht zu Ende gekommen sind. In den Szenarien 5 und 7 wird nur auf Replikaten des Pendelbereiches gesucht, die nicht der 5%-Berufsgruppe angehören, in Szenario 4 sind diese Stellenangebote sehr wohl integriert; beide Suchvorgänge werden über den Anteil $s2$ modelliert.

In Szenario 3 wird schließlich davon ausgegangen, daß die in den lokal noch nicht beendeten Suchen vorkommenden Berufe anteilmäßig der Verteilung der Stellen nach Berufen, zentral oder lokal entsprechen. Dies geht aus dem Faktor $(100 - X)$ hervor. Plausibler ist hingegen, daß genau diejenigen Berufe zentral gelagert werden, bei denen eine entfernte Suche höher wahrscheinlich ist. In bestimmten Berufen sind die regionalen Wünsche von Seiten der Bewerber häufig schon so begrenzt, daß allein deshalb eine zentrale Suche gar nicht erst be-

gonnen werden muß. Bei den weitergeleiteten s3% Suchen wird ja auch von vornherein angenommen, daß sie nur die 5%-Stellen der Szenarien 4 bis 7 betreffen!

Um diese Unklarheiten in der Modellierung auszuräumen, wäre eine deutlich feinere Aufteilung der Suchsituationen, auch nach Berufsgruppen erforderlich. Es ist fraglich, ob sich dieser Aufwand, für den zudem weitere Eingabewerte ermittelt werden müßten, lohnt, da insgesamt ausschließlich gemittelte Werte in das Modell eingehen. Die Unterschiede zwischen den Szenarien 4, 5 und 7 sind nicht so gravierend, da sie sich nur auf geringe Stellenanteile beziehen. Für Szenario 3 wäre zu überlegen, ob hier nicht eine höhere Zahl an zentralen Transaktionen anzusetzen wäre.

Erstkontakt - Update

Löst ein Sachbearbeiter aufgrund einer bewerberorientierten Stellensuche Vermittlungsvorschläge aus, so werden Einträge in den jeweiligen Stellen- und Bewerberdaten erforderlich.

Der Eintrag in die Bewerberdaten ist für Folgegespräche und für das Ausschließen identischer Vermittlungsvorschläge von Bedeutung. Der Speichervorgang in den Bewerberdaten soll auch das Abspeichern der Suchdaten umfassen.

Benötigt wird:

transUpPers	durchschnittliche Anzahl an Transaktionen zur Speicherung der für Vermittlungsvorschläge akzeptierten Stellendaten und der abgestimmten Suchanforderung
<i>TransUpzenPers</i>	durchschnittliche tägliche Anzahl an Transaktionen für das Update der Personendaten nach einem akzeptierten Suchvorschlag in SZ 1
<i>TransUpmPers</i>	durchschnittliche tägliche Anzahl an Transaktionen für das Update der Personendaten nach einem akzeptierten Suchvorschlag in einem „mittleren“ Arbeitsamt

In allen Szenarien, auch im siebten werden diese Informationen lediglich lokal gehalten. Daher ergeben sich die folgenden beiden einfachen Formeln:

$$TransUpzenPers = \mathbf{transUpPers} \times TAS+ \quad (\mathbf{SZ\ 1})$$

$$TransUpmPers = \mathbf{transUpPers} \times TmAS+ \quad (\mathbf{SZ\ 2 - 7})$$

Zur Beschreibung der Aktualisierungen der Stellendaten werden zunächst wieder einige Variablen eingeführt, für SZ 3 **KR**, für die restlichen Szenarien **K** als bewußt unscharf angenommen „gleichen“ Wert für die durchschnittliche Anzahl der gefundenen Stellen. Darüber hinaus werden **K1**, **K2**, **K3** sowie **KR1**, **KR2** und **KR3** für die prozentualen Anteile lokal, im Pendelbereich und zentral gefundener Stellen benötigt.

Für die Pflege der Stellendaten ist zu berücksichtigen, daß der ausschließlich lokal gehaltene Bestand die Zuordnungsinformation nur beim zuständigen AA-Rechner eingetragen bekommt und Replikate nicht berücksichtigt werden müssen. Die auf LAA- oder Bundesebene vorliegenden Replikate werden hingegen aktualisiert. Bei der Zählung der anfallenden Transaktionen ist zu berücksichtigen, daß das Auslösen der Anpassung eines Replikates nicht extra gerechnet wird, sofern bereits ein Update berechnet wird.

Die Formeln lauten hier⁵³:

$TransUPzenN = TAS+ \times transUpN \times K$	zentral (SZ 1)
$TransUpmN = TmAS+ \times transUpN$ $\times (K \times [I_2 + (1 + (K2 + K3)\%) \times I_{4,5,6,7}]$ $+ KR \times I_3)$	lokal
$TransUpzenN = TAS+ \times transUpN$ $\times \{ K \times [I_2 + K3\% \times I_{4,5,7}$ $+ ((100 - X^*) \times (K2 + K3)\% + K3\%) \times I_6]$ $+ KR \times (100 - KR1)\% \times I_3 \}$	zentral (SZ 2 - 7)
$TransUPLAAN = TAS+ \times transUpN \times K / LAA \times (1 + K3\%)$	LAA

Im zentralen Fall ist die Formel unmittelbar einsichtig, dieselbe Formel ist für den Zentralrechner in Szenario 2 zutreffend.

Lokal werden im zweiten Szenario $K1\% \times K$ Transaktionen für lokale Updates erforderlich. Während diese erfolgen, wird zugleich das zentrale Update - ohne dies doppelt zu zählen - veranlaßt. Außerdem werden $(100 - K1)\% \times K$ weitere Update-Anforderungen an den Zentralrechner gesandt, so daß dieser an sich durchschnittlich weitere $100 - K1$ Updates an jeden Rechner auf Amtsebene zurücksenden müßte und weitere $100 - K1$ Anforderungen aus der Ferne lokal hinzukämen. Damit würden genauso viele Transaktionen entstehen wie in den Fällen 4 bis 7. Diese werden jedoch hier nicht gezählt. Ganz analog zur Formel für Szenario 2 kann man sich die für Szenario 3 erklären. Hier ist lediglich an Stelle von K die Variable KR eingegangen, über die berücksichtigt wird, daß in SZ 3 aufgrund der reduzierten Zahl zentral gehaltener SteAs die Anzahl der Lösungen von K größer abweichen kann. In den restlichen Szenarien werden die eingehenden Updates berücksichtigt.

Alternativ wäre die Formel (F****) (SZ 2 - 7)

$$TransUpmN = TmAS+ \times transUpN \times [K \times (1 + (K2 + K3)\%) \times I_{2,4,5,6,7} + KR \times (1 + (KR2 + KR3)\%) \times I_3] \text{ denkbar.}$$

Die zentral anfallenden Transaktionen in Szenario 3 umfassen von den KR täglich gefundenen SteAs genau diejenigen, die zentral gehalten werden. Dies liefert den Anteil $100 - KR1$. Während deren Überarbeitung werden auch die weiteren lokalen Rechner zum Nachziehen der Änderungen veranlaßt. In den Szenarien 4, 5 und 7 fallen nur die zentral zu aktualisierenden Stellenangebote an. Andere Aktualisierungen erfolgen hier direkt über den Pendelbereich. Zudem werden in den Replikaten der Szenarien 5 und 7 keine Einträge über Zuordnungsinformationen vorgenommen.

Das Update im Szenario 6 erfolgt ausgehend von dem lokalen Rechner über den LAA-Rechner. Sofern es sich um ein zentral gehaltenes SteA handelt, geht der Update-Auftrag an den Zentralrechner weiter. Außerdem veranlaßt die niedrigst mögliche Ebene, sofern der Anstoß nicht vom Original ausging, das Update der lokalen Originaldaten, die als Zerschlagung über die Amtsrechner verteilt sind (vgl. graphische Darstellung, Anlage 1, Teil 2, S. 21). Entsprechend entstehen im LAA für jede LAA-weit gefundene Lösung Einträge. Hinzu kommen

⁵³ Sofern man diese der Studie entnommenen Formeln (Anlage 1, Teil 2, S. 15 - 19) mit den angegebenen Werten für die Modellrechnung (ebenda, S. 22) überprüft, ergeben sich für die zentralen Transaktionen der Szenarien 3 - 7 größere, unterschiedliche Abweichungen, die bis auf einen verbleibenden Unterschied beim zentralen Wert von SZ 6 verschwinden, wenn man anstelle der zur Modellrechnung angegebenen Werte für $K2$ 30% und für $K3$ 10% ansetzt, wie es auch in der darauffolgenden Modellrechnung der Studie erfolgt und zudem für $KR1$ 50% festlegt.

K3%. Damit werden die von hier aus weiter ausgelösten Updates auf dem Zentralrechner nun gezählt. Zentral werden in diesem Szenario $(100 - X^*) \times (K2 + K3)\% + K3\%$ der bundesweit anfallenden Lösungen angesetzt. Hier hätten meines Erachtens ebenso lediglich K3% zur Anrechnung kommen müssen.

Die weiter hinzugekommenen Variablenbezeichner sind:

<i>TransUpzenN</i>	durchschnittliche Anzahl täglich zentral erforderlicher Transaktionen für Stellenupdates in SZ 1 in Folge von bewerberorientierten Erstsuchen
<i>TransUpmN</i>	durchschnittliche Anzahl täglich in einem „mittleren“ Arbeitsamt erforderlicher Transaktionen für Stellenupdates in Folge von bewerberorientierten Erstsuchen
<i>TransUpLAAN</i>	durchschnittliche Anzahl täglich auf einem LAA-Rechner entstehender Transaktionen für Stellenupdates in Folge von bewerberorientierten Erstsuchen
transUpN	durchschnittliche Anzahl an Transaktionen zur Ausführung eines Stellen-Updates
K	durchschnittliche Anzahl gefundener Stellen je Bewerber, wobei angenommen wird, daß K für die Szenarien 1, 2 und 4 bis 7 (bis auf zu vernachlässigende Abweichungen durch die eingeschränkten Pendelbereiche an den LAA-Grenzen) identisch ist
KR	durchschnittliche Anzahl gefundener Stellen je Bewerber im Szenario 3
K1	durchschnittlicher prozentualer Anteil lokal gefundener Stellen je Bewerber von K
K2	durchschnittlicher prozentualer Anteil der im Pendelbereich gefundener Stellen je Bewerber von K
K3	durchschnittlicher prozentualer Anteil zentral gefundener Stellen je Bewerber von K
KR1	durchschnittlicher prozentualer Anteil lokal gefundener Stellen in SZ 3 von KR
KR2	durchschnittlicher prozentualer Anteil der im Pendelbereich gefundenen Stellen in SZ 3 von KR
KR3	durchschnittlicher prozentualer Anteil zentral gefundener Stellen in SZ 3 von KR

Folgeberatungen

Auch bewerberorientierte Suchen können, ohne unmittelbar durch den Sachbearbeiter initiiert zu werden, starten. In allen solchen Fällen wird in der Studie von „Folgeberatungen“ gesprochen, auch wenn nicht in jedem Fall eine Beratungssituation eintritt. Es werden die folgenden Fälle unterschieden:

- asynchron durchgeführte, periodisch stattfindende erneute Suchläufe nach noch nicht vorgeschlagenen Stellen (Fall 1),
- durch den Sachbearbeiter festgelegte sonstige Termine für einen Suchlauf, z. B. zu vereinbarten Beratungsgesprächen (Fall 2),
- durch Wunsch eines Arbeitssuchenden ausgelöste Suchläufe, z. B. bei spontanen Vorsprachen (Fall 3).

Die Ergebnisse dieser Suchen werden wie bei den stellenorientierten Suchen in die Teamdateien eingetragen.

Grundsätzlich erfordern diese Suchen ein Lesen der BewA-Daten, die Durchführung der abgestimmten Suchbedingungen sowie einen Vergleich der Ergebnismenge mit bereits angebotenen Stellen aufgrund dessen für den Bewerber die noch nicht vorgeschlagenen Stellen in die Teamdatei geschrieben werden.

Für das Abarbeiten der Teamdatei-Einträge kann in den beiden letzten Fällen das Ergebnis direkt mit dem Arbeitssuchenden abgeklärt werden. Es wird davon ausgegangen, daß auf diese Weise 12 Prozent der in die Teamdatei eingetragenen Stellen für eine Bewerbung in Frage

kommen. In Fall 1 sind zwei Situationen vorgesehen: Entweder trifft der Sachbearbeiter anhand der Datenlage eine Entscheidung (dies führe zu **11** Prozent Vorschlägen) oder er lädt den Arbeitsuchenden zur Abklärung ins Amt ein.

Insgesamt sind folgende Variablen zur Beschreibung der Folgeberatungen erforderlich:

fc	durchschnittliche jährliche Anzahl der bewerberorientierten Folge-Suchläufe in Fall c, $c \in \{1, 2, 3\}$ wobei Fall 1 alle periodisch stattfindenden Suchläufe umfaßt und die Fälle 2 und 3 alle sonstigen (entweder durch den Arbeitsvermittler oder Arbeitsuchenden ausgelöst), sofern ein Bewerber ein komplettes Jahr arbeitssuchend bleibt
f	durchschnittliche jährliche Anzahl der bewerberorientierten Folge-Suchläufe falls ein Bewerber ein komplettes Jahr arbeitssuchend bleibt
KNc	durchschnittlich jährlich gefundene Anzahl an Stellen bei bewerberorientierten Folge-Suchläufen im Fall c, falls ein Bewerber ein komplettes Jahr arbeitssuchend bleibt, $c \in \{1, 2, 3\}$. Fall 1 umfaßt alle periodisch stattfindenden Suchläufe, die Fälle 2 und 3 alle sonstigen, entweder durch den Arbeitsvermittler oder Arbeitsuchenden ausgelösten Fälle; KNc gilt nicht für SZ 3.
KRNc	durchschnittlich jährlich gefundene Anzahl an Stellen bei bewerberorientierten Folge-Suchläufen im Fall c, falls ein Bewerber ein komplettes Jahr arbeitssuchend bleibt, $c \in \{1, 2, 3\}$. Fall 1 umfaßt alle periodisch stattfindenden Suchläufe, die Fälle 2 und 3 alle sonstigen, entweder durch den Arbeitsvermittler oder Arbeitsuchenden ausgelösten Fälle; KRNc gilt nur für SZ 3.
KNT	durchschnittliche Anzahl gefundener Stellen je bewerberorientiertem Folge-Suchlauf, außer bei SZ 3
KRNT	durchschnittliche Anzahl gefundener Stellen je bewerberorientiertem Folge-Suchlauf bei SZ 3
I1	durchschnittlicher prozentualer Anteil der nur durch den Sachbearbeiter für einen Vermittlungsvorschlag akzeptierten Stellenangebote in bewerberorientierten Folgesuchläufen
I2	durchschnittlicher prozentualer Anteil der durch Sachbearbeiter und Arbeitsuchenden gemeinsam akzeptierten Stellenangebote in bewerberorientierten Folgesuchläufen
lk	durchschnittlicher prozentualer Anteil für einen Vermittlungsvorschlag akzeptierter Stellenangebote in bewerberorientierten Folgesuchläufen
lb	durchschnittlicher prozentualer Anteil der gefundenen Bewerber, denen Stellen aufgrund von bewerberorientierten Folgesuchläufen angeboten werden
FS	durchschnittliche Gesamtanzahl an täglich bundesweit stattfindenden bewerberorientierten Folgesuchläufen
FmS	durchschnittlich täglich in einem „mittleren“ AA stattfindende bewerberorientierte Folgesuchläufe

Mit den so festgelegten Variablen sind die folgenden Formeln sofort ersichtlich:

$$f = \mathbf{f1} + \mathbf{f2} + \mathbf{f3} \quad (\text{jährliche Folgesuchläufe für ganzjährig Arbeitssuchende})$$

$$\mathbf{KNT} = (\mathbf{KN1} \times \mathbf{f1} + \mathbf{KN2} \times \mathbf{f2} + \mathbf{KN3} \times \mathbf{f3}) / f \quad (\mathbf{SZ 1, 2, 4 - 7})$$

$$\mathbf{KRNT} = (\mathbf{KRN1} \times \mathbf{f1} + \mathbf{KRN2} \times \mathbf{f2} + \mathbf{KRN3} \times \mathbf{f3}) / f \quad (\mathbf{SZ 3})$$

$$\mathbf{lk} = (\mathbf{I1} \times \mathbf{KN1} + \mathbf{I2} \times (\mathbf{KN2} + \mathbf{KN3})) / (\mathbf{KN1} + \mathbf{KN2} + \mathbf{KN3}) \quad (\mathbf{SZ 1, 2, 4 - 7})$$

$$\mathbf{lk} = (\mathbf{I1} \times \mathbf{KRN1} + \mathbf{I2} \times (\mathbf{KRN2} + \mathbf{KRN3})) / (\mathbf{KRN1} + \mathbf{KRN2} + \mathbf{KRN3}) \quad (\mathbf{SZ 3})$$

Darüber hinaus ergibt sich aus der Zahl der bundesweit gespeicherten Arbeitssuchenden und der durchschnittlich für einen Arbeitssuchenden jährlich durchgeführten Folgesuchläufe unter der Annahme einer gleichmäßigen Betreuung über alle Bewerber als **tägliche Gesamtzahl an Folgesuchläufen bundesweit und je Amt**:

$$FS = AS \times f / AT \quad (*)$$

$$FmS = FS / RAA \quad (**)$$

Hiermit ergeben sich erneut für die diversen Szenarien eine Reihe von im folgenden erläuterten Formeln (s. Anlage 1, Teil 2, S. 24 - 29). In der Studie wurden diese aufgeteilt in die folgenden Vorgänge:

- Folgesuchen
- Teamdatei-Einträge
- Teamdatei lesen
- BewA-Daten lesen
- lokale SteA lesen
- entfernte SteA lesen
- Update BewA-Daten
- Update SteA-Daten lokal
- Update SteA-Daten entfernt
- Update SteA-Daten aus der Ferne

Die Formel für die Folgesuchen sind analog zu den Formeln für die Erstsuche (vgl. S. 94) entwickelt und berücksichtigen lediglich zusätzlich die tägliche Gesamtzahl an Folgeberatungen bundesweit und je Amt (s. o. unter [*] und [**]) sowie die Variable **transSufol**, die angibt, wie viele Transaktionen lokal bei einer Folgesuche anfallen. Da hier kein Abstimmungsprozeß mehr erfolgt, wird sie im allgemeinen mit **transSuremV** übereinstimmen.

$$TransSufolzen = transSufol \times FS \quad (SZ 1)$$

$$TransSufolm = transSufol \times FmS \quad (SZ 2 - 7, lokal)$$

$$+ transSuremV \times FmS \times [(100 - s1)\% \times I_{2,6}$$

$$+ (100 - s1)\% \times (100 - X) \times I_3$$

$$+ s3\% \times I_{5,7}$$

$$+ ((1 + p) \times s2\% + s3\%) \times I_4]$$

$$TransSufolzen = transSuremV \times FS \quad (SZ 2 - 7, zentral)$$

$$[(100 - s1)\% \times (100 - X)\% \times I_{2,3}$$

$$+ s3\% \times I_{4,5,6,7}]$$

$$TransSufolLAA = transSuremV \times FS / LAA \times (100 - s1)\% \quad (SZ 6, LAA)$$

mit

transSufol	durchschnittliche Anzahl der zum lokalen Suchen bei bewerberorientierten Folgesuchläufen benötigten Transaktionen
<i>TransSufolzen</i>	durchschnittliche Anzahl der täglich zentral entstehenden Transaktionen durch bewerberorientierte Folgesuchläufe
<i>TransSufolm</i>	durchschnittliche Anzahl der täglich auf einem Rechner eines „mittleren“ Arbeitsamtes entstehenden Transaktionen durch bewerberorientierte Folgesuchläufe
<i>TransSufolLAA</i>	durchschnittliche Anzahl der täglich auf einem LAA-Rechner entstehenden Transaktionen durch bewerberorientierte Folgesuchläufe

Einfach gestalten sich die Berechnungen für die Einträge und das Lesen der Teamdateien sowie der Bewerberdaten. Im zentralen Fall entsteht für jede Folgesuche jeweils eine Transaktion, also täglich FS Transaktionen. In den anderen Szenarien wird für diese Vorgänge stets nur lokal und ebenfalls mit je einer Transaktion zugegriffen, so daß sich hier jeweils FmS Transaktionen ergeben.

Das Lesen der Stellendaten erfordert im zentralen Fall $KNT \times FS$ Transaktionen, in den anderen Szenarien wird unterschieden, ob sich die Stellendaten lokal oder entfernt befinden und für den entfernten Fall werden dann auch Formeln für die betroffenen entfernten Rechner aufgestellt. Insgesamt ergibt sich für die Szenarien 2 bis 7 auf den lokalen Rechnern:

Lesen lokaler Stellendaten $TransRlmS = FmS \times [K1\% \times KNT \times I_{2,4,6} + KR1\% \times KRNT \times I_3]$ (SZ 2 - 7)

Lesen entfernter Stellendaten $TransRrmS = FmS \times [(100 - K1)\% \times KNT \times I_{2,6} + (100 - KR1)\% \times KRNT \times I_3 + ((100 - K1)\% + K2\%) \times KNT \times I_4 + K3\% \times KNT \times I_{5,7}]$ (SZ 2 - 7)

Die erste Formel läßt deutlich erkennen, daß in den Szenarien 5 und 7 die Lösungen des Pendelbereichs durch Zugriffe auf die lokal vorliegenden SteA-Replikate erhalten werden können. Entfernte Transaktionen müssen dort daher nur für die verbleibenden $K3\%$ angestoßen werden, in den restlichen Szenarien hingegen für alle verbleibenden $(100 - K1)\%$ bzw. $(100 - KR1)\%$. Im Szenario 4 kommen dann noch $K2\%$ Zugriffsanforderungen durch Anfragen von anderen Rechnern des Pendelbereiches hinzu.

In den Szenarien 2 und 3 werden alle entfernten Lösungen vom zentralen Rechner gelesen, in den restlichen Szenarien wird hingegen ein Anteil von $K2\%$ entweder aus dem Pendelbereich, aus Replikaten oder über den LAA-Rechner gewonnen. Dabei erfolgt der Lesevorgang in Szenario 6 stets über den LAA-Rechner. So ergeben sich die Formeln für das Lesen entfernter Stellendaten auf zentralen Rechnern:

Bundesebene $TransRrzenS = FS \times [(100 - K1\%) \times KNT \times I_2 + (100 - KR1)\% \times KRNT \times I_3 + K3\% \times KNT \times I_{4,5,6,7}]$ (SZ 2 - 7)

LAA-Ebene in SZ 6 $TransRrLAAS = FS / LAA \times (K2 + K3)\% \times KNT$ (SZ 6)

Zur Bezeichnung der diversen Formeln wurden erneut folgende Bezeichner erforderlich:

<i>TransRlmS</i>	durchschnittliche Anzahl der täglich lokal in einem „mittleren“ Amt entstehenden Transaktionen durch das Lesen von SteAs des eigenen Datenbestandes nach bewerberorientierten Folgesuchläufen
<i>TransRrmS</i>	durchschnittliche Anzahl der täglich lokal in einem „mittleren“ Amt entstehenden Transaktionen durch Leseanforderungen auf SteAs von entfernten Rechnern
<i>TransRrzenS</i>	durchschnittliche Anzahl der täglich zentral entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen
<i>TransRrLAAS</i>	durchschnittliche Anzahl der täglich auf LAA-Rechnern entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen

Nun sind abschließend noch die Transaktionsanforderungen für Updates zu berücksichtigen. Von den gelesenen Stellendaten werden jeweils $lk\%$ akzeptiert, $lb\%$ Bewerberangebote sind betroffen. Auf Bewerberseite ist die Lage wieder einheitlich einfach. Wenn jedes Update eine Transaktion erfordert, sind im zentralen Fall $FS \times lb\%$ Transaktionen durchschnittlich täglich

erforderlich und ansonsten in einem „mittleren“ Arbeitsamt $lk\% \times FmS$. Entsprechend entstehen im zentralen Fall $lk\% \times KNT \times FS$ Updates in den Stellendaten.

Für die Updates der Stellen in den anderen Szenarien können diverse Situationen eintreten:

- a) selbst ausgelöste Updates auf den eigenen Datenbeständen,
- b) Updates auf entfernten Originalen von Stellendaten,
- c) Updates auf Replikaten,
- d) eingehende Update-Anforderungen auf den Originaldaten.

Update lokaler Stellendaten	$TransUplmS = lk\% \times FmS \times [KNT \times I_4$ $+ KR1\% \times KRNT \times I_3$ $+ K1\% \times KNT \times I_{2,5,6,7}]$	(SZ 2 - 7)
--------------------------------	--	------------

Update entfernter Stellendaten	$TransUPemS = lk\% \times FmS \times [KNT \times I_6$ $+ (100 - K1)\% \times KNT \times I_{2,4,5,7}$ $+ (100 - KR1)\% \times KRNT \times I_3]$	(SZ 2 -7)
-----------------------------------	--	-----------

Updates aus der Ferne	$TransUPamS = lk\% \times FmS \times$ $[(100 - K1)\% \times KNT \times I_{2,4,5,6,7}$ $+ (100 - KR1)\% \times KRNT \times I_3]$	(SZ 2 -7)
--------------------------	--	-----------

Am leichtesten lassen sich die beiden unteren Formeln nachvollziehen. $(100 - K1)\%$ (bzw. $[100 - KR1]\%$) aller gefundenen Stellendaten stammen aus dem Pendelbereich oder vom Zentralrechner und der hieraus akzeptierte Anteil muß dort aktualisiert werden. In Szenario 6 sind von allen SteAs Replikate mit Zuordnungsinformationen auf LAA-Ebene vorhanden. Deshalb ist dort auch der Anteil **K1** berücksichtigt. Für die Szenarien 5 und 7 ergibt sich trotz der Replikate der Pendelbereichs-Angebote keine Abweichung, da hier die Zuordnungsinformationen nur bei den Originaldaten eingetragen werden. Die Gestaltung der Formel macht deutlich, daß keine Abhängigkeit von der regionalen Ebene, auf der ein Stellenangebot gefunden wurde und der Zustimmung für einen Vermittlungsvorschlag modelliert wurde.

Die Updates aus der Ferne ergeben sich dadurch, daß andere Rechner das Update ihnen nicht vorliegender Originaldaten verursachen. Dies sind in allen Szenarien $(100 - K1)\%$ (in Szenario 3 analog mit **KR1**) der akzeptierten Lösungen.

Lokal entstehen in allen Szenarien **K1%** (bzw. **KR1%**) $\times lk\% \times Fms$ Update-Anforderungen. Warum in Szenario 4 mehr Updates veranschlagt werden, ist mir unklar; hier müßte meines Erachtens dieselbe Formel wie in den restlichen Szenarien gelten. | ①

Es bleiben die Transaktionen zu zählen, die in den Szenarien 2 bis 7 zentral oder auf LAA-Ebene dafür sorgen, den dortigen Bestand zu aktualisieren. Diese entfallen in den Szenarien 2 und 3 auf den prozentualen Anteil $100 - K1$ bzw. $100 - KR1$, in den anderen Szenarien werden nur jeweils $K3\%$ zentral gehalten. In Szenario 6 wird auf LAA-Ebene erneut berücksichtigt, daß dort alle Daten repliziert vorgehalten werden. Somit ergibt sich für diese Rechner:

Updates aus der Ferne	$TransUPaLAAS = lk\% \times FS / LAA \times KNT \quad (SZ\ 6)$ $TransUPazenS = lk\% \times FS \times [KNT \times K3\% \times I_{4,5,6,7} + KNT \times (100 - K1)\% \times I_2 + KRNT \times (100 - KR1)\% \times I_3]$	<p>(SZ 2 - 5, 7)</p>
--------------------------	--	----------------------

Der Vollständigkeit halber seien auch hier erneut die hinzugekommenen Variablenbezeichner angeführt:

<i>TransUpImS</i>	durchschnittliche Anzahl der täglich lokal entstehenden Transaktionen durch die Aktualisierung von SteAs des lokalen Datenbestandes nach bewerberorientierten Folgesuchläufen auf einem Rechner in einem „mittleren“ AA
<i>TransUpemS:</i>	durchschnittliche Anzahl der täglich lokal entstehenden Transaktionen durch Updateanforderungen auf SteAs auf entfernten Rechnern nach bewerberorientierten Folgesuchläufen auf einem Rechner in einem „mittleren“ AA
<i>TransUpamS:</i>	durchschnittliche Anzahl der täglich lokal entstehenden Transaktionen durch Updateanforderungen auf SteAs von entfernten Rechnern nach bewerberorientierten Folgesuchläufen auf einem Rechner in einem „mittleren“ AA
<i>TransUpzenS</i>	durchschnittliche Anzahl der täglich zentral entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen
<i>TransUpLAAS</i>	durchschnittliche Anzahl der täglich auf LAÄ-Rechnern entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen

Damit sind nun sämtliche Transaktionen, die aus bewerberorientierten Folgesuchläufen entstehen, modelliert. Zu ihnen gibt es eine gesonderte Beispielrechnung in der Studie (Anlage 1, Teil 2, S. 32), auf die oben bereits hingewiesen wurde. Wenn man diese nachvollzieht, so entdeckt man, daß abweichend zu den vorhergehenden Modellrechnungen **KR1** = 70% festgelegt wurde (vorher war 60% angegeben und mit 50% gerechnet worden) und im Gegensatz zu den Angaben in den Annahmen mit **K3** = 10 (statt 5) gerechnet wurde, allerdings ohne **K2** anzupassen. Dies erklärt zum Teil auch die Abweichungen der Beispielrechnungen zum Erstkontakt (vgl. Fußnote 53). Ebenso haben sich Tippfehler ergeben und wurden leider weiter übernommen. So stimmen etwa die Summen der Zahlenkolonnen von Szenario 3, 5 und 7 nicht. In Szenario 4 sind vermutlich durch fälschlicherweise entstandenes Kopieren von Formeln aus Szenario 3 für SteD und LISt (in der Studie verwandte Beschriftungen der Tabelle, vgl. ebenda) falsche Werte (insgesamt unwesentlich) entstanden. Deshalb habe ich mich entschieden, alle Modellrechnungen mit den für die Variablen gesetzten Werten im Anschluß an die zur bewerberorientierten Suche lediglich noch fehlende Behandlung des Löschens von BewA-Daten erneut abzubilden (s. Abbildung 17, S. Daten 104 ff.). Gegenübergestellt werden zudem die Gesamtergebnisse, die in den Übersichten der Studie auftreten, so daß es möglich ist, den Umfang der Abweichungen einzuschätzen. Im unteren Teil der Abbildung wurden die von mir auf den Seiten 80, 82, 84 und 97 angeführten alternativen Formeln, die zentrale Kennzeichnung (s. S. 90) sowie Anmerkung ① von Seite 102 berücksichtigt. Die Felder, an denen sich dies auswirkt, sind hier grün hervorgehoben. Die Eingabedaten wurden so gewählt, daß im Vergleich zu den Werten der Studie zumindest weitestgehend Übereinstimmung erzielt werden konnte.

Löschen von Bewerberdaten

Die Studie spricht nur von einer Kennzeichnung der zu löschenden Bewerberdaten. Lediglich für die Replikate im Pendelbereich von Szenario 7, zu dem durchschnittlich **p** Rechner gehören, ist ein echtes Entfernen vorgesehen⁵⁴. Mit den in der Studie angesetzten Formeln kann jedoch auch ein tatsächliches Löschen an allen Standorten widergespiegelt werden.

Die Formeln für das Löschen von Bewerberdaten, die außer in Szenario 1 ausschließlich lokal vorkommen, sind:

$$\text{TransLözenB} = \text{transLöB} \times \text{TAS} + \quad (\text{SZ 1})$$

$$\text{TransLömB} = \text{transLöB} \times \text{TmAS} + I_7 \times p \times \text{transLöB} \quad (\text{SZ 2 - 7})$$

mit:

<i>TransLözenB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zum Löschen von Bewerberdaten bzw. zu deren „gelöscht“-Kennzeichnung im zentralen Szenario 1
<i>TransLömB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zum Löschen von Bewerberdaten bzw. zu deren „gelöscht“-Kennzeichnung in einem „mittleren“ AA
transLöB	durchschnittliche Anzahl an Transaktionen zum Löschen von Bewerberdaten bzw. zu deren „gelöscht“-Kennzeichnung

Modellrechnungen - Gesamtübersicht

Mit der folgenden Variablenbelegung ergeben sich die Daten zu Abbildung 17 auf den Seiten 106 und 107.107

gesetzt						ermittelt			
arbD	7.500	KR1	70	s1	50	transSufi	1	f	20
AS+	2.500.000	KRNT	1,5	s2	45	transSufol	1	FmS	2.000
AS	5.000.000	L	3	TAZ	450	transSulokB	5	FS	400.000
AT	250	l	60	TN	10.000	transSulokV	5	K3	10
ATS	200	L1	70	transErfB	3	transSuremB	1	r2	50
f1	12	L2	30	transErfBP	1	transSuremV	1	s3	5
f2	4	LAA	10	transLöB	2	transUpN	1	TAS+	10.000
f3	4	lb	55	transLöBP	1	transUpPers	1	TEber	1,67
fa	2	lk	50	transLön	2	TZopen	240	TmAS+	50
K	3	N	25.000.000	translön*	1	X	50	TmN	50
K1	60	p	3	transSpBer	1	X*	95	transSpN	2
K2	30	q	5	transSpN*	1	ZEber	30	TZaFber	210
KNT	2	r1	50	transSpPers	1			TZEber	50
KR	3	RAA	200	transSuAS	1			TZFber	190

zu **Abbildung 17**: Gesamtübersicht über die auftretenden Transaktionen, Variablenbelegung

Wie man aus der Übersicht dieser Abbildung ersieht, spielen die Abweichungen bei den gesetzten Werten eine nur unwesentliche Rolle. Verwendet man hingegen die Testdaten, die

⁵⁴ Auch für die Löschung von Personendaten nach Aufgabenerledigung gibt es selbstverständlich behördliche Vorschriften und gesetzliche Vorgaben (insb. SGB X, § 67c ff). Eine unbegrenzte unbegründete Speicherung ist zumindest nicht zulässig. Eine erste Kennzeichnung vor dem wirklichen Löschen ist jedoch grundsätzlich üblich, z. B. um das Löschen asynchron durchführen zu können. Da die Modellierung den erforderlichen Speicherplatz nicht berücksichtigt, wird durch die von der Studie gewählte Modellierung somit lediglich die Zahl der Transaktionen und Datenübertragungen für das endgültige Entfernen der Daten nicht abgebildet.

auch bereits für Abbildung 16 auf Seite 91 in der zweiten Berechnung verwandt wurden, so liegen die zentralen Transaktionszahlen der Szenarien 2 und 3 immerhin ca. 20% über dem Wert der Studie. Dies ist vornehmlich auf die Berücksichtigung der zentralen Kennzeichnung der SteA-Replikate zurückzuführen. Niedrigere Werte, bis zu 9% niedriger als mit den Formeln der Studie, finden sich bei den lokalen Transaktionsanforderungen der Szenarien 4, 5 und 7. Hier wirkt sich insbesondere aus, daß die alternativen Formeln die Variablen L, L1 und L2 anders interpretieren.

Der Arbeit ist als Anlage 4 eine Diskette beigelegt, auf der eine mit dem Kalkulationsprogramm Microsoft Excel für Windows 95, Version 7.0a (der amerikanischen Firma Microsoft Corporation) erstellte Datei „snistudieübersicht.xls“ abgelegt ist, die eine Überprüfung der aufgestellten Tabelle ermöglicht. Hier findet man auf dem ersten Tabellenblatt die Ausgangsdaten der Beispielberechnung der Studie, auf Blatt 2 die Kalkulation, auf Blatt 3 erneut die Ausgangsdaten (hier aber alphabetisch angeordnet und nur zur Orientierung des Anwenders, nicht als Kalkulationsgrundlage) und darüber hinaus auf Blatt 4 Ausführungen zu Szenario 2 in Hinblick auf den folgenden Abschnitt 4.4.2.3. Diese Blätter sind (ohne Paßwort) gegen unabsichtliches Verändern geschützt worden. Darüber hinaus befinden sich zwei Kopien hiervon unter den Dateinamen „kalkbeispiel1.xls“ und „kalku.xls“ um veränderte Ergebnisse unter anderen Testwerten aufrufen zu können. Gegenüber dem Original wurde der Vergleich mit den Daten der Studie entfernt, da sich die Werte der Studie anderen Eingabewerten nicht anpassen können. Die Datei „kalkbeispiel1“ zeigt die Ergebnisse, die mit den oben erwähnten Alternativ-Eingabewerten erzielt werden. In „kalku.xls“ wurde der Speicherschutz für Blatt 1 entfernt, so daß man dort mit beliebigen Ausgangswerten Modellrechnungen durchführen kann. Damit leicht zu verifizieren ist, daß die Formeln korrekt eingegeben wurden, wurde so gut wie nicht mit Feldbezügen gearbeitet, sondern den Ausgangsdatenfeldern Namen vergeben und die Formeln mit diesen Namen formuliert. Dabei mußte darauf Rücksicht genommen werden, welche Namensbezeichner in diesem Anwendungsprogramm zugelassen sind. So werden Namen wie K1 oder L1 von Excel abgelehnt, da sie einer Feldzuordnung entsprechen. In solchen Fällen wurde zwischen Buchstabe und Ziffer der Kleinbuchstabe o - wenn nötig auch zwei - eingefügt. Statt einem „*“ wurden betreffenden Variablennamen die Zeichenfolge „stern“ angehängt. Der Variablenname „l“ wurde der besseren Lesbarkeit halber in „kleinl“ umbenannt.

Im folgenden wird aufbauend auf den Formeln der Studie weiter verfahren; Berechnungen beziehen sich dann auf die in Abbildung 17 oben abgebildeten (erneuerten) Daten. Eine Auswertung der Daten in Hinblick auf die Bewertung der diversen Szenarien erfolgt erst in Abschnitt 0.

Abbildung 17:

Gesamtübersicht über die auftretenden Transaktionen

Die Variablenbelegung ist eine Seite zuvor wiedergegeben

	Szenario 1		Szenario 2		Szenario 3	
	Zentral	lokal	zentral	lokal	zentral	
Speichern Stellen	20.000	100	10.000	100	5.000	
Löschen Stellen	20.000	100	10.000	100	5.000	
Suchen BewAs	50.000	250		250		
Teamdatei-Einträge	30.000	105		105		
Sachbearbeitertätigkeiten	126.000	441		441		
Folgeaktivitäten	412.000	1.592		1.592		
Zusammen (StorBewV)	658.000	2.588	20.000	2.588	10.000	
Erfassen Personen	30.000	150		150		
Suchen Stellen	50.000	275	5.000	263	2.500	
Update BewAs	10.000	50		50		
Update SteAs	30.000	150	30.000	150	9.000	
Zusammen (BewStVE)	120.000	625	35.000	613	11.500	
Folgesuchen	400.000	3.000	200.000	2.500	100.000	
Teamdatei-Einträge	400.000	2.000		2.000		
Lesen Teamdatei	400.000	2.000		2.000		
Lesen BewA-Daten	400.000	2.000		2.000		
lokales Lesen STeA-Daten	800.000	2.400		2.100		
Lesen entfernter STeA-Daten		1.600	320.000	900	180.000	
Update BewA-Daten	220.000	1.100		1.100		
Update STeA-Daten, lokal	400.000	1.200		1.050		
Update STeA-Daten, entfernt		800	160.000	450	90.000	
Update STeA-Daten v.entfernt		800		450		
Zusammen (BewStVF)	3.020.000	16.900	680.000	14.550	370.000	
StorBewV	658.000	2.588	20.000	2.588	10.000	
BewSTVE	120.000	625	35.000	613	11.500	
BewSTVF	3.020.000	16.900	680.000	14.550	370.000	
LöBew	20.000	100		100		
Zusammen Studie (erneuert)	3.818.000	20.213	735.000	17.851	391.500	

Daten der Studie	3.818.000	20.213	735.000	16.850	397.500
Differenzen	0	0	0	-1.001	6.000
prozentual	0%	0%	0%	6%	2%

Daten mit alternativen Formeln	3.818.000	20.273	742.800	17.896	401.400
Differenz zur Studie (erneuert)	0	60	7.800	45	9.900
prozentual	0%	0%	1%	0%	3%

StorBewV	658.000	2.588	57.800	2.588	28.900
BewSTVE	120.000	685	5.000	658	2.500
BewSTVF	3.020.000	16.900	680.000	14.550	370.000
LöBew	20.000	100		100	
Zusammen Alternativ	3.818.000	20.273	742.800	17.896	401.400
Speichern Stellen	20.000	100	10.000	100	5.000
Löschen Stellen	20.000	100	10.000	100	5.000
Suchen BewAs	50.000	250		250	
Teamdatei-Einträge	30.000	105		105	
Sachbearbeitertätigkeiten	126.000	441		441	
zentrale Kz			12.600		6.300
Folgeaktivitäten	412.000	1.592	25.200	1.592	12.600
Zusammen (StorBewV)	658.000	2.588	57.800	2.588	28.900
Erfassen Personen	30.000	150		150	
Suchen Stellen	50.000	275	5.000	263	2.500
Update BewAs	10.000	50		50	
Update SteAs	30.000	210		195	
Zusammen (BewStVE)	120.000	685	5.000	658	2.500
Folgesuchen	400.000	3.000	200.000	2.500	100.000
Teamdatei-Einträge	400.000	2.000		2.000	
Lesen Teamdatei	400.000	2.000		2.000	
Lesen BewA-Daten	400.000	2.000		2.000	
lokales Lesen STeA-Daten	800.000	2.400		2.100	
Lesen entfernter STeA-Daten		1.600	320.000	900	180.000
Update BewA-Daten	220.000	1.100		1.100	
Update STeA-Daten, lokal	400.000	1.200		1.050	
Update STeA-Daten, entfernt		800	160.000	450	90.000
Update STeA-Daten v.entfernt		800		450	
Zusammen (BewStVF)	3.020.000	16.900	680.000	14.550	370.000

Szenario 4		Szenario 5		Szenario 6			Szenario 7	
lokal	zentral	lokal	zentral	lokal	LAA	zentral	lokal	zentral
100	500	350	500	100	1.000	500	350	500
100	500	350	500	100	1.000	500	350	500
400		400		250			250	
195		195		105			195	
702		657		441			657	
2.594		2.504		1.592			2.204	
4.091	1.000	4.456	1.000	2.588	2.000	1.000	4.006	1.000
150		150		150			300	
343	500	253	500	275	500	500	253	500
50		50		50			50	
210	3.000	210	3.000	210	3.300	3.600	210	3.000
753	3.500	663	3.500	685	3.800	4.100	813	3.500
5.700	20.000	2.100	20.000	3.000	20.000	20.000	2.100	20.000
2.000		2.000		2.000			2.000	
2.000		2.000		2.000			2.000	
2.000		2.000		2.000			2.000	
2.400		3.600		2.400			3.600	
2.800	80.000	400	80.000	1.600	32.000	80.000	400	80.000
1.100		1.100		1.100			1.100	
2.000		1.200		1.200			1.200	
800	40.000	800	40.000	2.000	40.000	40.000	800	40.000
800		800		800	4000		800	
21.600	140.000	16.000	140.000	18.100	96.000	140.000	16.000	140.000
4.091	1.000	4.456	1.000	2.588	2.000	1.000	4.006	1.000
753	3.500	663	3.500	685	3.800	4.100	813	3.500
21.600	140.000	16.000	140.000	18.100	96.000	140.000	16.000	140.000
100		100		100			250	
26.544	144.500	21.219	144.500	21.473	101.800	145.100	21.069	144.500

Farblegende:

Die Farben in den Textfeldern dienen der leichteren Orientierung. Die zugeordneten Zahlenfelder gehen in die Summenbildung ein.

Die **rote** Markierung weist auf die Gegenüberstellung der diversen Ergebnisse hin.

Hellgelbe Zahlenfelder weisen auf Abweichungen von den Daten der Studie hin, dunkelgelbe auf Folgeauswirkungen hiervon.

Das **blaue** Zahlenfeld wurde in den Übersichten der Studie lediglich vergessen.

In den **grün** gekennzeichneten Zahlenfeldern wird eine alternative oder zusätzlich eingebrachte Formel wirksam

25.944	144.500	21.504	144.500	22.060	105.800	145.850	21.304	144.500
-600	0	286	0	587	4.000	750	236	0
2%	0%	1%	0%	3%	4%	1%	1%	0%

25.744	144.200	21.194	144.200	21.473	102.280	143.390	21.044	144.200
-800	-300	-25	-300	0	480	-1.710	-25	-300
-3%	0%	0%	0%	0%	0%	-1%	0%	0%

4.091	3.700	4.431	3.700	2.588	5.780	2.890	3.981	3.700
753	500	663	500	685	500	500	813	500
20.800	140.000	16.000	140.000	18.100	96.000	140.000	16.000	140.000
100		100		100			250	
25.744	144.200	21.194	144.200	21.473	102.280	143.390	21.044	144.200
100	500	338	500	100	1.000	500	338	500
100	500	338	500	100	1.000	500	338	500
400		400		250			250	
195		195		105			195	
702		657		441			657	
	900		900		1.260	630		900
2.594	1.800	2.504	1.800	1.592	2.520	1.260	2.204	1.800
4.091	3.700	4.431	3.700	2.588	5.780	2.890	3.981	3.700
150		150		150			300	
343	500	253	500	275	500	500	253	500
50		50		50			50	
210		210		210			210	
753	500	663	500	685	500	500	813	500
5.700	20.000	2.100	20.000	3.000	20.000	20.000	2.100	20.000
2.000		2.000		2.000			2.000	
2.000		2.000		2.000			2.000	
2.000		2.000		2.000			2.000	
2.400		3.600		2.400			3.600	
2.800	80.000	400	80.000	1.600	32.000	80.000	400	80.000
1.100		1.100		1.100			1.100	
1.200		1.200		1.200			1.200	
800	40.000	800	40.000	2.000	40.000	40.000	800	40.000
800		800		800	4000		800	
20.800	140.000	16.000	140.000	18.100	96.000	140.000	16.000	140.000

4.4.2.3 Tageszeitabhängigkeiten und Netzbelastung

Für die Bewertung der diversen Szenarien findet in der Studie noch eine gemeinsame Auswertung der Transaktionszahlen in Hinblick auf tageszeitliche Abhängigkeiten und entstehende Netzbelastungen statt. Diese wird in der Studie an einer weiteren Beispielberechnung (vgl. Anlage 1, Teil 2, S. 33 - 39) dargestellt. Die Grundlagen dieser Berechnung werden jedoch nur sehr knapp erwähnt. Die zugrunde liegenden Formeln sind leider in der Studie nicht angegeben.

Es hätte den Zeitrahmen dieser Arbeit gesprengt, diese Formeln neu für alle Szenarien herzuleiten, insbesondere da erste Versuche in diese Richtung im Vergleich zu den Werten der Studie teilweise zu erneuten Abweichungen führten. Ich habe mich daher entschlossen, nach Darstellung der in der Studie dargelegten Berechnungsgrundlagen eine Formelherleitung nur exemplarisch für das Szenario 2 zu gestalten. Die Wahl fiel deshalb auf Szenario 2, weil es unter den Szenarien mit Datenverteilung das am einfachsten strukturierte Modell darstellt. Teils können diese Ergebnisse unmittelbar auch auf die anderen Szenarien übertragen werden, teils müssten sie erneut entwickelt werden.

Tageszeitabhängigkeiten

Um Aussagen darüber machen zu können, wie viele Transaktionen je Sekunde in den diversen Szenarien bewältigt werden müssen, berücksichtigt die Studie tageszeitlich bedingte Unterschiede in der Sachbearbeitung (vgl. Anlage 1, S. 52 - 53). Es werden drei Tagesbereiche (TB) unterschieden:

 Tagesbereich 1	Anforderungen während des Publikumsverkehrs
 Tagesbereich 2	Aktivitäten der Sachbearbeiter während der Arbeitszeit, unabhängig vom Publikum, kann auch während der Publikumszeiten stattfinden, sofern dort Zeit verbleibt
 Tagesbereich 3	Aktivitäten, die unabhängig von der Arbeitszeit eines Sachbearbeiters stattfinden können

Abbildung 18: Einteilung in Tagesbereiche

Die Dauer des ersten Tagesbereiches entspricht der Setzung der Variablen **TZopen**. für die beiden anderen Bereiche werden die Variablen **TZTB2** und **TZTB3** eingeführt:

TZTB2	durchschnittlich tägliche Zeit, die für Aktivitäten des Tagesbereiches 2 (Arbeiten ohne Publikum) zur Verfügung steht
TZTB3	durchschnittlich tägliche Zeit, die für Aktivitäten des Tagesbereiches 3 (ohne Sachbearbeiteranwesenheit) zur Verfügung steht

Wie die im Modell bisher betrachteten Vorgänge zu den drei Tagesbereichen zugeordnet werden, ist in der folgenden Tabelle wiedergegeben.

Stellenorientierte Bewerbervermittlung		
Speichern, Löschen, Erstsuche, Speichern in die Teamdatei	durch den Sachbearbeiter	gleichmäßig über TB1 und TB 2
Entscheidungsfindung	durch den Sachbearbeiter	TB 2
Folgeaktivitäten -Suchen	asynchron	TB 3
Folgeaktivitäten -Entscheiden	durch den Sachbearbeiter	TB 2
Bewerberorientierte Stellenvermittlung		
Erstkontakt	Sachbearbeiter mit Bewerber	TB 1
Folgeaktivitäten - Suchvorgänge bei unangemeldeten Besuche (Faktor f1)	Sachbearbeiter mit Bewerber	TB 1
- Suchvorgänge für angemeldete Besucher (Faktor f2)	asynchron	TB 3
- Nachbearbeitung der Folgesuchen	Sachbearbeiter	TB 2

Tabelle 10: Zuordnung der modellierten Vorgänge zu Tagesbereichen

Etwas überrascht hat mich hierbei die fehlende Zuordnung des „Speichern in die Teamdatei“ bei den asynchronen Folgeaktivitäten. Bis zu den hier angegebenen Zuordnungen hatte ich angenommen, daß nach einer stellenorientierten Bewerbersuche das System automatisch diese Einträge vornimmt. Nach einer vom Sachbearbeiter ausgelösten Suche würde dieser dann sofort seine Teamdatei-Einträge überprüfen. Die hier beschriebene Zuordnung läßt jedoch vermuten, daß das Modell an dem Rechner, an dem die Suche gestartet wird, eine komplette Präsentation des Ergebnisses vorsieht. Dies wäre ja in der Tat auch sinnvoll. Das Ergebnis würde dann anschließend vom Sachbearbeiter ausgewertet, wobei er die notwendigen Teamdatei-Einträge nun selbst erledigt.

Netzbelastung

Für die Bewertung der Netzbelastung werden die Nachrichtenlängen der übertragenden Nachrichten und die Entfernung, die sie zurücklegen müssen, berücksichtigt. In Abhängigkeit von der Anzahl der zu den verschiedenen Tagesbereichen entstehenden Datenfernübertragungen wird hier das Produkt aus Nachrichtenlänge (in Gigabyte, (GB)) und Entfernung (in km) gebildet.

Hierzu werden ein letztes Mal weitere Variablenbezeichner erforderlich:

mNRL1	mittlere Nachrichtenlänge eines Auftrages
mNRL2	mittlere Nachrichtenlänge des Ergebnisses eines Auftrages
mNRLg	mittlere Nachrichtenlänge aus Auftrag inklusive Ergebnis, $mNRLg = mNRL1 + mNRL2$
mEntfAZ	mittlere Entfernung eines AA-Rechners zum Zentralrechner
mEntfAA	mittlere Entfernung zwischen den AA-Rechnern
mEntfAL	mittlere Entfernung des AA-Rechners zum LAA-Rechner
mEntfLZ	mittlere Entfernung eines LAA-Rechners zum Zentralrechner

Die Modellberechnungen für Szenario 2

Für die Beispielrechnungen im Anlage 1, Teil 2, S. 33 - 39 wurden folgende Werte als Grundlagen angegeben:

Belegungen für die Tagesbereichzuordnungen in Stunden:

TZopen 4 **TZTB2** 4 **TZTB3** 15

Belegungen zur Ermittlung der Netzbelastung

mNRL1 = 0,5 (in KB)
mNRL2 = 1 (in KB)
 ⇒ **mNRLg** = 1,5 (in KB)

mEntfAZ = 500 (in km) **mEntfAA** = 50 (in km)
mEntfAL = 300 (in km) **mEntfLZ** = 300 (in km)

Wie bereits oben erwähnt sollen nur für Szenario 2 die in der Studie ermittelten Werte nachvollzogen werden. Diese sind in der folgenden Tabelle nachgebildet worden. Dabei steht TA/sec für Transaktionen pro Sekunde.

Aktivität	Tagesbereich 1 mit Publikum				Tagesbereich 2 ohne Publikum				Tagesbereich 3 ohne Sachbearbeiter				Netzlast insgesamt	
	TA/sec		Netzlast.		TA/sec		Netzlast.		TA/sec		Netzlast.		Anzahl	GB * km
	lokal	zentral	Anzahl	GB * km	lokal	zentral	Anzahl	GB * km	lokal	zentral	Anzahl	GB * km		
SZ2														
SteA->BewA	0,016	0,69	10.000	7,5	0,016	0,69	10.000	7,5	0,013				20.000	15,0
Bewerberbearbeitung					0,091									
BewA->SteA	0,043	2,43	35.000	26,25									35.000	26,25
Folgeberatungen	0,253	9,43	168.000	126	0,664	26,67	512.000	384,0	0,07	3,7	160.000	120,0	840.000	630,0
gesamt	0,312	12,55	213.000	159,75	0,851	27,36	522.000	391,5	0,083	3,7	160.000	120,0	895.000	671,00

Tabelle 11: Transaktionsraten und Netzbelastung nach Tagesbereichen für SZ 2

Um die Auswertung dieses Szenarios zu verstehen, werden zunächst die Angaben aus der stellenorientierten Bewerbersuche (s. rechts) hinzugezogen, die zu den Daten der ersten beiden Reihen der obigen Tabelle führen müssen.

SteA->BewA	Szenario 2	
	lokal	zentral
Speichern Stellen	100	10.000
Löschen Stellen	100	10.000
SuchenBewAs	250	
Teamdatei-Einträge	105	
Sachbearbeitertätigkeiten	441	
Folgeaktivitäten	1.592	
Zusammen (StorBewV)	2.588	20.000

Tabelle 12: Transaktionszahlen SZ 2 (1)

Zunächst soll die Anzahl der Datenübertragungen erklärt werden. Im Rahmen der stellenorientierten Bewerbersuche fallen insgesamt 22.588 Transaktionen an, davon 20.000 zentral. Diese 20.000 Transaktionen erfordern je einen entfernten Zugriff zwischen Arbeitsamt und Zentralrechner. Damit entstehen 20.000 Datenübertragungen. Da sie beim Speichern und Löschen von Stellen auftreten, gehören sie zu den Vorgängen, die in der Zuordnung zu den Tagesbereichen gleichmäßig über TB 1 und TB 2 verteilt werden. Damit sind die beiden Felder mit den 10.000 Datenübertragungen erklärt. Unmittelbare Konsequenz hiervon ist die Summe 20.000 in der Darstellung der Gesamtbelastungen für das Netz.

Mit der Anzahl der Datenübertragungen lassen sich dann die in den jeweils rechts benachbarten Feldern dargestellten Netzbelastungen (NB⁵⁵) über die folgende Formel ermitteln:

$$\begin{aligned} \text{NB} &= \text{Anzahl} \times \text{mNRLg} \times \text{mEntfAZ} / 1.000.000 \\ &= 10.000 \times 1,5 \times 500 / 1.000.000 = 7,5 \end{aligned}$$

Hier wird jede Datenübertragung mit ihrer durchschnittlichen Datenmenge und der durchschnittlichen Entfernung zwischen Arbeitsamt und Zentralrechner multipliziert. Die abschließende Division sorgt für die Umrechnung in Gigabyte. Für die Spalte „Netzlast insgesamt“ werden die Angaben aufsummiert.

Die Angaben zu den zentralen Transaktionsraten in der ersten Zeile gehen ebenso auf die 20.000 zentralen Transaktionen zurück. Aufgrund der Tageszeitzuordnung ist von einer gleichmäßigen Verteilung dieser Transaktionen über die Tagesbereiche 1 und 2 auszugehen. Dementsprechend ermittelt man die Transaktionsrate (TR, beachte Fußnote 55) je Sekunde durch Division durch die Gesamtdauer dieser Tagesbereiche, die hierfür in Sekunden umzurechnen ist:

$$\begin{aligned} \text{TR} &= \text{Anzahl_zentraler_Transaktionen} / ((\text{TZopen} + \text{TZTB2}) \times 60) \\ &= 20.000 / (480 \times 60) \approx 0,694 \approx 0,69 \end{aligned}$$

Aus den beiden oberen Zeilen von Tabelle 11 fehlen nun nur noch die Formeln für die lokal angegebenen Transaktionsraten. Diese sind nach Stellen- und Bewerberbearbeitung aufgeteilt. Da für die stellenorientierte Bewerbervermittlung der Tagesbereich 1 nur in Verbindung mit Speichern, Löschen, Erstsuche und Speichern in die Teamdatei vorkommt (im Rahmen der schon eben betrachteten gleichmäßigen Verteilung über die Tagesbereiche 1 und 2), müssen die Angaben der ersten Zeile auf den lokalen Transaktionen dieser Vorgänge basieren. Dies sind 555. Damit ergibt sich:

$$\text{TR} = 555 / (480 \times 60) \approx 0,019$$

Der Tagesbereich 3 wiederum taucht in der Zuordnungstabelle für die stellenorientierte Bewerbervermittlung nur in Verbindung mit den Folgeaktivitäten auf, hier aber nur in Hinblick auf das Suchen. Für die Folgebearbeitung galt ganz allgemein in den Szenarien 2 bis 7:

$$\text{TransFolmN} = \text{fa} \times (\text{TransSumB} + \text{TransWrmTD} + \text{TransSachmB})$$

Daher muß für die Suche hieraus $\text{fa} \times (\text{TransSumB})$ ausgegliedert werden und es ergibt sich hier:

$$\text{TR} = \text{fa} \times \text{TransSumB} / (\text{TZTB3} \times 60) = 2 \times 250 / (900 \times 60) \approx 0,009$$

Nun bleiben lediglich die lokalen Transaktionen für die Sachbearbeitertätigkeiten und die noch offenen Folgeaktivitäten (erneut Sachbearbeitertätigkeiten und Teamdatei-Einträge) übrig. Diese müssen dem Tagesbereich 2 zugeordnet werden.

$$\text{TR} = (\text{fa} \times (\text{TransWrmTD} + \text{TransSachmB}) + \text{TransSachmB}) / \text{TZTB2} \times 60$$

⁵⁵ NB und TR wurden nicht mehr in die Variablenübersicht in Anlage 3 aufgenommen.

$$= (2 \times (441+105) + 441) / (240 \times 60) \approx 0,1065$$

Wie man leicht sieht, decken sich diese Angaben für die lokalen Transaktionsraten **nicht** mit denen, die in der Studie angegeben sind. Trotz umfangreicher Überlegungen und Tests ist es mir leider nicht gelungen, hierfür die Ursache herauszufinden.

Es verbleiben noch die beiden Zeilenangaben zur bewerberorientierten Stellensuche. Die erste Zeile bezieht sich ausschließlich auf die Vorgänge zu Erstkontakten. Die Anzahlen der hierbei auftretenden Transaktionen sind nebenstehend wiedergegeben. Die Zuordnung zu den Tagesbereichen gestaltet sich einfach: Alle Vorgänge fallen in den Tagesbereich 1.

BewA->SteA, Erstkontakt	Szenario 2	
Erfassen Personen	150	
Suchen Stellen	275	5.000
Update BewAs	50	
Update SteAs	150	30.000
Zusammen (BewStVE)	625	35.000

Tabelle 13: Transaktionszahlen SZ 2 (2)

Über die zentralen Transaktionen, die wieder jede mit genau einer Datenübertragung verbunden sind, erhält man erneut die Zahl der Datenübertragungen, nämlich 35.000. Mit denselben Formeln wie zuvor gelangt man zu den Angaben für die GB × km-Angabe (26,25). Damit sind automatisch auch die Netzgesamtelastungen ermittelt, da hier kein weiterer Tagesbereich betroffen ist. Ebenso ermitteln sich die zentralen Transaktionsraten wie gehabt und führen zu den Angaben der Studie. Und auch die lokale Transaktionsrate läßt sich für die Erstkontakte entsprechend dem in der Studie angegebenen Wert ermitteln. Sie ergibt sich aus der Summe über alle lokalen Transaktionen bezogen auf den Tageszeitraum des Tagesbereiches 1 in Sekunden:

$$TR = (150 + 275 + 50 + 150) / (240 \times 60) \approx 0,0434$$

Für die Folgeberatung treten Netz- und Transaktionsanforderungen zu allen drei Tagesbereichen auf. Die Gesamtzahl aller Datenübertragungen deckt sich hier erstmals **nicht** mit der Zahl der zentralen Transaktionen. Zu den in der Anzahl den zentralen Transaktionen entsprechenden Datenübertragungen zum Zentralrechner kommen in der Folgeberatung erstmals auch Übertragungen in der Rückrichtung zum Tragen. Hier werden an alle AA-Rechner 800 Update-Aufträge für entfernte Original-SteA-Daten versandt. Es kommen daher weitere $200 \times 800 = 160.000$ Datenübertragungen hinzu. Damit ist die Gesamtzahl von $680.000 + 160.000 = 840.000$ Datenübertragungen erklärt.

BewA->SteA, Folgesuchläufe	Szenario 2	
Folgesuchen	3.000	200.000
Teamdatei-Einträge	2.000	
Lesen Teamdatei	2.000	
Lesen BewA-Daten	2.000	
lokales Lesen STeA-Daten	2.400	
Lesen entfernter STeA-Daten	1.600	320.000
Update BewA-Daten	1.100	
Update STeA-Daten, lokal	1.200	
Update STeA-Daten, entfernt	800	160.000
Update STeA-Daten v.entfernt	800	
Zusammen (BewStVF)	16.900	680.000

Tabelle 14: Transaktionszahlen SZ 2 (3)

Zunächst sollen die ersten 200.000 Übertragungen in Richtung Zentralrechner für die Folgesuchen zu den Tagesbereichen zugeordnet werden. Hierzu müssen die Faktoren **f1**, **f2** und **f3** berücksichtigt werden, über die modelliert wurde, wie häufig eine solche Suche periodisch, durch den Arbeitsvermittler alleine (z. B. zu einem geplanten Besuchstermin) oder mit einem

Bewerber gemeinsam ausgelöst wird. Die in der Beispielberechnung der Studie angenommenen Werte hierfür sind:

$$f_1 = 12, \quad f_2 = 4, \quad f_3 = 4$$

Der Anteil f_1/f ($3/5$) fällt in die Zeitspanne 3. Hiermit werden die periodischen Suchen zugeordnet. Ebenso können durch den Sachbearbeiter terminierte Suchen innerhalb dieses Zeitraumes abgewickelt werden. Damit hat man $4/5$ der 200.000 Transaktionen, also 160.000 dem Zeitbereich 3 zugeordnet und die hier erforderliche Zahl der Fernübertragungen bereits zusammen. Mit Besuchern gemeinsam wird in den verbleibenden $1/5$ der Fälle gesucht. Daher müssen die aus diesen Transaktionen resultierenden 40.000 Fernübertragungen im Tagesbereich 1 berücksichtigt werden.

Damit ergibt sich:	TB 1:	40.000	(bisher)
	TB 2:	0	(bisher)
	TB 3:	160.000	(fertig)

Zu verteilen bleiben noch die folgenden Datenübertragungen:

für Lesen entfernter SteA-Daten:	320.000	} 480.000
für Update SteA-Daten, entfernt:	160.000	
<u>für Update SteA-Daten von entfernt:</u>	<u>160.000</u>	
insgesamt:	640.000	

Hierbei handelt es sich in allen drei Fällen um Vorgänge, die durch den Sachbearbeiter initiiert werden. Wieder kommen drei Situationen in Frage. Der Vorgang kann erforderlich werden im Anschluß an eine periodisch stattfindende Suche ($3/5$ aller Fälle), im Rahmen einer vom Sachbearbeiter allein ausgelösten Suche ($1/5$) oder während eines Besuchergesprächs ($1/5$). Die beiden ersten Situationen fallen in den Tagesbereich 2, die letzte in den Tagesbereich 1. Damit hat man:

$$640.000 \times 4/5 = 512.000 \quad \text{Datenübertragungen für den Tagesbereich 2}$$

$$640.000 \times 1/5 = 128.000 \quad \text{Datenübertragungen für den Tagesbereich 1}$$

Da für den Tagesbereich 1 bereits 40.000 Übertragungen verzeichnet wurden, ergeben sich hier zusammen 168.000. Somit sind die Werte der Studie nachvollzogen. Entsprechend ergeben sich die rechts daneben stehenden Zahlen wieder als Produkt aus der Anzahl der Übertragungen und der durchschnittlichen Entfernung zwischen Ämtern und Zentralrechner unter Berücksichtigung der Umrechnung in die geforderte Einheit Gigabyte.

Zu untersuchen sind nun nur noch die Angaben zu den Transaktionsraten. Von den 680.000 zentralen Transaktionen fanden $1/5$ im Tagesbereich 1 statt, das sind 136.000. Teilt man diese durch die in Sekunden gemessene Dauer des TB 1, so erhält man auch hier den gesuchten Wert der Studie. In Tagesbereich 2 liegen $4/5$ der 480.000 zentralen Transaktionen für das Lesen und Aktualisieren entfernter SteA-Daten, ansonsten keine weiteren. Damit erhält man hier wieder übereinstimmend mit der Studie:

$$TR = 4/5 \times 480.000 \times \mathbf{TZTB2} / 60 = 26,67$$

Im Tagesbereich 3 wurden 160.000 Transaktionen zentral für Folgesuchen abgewickelt. Dies ergibt eine Transaktionsrate von 2,96; ein Wert der leider von dem der Studie abweicht.

Die Einteilung der lokalen Transaktionen gestaltet sich wie folgt:

Die Folgesuchen werden in **f3** aller **f** Fälle während des Publikumverkehrs (TB 1), in **f1 + f2** aller **f** Fälle asynchron abgewickelt (TB 3). Asynchron ausgelöste Folgesuchen sorgen auch für ein automatisches Schreiben in die Teamdatei. Daher werden dem TB 3 weitere 1.600 Transaktionen zugeordnet.

	<i>gesamt</i>	<i>TB 1</i>	<i>TB 2</i>	<i>TB3</i>
Folgesuchen	3.000	600	0	2.400
Teamdatei-Einträge	2.000	400	0	1.600
Lesen Teamdatei	2.000	400	1600	
Lesen BewA-Daten	2.000	400	1600	
lokales Lesen STeA-Daten	2.400	480	1920	
Lesen entfernter STeA-Daten	1.600	320	1280	
Update BewA-Daten	1.100	220	880	
Update STeA-Daten, lokal	1.200	240	960	
Update STeA-Daten, entfernt	800	160	640	
Update STeA-Daten v.entfernt	800	160	640	
Zusammen (BewStVF)	16.900	3380	9520	4000
Transaktionsraten - Version 1		0,2347	0,6611	0,0741
		156		
Transaktionsraten - Version 2		0,246	0,650	
Werte der Studie		0,253	0,664	0,07

Tabelle 15: Transaktionszahlen SZ 2 (4)

In Verbindung mit einem Besucher ausgelöste Suchen werden inklusive aller Folgeaktivitäten während des Besuchszeitraumes erledigt. Damit erklärt sich die gesamte Spalte zum Tagesbereich 1.

Berechnet man dann wie zuvor die Transaktionsrate, so erhält man ausschließlich für TB3 direkt den Wert der Studie (vgl. Zeile: „Transaktionsraten - Version 1“). Als denkbare Ursache hierfür nahm ich an, daß bei den Arbeitsgängen im TB 2 Situationen entstehen können, die zu einer Einladung eines Arbeitssuchenden führen und dann bezüglich der in TB 2 fett markierten Transaktionen zu einer Verlagerung in den Tagesbereich 1 führen können. Aber auch eine Berücksichtigung von 5% solcher Fälle (+156 in der linken Spalte, -156 in der rechten Spalte) führte nicht zu Deckungsgleichheit. Ebenso habe ich die Möglichkeit in Betracht gezogen, daß die lokalen Transaktionen zum Löschen von Bewerbern bei bewerberorientierten Stellensuchen, dies sind in SZ 2 genau 100, mit berücksichtigt wurden. Dies scheint jedoch auch nicht der Fall zu sein, so daß sie möglicherweise in der Studie völlig vergessen wurden. Insgesamt bleibt also auch bei diesen beiden Feldern die genaue Ursache für die Abweichung von der Vorgabe unklar.

Die Übersicht schließt ab mit einer Summenbildung über alle Spalten. Hier findet sich in Spalte 1 des Tagesbereich 2 in den Daten der Studie ein Rechen- oder Tippfehler.

Auch wenn einige Feinheiten offen blieben, im wesentlichen dürfte durch diese Erklärungen der Berechnungen zu Szenario 2 deutlich geworden sein, wie in der Studie verfahren wird, um eine etwas exaktere Beurteilungsgrundlage für die Bewertung der verschiedenen Szenarien zu erhalten. Dabei entstehen natürlich in den Varianten mit Pendelbereichzugriffen oder mit Zugriffen auf LAA-Daten etwas komplexere Situationen und es sind unterschiedliche Entfernungen zu berücksichtigen.

Eine Analyse der Übersichten wird in der Studie selbst vorgenommen und wird in knapper Form im folgenden Abschnitt erwähnt.

4.4.3 Lösungsvorschlag der Studie

Für die Bewertung der diversen Szenarien werden in der Studie eine Reihe von Kriterien herangezogen (vgl. Anlage 1, S. 55 - 58). So wird verglichen, auf welchen Datenbestand beim jeweiligen Szenario Zugriff besteht. Dieser wird als die Auswahlmenge bezeichnet. Ebenso wird betrachtet, welche Einschränkungen sich bei Netzausfall in Hinblick auf unmittelbar erforderliche Anforderungen ergeben. Und schließlich werden die Transaktionsraten in Abhängigkeit zu unterschiedlichen Tageszeitabschnitten sowie die Netzbelastung berücksichtigt (vgl. Tabelle 16). Nicht betrachtet wird der unterschiedliche Speicherbedarf der Rechner in den verschiedenen Szenarien aufgrund der Replikhaltung.

Auswahlmenge

Bei der bewerberorientierten Stellensuche, von der in der Modellierung angenommen wird, daß sie den Alltag stärker dominiert als die stellenorientierte Bewerbersuche, wird nach einer noch nicht vollständig erfolgreichen lokalen Suche entfernt weiter gesucht. Hierbei besteht in allen Szenarien Zugriff auf sämtliche Stellen ausgewählter Berufsgruppen. Dies kann also von den folgenden Betrachtungen ausgenommen werden.

Ansonsten besitzt man für die bewerberorientierten Suchen nur in Szenario 1 und 2 Zugriff auf alle möglichen Lösungen. In Szenario 3 wird nur ein Teil aller Stellen zentral gehalten und die Ämter können nicht untereinander kommunizieren. Deshalb herrschen in den unterschiedlichen Ämtern unterschiedliche Sichten auf den Datenbestand. Alle rein lokalen Stellenangebote anderer Ämter sind unsichtbar, auch die der Nachbarämter im Pendelbereich. Demgegenüber sind in den Szenarien 4, 5 und 7 alle Stellen des Pendelbereiches verfügbar. In Szenario 6 schließlich hat man über die LAA-Rechner Zugriff auf alle Stellen des eigenen LAA-Bezirktes. Stellen aus Pendelbereichen, die über den eigenen Bezirk hinausgehen, sind unsichtbar.

Bei der stellenorientierten Bewerbersuche wird außer bei den Szenarien 4 und 5 nur auf dem Datenbestand des Rechners, von dem die Suche startet, gesucht. Entsprechend kann nur in Fall 1 auf alle Stellen zugegriffen werden. In den Szenarien 2, 3, und 6 steht nur der jeweilige lokale Bestand des zuständigen Arbeitsamtes zur Verfügung. Für die Szenarien 4, 5 und 7 hingegen können auch Bewerber aus dem Pendelbereich berücksichtigt werden, bei Szenario 7 durch hier vorliegende Replikate, in den beiden anderen Szenarien durch entfernte Zugriffe.

SZ	Auswahlmenge		Netzausfall	Transaktionsrate	Netzbelastungen
	BewA→SteA	SteA→BewA			
1	alle Daten	alle Daten	kein Arbeiten mgl.	sehr hoch	sehr hoch
2	alle Daten	nur AA-Daten	nur AA-Daten	zentral hoch	hoch
3	unvollständig, auch im PB	nur AA-Daten	nur AA-Daten	zentral hoch	mittel
4	unvollständig nur außerhalb des PB	PB-Daten	je nach Ausfall über den AA-Bestand hinaus	gering	hoch
5	unvollständig nur außerhalb des PB	PB-Daten	AA-Daten und SteAs des PB	gering	gering
6	künstliche Einteilung in LAA-Bezirke, unvollständig außerhalb des LAA-Bezirks	nur AA-Daten	nur AA-Daten	auf dem LAA-Rechner hoch	zum LAA-Rechner hoch
7	unvollständig nur außerhalb des PB	PB-Daten	nur zentrale Daten fehlen	gering	gering

Tabelle 16: Bewertungsaspekte

Netzausfall

In Szenario 1 gibt es bei einem Verbindungsverlust zum Zentralrechner keinerlei Möglichkeit mehr, Vermittlungen durchzuführen. In den Szenarien 2, 3 und 6 kann bei Netzausfall nur noch lokal weitergearbeitet werden. Dies wird auch zu Szenario 4 angegeben. Ein Totalausfall des Netzes für Szenario 4 ist aber eigentlich allenfalls mit dem Ausfall der Verbindung zum zentralen Rechner nicht mit dem Ausfallen einer Verbindung in den anderen Szenarien zum Zentralrechner oder LAA-Rechner zu vergleichen, so daß davon auszugehen ist, daß neben den lokalen Daten zumindest eingeschränkt weiter der Pendelbereich zur Verfügung steht. Szenario 5 hat zudem noch den Vorteil, daß lokal Replikate von SteAs des Pendelbereichs zur Verfügung stehen und in Szenario 7 schließlich liegen sogar die BewAs des Pendelbereichs als Replikate vor.

Transaktionsrate

Für Szenario 1 können extreme Transaktionsraten erforderlich werden. Es können Bereiche auftreten, „die nur von einem Hochleistungszentralsystem bewältigt werden können (> 100 TA/sec)“ (Anlage 1, S. 56). Die Zentralrechner der anderen Szenarien werden nur in den Szenarien 2 und 3 mit mehr als 10 Transaktionen je Sekunde belastet, wofür sich mittlere bis große UNIX-Rechner eignen würden. Für Szenario 6 müßten die Rechner in den LAÄ ausgebaut werden. Die lokalen Belastungen sind in allen Szenarien um 1 TA/sec oder geringer und stellen daher keine außergewöhnlichen Anforderungen.

Netzbelastung

Die Netzbelastung ist naturgemäß in Szenario 1 am höchsten. Sie wird in den Szenarien 2 und 3 zunehmend auf die lokalen Rechner verlagert, wobei dies in Szenario 3 gegenüber Szenario 2 auf Kosten der möglichen Lösungen erfolgt. In Szenario 6 werden die Verbindungen zwischen den AÄ und den zugehörigen LAÄ stark belastet. Die Szenarien 4, 5 und 7, die bezüglich der Auswahlmenge vergleichbar sind, weisen unter Replikatnutzung (SZ 5, SZ 7) eine deutlich geringere Netzbelastung auf. Zwischen Szenario 5 und 7 ist der Unterschied nicht so groß. Er würde aber mit zunehmender Bedeutung der stellenorientierten Bewerbervermittlung zugunsten von Szenario 7 ausschlagen, weil dort auch Replikate der Bewerber vorgehalten werden.

Vorschlag der Studie

Aufgrund dieser Kriterien wird von der Studie empfohlen, Szenario 5 oder 7 zu wählen, da hier eine vollständige Suche im Pendelbereich möglich ist und zudem die niedrigsten Transaktionsraten und die geringsten Netzbelastungen vorliegen. Zudem ermöglichen sie auch bei Netzausfall bewerberorientierte Suchen im Pendelbereich. In Szenario 7 wären zudem stellenorientierte Suchen im Pendelbereich möglich.

4.5 Grenzen des SNI-Modells

Bevor die Grenzen des SNI-Modells aufgezeigt werden, sei zunächst noch einmal daran erinnert, daß die Studie von SNI zu Präsentationszwecken und nicht mit dem Anspruch auf Vollständigkeit entstanden ist (vgl. S. 70). Für die von SNI erstellte Ausarbeitung bestand kein Auftrag von Seiten der BA, d.h. sie wurde auf eigene Kosten angefertigt. In diesem Kontext ist klar, daß der Detailtiefe und der Betrachtung möglicher Fälle schon vom Rahmen her Grenzen gesetzt waren. An welche Grenzen die Studie letztendlich stößt, das soll im folgenden herausgearbeitet werden.

Dazu möchte ich das SNI-Modell zunächst in den Kontext der in Kapitel 3 beschriebenen Modelle rücken, um eine Grundlage zu schaffen, Ergebnisse dort untersuchter Aspekte auf die SNI-Studie übertragen zu können. Ebenfalls sollen die schon während der Darstellung der Studie getroffenen Anmerkungen zu Formeln wieder aufgegriffen und zusammengestellt werden. Dies führt schließlich zu allgemeineren Überlegungen zum konzeptionellen Ansatz des SNI-Modells.

Das SNI-Modell im Kontext theoretischer Betrachtungen

Wie für die in Kapitel 3 beschriebenen theoretischen Modelle erforderlich, werden im SNI-Modell zunächst zu den stattfindenden Anforderungen auf dem Datenbestand grundlegende Annahmen getroffen (allgemeine Modellbildung). Hierauf basierend erfolgt eine genaue Analyse der Anforderungen, die zunächst zu verschiedenen Typen von Anforderungen führt. Diese werden einer Formalisierung in Hinblick auf ihr Zugriffsverhalten unterzogen. In Verbindung mit einigen Annahmen über das Netz und tageszeitliche Gegebenheiten folgt eine Auswertung in Hinblick selbst gesetzter Bewertungsmerkmale. In dieser Vorgehensweise ähnelt die SNI-Studie zunächst den Modellbildungen der ersten drei Abschnitte aus Kapitel 3.

Dabei werden nicht nur wie bei Ceri und Pelagatti die entfernten Zugriffe berücksichtigt, sondern auch lokales Zugriffsverhalten, Mittelwerte für die Übertragungsmenge sowie Entfernungen. Eine einzige Zielfunktion wird nicht aufgestellt, statt dessen werden für jedes Szenario gesondert Formeln zur Beschreibung der Gesamtbelastungen in Transaktionsraten je Sekunde auf den verschiedenen Rechnerebenen (AA, LAA, zentral) sowie für die Netzbelastungen entwickelt. Niedrige Werte bei den Transaktionsraten gehen in den Ergebnissen einher mit höheren Netzbelastungen. Die fehlende gemeinsame Formel zur Entscheidungsfindung, die zwischen diesen beiden gegensätzlichen Aspekten gewichten würde, kann als Vorteil ausgelegt werden (s. S. 58). So kann kein Szenario über eine am Endergebnis nicht mehr erkennbare, intern in einer Gesamtformel versteckte Gewichtung der Teilaspekte besonders bevorzugt oder benachteiligt werden. Aus den Ergebnissen läßt sich zunächst nur ablesen, ob die durchschnittlich auftretenden Anforderungen technisch gegenwärtig realisierbar sind (Machbarkeitsstudie). Eine darüber hinausgehende Beurteilung bleibt den vergleichenden Betrachtungen vorbehalten.

In Hinblick auf die Ausgestaltung der Szenarien stehen im SNI-Modell bestimmte Allokationen gar nicht erst in Frage. Damit weicht die Modellierung von den theoretischen Ansätzen ab, in denen alle denkbaren Allokationen (bei einer vorgegebenen Fragmentierung) durchgegangen werden und eine optimale gefunden werden soll. Auch in Hinblick auf die Zugriffe der Rechner untereinander gibt es je nach Szenario des SNI-Modells starke Einschränkungen. Keineswegs herrscht von jedem Rechner Zugriffsmöglichkeit auf jeden anderen Rechner im Netz. Dieser gravierend andere Ausgangspunkt sorgt allerdings für eine überschaubare Modellierung mit geringem Berechnungsaufwand bei gleichzeitiger Vermeidung isolierter Allo-

kationsentscheidungen für einzelne Fragmente. Ins Auge gefaßt werden hierbei Szenarien, wie sie zum Entscheidungszeitpunkt der Studie unter Ausnutzung der vorhandenen Rechenzentren sinnvoll erscheinen könnten.

Wenn man einmal von dem zentralen Szenario absieht, so werden in allen Szenarien die den Ämtern zugeordneten Bewerber und Stellendaten lokal vor Ort allokiert. Diese Allokationen stehen überhaupt nicht in Frage. Letztendlich wird nur ermittelt, welche Auswirkung eine geschickte Replikatbildung - und hier vornehmlich der Stellenangebote - auf die Auswahlmenge sowie Transaktionsrate und Netzlast hat. Antwortzeiten bleiben unberücksichtigt, somit auch Aspekte der Fragmentierung und Replikatbildung zur Erhöhung des Datendurchsatzes. Dabei werden bezüglich der Stellenangebote ausschließlich horizontale Fragmentierungen nach Berufsgruppen in Betracht gezogen. Das Verteilen von Replikaten der vermittlungsrelevanten Daten der Bewerber in Szenario 7 stellt demgegenüber eine vertikale Fragmentierung dar. Abhängigkeiten zwischen Zugriffen und Fragmenten werden lediglich über Entfernungen zwischen den jeweils zuständigen Arbeitsämtern des Stellen- oder Bewerberangebotes, von dem die Suche ausgeht, und den Lösungen (lokal, Pendelbereich, zentral) modelliert. Für detaillierte Allokationsüberlegungen wären echte Zerschlagungen des Datenbestandes auch nach anderen Merkmalen, naheliegender Weise zumindest nach Berufen, erforderlich. Auf diese Weise hätte dann auch die 5%-Ausgliederung von speziellen Berufen in den Szenarien 4 bis 7 exakt aus dem SNI-Modell hervorgehen müssen. Man bedenke aber, daß hierfür bei 181 AA-Rechnern, 10 LAÄ-Rechnern und einem Zentralrechner bei nur 3 Berufsgruppen bereits mit einer Anzahl von 201 Rechnerknoten und 543 Fragmenten zu rechnen gewesen wäre (pro AA-Rechner 3 Fragmente). Bei diesen Werten erzielt man bereits für die Basis der Formel zur Ermittlung der Zahl der Allokationsmöglichkeiten bei zugelassener Redundanz eine 60-stellige Zahl! Der Ansatz über einige wenige denkbare Szenarien vorzugehen, müßte also auch bei einer feineren Fragmentierung beibehalten werden.

Recht genau werden im SNI-Modell die betrachteten Anforderungen in Transaktionen umgesetzt: Schreibende und lesende Zugriffe werden gesondert modelliert und auch das Aktualisieren der Replikate ist berücksichtigt. Wie auch in den theoretischen Modellen werden hingegen Anforderungen, die aus dem Transaktionsmanagement, aus automatischen Integritätsprüfungen oder der Gewährleistung eines korrekten Mehrbenutzerzugriffs entstehen, nicht berücksichtigt. Damit unterliegen sie denselben Kritikpunkten, die Özsu und Valduries für Modelle der von Dadam sowie Ceri und Pelagatti vorgestellten Form geäußert haben (s. S. 46). Eine weitere Gemeinsamkeit mit den theoretischen Modellen besteht darin, daß der Aufwand, der durch Netzübertragungen entsteht, nicht eigens modelliert wurde. Aspekte wie Protokoll- und Adreßinformationen könnten allenfalls in dem Durchschnittswert der zu übertragenden Datenmenge eingeflossen sein.

Formelanalyse

Weitere Aspekte lassen sich aufzeigen, wenn man, wie bei der Beschreibung der Studie schon erfolgt, die gewählten Formeln einer genauen Analyse unterwirft. Dann werden zusätzliche teils grundlegende Annahmen des SNI-Modells, teils für die Modellbildung vorgenommene Vereinfachungen, deutlich.

Besonders offensichtlich ist, daß nahezu alle Variablen mit Mittelwerten belegt werden. Damit können Unterschiede durch spezielle regionale Gegebenheiten, die durchaus existieren (vgl. [DR 97], S. 95) keine Berücksichtigung finden. Auch die fehlende, oben bereits angesprochene Einteilung in Berufsgruppen wird, mit Ausnahme auf die berücksichtigten Prozentsätze für zentral gehaltene Stellen, deutlich (vgl. S. 82, 83). So ist kaum vorstellbar, daß in

allen Berufen die Obergrenze **O** anzuzeigender Lösungen erzielt werden kann. Ebenso dürften Abhängigkeiten zwischen der Bereitschaft zu regionaler Mobilität und Berufen existieren. Diese werden insbesondere in Szenario 3 (vgl. Hinweis, S. 96), aber auch allgemein durch das enge Konzept der Pendelbereiche (s. u.) nicht berücksichtigt. Über die Möglichkeiten des gegenwärtig genutzten Verfahrens coArb zu statistischen Auswertungen könnte man seitens der BA exaktere Modellierungsgrundlagen herstellen. Mit den berücksichtigten Mittelwerten entsteht die bereits zu den theoretischen Modellen aufgekommene Frage nach dem Verhältnis der Genauigkeit der Modellierung der Transaktionen und der Eingabedaten.

Ebenso werden, z. B. an Stellen, an denen von mir Alternativformeln angegeben wurden, gewisse Ungenauigkeiten in Kauf genommen (vgl. S. 80, 82, 84, 88, 97). Einige dieser Ungenauigkeiten basieren darauf, daß eine nur nahezu einheitliche Fragmentbildung über die verschiedenen Szenarien gelegt wird, dennoch aber mit gleichen Variablen auf die diversen Stellenangebotsbestände Bezug genommen wird.

Weitere bewußt in Kauf genommene Vereinfachungen liegen zum einen in der Modellierung der Folgetätigkeiten bei den stellenangebotsorientierten Suchen. Diese gehen nur pauschalisiert über den Faktor **fa** ein. Zum anderen wird in den Formeln, wie bei Modellbildungen üblich, stets von einer vollen Verfügbarkeit des Netzes und der daran angeschlossenen Rechner ausgegangen (vgl. S. 83). Immerhin wird der Aspekt der Verfügbarkeit in der abschließenden Gesamtbewertung der verschiedenen Szenarien zumindest grob berücksichtigt.

Außerdem entnimmt man den Formeln (s. S. 81), daß von einer ausgeglichenen Vermittlungssituation ausgegangen wird, da die zu löschende Anzahl an Stellen mit der Variablen für die bundesweit täglich neu zu erfassenden Stellen gleich gesetzt wird. Somit werden Schwankungen im Stellenangebot nicht modelliert. Bezüglich der Löschvorgänge gilt ferner, daß hier sowohl auf Stellen- als auch auf Bewerberseite nur eine Kennzeichnung als bereits vermittelt modelliert wurde, nicht jedoch das tatsächliche Löschen. Eine erste Kennzeichnung ist auch in dem arbeitenden System üblich. Die wirklichen Löschzugriffe kämen also in Wirklichkeit noch hinzu und sind dann im SNI-Modell nicht berücksichtigt (vgl. Fußnote 51, Seite 104).

Zudem geht das SNI-Modell von einer deutlichen Dominanz der bewerberorientierten Stellensuche aus. Die Zuständigkeit für die Bewerber und deren Betreuung steht gegenüber einer optimalen Betreuung der Arbeitgeberanliegen im Vordergrund. Freie Suchen, also weder stellen- noch bewerberorientierte Suchen, die ebenfalls in den Ämtern regelmäßig erfolgen, beispielsweise um Kenntnisse über die Arbeitsmarktlage der Region zu erhalten sowie Abfragen an die Datenbank zu Auskunftszwecken wurden in die Modellierung nicht mit einbezogen.

Konzeptueller Ansatz

Zuletzt sind bereits stärker konzeptuelle Annahmen in den Vordergrund gerückt. In diese Richtung soll nun weitergedacht werden. Während bisher ausgehend von der Studie nach Grenzen gesucht wurde, wird nun auf einer höheren Abstraktionsstufe überlegt, ob die zugrunde gelegten Annahmen des SNI-Modells selbst den (heutigen) Anforderungen der Arbeitsvermittlung genügen.

Konzeptuelle Modelle sollen sich möglichst durch Langlebigkeit auszeichnen. Dies ist nur möglich, wenn sie die Anforderungen zunächst unabhängig von realen Begrenzungen oder Möglichkeiten modellieren.

Demgegenüber haben in der Studie nicht nur die bereits existierenden Rechenzentren, sondern auch bestehende Verbindungsstrukturen zwischen ihnen Berücksichtigung gefunden. Auch die

Vorgänge selbst wurden stark dem bisherigen Verfahren entsprechend betrachtet. Dies spiegelt sich insbesondere am Pendelbereichskonzept und der Suchstrategie wider. Ich hatte daher den Eindruck, daß sich die Studie allenfalls für eine Aussage der Form, ob und wie mit den vorhandenen Rechnern und Verfahren die Vermittlungsaufgabe in der BA günstiger als bisher gestaltet werden kann, eignet. Eine solche Aussage mag im Rahmen einer Umorganisation des Datenbestandes und einer Anpassung vorhandener Software sinnvoll und wichtig sein, im Rahmen einer Neuentwicklung liefert sie meines Erachtens keine hinreichende Antwort. Allenfalls das gesamte Transaktionsaufkommen aus Szenario 1 kann hierfür interessant sein. Im folgenden sollen daher Punkte aufgezeigt werden, die für die Neuentwicklung im konzeptionellen Ansatz nicht beachtet wurden.

Über das Internet besitzt heute - zum Entstehungszeitpunkt der Studie noch nicht - bereits jeder Nutzer unabhängig vom Ort seines Zugangsrechners eine Sicht auf den bundesweit freigegebenen Stellen und Bewerberdatenbestand. Zudem kann er regionale Wünsche eingeben, die ebenso unabhängig von seinem gegenwärtigen Standort oder Wohnsitz sind. Schon allein dieser Aspekt sollte Grund genug sein, für die neue Softwaregeneration ausschließlich Lösungen in Betracht zu ziehen, bei denen bundesweit auf einen einheitlichen Datenbestand zugegriffen wird und bei denen weitgehend **beliebige räumliche** Teilmengenbeziehungen festlegbar sind. Dabei soll das "weitgehend" lediglich in Hinblick auf die Feinheit der Zerschlagung der deutschen Landkarte verstanden werden. So sollte ein Arbeitsvermittler aus Hamburg für einen dort lebenden Arbeitsuchenden problemlos auch Stellen weiter entfernter Regionen abfragen können. Dieser Aspekt ist in einer Gesellschaft, in der zunehmende Mobilität zu verzeichnen ist, von steigender Bedeutung, so daß die Modellierung über Pendelbereiche innerhalb der Studie aus heutiger Sicht nicht mehr genügt. Dabei wäre auch eine Einstufung der Wunschgebiete nach Prioritäten (bevorzugt, möglich) denkbar.

Das Beispiel führt zu einem anderen Punkt der Modellbildung, der mir vom konzeptuellen Ansatz her überdenkenswert erscheint. So müßte man meines Erachtens die starre Bindung von Stellen und Bewerbern zu einem Amt, die auch in der gegenwärtigen Lösung vorhanden ist, in Frage stellen. Zumindest wäre offen zu legen, welche Aufgaben zusätzlich zur Vermittlung mit dieser Modellierung eigentlich erzielt werden sollen. Hier sind eine gleichmäßige Arbeitsverteilung, Grundlagen für Statistiken usw. gegenwärtig nur zu vermuten. Auch das Verfahren, zunehmend entfernt zu suchen, verwundert. Vom theoretischen Standpunkt irritiert, daß gleiche Anfragen von unterschiedlichen Rechnern zu unterschiedlichen Lösungen führen - letztendlich liegt keine transparente Datenhaltung vor -, vom praktischen Standpunkt her wird die Beurteilung der Ergebnismenge problematisch. Eine Einschränkung der Ergebnismenge muß meines Erachtens über die im Bewerber- oder Stellenangebot festlegbaren Kriterien, darunter durchaus auch regionale Wünsche erzielt werden. Nur dann ist auch sichergestellt, daß man tatsächlich nach den passendsten Angeboten im Datenbestand und nicht nach den erst besten sucht.

Ein alternativer Ansatz würde die mögliche Zuordnung von Suchergebnissen zu allen Vermittlern der BA unterstützen. Für die Zuordnung könnten Auslastung (Fülle des Eingangskorbes) **und** Lokalität als Kriterium herangezogen werden. Nicht zu gerade stattfindenden oder geplanten Besuchen erfolgende Recherchen könnten von einem beliebigen Vermittler bearbeitet werden. Damit wäre auch ein regionaler Ausgleich des Arbeitsanfalls zu realisieren. Zudem käme jeder Mitarbeiter automatisch als möglicher Ansprechpartner für jeden Kunden in Frage. Ein solcher Lösungsansatz setzt allerdings ein Arbeitsamtsgrenzen überschreitendes Denken der Mitarbeiter voraus. Eine für einen solchen Ansatz passende Statistik zur Beurtei-

lung des Arbeitsvolumens müßte vorgangsorientiert sein. Andere Statistiken zur Auswertung des Datenbestandes blieben unberührt.

Ein weiterer, in der Studie nicht im Vordergrund stehender Modellierungsaspekt, ist die heute noch übliche vorrangige Suche über Berufe. Nachrangig kommen weitere Kriterien hinzu. für jeden Beruf gibt es eine Berufskennziffer. Durch Weglassen letzter Angaben aus der Ziffer vergrößert man eine Suche auch in "ähnliche" Berufe. Eine automatische Suche über diese Kennziffern ist sehr gut unterstützbar. Nachteilig ist, daß sich die beruflichen Einteilungen nicht immer mit den Bewerbungsprofilen decken. Deshalb kann man heute Bewerber unter mehreren Berufskennziffern führen. Konzeptuell ist jedoch etwas ganz anderes gewollt. Man möchte für eine Person mit bestimmten Qualifikationen eine möglichst adäquate Stelle und umgekehrt. Deshalb werden für die neue Softwarelösung zusätzlich auch Profilsuchen angestrebt, bei denen vorrangig über berufliche Fähigkeiten gesucht werden kann. Wie man sich leicht vorstellen kann, dürften Profilsuchen wesentlich schwieriger zu unterstützen sein, da sich eine maximale Übereinstimmung in Qualifikationsmerkmalen nur über die Suche mehrere gleichberechtigte Felder gestalten läßt und sich bei einer solchen Suche die gefundene Datenmenge nicht so schnell reduziert wie bei einer Suche über Berufskennziffern. Nicht nur eine geeignete Verschlüsselung der zahlreichen beruflichen Merkmale, auch eine geeignet gestaltete Suchstrategie und Indizierung wären hier anzudenken. Möglicherweise könnten sich aus dieser Anforderung heraus doch noch einmal Allokationsüberlegungen zur Erhöhung des Durchsatzes ergeben.

Zur Berücksichtigung solcher (und weiterer) Aspekte wäre jedoch eine neue (semantische) Modellierung der Datenbank bzw. der Informationssysteme der BA notwendig gewesen. Dies konnte im Entstehungszusammenhang der SNI-Studie nicht geleistet werden. Das Festhalten an den vorhandenen Rechenzentrumsstrukturen im SNI-Modell erfolgte auf Basis der SNI zur Verfügung stehenden statistischen Daten. 90% aller Anfragen des heutigen Verfahrens betreffen lokale Angebote. Zudem könnte die Zukunft Wünsche an eine digitalisierte Aktenhaltung mit sich bringen. Dies aber würde in Hinblick auf die hiermit verbundenen Speichermengen wieder für eine lokale Datenhaltung, zumindest der Akten sprechen (Gesprächsinformation [Kr99]).

5 Zusammenfassung und Ausblick

Im Rahmen der Modernisierung der Arbeitsverwaltung wurden in den letzten Jahren sowohl im organisatorischen als auch im informationstechnischen Bereich umfangreiche neue Lösungen angedacht, entwickelt und zum Teil auch bereits umgesetzt.

In diesem Kontext ist auch die in dieser Arbeit im Mittelpunkt stehende Allokationsstudie der Firma Siemens Nixdorf entstanden. Sie wurde 1996 zu Beginn der Planungsphase einer neuen IT-Generation zur Unterstützung der Aufgabenerfüllung in den Ämtern der BA angefertigt. Dabei berücksichtigt die Studie die wesentlichen Vorgänge der Arbeitsvermittlung, die eine der zentralen Aufgaben der Mitarbeiter in den AÄ darstellt. Dieser Aufgabe wurde bei der Neugestaltung von Software neben der Aufgabe der Leistungsgewährung besonders hohe Priorität eingeräumt. Für die Arbeitsvermittlung ist der Zugriff auf einen umfangreichen Datenbestand, insbesondere Bewerber- Stellen- und Betriebsdaten erforderlich. Die bisherigen computerunterstützten Verfahren zur Arbeitsvermittlung arbeiten auf eigenständigen, homogenen zentralen Datenbanken in allen Ämtern. Ein Informationsaustausch ist oberhalb der Datenbankebene, also konventionell, geregelt. Diese Situation soll abgelöst werden. Eine Überlegung bestand darin, eine gemeinsame verteilte Datenbank für die Vermittlung aufzubauen. Die von Siemens Nixdorf angefertigte Studie zielte darauf ab, exemplarisch zu zeigen, wie mit Hilfe einer mathematischen Modellbildung für eine gewählte Rechnersituation die in der Vermittlung anfallenden Datenbankanwendungen theoretisch so abgebildet werden können, daß hierdurch planerische Hilfestellung für die Verteilung des Datenbestandes entsteht.

Eine Auseinandersetzung mit dieser Studie bedeutete daher, wesentliche Aspekte der Datenbanktheorie, insbesondere verteilter Datenbanken und hier wiederum Aspekte der Fragmentierung und Allokation von Daten und der damit zusammenhängenden Datenbankmanagementsystem-Aufgaben zu berücksichtigen. Daß dabei vor allem mathematische Modelle zur Datenallokation eine genauere Beachtung verdienen, versteht sich von selbst.

Verteilte Datenbanken werden in der Literatur unterschiedlich definiert. Für diese Arbeit wurde eine Definition der Autoren Ceri und Pelagatti gewählt, die dann von einer verteilten Datenbank ausgeht, wenn eine Menge von Daten über verschiedene Rechnerknoten eines Netzes verteilt sind und an jedem dieser Knoten autonom gearbeitet werden kann. Darüber hinaus ist aber auch jeder der Knoten an der Ausführung mindestens einer globalen Anforderung beteiligt, die den Zugriff an einigen Knoten des Netzes erfordert.

In jedem Unternehmen, das sich über größere regionale Entfernungen erstreckt, stellt sich beim Aufbau einer gemeinsamen Datenbank die Frage, wo die Daten, auf die an jedem Unternehmensstandort Zugriff bestehen soll, allokiert werden sollen. Dabei existieren vielfältige Möglichkeiten, beginnend mit einer zentralen Lösung über Multiprozessorysteme bis hin zu verteilten Datenbanken.

Gründe, die für eine verteilte Lösung sprechen, liegen in einer Erhöhung der Ausfallsicherheit, in der Verbesserung des Datendurchsatzes und in anpassungsfähigeren Strukturen. Auch die Unterstützung dezentraler Organisationsstrukturen und eine allgemeine Kostenreduzierung können eine Rolle spielen. Dabei ist allerdings zu berücksichtigen, daß sich die einzelnen Vorteile zum Teil gegenseitig widersprechen. So kann eine zur Erhöhung der Ausfallsicherheit getroffene Entscheidung zu mehr Replikhaltung zwar einhergehen mit einem beschleunigten lesenden Zugriff, da die schreibenden Zugriffe hierdurch aber aufwendiger werden,

kann der Datendurchsatz bei diesen Vorgängen sinken. Argumente, die gegen eine Verteilung vorgebracht werden, sind die kompliziertere Verwaltung der Daten und die damit möglicherweise einher gehende höhere Fehleranfälligkeit.

Im Rahmen der Neuentwicklung einer Datenbank für die BA war es daher folgerichtig, daß verteilte Lösungen auf ihre Realisierbarkeit hin durchdacht werden sollten. Verteilte Lösungen sind heute noch weitgehend an das relationale Datenbankmodell gebunden. Hier ist es vergleichsweise einfach, den Datenbestand zu fragmentieren. Diese Voraussetzung traf für die Datenmodellierung des Datenbestandes der BA zu. Will man dann auf methodisch exakte Weise prüfen, ob eine Verteilung des Datenbestandes gegenüber anderen Lösungen vorteilhaft ist, so kann man hierzu mathematische Modelle heranziehen.

Solche Modelle zur Allokationsplanung in verteilten Datenbanken gehen von einer vernetzten Rechnersituation und zumeist auch von einer gegebenen Fragmentierung des Datenbestandes aus. Darüber hinaus modellieren sie die auf diesen Fragmenten stattfindenden Anforderungen, je nach Variante unterschiedlich exakt. So können entfernte Zugriffe und die transportierten Datenmengen betrachtet werden oder aber die Transaktionszahlen auf den diversen Rechnern. Auch der Speicherbedarf an den verschiedenen Rechnern des Netzes und Verarbeitungskosten werden in einigen Modellen berücksichtigt. Dazu werden mathematische Formeln aufgestellt, die das Aufkommen der jeweils betrachteten Aspekte bezüglich der auf dem Datenbestand stattfindenden Anforderungen beschreiben. Eine Summation über das so modellierte Aufkommen und die Suche nach einem Minimum für diese Summe ist das Grundprinzip dieser Modelle. Werden mehrere Aspekte betrachtet, wird häufig mit Nebenbedingungen gearbeitet, da das Aufsummieren unterschiedlicher zu minimierender Größen problematisch ist, weil es zwischen diesen zu einer Gewichtung führen würde. So können z. B. für die zur Verfügung stehenden Speicherkapazitäten an den Rechnerknoten obere Schranken als Nebenbedingung formuliert werden und darüber hinaus sind auch Angaben zu maximalen Transaktionsraten denkbar, während die eigentliche Zielfunktion das Minimum über die Anzahl entfernter Zugriffe, ggf. unter Berücksichtigung der übertragenen Datenmenge, bildet.

Wie bei Modellbildungen üblich, wird die reale Situation aber vereinfacht wiedergegeben. So werden normalerweise Aspekte wie die Anfragebearbeitung, die Anfrageoptimierung und der Aufwand, der durch Synchronisationsverfahren und Kopienbehandlung sowie Netzmanagement entsteht, nicht modelliert. Trotz dieser Vereinfachungen ist die Menge aller denkbaren Allokationen schon bei einer relativ geringen Zahl an Rechnern und Fragmenten so immens, daß ein Durchlaufen aller Lösungen zur Feststellung der minimalen Lösung nicht effektiv durchzuführen wäre. Zugleich weiß man aber, daß das Finden eines geschickteren Lösungsalgorithmus außerordentlich schwierig, wenn nicht unmöglich sein dürfte, da die Fragestellung zu einer Klasse von Problemen gehört, von denen diese Eigenschaft bekannt ist. Deshalb kommen für die in der Praxis auftretenden Anzahlen von Rechnern und Fragmenten eigentlich nur heuristische Ansätze in Frage, die menschlichen Sachverstand nutzen, um wenig sinnvolle Konstellationen von vornherein auszuschließen und zu einer begrenzteren Zahl an Prüfungen führen. Dabei ist ein solcher Ansatz um so besser, je höher zu erwarten ist, daß damit trotzdem die beste oder zumindest nahezu beste Lösung gefunden wird.

Sowohl dem Problem, alle Fälle durchlaufen zu müssen, als auch dem Aspekt, die Anfragebearbeitung nicht zu beachten, ist Apers in seinem von den oben beschriebenen mathematischen Modellen weit abweichenden Modell begegnet. Einer seiner Lösungsalgorithmen kann so angewandt werden, daß er bei jedem Schritt den wirklich zum Einsatz kommenden Anfragealgorithmus nutzt. Zudem erfordert sein Ansatz nicht das Ausprobieren aller Allokationsfälle.

Dennoch bleibt auch bei dieser Lösung der Berechnungsaufwand immens, zumal in seinem Verfahren nicht von der Zahl der Rechner ausgegangen wird, sondern von einer weit höheren Anzahl, die sich aus der Zahl der Rechner und der Anzahl der Anforderungen an die Fragmente zusammensetzt.

Ein Problem bleibt hingegen für alle Modelle bestehen. Sie können nur so gut sein, wie die zugrunde gelegten vermuteten oder statistisch erhobenen Daten. Jeder Detailaufwand in der Modellierung muß hiergegen abgewägt werden. Bei Modellierungen, die nicht aus wissenschaftlichem Interesse entstehen, kommen Aspekte der Wirtschaftlichkeit in der Anfertigung solcher Betrachtungen hinzu.

In diesem Zusammenhang darf es nicht verwundern, daß in der Studie von SNI lediglich eine feste Zahl denkbarer, möglicher Lösungen untereinander verglichen wurde. Dies sichert zum einen ein effektives Ermitteln einer Lösung, zum anderen unterstreicht es den exemplarischen Ansatz. Ansonsten weist das mathematische Modell der Studie von Siemens Nixdorf eine Reihe der Charakteristika der ersten Gruppe der beschriebenen mathematischen Modelle auf und zeichnet sich durch eine sehr genaue Modellierung der Transaktionsanforderungen aus. Darüber hinaus werden entfernte Zugriffe, Übertragungsmengen und -Entfernungen sowie tageszeitliche Belastungsschwankungen berücksichtigt. In dieser genauen, erzwungenermaßen formalen Modellierung steckt ein positiver Nebeneffekt. Die genaue Abbildung der diversen Vermittlungsvorgänge dürfte einer exakten Anforderungsanalyse entgegenkommen.

Unter den BA-spezifisch modellierten sieben Szenarien werden zwei aufgrund von Modellberechnungen zu den erstellten Formeln favorisiert, da bei ihnen die an den beteiligten Rechnern auftretenden Transaktionsraten sowie die Netzübertragungen die günstigsten Werte annehmen. Die ausgewählten Lösungen zeichnen sich durch gegenseitige Datenzugriffe relativ benachbarter Rechner sowie durch Datenzugriffe zu einem zentralen Rechner bei der Suche nach speziellen Berufen aus. Der Unterschied zwischen den Lösungen besteht darin, daß das eine Szenario Replikate vertikaler Fragmente zu Bewerberangeboten in deren Pendelbereich verteilt, während die andere Lösung für ein Zugreifen auf diese Daten die Rechner mit den Originaldaten benötigt. Replikate von Stellenangeboten des Stellenpendelbereiches weisen beide Lösungen auf.

Die Einordnung dieser Lösungen im Rahmen der in dieser Arbeit eingeführten Begrifflichkeiten stellt sich als schwierig heraus. Zwar sind die Daten über die Rechner des Netzes verteilt, aber der zentrale Rechner wird nicht für ein autonomes Arbeiten genutzt, wenngleich man sich eine solche Nutzung vorstellen könnte. Problematischer zu beurteilen ist, daß nur der zentrale Datenbestand für alle Rechner sichtbar ist und Suchen von der Lokalität der Daten abhängig sind. Bei der Modellierung stand der Aspekt im Vordergrund, das Zugriffsverhalten insgesamt so lokal wie nur möglich zu gestalten. Für eine Entscheidung, wie die Daten bei bundesweit einheitlicher Sicht zu verteilen wären, wäre daher eine weitere Modellierung erforderlich gewesen.

Laut [En99] zeichnet sich nun als Lösung ein zentrales Rechenzentrum mit circa zehn bis zwölf über sehr schnelle Kommunikationswege miteinander verbundenen Rechnern ab. Ein gemeinsamer Externspeicherzugriff ist genauso wenig vorgesehen wie ein gemeinsamer Hauptspeicherzugriff. Es dürfte sich daher in der Terminologie von Rahm um eine partitioniertes, lokales Mehrrechnerdatenbanksystem vom Typ Shared-Nothing mit loser Kopplung handeln (vgl. [RA94], S. 29). Aus Ausfallsicherheitsgründen wird ein zweites, räumlich entferntes kreuz-gespiegeltes Back-Up-Rechenzentrum eingerichtet. Mit dieser Entscheidung

stehen erneut Allokationsüberlegungen an. So wird zu entscheiden sein, wie der bundesweite Datenbestand auf die 10 Rechner zu verteilen ist. Zeitlimits setzen vor allem Vorgänge wie das Laden und Entladen der Datenbanken. Auf den angebundenen Client-Server-Netzen in den Ämtern werden lediglich Applikationsserver eingerichtet. Damit hat sich eine räumlich nicht verteilte Mehrrechnerdatenbank gegenüber einer möglichen geographisch verteilten Datenbank durchgesetzt. Verteilte Datenhaltung ist dennoch erforderlich, wenn auch keine verteilte Datenbank vorliegt, da überhaupt keine lokale Autonomie existiert. In jedem Falle wird aber nach außen die Sicht einer gemeinsamen Datenbank erzeugt und diese Datenbank steht den Applikationsservern als Datenbankserver zur Verfügung.

Das Thema Allokationsplanung sowie verteilte Datenbanken wird in jedem Fall weiter spannend bleiben. Neben den in 4.5 angeschnittenen Themen der digitalen Aktenhaltung und der erhöhten Durchsatzanforderungen bei Profilsuchen sind auch Entwicklungen in Hinblick auf das Zusammenwachsen europäischer oder weltweit existierender Arbeitsverwaltungen zu vermuten. Bisher existieren nur rudimentäre Möglichkeiten gegenseitiger Stellenangebotsbekanntmachungen. Hier wäre es sicher interessant zu untersuchen, inwieweit man durch geeignete Datenbank- und Oberflächenstrukturen eine möglichst reibungslose Zusammenarbeit unterstützen könnte.

Gegenwärtig dürfte jedoch innerhalb der BA das Anliegen Priorität besitzen, die eigene neue Lösung zu implementieren und erfolgreich in die Fläche zu bringen. Damit das damit verbundene Potential dann aber auch voll genutzt wird, sind nach Brinckmann (vgl. S. 5 ff.) die herrschenden Organisationsstrukturen nicht unbedeutend, auf die deshalb im Rahmen dieser Arbeit ebenfalls kurz eingegangen wurde. Es bleibt daher zu hoffen, daß das Projekt „IT2000“ in Zusammenhang mit den Umstrukturierungsmaßnahmen im Rahmen von „AA2000“ sich wirkungsvoll entfalten wird.

A Verzeichnisse

Abkürzungsverzeichnis

Abkürzungen, die innerhalb der Bundesanstalt für Arbeit gebräuchlich sind, werden gesondert aufgeführt. Zu Abkürzungen zu denen im Text Erläuterungen zu finden sind, wird zusätzlich zur Langform auf die Seite verwiesen, auf der die Abkürzung erstmalig verwandt wird.

Im Bereich der Bundesanstalt für Arbeit übliche Abkürzungen

AA

Arbeitsamt 60

AÄ

Arbeitsämter 60

ACID

Atomarität, Konsistenz, Isolation Dauerhaftigkeit 27

AIS

Software, Arbeitgeberinformationssystem 70

BA

Bundesanstalt für Arbeit III

BewA

Bewerberangebot 67

coArb

Software zur computerunterstützten Arbeitsvermittlung 67

DBMS

Datenbankmanagementsystem 2

DBS

Datenbanksystem 2

DDBMS

distributed database system, engl. Abk. für vDBMS 7

EDV

Elektronische Datenverarbeitung III

GB

Gigabyte 109

GKS

Globales konzeptionelles Schema 21

IBAN

Integriertes BA-Netz 13

IEEE

Institute of Electrical and Electronics Engineers 13

IMP

Intermediate Message Processor, Kommunikationsrechner 15

IT

Informationstechnik 5

IT2000

Informationstechnik 2000, Projektbezeichnung III

IV

Informationsverarbeitung 64

LAA

Landesarbeitsamt 60

LAÄ
Landesarbeitsämter 60

LAN
Lokal Area Network, lokales Netz 13

LDB
Lokale Datenbank 7

LDBMS
Lokales Datenbankmanagementsystem 7

LKS
Lokales konzeptuelles Schema 21

NT
NewTechnologie *Siehe* Windows-NT

OODBS
Objektorientierte Datenbanksysteme 3

PC
Personal Computer 3

PlaIT2000
Planung der Informationstechnik 2000, Projektbezeichnung III

POP
Point of Present, Kommunikationsrechner 16

PSG
Processing-schedules-graph 54

SIS
Software, Stelleninformationssystem 69

SNI
Siemens Nixdorf Informationssysteme Aktiengesellschaft III

SQL
verbreitete Datenbankabfragesprache, Structured Query Language 20

SteA
Stellenangebot 67

SZ
Szenario 74

TB
Tagesbereich 108

vDB
verteilte Datenbank 7

vDBMS
verteiltes Datenbankmanagementsystem 7

vDBS
verteiltes Datenbanksystem 7

VDBS *Siehe* vDBS

WAN
Wide Area Network, geographisch verteiltes Netz 13

Windows-NT
Netzwerkfähiges Betriebssystem, kurz
NT 65

zDBMS
besonderer Hinweis auf zentrales DBMS 10

Abbildungsverzeichnis

Abbildung 1:	Das integrierte BA-Netz (BA-Dokument)	14
Abbildung 2:	Aufbau eines Rechnernetzes (entnommen aus [WE95], S. 349)	15
Abbildung 3:	Fragmentierung und Allokation (nach [CP84], S. 41)	22
Abbildung 4:	Fragmentierungsarten (vgl. Rahm S. 63, 64)	24
Abbildung 5:	Summenschreibweisen	34
Abbildung 6:	Fälle zu $B_{nr \rightarrow sv}$ bei vertikaler Fragmentierung nach Ceri und Pelagatti	38
Abbildung 7:	Fälle zu $B_{nr \rightarrow sv=r}$ bei vertikaler Fragmentierung nach Ceri und Pelagatti	39
Abbildung 8:	Horizontale Anfangsfragmentierung bei Apers	53
Abbildung 9:	PSG (in Anlehnung an Figur 10, S. 287 aus [Ap88])	54
Abbildung 10:	Organisation der BA (nach [IT98], Teil A 1, S. 2)	60
Abbildung 11:	Räumliche Gliederung der LAÄ und AÄ (vgl. S. 13 in [BA97])	61
Abbildung 12:	Arbeitsamt 2000 - Einflußfaktoren, nach ([AA2000] / Anlage 2; dort Anhang II, Abb. 1)	63
Abbildung 13:	Beispiele für besonders häufig benötigte „Bilder“ des Alt-Verfahrens coArb	68
Abbildung 14:	Suchbilder in coArb	
Abbildung 15:	Arbeitsämter - Pendelbereiche	80
Abbildung 16:	Stellenorientierte Bewerbervermittlung, Beispielrechnungen	91
Abbildung 17:	Gesamtübersicht über die auftretenden Transaktionen	106
Abbildung 18:	Einteilung in Tagesbereiche	108

Tabellenverzeichnis

Tabelle 1:	Realisierungsarten verteilter Systeme	12
Tabelle 2:	Relationenalgebra - Operationen	21
Tabelle 3:	Allokationsmöglichkeiten und ihre Handhabung, (aus: [ÖV91], S. 104)	26
Tabelle 4:	Anzahl möglicher Allokationen im partitionierten Fall ($\#K^{\#P}$)	44
Tabelle 5:	Anzahl möglicher Allokationen, falls Replikate zugelassen sind ($(2^{\#K}-1)^{\#P}$)	44
Tabelle 6:	Grundlegende Objekte der SNI-Studie und ihre Beziehungen	72
Tabelle 7:	Szenarien-Übersicht	77
Tabelle 8:	Durchschnittlich täglich hervorgerufene Transaktionen aufgrund von Sachbearbeiterentscheidungen nach stellenorientierten Bewerbersuchen	86
Tabelle 9:	Formeln für die stellenorientierte Bewerbersuche	89
Tabelle 10:	Zuordnung der modellierten Vorgänge zu Tagesbereichen	109
Tabelle 11:	Transaktionsraten und Netzbelastung nach Tagesbereichen für SZ 2	110
Tabelle 12:	Transaktionszahlen SZ 2 (1)	110
Tabelle 13:	Transaktionszahlen SZ 2 (2)	112
Tabelle 14:	Transaktionszahlen SZ 2 (3)	112
Tabelle 15:	Transaktionszahlen SZ 2 (4)	114
Tabelle 16:	Bewertungsaspekte	115

Literaturverzeichnis

Bücher

- [CP84] Ceri, Stefano; Pelagatti, Giuseppe (1984): **Distributed Databases - Principles and Systems**. Mc Graw Hill, New York
- [Da96] Dadam, Peter (1996): **Verteilte Datenbanken und Client/Server-Systeme**. Springer Verlag, Berlin
- [DR97] Statistisches Bundesamt (Hrsg.) (1997): **Datenreport 7 - Zahlen und Fakten über die Bundesrepublik Deutschland 1995/96**. Verlag Bonn Aktuell, Landsberg
- [Fr95] Friedrich, Jürgen; Herrmann, Thomas, Peschek, Max; Rolf, Arno (Hrsg.) (1995): **Informatik und Gesellschaft**. Spektrum Akademischer Verlag, Heidelberg
- [HLP92] Hesse, Joachim Jens; Lüder, Klaus; Pitschas, Rainer (1992): **Verbesserungsmöglichkeiten bei der Aufgabenwahrnehmung einer wirtschaftlicheren Gestaltung der Fachaufgaben in den Dienststellen der Bundesanstalt für Arbeit (§§ 7,9 BHO)**. Gutachten im Auftrag des Präsidenten der BA, Speyer und Oxford 1992. *Quellenangabe aus [Ho97]*
- [Ho97] Hoffjan, Andreas (1997): **Entwicklung einer verhaltensorientierten Controlling-Konzeption für die Arbeitsvermittlung**. Deutscher Universitäts-Verlag, Wiesbaden, insbesondere Kapitel 2
- [HS95] Heuer, Andreas; Sake Gunter (1995): **Datenbanken Konzepte und Sprachen**. Internat. Thomson Publ., Bonn, Albany [u. a.], insbesondere S. 1 - 23 u. 441 - 453
- [Ki81] Kirsch, Werner (1981): **Betriebswirtschaftspolitik und geplanter Wandel betriebswirtschaftlicher Systeme, in: Unternehmenspolitik: Von der Zielforschung zum strategischen Management**, hrsg. Von Werner Kirsch, München 1981, S. 121 - 152. *Quellenangabe aus [Ho97]*
- [KM94] Karer, Albert; Müller Bernd (1994): **ClientServerTechnologie in der Unternehmenspraxis**. Springer Verlag, Berlin
- [KS91] Korth, Henry; Silberschatz Abraham (1991): **Database System Concepts**. Mc Graw Hill Company, New York
- [Ku92] Kudlich, Herrmann (1992): **Verteilte Datenbanken**. Siemens Nixdorf Informationssysteme AG, München
- [Lu97] Lusti, Markus (1997): **Dateien und Datenbanken**. Springer Verlag, Berlin
- [Oß90] Oßwald, Rüger; Försterling Frank u. a.(1990): **Verteilte Datenbanken**. Berlin
- [ÖV91] Özsü, M. Tamer, Valduriez, Patrick (1991): **Principles of Distributed Database Systems**. Prentice Hall, Englewood Cliffs
- [Pa85] Papke, Thomas (1985): **Datenallokation in verteilten Informationssystemen**. Minerva Publikationen, München
- [Ra94] Rahm, Erhard (1994): **Mehrrechner-Datenbanksysteme**. Addison Wesley, Bonn

- [Schw92] Schwinn, Hans (1992): **Relationale Datenbanksysteme**. Carl Hanser Verlag, Wien
- [Se91] Sedgewick, Robert (1991): **Algorithmen**. Addison Wesley, Bonn, 3. Auflage (1994)
- [We95] Weske, Mathias (1995): **Deadlocks in Computersystemen**. Intern. Thomson Publ., Bonn, Albany u. a.
- [Wei83] Weigert, Peter (1983): **Benutzerorientierte Datenallokation in verteilten Informationssystemen**. Minerva Publikationen, München
- [Wi92a] Williams, Richard (Hrsg.), (1992): **Data Distribution: Managing the Environment**. Ashgate Publishing, Aldershot (Engl.), Beitrag von: Holloway, Simon; Warren, Tim: **Distribution - Myth or Reality**, (S. 7- 24)

Tagungsbände

- [See97] Seeger, Otto (1997): **Arbeitsamt 2000. Modernisierung der Bundesverwaltung, Aktueller Stand und Perspektiven**, Schriftenreihe der Fachhochschule des Bundes für öffentliche Verwaltung Bd. 28, S. 99 - 113, Brühl
- [Sel97] Sell, Stefan (1997): **Organisationsentwicklung in der Arbeitsverwaltung**. Modernisierung der Bundesverwaltung, Aktueller Stand und Perspektiven, Schriftenreihe der Fachhochschule des Bundes für öffentliche Verwaltung Bd. 28, S. 169 - 180, Brühl
- [SNI96] Siemens Nixdorf Informationssysteme AG (1996): **Das mathematische Modell für die verteilte Verarbeitung im Bereich Vermittlung**. 64 Seiten, zuzüglich 40 Abbildungen, siehe Anlage 1

Artikel

- [Ap88] Apers, P.M. G. (1988): **Data Allocation in Distributed Database Systems** in: ACM Transactions on Database Systems, Vol. 13, No. 3, 1988, S. 263 - 304
- [Bo98] Bode, Ulrich (1998): **Ein neues Modell von Unternehmensorganisation** in: it Management März 98, S. 14 - 19
- [La96] Lang, Kurt (1996): **Verteilte Datenhaltung Wie Geschäftsobjekte an Freiheit gewinnen** in: Focus 2/96, Beilage zur Computerwoche vom 3. 5. 96, S. 12 - 14

Erlasse, Schriften der Bundesanstalt für Arbeit sowie Gesprächspartner

- [AA96] Munker, Funke (1996): **Kurzfassung Arbeitsamt 2000**. Hauptstelle der BA (Der Präsident, Ref. IVc3-1083), Nürnberg. Siehe **Anlage 2**
- [BA97] Bundesanstalt für Arbeit, Referat Öffentlichkeitsarbeit (Hrsg.), (1997): **Die BA stellt sich vor**, Nürnberg
- [Dialog] Bundesanstalt für Arbeit (Hrsg.), Jahrgänge 1 - 6 (1994 - 1999): Zeitschrift für Mitarbeiterinnen und Mitarbeiter der Bundesanstalt für Arbeit

- [En99] Telefonisches Gespräch mit Ingo Engelhardt, verantwortlich für die technische Architektur der Realisierung IT2000, Stufe 1, Nürnberg 12.01.1999
- [HR94] Hauptstelle, Ref. Vb4 (Hrsg.), (1994): **Handreichungen für die Aus- und Fortbildung** Kompendium der Fachaufgaben der Bundesanstalt für Arbeit - Organisation und Verwaltung (einschließlich Finanzwirtschaft), Nürnberg
- [IT95] Hauptstelle, Ref. VIb4 (Hrsg.), (1995): **IT-Rahmenkonzept 1995**. Nürnberg
- [IT96] Hauptstelle, Ref. IIIc4 (Hrsg.), (1996): **IT-Rahmenkonzept 1996**. Nürnberg
- [IT98] Hauptstelle, Ref. IIIc4 (Hrsg.), (1998): **IT-Rahmenkonzept 1998**. Nürnberg
- [Kr99] Telefonisches Gespräch mit Karl Kronewald, Leiter des Business Support Team for System Design bei Siemens Business Services, ehemals Siemens Nixdorf Informationssysteme (SNI), Nürnberg 14.01.1999
- [RdErl 24/94]) Hauptstelle (1994): Runderlaß 24, Nürnberg

B ANLAGEN

1 SNI-Studie

Die folgend wiedergegebene Studie besteht aus 2 Teilen.

Im ersten Teil befindet sich eine textuelle Beschreibung des von der Siemens-Nixdorf-Aktiengesellschaft angefertigten mathematischen Modells, im zweiten Teil Abbildungen, Tabellen und Modellrechnungen.

Beide Teile werden in den Originalen beginnend mit 1 durchgezählt. Dies ist bei den Verweisen auf die Studie zu beachten.

ENTWURF

Mathematisches Modell für die verteilte Verarbeitung im Bereich Vermittlung

Arbeitsamtskopie

Inhaltsverzeichnis

1 EINFÜHRUNG	4
2 ALLGEMEINE ANNAHMEN	5
3 ALLGEMEINE SUCHSTRATEGIE	6
3.1 Suchstrategie	6
3.2 Zusammenspiel der Arbeitsamtsrechner im Pendelbereich	7
4 STELLENORIENTIERTE BEWERBERVERMITTLUNG	8
4.1 Aufnahme einer neuen Stelle	8
4.2 Suche nach Bewerbern	8
4.3 Folgebearbeitungen	10
4.4 Löschen einer Stelle	10
5 BEWERBERORIENTIERTE STELLENVERMITTLUNG	11
6 VERFAHREN DER VERTEILTEN VERARBEITUNG	13
6.1 Szenario1 (zentrales Verfahren)	15
6.2 Szenario2 (zentrale Stellenverwaltung, dezentrale Bewerberverwaltung)	16
6.3 Szenario3 (teilweise zentrale Stellenverwaltung, dezentrale Bewerberverwaltung)	17
6.4 Szenario4 (Suche im Pendelbereich)	18
6.5 Szenario5 (Speicherung der Stellen im Pendelbereich)	19
6.6 Szenario6 (Speicherung der Stellen in den Landesarbeitsämtern)	20
6.7 Szenario7 (Speicherung der Stellen- und Bewerberdaten im Pendelbereich)	21
6.8 Das Speicherkonzept	22

Arbeitsamtskopie

7 DAS MATHEMATISCHE MODELL

7.1 Die Basisvariablen	23
7.2 Stellenorientierte Bewerbervermittlung	24
7.2.1 Neuaufnahme einer Stelle	27
7.2.2 Löschen einer Stelle	28
7.2.3 Suchen nach geeigneten Bewerbern	30
7.2.4 Schreiben in die Team-Datei	31
7.2.5 Entscheidung der Sachbearbeiter	33
7.2.6 Diskussion für die einzelnen Szenarien	34
7.2.7 Folgebearbeitung	35
7.2.8 Gesamtanzahl der Transaktionen	37
7.2.9 Beispielrechnung	38
7.3 Bewerberorientierte Stellenvermittlung	39
7.3.1 Erstkontakt	40
7.3.2 Folgeberatungen	40
7.3.3 Löschen von Bewerberdaten	46
7.4 Zusammenfassung	50
7.5 Bewertung der Ergebnisse der einzelnen Szenarien	51
7.5.1 Aufteilung der Transaktionen auf Tageszeiten	52
7.5.2 Netzbelastung	52
7.5.3 Bewertung der einzelnen Szenarien	54
7.5.4 Empfehlung	55
8 GLOSSAR	58

1 Einführung

Im folgenden wird ein mathematisches Modell entwickelt, das verschiedene Szenarien der verteilten Verarbeitung im Bereich der Arbeitsvermittlung beleuchten soll.

Im Rahmen von der Vermittlung sind in den Arbeitsämtern im wesentlichen folgende Aufgabenschwerpunkte durchzuführen:

- Stellenorientierte Bewerbervermittlung

Bei diesem Geschäftsprozess werden für Stellen, die von Firmen an das Arbeitsamt gemeldet werden, geeignete Bewerber ermittelt.

- Bewerberorientierte Stellenvermittlung

Bei diesem Geschäftsprozess werden für Arbeitssuchende in Frage kommende Stellen gesucht.

Bei beiden Geschäftsprozessen wird davon ausgegangen, daß die zur Durchführung der Aufgabe notwendigen Daten in Rechner gespeichert sind und daß zur Ermittlung von geeigneten Bewerbern oder Stellen Datenbank-basierende EDV - Verfahren eingesetzt werden.

Im folgenden werden

- zunächst die beiden Geschäftsprozesse allgemein beschrieben
- danach die verschiedenen Verfahren der verteilten Verarbeitung erläutert
- ein mathematisches Modell für die Geschäftsprozesse entwickelt, wobei vor allem die zu erwartenden Transaktionsraten betrachtet werden.
- die Auswirkung der unterschiedlichen Verfahren der verteilten Verarbeitung dargestellt
- ein Beispiel im einzelnen durchgerechnet

Zunächst sollen jedoch noch einige allgemeine Annahmen des Modells beschrieben werden.

2 Allgemeine Annahmen

- jede zu besetzende Stelle ist eindeutig einer Firma zugeordnet
- jede Firma oder Niederlassung ist eindeutig einem Arbeitsamt zugeordnet, damit ist auch jede Stelle eindeutig verantwortlich einem Arbeitsamt zugeordnet
- jeder Arbeitssuchende wird zu einem Zeitpunkt nur in einem Arbeitsamt verantwortlich geführt
- jede Stelle und jeder Arbeitssuchende wird eindeutig einem Berufsbereich zugeordnet
- in einem Arbeitsamt ist ein Team Anzahl von Sachbearbeitern für einen Berufsbereich zuständig.
- von diesem Team können sowohl stellenorientierte als auch bewerberorientierte Vermittlungen durchgeführt werden
- jedem Team kann eine Datei zugeordnet werden, in die zu bearbeitende Vorgänge gespeichert werden können. Wie diese Datei verwendet wird, wird bei der Beschreibung der einzelnen Vorgänge erläutert.
- nicht berücksichtigt sind die Auswirkungen eines Workflow-Systems und einer elektronischen Archivierung von Dokumenten.

3 Allgemeine Suchstrategie

Im folgenden wird allgemein beschrieben, in welcher Reihenfolge welche rechnerunterstützte Suchoperationen durchgeführt werden.

3.1 Suchstrategie

Bei der Suche nach geeigneten Bewerbern oder geeigneten Stellen wird nach folgenden Kriterien gesucht

- Übereinstimmung von berufsbezogenen Eigenschaften (Berufsschlüssel, BKZ, berufsfachliche Merkmale) innerhalb eines bestimmten Pendelbereiches (räumliche Umgebung einer Firma, aus der Arbeitssuchende in Frage kommen oder Umgebung eines Arbeitssuchenden, in der ihm zuzumuten ist, eine Stelle anzunehmen). Es wird im folgenden zwischen dem Pendelbereich einer Stelle und dem Pendelbereich eines Arbeitssuchenden unterschieden (siehe. ABB.1)
- wobei davon ausgegangen wird, daß im allgemeinen der erste Bereich größer ist als der zweite. (Es wird dabei vorausgesetzt, daß der Pendelbereich als Suchbedingung formuliert werden kann)

Grundsätzlich wird bei der Suche so vorgegangen, daß

- zunächst versucht wird, die Suche im eigenen Arbeitsamt zu befriedigen
- und nur dann, wenn dies nicht möglich ist, andere Arbeitsämter oder ein zentrales System einzubeziehen.

3.2 Zusammenspiel der Arbeitsamtsrechner im Pendelbereich

In Abb.2 ist anhand eines einfachen Beispiels der stellenorientierten Bewerbervermittlung das Zusammenspiel der einzelnen Arbeitsamtsrechner dargestellt.

Ein bestimmter Prozentsatz der Suchfragen kann lokal erfüllt werden, im Rest der Fälle wird die Suchfrage an die Arbeitsämter im Pendelbereich geschickt. Aus dem Beispiel ergibt sich, daß ein Arbeitsamtsrechner im Durchschnit $p \cdot *$ so viele Suchfragen von außen erhält, wie er selbst nach außen schickt (p = die Anzahl der AA-Rechner im Pendelbereich).

Es werden im Durchschnit aus einem Arbeitsamtsgebiet so viele Stellenangebote bzw. Bewerbervorschläge an andere Arbeitsämter geliefert wie umgekehrt das Arbeitsamt erhält.

In Bereichen mit einer hohen Arbeitslosigkeit wird dies natürlich ganz anders sein.

4 Stellenorientierte Bewerbervermittlung

4.1 Aufnahme einer neuen Stelle

Jede dem Arbeitsamt gemeldete neu zu besetzende Stelle wird erfäßt und in dem EDV-System des Arbeitsamtes gespeichert. Gespeichert werden u.a

- der Betrieb, der die Stelle anbietet
- die beruflichen Anforderungen an einen in Frage kommenden Bewerber

4.2 Suche nach Bewerbern

Nach der Speicherung der Daten wird ein erster rechnerunterstützter Suchlauf nach geeigneten Bewerbern durchgeführt. Dabei wird davon ausgegangen, daß es eine obere Grenze an Bewerbern für eine Stelle gibt. Diese Grenze ist entweder für alle Stellen gleich oder wird pro Stelle (evtl. durch den Betrieb) festgelegt. Die Suche geschieht nach den in Kapitel 3 beschriebenen Prinzipien d. h.

- es wird zunächst lokal gesucht und nur wenn nicht genügend Bewerber lokal gefunden werden, wird
- im Pendelbereich der Stelle weitergesucht.

Das Ergebnis des Suchlaufes sind Bewerbersätze, die entweder im lokalen Arbeitsamt oder in den Arbeitsämtern des Pendelbereiches verwaltet werden. Die entgültige Entscheidung, ob ein Arbeitssuchender als Bewerber für die Stelle in Frage kommt, wird den Sachbearbeitern des Teams überlassen, zu der der ermittelte Arbeitssuchende gehört. Die Sachbearbeiter müssen nach einer positiven Entscheidung den Arbeitssuchenden über die Stelle informieren, so daß dieser sich mit dem Betrieb in Verbindung setzen kann.

Es besteht also die Aufgabe, die Teams über die in Frage kommenden Vermittlungen zu informieren. Dies soll über die den Teams zugeordneten Team-Dateien geschehen. D. h. aus den Daten der gefundenen Arbeitssuchenden muß eindeutig die Team-Datei ermittelt werden können, in die die Informationen der in Frage kommenden Arbeitssuchenden zu schreiben sind.

Dies bedeutet, daß pro gefundenem Arbeitssuchenden ein Eintrag in die entsprechende Team-Datei geschrieben wird, der einen Verweis auf die Stellenbeschreibung und den Arbeitssuchenden enthält. Betrifft der Eintrag einen Arbeitssuchenden aus einem anderen Arbeitsamt, so muß in eine Datei eines anderen Rechners geschrieben werden.

Wie die einzelnen Einträge der Team-Datei auf die Sachbearbeiter aufgeteilt werden, wird hier nicht weiter betrachtet.
Es wird davon ausgegangen, daß ein Teil der gefundenen Arbeitssuchenden von den Sachbearbeitern wieder ausgeschlossen wird. Dies geschieht aufgrund von zusätzlichen Informationen über die Arbeitssuchenden, die nur den Sachbearbeitern bekannt sind (z. B. Aktennotizen, Informationen aus persönlichen Gesprächen).
Nachdem ein Sachbearbeiter entschieden hat, daß ein Arbeitssuchender für eine Stelle in Frage kommt, werden folgende Aktionen durchgeführt:

- schriftliche oder telefonische Benachrichtigung des Arbeitssuchenden. (Es wird nicht davon ausgegangen, daß der Arbeitssuchende zunächst einberufen, ein Gespräch über die Stelle mit ihm geführt und danach erst entschieden wird, ob die Stelle in Frage kommt.)
- Update der Stellendaten (Zähler für Anzahl angebotener Bewerber, Verweis auf den Arbeitssuchenden)
- Update der Daten des Arbeitssuchenden (Verweis auf die Daten der Stelle)

Damit ist es jederzeit möglich, pro Stelle eine Übersicht über die vorgeschlagenen Arbeitssuchenden zu erstellen. Da die Stellen an den Betrieben gekoppelt sind, ist es auch jederzeit möglich, pro Betrieb die Vorschläge für die einzelnen von ihr angebotenen Stellen zu ermitteln.

4.3 Folgebearbeitungen

Im weiteren Zusammenhang mit den gemeldeten Stellen werden einige Folgeaktivitäten

- entweder automatisiert oder
- auf Initiative der Sachbearbeiter

durchgeführt.

Als Folgebearbeitungen sind folgende Verfahren denkbar:

1. Ein Folgesuchlauf nach z. B. 8 Tagen, da von den Sachbearbeitern ein Teil der ermittelten Bewerber ausgesondert wird. (Dies kann jedoch auch dadurch abgefangen werden, daß beim ersten Suchlauf mehr Bewerber als benötigt gesichtet werden). Dieses Verfahren setzt jedoch voraus, daß bei den Stellen vermerkt wird, welche Bewerber ausgeschlossen wurden, da sonst bei dem erneuten Suchlauf u. a. wieder diese Personen gefunden werden.
2. Weiterhin erscheint es sinnvoll periodisch automatisierte Suchläufe durchzuführen, die die Stellen ausgeben, die länger als bsp. 1 Monat, nicht besetzt wurden (sortiert nach bsp. Berufsgruppen und Betrieben). Dies kann gekoppelt werden mit Suchläufen nach neuen in Frage kommenden Bewerbern.
3. Weiterhin werden in bestimmten Zeitabständen firmenspezifische Auswertungen durchgeführt, die eine Übersicht über die Besetzung von angebotenen Stellen und noch nicht besetzte Stellen geben. Diese Übersichten sind die Grundlage für Gespräche mit den Firmen.
Diese Aktivitäten werden nicht einzeln in das mathematische Modell aufgenommen, es werden die Daten bestimmter Aktivitäten der Neuaufnahme einer Stelle um eine wählbaren Faktor f erhöht (s. h. Abschnitt 7.2.6).

4.4 Löschen einer Stelle

Stellen, die besetzt worden sind, werden aus dem Vermittlungsprozess herausgenommen. Sie werden unter anderem jedoch für Statistikauswertungen benötigt. Deshalb bleiben sie im Rechner des Arbeitsamtes gespeichert, dem die Stelle zugeordnet ist, sie werden aber beim Suchverfahren nicht mehr berücksichtigt. Da Stellen evtl. in den Datenbanken mehrerer Arbeitsämter gespeichert werden, muß das Löschen über eine verteilte Verarbeitung durchgeführt werden.

5 Bewerberorientierte Stellenvermittlung

Bei der bewerberorientierten Stellenvermittlung wird versucht, für einen Arbeitssuchenden eine Reihe von Stellen zu ermitteln, die für ihn in Frage kommen, so daß er sich bei den Betrieben die die Stellen anbieten, bewerben kann.

In dem Modell wird unterschieden zwischen

- Dem Erstkontakt, bei dem zunächst die Bewerberdaten erfaßt werden. Zu diesen Bewerberdaten gehören auch die Ausbildungsdaten, die Wünsche des Bewerbers, seine Fähigkeiten. Weiterhin wird der Pendelbereich für den Bewerber ermittelt. Aufgrund dieser Angaben wird die Suchfrage für den Bewerber festgelegt, dies geschieht u. a. dadurch, daß erste Suchfragen lokal durchgeführt werden und anhand der Treffermenge festgestellt wird, ob das Ergebnis den Wünschen des Bewerbers entspricht. Andernfalls kann mit einer an die Wünsche des Bewerbers angepaßten Suchfrage nochmals gesucht werden. Dieser Prozess kann so lange wiederholt werden, bis das Ergebnis befriedigend ist. Die endgültige Suchfrage wird gespeichert, so daß sie jederzeit auch ohne die Gegenwart des Arbeitssuchenden wiederholt werden kann.

Werden Stellen gefunden, die für den Bewerber in Frage kommen, so werden

die Stelleninformationen in die Daten des Bewerbers aufgenommen. Damit kann u. a. in einem späteren Suchlauf verhindert werden, daß die gleichen Stellen nochmals angeboten werden. Dies geschieht dadurch, daß man die gefundenen Stellen mit den schon vorher angebotenen Stellen vergleicht.

- Verweise über Bewerber in die Stellen aufgenommen

Bei der Suche nach Stellen muß berücksichtigt werden, daß die Maximalzahl der Bewerber pro Stelle nicht überschritten wird. Andererseits werden einem Bewerber auch nur eine bestimmte Anzahl von Stellen angeboten.

- Folgeberatungen. Hier werden folgende 3 Verfahren in das mathematische Modell aufgenommen

1. es wird periodisch (z. B. monatlich) die zuletzt für den Bewerber gespeicherte Suchfrage asynchron durchgeführt und nach Stellen gesucht, die dem Arbeitssuchenden noch nicht angeboten wurden. Das Ergebnis wird in die Team-Datei gespeichert. Die Weiterbearbeitung geschieht analog der stellenorientierten Bewerbervermittlung (s. h. Abschnitt 4.2).
2. der Sachbearbeiter legt Termine für Folgeberatungen fest. Dies kann

z. B. dadurch geschehen, daß er bei einer Beratung den Folgetermin schon bestimmt. Es besteht dann die Möglichkeit, automatisch einen asynchronen Suchlauf vorher ablaufen zu lassen und das Ergebnis in der Team-Datei zu speichern. Diese Informationen liegen dann beim Besuch des Bewerbers schon vor.

3. Arbeitssuchende kommen von sich aus (evtl. unangemeldet) und hätten gerne eine Folgeberatung. Auch in diesem Fall kann bei der Empfangsstelle quasi per Knopfdruck zunächst ein Asynchronlauf gestartet werden, so daß der Sachbearbeiter bei der Beratung schon ein Ergebnis über in Frage kommende Stellen vorliegen hat.

6 Verfahren der verteilten Verarbeitung

In den vorherigen Abschnitten wurden die Vorgehensweisen bei

- der stellenorientierten Bewerbervermittlung und
- der bewerberorientierten Stellenvermittlung beschrieben.

Zur Erfüllung der Aufgaben werden u. a. die Stellendaten und die Bewerberdaten in Datenbanken gespeichert und in diesen Daten nach Stellen oder Bewerbern gesucht.

Es gibt nun mehrere Möglichkeiten die Daten auf die verschiedenen Rechner zu verteilen. Die Art der Verteilung bestimmt das Zusammenwirken der einzelnen Rechner und u. a. auch

- die Verfahren der Speicherung der Daten
- die Methoden des Suchens in den Daten
- die Treffmengen
- die Anforderungen an die Verfügbarkeit der Verbindungen zwischen den Rechnern
- die Performance des Gesamtsystems

Die unterschiedlichen Methoden und ihre Auswirkungen werden im mathematischen Modell im nächsten Kapitel beschrieben.

Es wird grundsätzlich davon ausgegangen, daß die rechnerübergreifende Bearbeitung innerhalb einer Transaktion aus der Sicht des lokalen Rechners (Auftraggeber) geschieht. Andernfalls müßte

- entweder eine separate Transaktion vom Sachbearbeiter gestartet werden oder
- eine zusätzliche asynchrone Transaktion initiiert und verwaltet werden

In beiden Fällen würde dadurch eine zusätzliche Belastung der Rechner entstehen.

Nur wenn keine Verbindung zum Partnerrechner besteht wird lokal eine zusätzliche asynchrone Transaktion notwendig. Es wird im folgenden davon ausgegangen, daß dies ein seltener Fall ist und daher kaum ins Gewicht fällt. Daher wird dieser Fall in dem Modell vernachlässigt. Auf der Auftragnehmerseite (entfernter Rechner) entstehen jedoch neue Transaktionen die in den einzelnen im folgenden beschriebenen Szenarien berücksichtigt werden müssen.

Im folgenden werden zunächst mehrere unterschiedliche Szenarien betrachtet. Dabei werden vor allem die Voraussetzungen der einzelnen Szenarien beschrieben.

6.1 Szenario1 (zentrales Verfahren)

Bei diesem Szenario (s. h. ABB.3) werden alle notwendigen Daten nur auf einem Zentralrechner gespeichert und die PC der Sachbearbeiter kommunizieren direkt mit diesem Zentralrechner.

Dabei wird von folgenden Voraussetzungen ausgegangen:

1. es gibt einen leistungsfähigen zentralen Rechner mit einem Datenbanksystem
2. alle PC's der Sachbearbeiter kommunizieren direkt mit diesem zentralen Rechner
3. alle Betriebsdaten und Stellendaten werden zentral gespeichert
4. alle Arbeitssuchende werden zentral gespeichert
5. alle EDV-Aktivitäten für die Vermittlung finden auf dem Zentralrechner statt
6. die einzelnen Team-Dateien liegen ebenfalls auf dem Zentralrechner

Dieses Verfahren soll im wesentlichen Aussagen über die Last der Transaktionen machen, wenn keine verteilte Verarbeitung zwischen Rechnern stattfindet.

6.2 Szenario2 (zentrale Stellenverwaltung, dezentrale Bewerberverwaltung)

Bei diesem Szenario (s. h. ABB.3) werden

- die Bewerberdaten nur auf den lokalen Rechnern der Arbeitsämter geführt
- die Stellendaten sowohl lokal als auch zentral gespeichert.

Die Suche nach

- geeigneten Bewerbern für eine Stelle erfolgt nur lokal geeigneten Stellen für Bewerber geschieht lokal und zentral. Dieses Verfahren geht von folgenden Voraussetzungen aus:

1. es gibt einen zentralen Server mit einem Datenbanksystem
2. es gibt pro Arbeitsamt einen dezentralen Server mit einem Datenbanksystem
3. alle PC der Sachbearbeiter kommunizieren mit einem der dezentralen Server
4. alle dezentralen Server sind mit dem zentralen Server verbunden
5. jeder Betrieb, der Stellen anbietet ist eindeutig einem Arbeitsamt zugeordnet. Damit ist auch jede Stelle einem Arbeitsamt zugeordnet
6. alle Stellen werden auf dem lokalen Server und dem zentralen Server gespeichert
7. alle Bewerberdaten werden nur lokal geführt
8. die Informationen über die Zuordnung von Bewerbern zu Stellen werden lokal und zentral geführt. D. h. die Updates der Sachbearbeiter in den Stellendaten führen zu einer verteilten Verarbeitung. Auch Updates in Bewerberdaten (Zordnung stellen) führen zu verteilter Verarbeitung.
9. die Suchstrategie sieht dann folgendermaßen aus:

stellenorientierte Bewerbervermittlung

es wird nur auf dem lokalen Rechner des Arbeitsamtes, das der Stelle zugeordnet ist, gesucht

bewerberorientierte Stellenvermittlung

es wird zunächst lokal gesucht. Werden genügend Stellen gefunden, wird abgebrochen; andernfalls wird zentral weitergesucht, wobei natürlich die Wünsche bzgl. Pendelbereich berücksichtigt werden müssen.

6.3 Szenario3 (teilweise zentrale Stellenverwaltung, dezentrale Bewerberverwaltung)

Dieses Szenario (s. h. ABB.3) unterscheidet sich von Szenario2 dadurch, daß $X\%$ der Stellen nur lokal geführt werden (bestimmte Berufsgruppen) und $(100 - X)\%$ weiterhin zentral verwaltet werden. Dies bedeutet, daß alle Vorgänge, die diese Berufsgruppe betreffen lokal behandelt werden. Dies betrifft auch die Suchstrategie für die bewerberorientierte Stellenvermittlung. Für die Gruppen wird dann nur noch lokal gesucht und auch ein zentraler Update entfällt.

6.4 Szenario4 (Suche im Pendelbereich)

Dieses Szenario (s. h. ABB.4) geht davon aus, daß bis auf einen kleinen Prozentsatz für ausgewählte Berufsgruppen auch die Stellen lokal verwaltet werden und daß die Suche nach Stellen auf die Arbeitsämter des Pendelbereichs eines Bewerbers ausgedehnt wird. Bei diesem Modell gelten folgende Voraussetzungen.

1. es gibt einen zentralen Server mit einem Datenbanksystem, in dem die Stellen der ausgewählten Berufsgruppen gespeichert werden
2. es gibt pro Arbeitsamt einen dezentralen Server mit einem Datenbanksystem
3. alle PC der Sachbearbeiter sind mit dem dezentralen Server verbunden
4. alle dezentralen Server kommunizieren mit dem zentralen Server
5. alle dezentralen Server haben Zugriff auf die Daten der umliegenden Arbeitsämter, die im Pendelbereich liegen
6. jeder Betrieb, der Stellen anbietet, ist eindeutig einem Arbeitsamt zugeordnet. Damit ist auch jede Stelle einem Arbeitsamt zugeordnet
7. $X^* \%$ aller Stellen werden lokal verwaltet und $(100 - X^*)\%$ lokal und zentral, wobei X^* nicht identisch mit X im Szenario3 ist. In diesem Modell wird davon ausgegangen, daß nur noch ein geringer Anteil (bestimmte Berufsgruppen) zentral verwaltet werden, d. h. X^* groß (ca. 95%) ist
8. alle Bewerberdaten werden lokal geführt
9. die Informationen über die Zuordnung von Bewerbern zu Stellen werden lokal geführt und nur für die bestimmten Berufsgruppen auch zentral. Dies bedeutet, daß nicht nur zur Suche auf die Daten der im Pendelbereich befindlichen Server zugegriffen wird sondern auch zum Update der Stellendaten (siehe Szenario 2)
10. die Suchstrategie im Pendelbereich und das Zusammenspiel der Server der einzelnen Arbeitsämter ist in Kapitel 3 dargestellt. Zusätzlich muß die zentrale Suche für die bestimmten Berufsgruppen berücksichtigt werden.

Dieses Verfahren hat den Nachteil, daß relativ viele Suchfragen in umgebende Arbeitsämter geschickt werden müssen. Eine Entlastung könnte durch das folgende Szenario erreicht werden.

6.5 Szenario5 (Speicherung der Stellen im Pendelbereich)

Bei diesem Szenario (s. h. ABB.5) werden die Stellendaten nicht nur lokal gespeichert sondern auch auf den Servern der Arbeitsämter, die im Pendelbereich für Stellen des Arbeitsamtes liegen. Diese Speicherung geschieht bei der Erstspeicherung der Stelle. Die Information, welche Bewerber für eine Stelle vorgeschlagen wurden, werden nur in dem Amt geführt, dem die Stelle zugeordnet ist. Die Replikate dienen nur der Vorauswahl, um zu ermitteln, ob ein Bewerber für die Stelle in Frage kommt.

Das Szenario4 wird folgendermaßen abgeändert:

1. alle Stellen, die nicht zentral verwaltet werden, werden auch auf den Rechnern des Pendelbereichs der Stelle gespeichert. Damit werden alle Suchfragen, die bisher in den Pendelbereich geschickt wurden, lokal abgehandelt
2. die Zuordnung der Bewerber zu den Stellen wird nur lokal geführt, d. h. der Stellenupdate geschieht über das Netz
3. Stellen, bei denen die Obergrenze der Bewerbervorschläge erreicht wurde, werden auf allen Rechnern als nicht zu berücksichtigende Stellen gekennzeichnet, sie können aber wieder aktiviert werden, wenn keiner der Bewerber sich als geeignet herausstellt. Dieser Update kann asynchron geschehen. Die Transaktionen hierzu werden in dem Modell nicht berücksichtigt.
4. die Stellen werden erst beim expliziten Löschen auf allen Rechnern entfernt.

6.6 Szenario6 (Speicherung der Stellen in den Landesarbeitsämtern)

Bei diesem Szenario (s. h. ABB.6) wird ein dreistufiges Konzept realisiert. Dabei werden zwischen die Rechner der einzelnen Arbeitsämter und dem zentralen Rechner die Rechner der Landesarbeitsämter geschaltet. Dabei gelten folgende Voraussetzungen:

1. jeder Arbeitsamtrechner ist mit einem Rechner eines LAA verbunden
2. jeder LAA-Rechner ist mit dem Zentralrechner gekoppelt
3. alle Bewerber werden nur in den lokalen Rechnern geführt
4. alle Stellen werden sowohl lokal als auch auf dem Rechner des LAA gespeichert
5. die Stellen der besonderen Berufsgruppen des Szenarios4 werden zusätzlich zentral gespeichert
6. die Zuordnung der Stellen zu den Bewerbern wird lokal, in den LAA-Daten und für die besonderen Berufsgruppen zentral geführt. Das hierzu notwendige Update-Verfahren ist in ABB.11 dargestellt
7. die Suchstrategie ist folgende:
 - zunächst wird lokal gesucht
 - bei nicht ausreichender Treffermenge wird in den LAA-Daten recherchiert
 - im Bedarfsfall wird für die speziellen Berufsgruppen auch zentral gesucht.

Dieses Modell hat den Nachteil, daß ein Pendelbereich zwischen den LAA nicht berücksichtigt wird. Eine Ausdehnung auf einen solchen Bereich müßte evtl. noch durchgeführt werden.

6.8 Das Speicherkonzept

In ABB. 8 ist das Speicherkonzept der einzelnen Szenarien dargestellt. Daraus geht hervor, daß

- die Bewerberdaten in fast allen Szenarien nur lokal gespeichert werden mit Ausnahme Szenario1 wo alle Daten in einer zentralen Datenbank enthalten sind und in Szenario7 wo zur Erleichterung der stellenorientierten Bewerbervermittlung die Bewerber auch im Pendelbereich verteilt werden
- die Stellendaten auch alle lokal gespeichert werden, Ausnahme Szenario 1, die einzelnen Szenarien aber in der unterschiedlichen Art der Speicherung der Replikate differieren. Dabei wird davon ausgegangen, daß auf dem Server, der Owner einer Stelle ist, immer alle Informationen vorhanden sind, während die Replikate oft nur die Daten beinhalten, die zur Auswahl der Stelle für einen Bewerber benötigt werden

6.7 Szenario7 (Speicherung der Stellen - und Bewerberdaten im Pendelbereich)

Dieses Szenario (s.h. ABB. 7) ist eine Erweiterung von Szenario 5. Es werden nicht nur die Stellendaten sondern auch die Bewerberdaten in den Servern des Pendelbereichs eines Arbeitsamtes gespeichert. Damit wird erreicht, daß bei der stellenorientierten Bewerbervermittlung die notwendigen Suchfragen in den Pendelbereich nicht mehr über das Netz geschickt werden müssen. Dies hat neben der Reduktion der Netzbelastung den Vorteil, daß eine komplette Suche auch dann möglich ist, wenn die Verbindung zu einem Nachbar-Server ausgefallen ist.

Das Szenario5 wird folgendermaßen abgeändert:

1. die Bewerberdaten werden nicht nur lokal sondern auch im Pendelbereich des Bewerbers gespeichert. Dies gilt jedoch nur für die Daten, die zur Auswahl des Bewerbers benötigt werden, die Erweiterungen, die auch die angebotenen Stellen beinhalten, werden weiterhin nur lokal geführt. Damit ist ein häufiger Update im Pendelbereich nicht notwendig
2. der Pendelbereich für Bewerber ist nicht immer identisch mit dem Pendelbereich für Stellen
3. Das Schreiben in die Team-Datei und der Update der Stellendaten geschieht wie in Szenario5 über das Netz
4. Die Bewerberdaten werden gelbtscht, sobald der Bewerber aus dem Prozess der Arbeitsuche ausscheidet. Im lokalen Rechner werden die Daten noch für einen bestimmten Zeitraum weitergeführt.

7 Das mathematische Modell

Im folgenden wird ein mathematisches Modell entwickelt, das

- die stellenorientierte Bewerbervermittlung und
 - die bewerberorientierte Stellenvermittlung
- umfaßt. Dieses Modell soll
- die verteilte Verarbeitung zwischen den einzelnen Rechnern für die verschiedenen Szenarien beleuchten
 - feststellen, welche Parameter die verteilte Verarbeitung bestimmen
 - es jederzeit ermöglichen, den Einfluß der einzelnen Parameter zu bestimmen (dies soll durch Speicherung der mathematischen Formeln in EXCEL geschehen)
 - durch die Ermittlung der anfallenden Transaktionen auf den einzelnen Rechnern untersuchen, welche Szenarien am wenigsten Last auf die zur Verfügung stehenden Rechner bringt

Das Modell soll also eine Entscheidungshilfe für die Auswahl der Art der verteilten Verarbeitung in einem zukünftigen Arbeitssystem sein.

7.1 Die Basisvariablen

Im folgenden werden zunächst einige Basisvariablen beschrieben.

Architektur

AA: Anzahl der Arbeitsämter

RAA: Anzahl der UNIX-Rechner mit INFORMIX-Online-Systemen in denen Stellen-oder Bewerberdaten gespeichert sind. Diese Anzahl ist nicht identisch mit AA, da in einigen Fällen mehrere Arbeitsämter mit einem Rechner zusammenarbeiten.

LAA :Anzahl der Rechner der Landesarbeitsämter

D :Gesamtanzahl der PC's aller Arbeitämter

mD =D/RAA: Anzahl solcher PC'S eines mittleren Arbeitsamtes

xD :die größte Anzahl solcher PC's eines Arbeitsamtes (? des größten Arbeitsamtes)

$xD = a * mD$ d.h. a ist der Faktor, um den sich das größte Arbeitsamt von einem mittleren unterscheidet.

Aufteilung der Sachbearbeiter

arbd: Anzahl der PC'S an denen Arbeitsvermittlung durchgeführt wird, dies ist gleichzeitig die Anzahl der Sachbearbeiter, die diese Tätigkeit tun

marbD = arbd/RAA Anzahl der Vermittlungs-Sachbearbeiter eines mittleren Arbeitsamtes

xarbD = a * marbD Anzahl der Vermittlungs-Sachbearbeiter des größten Arbeitsamtes

leiD, mleiD sind analoge Variable für die Leistungs-Sachbearbeiter

AT: Anzahl der Tage, in denen die Arbeitsämter im Jahr geöffnet sind

ATS: Anzahl der Arbeitstage eines Sachbearbeiters, an denen er seine Aufgabe durchführt

$xN = a * mN$: die Anzahl der Stellen des größten Arbeitsamtes

$TN = N/AT$: durchschnittliche Anzahl der neuen Stellen, die bundesweit täglich aufgenommen werden

TmN, TxN sind die analogen Definitionen für das mittlere und große Arbeitsamt

Pendelbereich

q: durchschnittliche Anzahl der Arbeitsämter für den Pendelbereich der Stellen

p: durchschnittliche Anzahl der Arbeitsämter für den Pendelbereich der Bewerber

(AT-Urlaubstage, Krankheitstage, Ausbildungstage)

t: durchschnittliche tägliche Arbeitszeit eines Sachbearbeiters

t1: tägliche Arbeitszeit für Publikumsverkehr

t2: Arbeitszeit ohne Publikumsverkehr

$t = t1 + t2$

Arbeitssuchende

AS: Gesamtanzahl der bundesweit registrierten und in Datenbanken gespeicherten Arbeitssuchenden

mAS = AS/RAA: durchschnittliche Anzahl registrierter Arbeitssuchender eines mittleren Arbeitsamtes

xAS = a * mAS größte registrierte Anzahl von Arbeitssuchenden, die in einer Datenbank gespeichert sind

AS+: Gesamtanzahl von Arbeitssuchenden, die bundesweit jährlich neu registriert werden

AS-: Gesamtanzahl der Arbeitssuchenden, die bundesweit jährlich aus den Datenbanken gelöscht werden

mAS+, xAS+, mAS-, xAS- sind analoge Variable für mittlere Arbeitsämter und das größte Arbeitsamt

TAS+ = AS+/AT: Anzahl Arbeitssuchender, die im Durchschnitt täglich neu hinzukommen

- TAS- = AS-/AT
- TmAS+ = mAS*/AT
- TmAS- = mAS-/AT
- TxAS+ = xAS+/AT
- TxAS- = xAS-/AT

analog gilt:

Stellen

N: Anzahl der Stellen, die bundesweit jährlich neu besetzt werden sollen. Diese Stellen werden gespeichert.

m.N: zu besetzende Stellen eines mittleren Arbeitsamtes

7.2 Stellenorientierte Bewerbervermittlung

Das Verfahren der stellenorientierten Bewerbervermittlung ist in Kapitel 4 beschrieben. Das Verfahren besteht demnach aus folgenden Aktivitäten:

- Neuaufnahme von Stellen
- Löschen von Stellen
- Suche nach geeigneten Bewerbern bei der Neuaufnahme
- Schreiben der Bewerber in die Team-Datei
- Entscheidung der Sachbearbeiter
- Folgebearbeitungen

7.2.1 Neuaufnahme einer Stelle

Bei der Neuaufnahme einer Stelle wird davon ausgegangen, daß die Informationen des Betriebes, der die Stelle anbietet schon im lokalen Rechner gespeichert sind. (Die Erfassung der Daten der Betriebe ist in dem Modell nicht berücksichtigt).

Zur Erfassung der Daten einer Stelle werden zunächst die Betriebsdaten gelesen.

transSuFi : durchschnittliche Anzahl der Transaktionen zum Lesen der Betriebsdaten

transSpN^* : die Anzahl der Transaktionen zur reinen Speicherung der Stellendaten, wobei ein Verweis auf den Betrieb enthalten ist, so ergibt sich als Gesamtanzahl der Transaktionen zur Speicherung einer neuen Stelle auf einem lokalen Rechner

$$\text{transSpN} = \text{transSuFi} + \text{transSpN}^*$$

Die tägliche Anzahl der Transaktionen zur Speicherung der lokalen Stellen in einem mittleren Arbeitsamt sind dann:

$$\text{TransSpmN} = \text{transSpN} * \text{TmN}$$

wobei TmN die durchschnittliche Anzahl neuer Stellen pro Tag ist. Berücksichtigt man, daß z. B. in Szenario5 die Stellendaten auch in den Rechnern der Arbeitsämter im Pendelbereich gespeichert werden, ergibt sich für einen lokalen Rechner

$$\text{TransSpmN} = (\text{transSpN} + q * \text{transSpN}^*) * \text{TmN}$$

Die Anzahl der Transaktionen zur Speicherung von Stellen setzt sich also wie folgt zusammen

- einem Anteil zur Speicherung der Stellen, die in dem betreffenden Arbeitsamt verwaltet werden (dieser Anteil tritt bei allen Szenarien auf) und
- einem Anteil zur Speicherung fremder Stellen (Szenario5 und 7)

Dies ist die allgemeine Formel für die Anzahl der täglichen Transaktionen zur Speicherung neuer Stellen auf dem lokalen Rechner.

Bei den Szenarien 2 bis 7 müssen Stellendaten zusätzlich auf einem zentralen Rechner oder auf Rechnern des Pendelbereichs gespeichert werden. Dies sind zusätzliche Transaktionen, wobei dabei nicht mehr das Lesen der Firmendaten notwendig ist.

$\text{TransSpz}N = \text{transSp}N * TN$ ist die Anzahl der Transaktionen, wenn nach Szenario 1 alles zentral abgehandelt wird.

$\text{TransSpz}N^* = \text{transSp}N^* * TN$: Anzahl der zentralen Transaktionen bei Szenario 2 (lokal u. zentral)

$\text{TransSp}X = (100 - X)/100 * \text{transSp}N^* * TN$: Anzahl der Transaktionen am Zentralrechner bei Szenario 3

$\text{TransSp}X^* = (100 - X^*)/100 * \text{transSp}N^* * TN$: Anzahl der Transaktionen am Zentralrechner, wenn nur bestimmte Berufsgruppen zentral verwaltet werden.

(X, X* sind der prozentuale Anteil der lokal behandelten Stellen, s. h. Abschnitt 6.3 und 6.4).

7.2.2 Löschen einer Stelle

- Stellen werden im allgemeinen aufgrund von Informationen der Betriebe wieder gelöscht. Man kann sich auch asynchrone Verfahren vorstellen, die in bestimmten Zeitabständen den Stellenpool durchforsten und entsprechend ermittelte Stellen dem Team der Sachbearbeiter melden, die dann die Stellen löschen.
- Der Löschvorgang besteht darin, daß lokal die Stelle als besetzt (zur Reaktivierung vorgemerkt) gekennzeichnet wird , sie steht unter anderen für Statistikzwecke weiter zur Verfügung
- die Replikate auf dem zentralen Rechner oder den Rechnern des Pendelbereichs werden gelöscht werden.

Es wird in dem Modell davon ausgegangen, daß beim Löschen die gleichen Vorgänge ablaufen wie beim Speichern, evtl. ist die Anzahl der Basistransaktionen unterschiedlich. Daher ist in den obigen Formeln grundsätzlich Sp durch Lö zu ersetzen. Also z. B.

$\text{TransLö}N = \text{transLö}N * TmN$ wobei z. B. $\text{transLö}N$ ungleich $\text{transSp}N$ sein kann, d. h. es wird davon ausgegangen, daß die durchschnittliche Anzahl der Transaktionen zum Speichern und zum Löschen unterschiedlich sein kann.

7.2.3 Suchen nach geeigneten Bewerbern

Die Suche nach geeigneten Bewerbern geschieht nach den in Abschnitt 3.1 beschriebenen Kriterien. ABB.2 zeigt beispielhaft den Ablauf einer solchen Suche.

In dem Modell wird davon ausgegangen, daß eine maximale Anzahl O von Bewerbern pro Stelle gesucht wird. Es wird vorausgesetzt, daß diese Zahl O für alle Berufsgruppen und Stellen gleich ist, man kann sie auch als einen Mittelwert ansehen.

Wie in dem Beispiel gezeigt, wird für einen bestimmten Prozentsatz ($r1$) der Suchfälle das Limit schon lokal erreicht, $r2$ ist der prozentuale Anteil der Fälle, in denen im Pendelbereich gesucht wird.

$$r1 + r2 = 100$$

transSulokB: durchschnittliche Anzahl der lokalen Suchvorgänge pro neuer Stelle, hierin sind auch enthalten Probeversuche zur Festlegung der entgeltigen Suchfrage

transSuremB: durchschnittliche Anzahl der Suchvorgänge in den Arbeitämtern des Pendelbereichs. Diese Zahl wird normalerweise 1 sein, da die Suchfrage schon bei der lokalen Suche festgelegt wurde.

Es wird davon ausgegangen, daß das Suchen in den anderen Rechnern aus der Sicht des lokalen Rechners eine verteilte Transaktion ist, es entstehen jedoch zusätzliche Transaktionen auf den anderen Rechnern. Dies bedeutet, daß am Anfang der Transaktion die Suchfrage an alle Server im Pendelbereich geschickt wird und danach auf die Ergebnisse von den einzelnen Servern gewartet wird. Besteht zu einem der Rechner keine Verbindung, so wird dort nicht gesucht. Es wird davon ausgegangen, daß eine nicht vorhandene Verbindung in Zukunft ein seltenes Ereignis ist.

Die Gesamtanzahl der Transaktionen zum Suchen auf einem Rechner eines mittleren Arbeitsamtes setzt sich folgendermaßen zusammen:

$$\text{TransSumB} = \text{TransSulokmB} + \text{TransSuremB} + \text{TransSuvremmB}$$

wobei

TransSulokmB: die Anzahl der Transaktionen, die durch das lokale Suchen entstehen

TransSuremB: Anzahl der Suchaufträge an die Rechner im Pendelbereich

TransSuvremmB: Anzahl der Transaktionen, die sich durch Suchaufträge aus dem Pendelbereich ergeben

Es gilt dann:

$$\text{TransSumB} = \text{TmN} * (\text{transSulokB} + r2/100 * \text{transSuremB} + q * r2/100 * \text{transSuvremmB})$$

Die Gesamtanzahl der Suchtransaktionen nach geeigneten Bewerbern setzt sich also zusammen aus

- lokalen Suchtransaktionen
- den Suchtransaktionen an andere Rechner (Anteil $r2/100$)
- den Aufträgen der q Server im Pendelbereich

7.2.4 Schreiben in die Team-Datei

Wie in Abschnitt 4.1 beschrieben, werden pro gefundenem Bewerber ein Eintrag in eine Team-Datei gemacht. Es wird davon ausgegangen, daß jeder Eintrag eine eigene Transaktion ist.

L ist die durchschnittliche Anzahl der gefundenen Bewerber pro Stelle. In vielen Fällen kann man heute davon ausgehen, daß $L = 0$ ist.

$L1$: prozentualer Anteil der lokal gefundenen Bewerber, dieser Anteil ist nicht identisch mit $r1$ wie das Beispiel in ABB.2 zeigt.

$L2$: prozentualer Anteil remote gefundener Bewerber

$$L1 + L2 = 100\%$$

$TransWrmTd = (L1/100 * L + L2/100 * L + L2/100 * L) * TmN$ ist die Anzahl der Transaktionen zum Schreiben in eine Team - Datei, die in einem mittleren Amt anfällt. Dabei ist

- der erste Anteil, der durch die lokal gefundenen Bewerber entsteht
- der zweite Anteil, das Schreiben in Team-Dateien der anderen Arbeitsämter
- der dritte Anteil, das Schreiben von den entfernten Rechnern in die lokale Team-Datei

Dabei entstehen im Durchschnitt in einem mittleren Arbeitsamt pro Tag in alle Team-Dateien

$$L * TmN \text{ Einträge}$$

7.2.5 Entscheidung der Sachbearbeiter

Im folgenden werden die Transaktionen betrachtet, die durch Aktionen der Sachbearbeiter entstehen.

Jeder Eintrag in der Team-Datei enthält Verweise zu der Stelle (die lokal oder remote sein kann) und zu dem Bewerber, dessen Daten sich lokal befinden. Diese Transaktionen setzen sich wie folgt zusammen:

- Lesen der Einträge der Team-Datei (täglich: $L * TmN$)
- Lesen der Bewerberdaten (lokal) (täglich: $L * TmN$)
- Lesen der lokalen Stellendaten (täglich: $L1/100 * L * TmN$)
- Lesen entfernter Stellendaten (täglich $L2/100 * L * TmN$)
- Lesen lokaler Stelleninformationen aus der Ferne (tägl: $L2/100 * L * TmN$)
- Update der akzeptierten Bewerber (täglich: $1/100 * L * TmN$, wobei 1 der prozentuale der akzeptierten Bewerber ist)
- Update der lokalen Stellendaten (täglich: $1/100 * L1/100 * L * TmN$)
- Update der entfernten Stellendaten (täglich: $1/100 * L2/100 * L * TmN$)
- Update lokaler Stelleninformationen aus der Ferne (tägl: $1/100 * L2/100 * L * TmN$)

Durch Zusammenfassung ergibt sich dann folgende Formel für die lokalen Transaktionen, die durch die Sachbearbeiter aufgelöst werden:

$$TransSachmB = (1 + (1 + 1/100) * (2 + L2/100)) * L * TmN$$

Die Transaktionen, die sich aus der Bearbeitung der Sachbearbeiterdatei ergeben, setzen sich zusammen aus

- dem Lesen der Informationen der Sachbearbeiterdatei
- dem Lesen der Informationen der Stellen und Bewerber
- dem Update der Informationen der Stellen und Bewerber

7.2.6 Diskussion für die einzelnen Szenarien

Szenario1:

Bei diesem Szenario werden alle Stellen- und Bewerberdaten zentral gespeichert. In diesem Fall entfällt der Anteil der sich durch das Suchen und dem Update in der Ferne ergibt.

$$\text{TransSachB1} = (1 + 2 * (1 + 1/100)) * L * \text{TN}$$

Szenario2:

Bei diesem Szenario werden die Bewerberdaten lokal und die Stellendaten zentral gespeichert. Es gibt keinen Zugriff auf Daten anderer Ämter.

Es wird in diesem Fall nur lokal nach Bewerbern für die Stellen gesucht, d. h. man erhält also auch nur den Anteil LI

$$\text{TransSachB2} = (1 + 2 * (1 + 1/100)) * L/100 * L * \text{TmN}$$

Szenario3:

Dieses Szenario unterscheidet sich von Szenario2 dadurch, daß ein Teil der Stellen nur lokal gespeichert wird. Dies hat aber keinen Einfluß auf die stellenorientierte Bewerbervermittlung sondern nur auf bewerberorientierte Stellenvermittlung.

Es gilt daher die gleiche Formel wie bei Szenario2, da auch nur lokal gesucht wird.

Szenario4:

Bei diesem Szenario wird lokal und im Pendelbereich nach geeigneten Bewerbern gesucht.

Die generelle Formel gilt genau für dieses Szenario.

Szenario5:

Dieses Szenario unterscheidet sich von Szenario4 dadurch, daß die Stellen zusätzlich im Pendelbereich gespeichert werden. Die beiden Szenarien unterscheiden sich bzgl. der stellenorientierten Bewerbervermittlung dadurch, daß die Stellendaten nicht in der Ferne gelesen werden müssen, d. h. es entfällt der Anteil des Lesens lokaler Stellendaten aus der Ferne.

Der Update der Stellendaten wird nur auf dem Rechner durchgeführt, der Owner der Stelle ist, d. h. der Stellenupdate für entfernte Stellen geschieht über verteilte Verarbeitung.

Szenario6:

Bei diesem Szenario werden die Stellendaten zusätzlich in den Landesarbeitsämtern gespeichert. Da hier wie bei Szenario2 und 3 nur lokal gesucht wird gilt die bei Verfahren2 angegebene Formel

Szenario7:

In Szenario7 werden in Abwandlung von Szenario 5 auch die Bewerberdaten in den Pendelbereich verteilt. Dies hat Auswirkungen auf das Suchen nach geeigneten Bewerbern, es müssen keine Suchaufträge in die Ferne geschickt werden. Dadurch vereinfacht sich die Formel für das Suchen. Alle anderen Aktivitäten, außer den Folgeaktivitäten, in denen das Suchen wieder enthalten ist, sind nicht betroffen.

7.2.7 Folgebearbeitung

Wie in Abschnitt 4.2 ausgeführt, wird die Folgebearbeitung nicht einzeln modelliert sondern durch einen wählbaren Faktor berücksichtigt. Dabei werden jedoch nur die Aktivitäten des Suchens nach Bewerbern, das Schreiben in die Team-Datei und die Sachbearbeitertätigkeiten berücksichtigt, das Speichern und Löschen von Stellen spielt bei den Folgebearbeitungen keine Rolle.

$$\text{TransFormN} = \text{fa} * (\text{TransSumB} + \text{TransWrmTd} + \text{TransSachmB})$$

wobei fa ein wählbarer Faktor ist.

Wie groß dieser Faktor gewählt werden kann, wird u. a. von der Auslastung der Sachbearbeiter und den Anforderungen der Betriebe abhängen.

7.2.8 Gesamtanzahl der Transaktionen

Die Gesamtanzahl der Transaktionen ergibt sich als Summe der Einzeltransaktionen, wobei die Summierung getrennt für die lokalen, zentralen und LAA-Rechner erfolgen muß.

Eine Zusammenfassung der Formeln ist in den Abbildungen 9, 10 und 11 dargestellt.

7.2.9 Beispielrechnung

In den Abbildungen 12 und 13 ist eine Beispielrechnung für die stellenorientierte Bewerbervermittlung dargestellt. Dabei enthält ABB.12 die Annahmen für die einzelnen Parameter und ABB.13 die Ergebnisse. Eine Diskussion der Ergebnisse wird in Zusammenhang mit den Ergebnissen der bewerberorientierten Stellenvermittlung durchgeführt.

7.3 Bewerberorientierte Stellenvermittlung

Das Verfahren der bewerberorientierten Stellenvermittlung ist in Kapitel 5 beschrieben. Dabei wird unterschieden zwischen

- dem Erstkontakt, bei dem die Personaldaten erfasst werden und
- den Folgeberatungen.

7.3.1 Erstkontakt

Bei einem Erstkontakt sind folgende Aktivitäten durchzuführen

- Erfassung der Personaldaten
- Suche nach geeigneten Stellen
- Update der Personaldaten um die Stellendaten
- Update der Stellendaten um die Personaldaten der Bewerber

7.3.1.1 Zeit der Sachbearbeiter für Erst- und Folgeberatungen

Die durchschnittliche tägliche Anzahl der Erstkontakte pro Sachbearbeiter ist:

$$\begin{aligned} \text{TEber} &= \text{TAS} + \text{arbD} * \text{AT/ATS} \\ &= \text{TmAS} + \text{marbD} * \text{AT/ATS} \end{aligned}$$

ZEber sie die durchschnittliche Zeit für einen Erstkontakt, dann ist

$$\text{TZEber} = \text{ZEber} * \text{TEber} \text{ die Zeit, die ein Sachbearbeiter insgesamt für Erstkontakte pro Tag benötigt.}$$

Ist TZopen die tägliche Zeit, in der Publikumsverkehr ist, dann

$$\text{TZFber} = \text{TZopen} - \text{TZEber} \text{ die Zeit, die für Folgeberatungen mit Publikumsverkehr zu Verfügung steht}$$

Diese Anzahl ist im Durchschnitt unabhängig von der Größe des Arbeitsamtes.

Eine Beispielrechnung für die Zeitaufteilung der Arbeitszeit eines Sachbearbeiters ist in ABB. 14 durchgeführt.

7.3.1.2 Erfassen der Personaldaten

Die Transaktionen zur Erfassung der Personaldaten setzen sich wie folgt zusammen:

- **transSuAS:** Transaktionen zur Suche der Person
- **transSpPers:** Transaktionen zur Speicherung der Personaldaten
- **transSpBer:** Transaktionen zur Erfassung der Berufsdaten

Zusammengefaßt ergibt sich eine durchschnittliche Transaktionsrate:

$$\text{transErfB} = \text{transSuAS} + \text{transSpPers} + \text{transSpBer}$$

Außer in Szenario1 und Szenario7 fallen diese Transaktionen nur lokal an, da Bewerberdaten in den anderen Szenarien nur lokal gespeichert werden.

In Szenario7 werden die Personaldaten auch im Pendelbereich gespeichert, so daß bei der stellenorientierten Bewerbervermittlung nur noch lokal gesucht werden muß. Bei der Übertragung der Daten wird davon ausgegangen, daß die Personal- und Berufsdaten zusammen gesendet werden. Für diese Übertragung werden im Durchschnitt

transErfBP Transaktionen benötigt.

7.3.1.3 Suche nach geeigneten Bewerbern

Wie bei der allgemeinen Beschreibung der Suchstrategie erläutert, wird zunächst lokal gesucht und falls nicht genügend Stellen gefunden werden, wird im Pendelbereich oder zentral gesucht

transSulokV sei die durchschnittliche Anzahl der lokalen Suchtransaktionen, einschließlich Probeversuche, die mit dem Bewerber abgesprochen werden.

transSuremV ist die Anzahl der Suchtransaktionen in die Ferne. Eine solche Suchtransaktion ist lokal eine verteilte Transaktion erzeugt aber in jedem Rechner, der die Suchfrage erhält, jeweils eine Transaktion.
 o^* sei die Maximalzahl der angebotenen Stellen pro Bewerber (o^* muß nicht gleich o sein)

s1: prozentualer Anteil der Suchfälle, in denen nur lokal gesucht wird (weil z. B. schon o^* erreicht wurde)

s2: prozentualer Anteil der Suchfälle in denen zusätzlich im Pendelbereich gesucht wird

s3: prozentualer Anteil mit zusätzlicher zentraler Suche (bestimmte Berufsgruppen)

$$s1 + s2 + s3 = 100\%$$

s1, s2, s3 wurde bewußt neben **r1, r2** eingeführt, da z. B. **s1** nicht gleich **r1** sein muß

p sei die durchschnittliche Anzahl der Arbeitsämter des bewerberorientierten Pendelbereichs.

Wie schon bei der Beschreibung des Pendelbereichs angeführt muß **p** nicht gleich **q** sein

X ist im Szenario3 der prozentuale Anteil der Stellen, die nur lokal geführt werden.

Man muß bei der Betrachtung der Suche im Pendelbereich immer berücksichtigen, daß ein lokaler Rechner nicht nur Aufträge absendet, sondern auch Aufträge von den Partnerrechnern erhält.

Die Formeln für das Suchen für die einzelnen Verfahren ist in den ABB.15-19 dargestellt.

7.3.1.4 Update der Bewerberdaten

Die von den Sachbearbeitern akzeptierten Stellenvorschläge werden bei den Personaldaten gespeichert, um

- bei Folgegesprächen mit dem Bewerber über die Bewerbungen sprechen zu können
- bei FolgeSuchvorgängen nicht wieder die gleichen Stellen anzubieten

Dieser Speichervorgang enthält unter anderem auch die Speicherung der Suchfrage für den Bewerber nach geeigneten Stellen.

transUpPers sei die Anzahl der Transaktionen zur Speicherung der Stellendaten.

Diese Speicherungen finden nur lokal statt. In Szenario7 werden nur die Personaldaten zusammen mit den Berufsinformationen im Pendelbereich verteilt,

nicht aber die Informationen über die angebotenen Stellen. Damit ergeben sich in dem Szenario keine zusätzlichen Transaktionen beim Update der Bewerberdaten.

7.3.1.5 Update der Stellendaten

K sei die durchschnittliche Anzahl der gefundenen Stellen

Es wird davon ausgegangen, daß außer bei Szenario3, wo für einen bestimmten Teil von Berufsgruppen nur lokal gesucht wird, diese K für alle Szenarien identisch ist. (Der Pendelbereich zwischen den LAA bei Szenario6 wird vernachlässigt). Bei Szenario3 wird eine Zahl KR angenommen.

K1 sei der prozentuale Anteil lokal gefundener Stellen (KRI bei Szenario3)

K2 sei der prozentuale Anteil der im Pendelbereich gefundenen Stellen

K3 ist der Anteil zentral gefundener Stellen

In ABB. 20 ist an einem Beispiel der Unterschied zwischen Szenario2 und Szenario3 beim Update der Stellendaten dargestellt. In diesem Beispiel wird auch der Unterschied zwischen K und KR bzw. K1 und KRI deutlich.

Bei der Berechnung der Transaktionen muß folgendes berücksichtigt werden:

- für die lokal geführten Stellen wird der Update nur auf den Rechner durchgeführt, die Owner der Stelle sind. Dies gilt auch für das Szenario5, wo zwar die Stellenbeschreibungen in den Pendelbereich verschickt werden, aber nicht die Besetzungsdaten.
- für die Stellen, die zusätzlich zentral oder auch in den LAA gespeichert sind werden die Daten ebenfalls auf diesen Rechnern gespeichert. Der Update geschieht dann so, daß zunächst die zentrale Speicherung erfolgt und danach auf dem Owner-Rechner der Stelle, denn nur von dort aus besteht eine Verbindung zum Owner-Rechner. Dies geschieht alles in einer Transaktion. Analoges gilt für die LAA-Rechner von Szenario6. Wie der Stellenupdate für Szenario6 im einzelnen geschieht ist in ABB.21 dargestellt.

Die Formeln für die einzelnen Verfahren sind in den ABB. 15 bis 19 dargestellt.

7.3.1.6 Beispielrechnung für Erstkontakte

In den Abbildungen 22 und 23 sind Beispielrechnungen durchgeführt, wobei ABB.22 die Annahmen beschreibt und ABB.23 die Ergebnisse für die 7 Szenarien. Eine Diskussion der Ergebnisse findet in Kapitel 7,5 statt.

7.3.2 Folgeberatungen

Bei den Folgeberatungen wird davon ausgegangen, daß alle Arbeitssuchende (Anzahl: AS, zu betreuende Arbeitssuchende eines mittleren Arbeitsamtes: mAS), die registriert sind, gleichmäßig durch Folgeberatungen betreut werden.

Wie schon in Kapitel 5 erwähnt, können Folgeberatungen folgendermaßen zustande kommen:

- es wird periodisch (z.B. monatlich) die gespeicherte Suchfrage asynchron durchgeführt und nach Stellen gesucht, die dem Arbeitssuchenden noch nicht angeboten wurden. Das Ergebnis wird in die Team-Datei gespeichert. Der Sachbearbeiter entscheidet dann, welche Stellen angeboten werden sollen. Aus der Sicht des Sachbearbeiters werden die gleichen Informationen zur Verfügung gestellt wie bei der Suche, die von den Stellen ausgeht. Ein solcher Suchlauf ist eine Transaktion, in der
 1. die Daten des Arbeitssuchenden gelesen werden,
 2. die gespeicherte Suchfrage durchgeführt wird
 3. die Ergebnismenge mit den schon angebotenen Stellen verglichen wird
 4. die noch nicht angebotenen Stellen in die Team-Datei geschrieben werden.
- **f1**: Anzahl dieser Suchläufe pro Jahr
- **KN1**: durchschnittliche Anzahl der gefundenen neuen Stellen
- der Sachbearbeiter legt Termine für Folgeberatungen fest. Dies kann z. B. dadurch geschehen, daß er bei einer Beratung den Folgetermin schon bestimmt. Es besteht dann die Möglichkeit, automatisch einen asynchronen Suchlauf vorher ablaufen zu lassen und das Ergebnis in der Sachbearbeiterdatei zu speichern. Diese Informationen liegen dann beim Besuch des Arbeitssuchenden schon vor.
- **f2**: Anzahl solcher Folgetermine pro Jahr **KN2**: durchschnittliche Anzahl gefundener Stellen (**KN2** muß nicht unbedingt identisch mit **KN1** sein, da die Zeitabstände unterschiedlich sein können)
- Arbeitssuchende kommen von sich aus (evtl. unangemeldet) und hätten gerne eine Folgeberatung. Auch in diesem Fall kann bei

Mathematisches Modell für die verteilte Verarbeitung im Bereich Vermittlung

einer Empfangsstelle zunächst ein Asynchronlauf gestartet werden, so daß der Sachbearbeiter schon ein Ergebnis einer Suchfrage vorliegen hat.

- f_3 : Anzahl solcher Besuche pro Jahr
- KN_3 : Anzahl der gefundenen Stellen

Es gibt nun 2 unterschiedliche Situationen

1. Folgeberatungen, bei denen der Arbeitssuchende im Arbeitsamt erschienen ist (Fall 2 und 3). In diesen Fällen wird mit dem Arbeitssuchenden zusammen beraten, welche der asynchron ermittelten Stellen in Frage kommen. I_2 ist der durchschnittliche Prozentsatz der von den Arbeitssuchenden akzeptierten Angebote.
2. Fälle, in denen durch periodisches Suchen oder ausgehend von den Stellen Informationen für Arbeitssuchende vorliegen. In diesen Fällen muß der Sachbearbeiter

- entweder selbst entscheiden, welche Stellen in Frage kommen und den Arbeitssuchenden telefonisch oder schriftlich benachrichtigen (prozentualer Anteil = I_1)
- oder er muß den Arbeitssuchenden zur Konsultation einbestellen

Daraus folgt

$f = f_1 + f_2 + f_3$ ist die durchschnittliche Anzahl der Folgeberatungen pro Bewerber pro Jahr, wenn der Bewerber das ganze Jahr arbeitslos ist.

$$KNT = (KN_1 * f_1 + KN_2 * f_2 + KN_3 + f_3) / f$$

KNT ist die durchschnittliche Anzahl der gefundenen Stellen pro asynchronem Suchlauf. Analog zu der Erstberatung gibt es für Szenario 3 eine durchschnittliche reduzierte Anzahl $KRNT$.

$$I_k = (I_1 * KN_1 + I_2 * (KN_2 + KN_3)) / (KN_1 + KN_2 + KN_3)$$

I_k ist der durchschnittliche prozentuale Anteil der akzeptierten Stellenangebote

I_b sei der prozentuale Anteil der Bewerber, denen Stellen angeboten werden. (I_b muß nicht gleich I_k sein.)

Mathematisches Modell für die verteilte Verarbeitung im Bereich Vermittlung

$FS = AS * f/AT$ ist die Gesamtzahl der Folgeberatungen, die täglich durchgeführt wird.

$FmS = FS/RAA$ ist die durchschnittliche Anzahl, die in einem mittleren Arbeitsamt anfällt.

Der Ablauf ist in allen 3 Fällen folgender (die sich ergebenden Formeln für die einzelnen Szenarien ist in den Abb. 24 bis 29 dargestellt):

- asynchrones Suchen. Es wird die in den Daten des Bewerbers gespeicherte Suchfrage asynchron durchgeführt, wobei auch hier eine maximale Anzahl der Stellen gesucht wird. Diese Suche geschieht lokal und je nach Szenario auch zentral oder im Pendelbereich.

transSufol sei die Anzahl der Transaktionen für das lokale Suchen, diese Zahl wird normalerweise identisch mit transSuremV sein.

- Schreiben in die Team-Datei: Die Ergebnisse werden zur endgültigen Entscheidung in die schon bei der stellenorientierten Bewerbervermittlung aufgeführten Datei geschrieben. Der Unterschied zur stellenorientierten Bewerbervermittlung besteht jedoch darin, daß ein Bewerbereintrag mehrere Stellen enthalten kann.
- Lesen aus der Team-Datei: Die Sachbearbeiter des Teams lesen die Daten der Team-Datei und
- Lesen der Bewerberdaten: Lesen aus der Datenbank die vollständigen Informationen des aufgeführten Bewerbers. Danach
- Lesen der Stelldaten: Holen sie sich die Informationen zu den einzelnen Stellen. Dies geschieht durch eine lokale Transaktion wenn die Stelle im gleichen Arbeitsamt verwaltet wird oder über eine verteilte Transaktion wenn die Informationen zu der Stelle auf einem anderen Rechner liegen.

Nachdem der Sachbearbeiter alle Informationen zur Verfügung hat, entscheidet er entweder alleine oder in Anwesenheit des Bewerbers, ob die Stelle in Frage kommt oder nicht. Für die Stellen, die akzeptiert werden, müssen

- ein Update der Bewerberdaten durchgeführt werden, dies geschieht immer lokal und

- die Stellendaten um die Bewerberinformationen erweitert werden. Dies geschieht
 - entweder lokal (für lokale Stellen)
 - oder auf einem Rechner im Pendelbereich
 - oder zentral für zentral geführte Stellen (wobei in der Zentrale wiederum ein Update beim Owner der Stelle initiiert wird)
 - oder im LAA-Rechner bei Szenario6 (s. h. Abb.21)

Abb. 30 und 31 zeigen Ergebnisse für die einzelnen Schritte an einem konkreten Beispiel für die 7 Szenarien. Dabei sind in Abb. 30 die Annahmen aufgeführt und in Abb. 31 die Ergebnisse. Eine Diskussion erfolgt nach der Zusammenfassung aller Ergebnisse für die coArb-Aktivitäten.

7.3.3 Löschen von Bewerberdaten

Bewerber werden aus dem Bewerbungsprozess herausgenommen, wenn sie beispielsweise

- eine neue Stelle gefunden haben
- Rente beziehen etc.

In diesen Fällen werden die gespeicherten Daten mit einem entsprechenden Kennzeichen versehen. Ein echtes Löschen der Daten von Bewerbern findet nur in Szenario7 statt und zwar auf den Rechnern im Pendelbereich, denn dort werden die Daten nicht mehr benötigt. Dadurch daß die Daten nicht gelöscht sondern nur gekennzeichnet werden können

- Statistiken durchgeführt werden, oder
- die Daten wieder verwendet werden, wenn ein Bewerber wieder arbeitslos wird.

transLöB sei die Anzahl der Transaktionen zur Kennzeichnung der Daten des Bewerbers. Die Formeln zum Löschen ergeben sich dann aus denen der Datenerfassung in den ABB. 15-19, indem man transErfB durch transLöB ersetzt.

7.5 Bewertung der Ergebnisse der einzelnen Szenarien

7.5.1 Aufteilung der Transaktionen auf Tageszeiten

Normalerweise ist es üblich, Performance-Aussagen in TA/sec zu machen. Die Aktivitäten bei der Arbeitsvermittlung sind nicht gleichmäßig über den gesamten Tagesverlauf verteilt. Daher ist es zunächst notwendig, die Aktivitäten auf folgende Tagesbereiche zu verteilen:

1. Aktivitäten, die während der Zeit des Publikumsverkehrs anfallen (in der folgenden Beispielrechnung werden hierfür 4 Std. pro Tag angenommen)
2. Aktivitäten der Sachbearbeiter in der Arbeitszeit, aber nicht notwendigerweise während der Publikumsverkehrszeit (Annahme: 4 Std. pro Tag). Sollte ein Sachbearbeiter in der ersten Zeitspanne nicht ausgelastet sein so können Teile dieser Aktivitäten in die erste Phase übertragen werden.
3. Aktivitäten außerhalb der Arbeitszeit eines Sachbearbeiters. Dies sind die Asynchronaufträge, die ein Rechner ohne Beteiligung des Sachbearbeiters durchführen kann.

Bezüglich der Aktivitäten werden folgende Annahmen getroffen:

Stellenorientierte Bewerbervermittlung

- Die Aktivitäten Speichern und Löschen von Stellen, Erstsuche nach geeigneten Bewerbern und Speichern in die Team-Datei werden von bestimmten ausgewählten Sachbearbeitern durchgeführt. Diese Aufgaben werden gleichmäßig auf die gesamte Arbeitszeit (8 Std., Zeitspanne 1 und 2) verteilt.
- Die Aktivitäten zur Entscheidung der Sachbearbeiter anhand der Team-Datei werden in die zweite Zeitspanne gelegt, da die Entscheidungen ohne die Gegenwart der Arbeitssuchenden stattfinden können
- Die Folgeaktivitäten beinhalten

einerseits wiederholte Suchvorgänge
andererseits die damit verbundenen Entscheidungen der Sachbearbeiter der Arbeitssuchenden

7.4 Zusammenfassung

Die Basis für ein mathematisches Gesamtmodell der die Aktivitäten im Bereich Arbeitsvermittlung erhält man dadurch, daß man folgende Einzelkomponenten zusammenfaßt:

- die stellenorientierte Bewerbervermittlung
- die Erstkontakte für die bewerberorientierte Stellenvermittlung
- die Folgeberatungen für die bewerberorientierte Stellenvermittlung
- das Löschen von Bewerberdaten

In ABB.32 sind die Ergebnisse der einzelnen Beispielrechnungen zusammengefaßt.

An diesen Berechnungen sieht man zunächst, daß die Transaktionsraten überwiegend durch die hohe Anzahl von Folgeberatungen bestimmt werden. Dies liegt u. a. an der Annahme, daß pro Jahr und Bewerber 20 Folgeberatungen (12 ohne Besuch des Bewerbers, 8 mit Besuch) durchgeführt werden.

Das mathematische Modell macht es nun möglich, diese Annahmen zu verändern und zu untersuchen, wie Änderungen von Annahmen sich auf die Ergebnisse auswirken. Es wird daher empfohlen, die entwickelten Formeln in Excel zu speichern und verschiedene Annahmen durchzuspielen.

Weiterhin ist es notwendig, die Vor- und Nachteile der einzelnen Szenarien genauer zu beleuchten und zu bewerten.

Die Suchvorgänge werden in die Zeitspanne 3 gelegt, da dieses Suchen in vielen Fällen asynchron erfolgen kann, die Entscheidungsprozesse fallen wieder in die Zeitspanne 2.

Bewerberorientierte Stellenvermittlung

- Die Aktivitäten für den Erstkontakt fallen in die Zeitspanne mit Publikumsverkehr
 - Die Folgeaktivitäten werden wie folgt aufgeteilt (s. h. 7.3.2)
1. unangemeldete Besuche (Faktor f_3) müssen während der Besuchszeit (Zeitspanne 1) abgewickelt werden
 2. die Suchvorgänge für geplante Besuche (Faktor f_2) und das periodische Suchen (Faktor f_1) können asynchron in der Zeitspanne 3 durchgeführt werden
 3. die daraus resultierenden Aktivitäten der Sachbearbeiter der Bewerber fallen wieder in die Zeitspanne 2

Unter diesen Annahmen ist es möglich, Aussagen über durchschnittliche Transaktionen in den einzelnen Zeitspannen und für die einzelnen Szenarien zu machen.

Die anfallenden Transaktionsraten für die bisherigen Beispielrechnungen sind in den ABB. 33-39 für die einzelnen Szenarien dargestellt. Eine Bewertung der Ergebnisse findet im Abschnitt 7.5.3 statt.

7.5.2 Netzbelastung

Neben den Transaktionsraten ist die Netzbelastung ein wichtiges Kriterium zur Beurteilung der verschiedenen Szenarien. Die Netzbelastung wird bestimmt

- durch die Nachrichtenlänge der zu übertragenden Nachrichten

$mNRL1$: mittlere Nachrichtenlänge eines Auftrages

$mNRL2$: mittlere Nachrichtenlänge des Ergebnisses eines Auftrages

$mNRLg = mNRL1 + mNRL2$

In den folgenden Beispielrechnungen wird angenommen, daß

$mNRL1 = 0.5$ KB und $mNRL2 = 1$ KB ist, d. h. für einen Auftrag müssen im Durchschnitt 1.5 KB über das Netz transportiert werden.

- durch die Entfernung, in die eine Nachricht geschickt werden muß

In den einzelnen Szenarien spielen folgende Entfernungen eine Rolle, für die mittlere Entfernungen angenommen werden

$mEntfAZ$: mittlere Entfernung eines Arbeitsamtsrechners zum Zentralrechner (Annahme: 500km)

im $mEntfAA$: mittlere Entfernung zwischen den Arbeitsamtsrechnern Pendelbereich(Annahme: 50 km)

$mEntfAL$: mittlere Entfernung eines Arbeitsamtsrechners mit dem zugehörigen Landesamtsrechner (Annahme: 300 km)

zum $mEntfLZ$: mittlere Entfernung eines Landesamtsamtsrechners Zentralrechner (Annahme: 300 km)

Als Maß für die Netzbelastung werden zwei Größen eingeführt

- die Anzahl der Nachrichten zwischen den einzelnen Partnern
- Nachrichtenlänge * Entfernung in der Maßeinheit Gigabyte*km

In den ABB. 33-39 sind neben den Transaktionsraten auch die Ergebnisse der Netzbelastung für die einzelnen Zeitspannen dargestellt.

7.5.3 Bewertung der einzelnen Szenarien

Für die Bewertung der einzelnen Szenarien sollen folgende Kriterien herangezogen werden:

- Auswahlmenge
- bzgl. Bewerbern bei der stellenorientierten Bewerbervermittlung
bzgl. Stellen bei der bewerberorientierten Stellenvermittlung
- Einschränkungen bzgl. der Auswahlmenge bei Netzausfall

Dabei wird davon ausgegangen, daß nur dann eine Einschränkung vorliegt, wenn Informationen direkt benötigt werden. Solche Fälle sind:

1. Suche nach Stellen beim Erstkontakt oder bei unangemeldetem Besuch eines Arbeitssuchenden
2. Bei Szenario1, wenn ein Sachbearbeiter nicht arbeiten kann

Die meisten Aktivitäten, die fremde Rechner betreffen, müssen nicht unbedingt online durchgeführt werden. Es reicht, wenn Aufträge asynchron hinterlegt werden, die bei Verbindungsaufbau im Hintergrund bearbeitet werden. Dies setzt natürlich voraus, daß für die verteilte Verarbeitung Verfahren und Produkte eingesetzt werden, die diese Vorgehensweise unterstützen.

- Transaktionsraten auf den einzelnen Rechnern in den verschiedenen Zeitspannen
- Netzbelastung zwischen den Rechnern in den einzelnen Zeitphasen

In ABB.40 sind die Ergebnisse der Bewertung tabellarisch dargestellt. Die Ergebnisse beruhen auf folgenden Überlegungen:

Vollständigkeit bzgl. der Suche

Stellensuche

- In Szenario1 und 2 stehen alle Stellen bei der Suche zur Verfügung
- In Szenario3 werden ein Teil der Stellen nur lokal gespeichert und stehen daher für eine lokale Suche nicht zur Verfügung

- In Szenario 4,5 und 7 werden alle Stellen im Pendelbereich und die zentral verwalteten Stellen in die Suche eingeschlossen. Es fehlen nur die Stellen außerhalb des Pendelbereichs.
- In Szenario6 stehen alle Stellen innerhalb eines LAA-Bereichs und die zentral verwalteten Stellen zur Verfügung. Es fehlen also vor allem die Stellen in den Grenzbereichen der Arbeitsämter, die von einem LAA verwaltet werden.

Bewerbersuche

- In Szenario1 stehen wieder alle Informationen zur Suche nach Arbeitssuchenden zur Verfügung
- In den Szenarien 2,3 und 6 kann bei der stellenorientierten Bewerbervermittlung nur auf die lokal gespeicherten Bewerberdaten zugegriffen werden.
- Für die Szenarien 4,5 und 7 gelten die gleichen Aussagen wie für die Stellen, jedoch mit der Ausnahme der zentralen Suche, da Bewerberdaten nicht zentral gespeichert werden.

Vollständigkeit der Suche nach Verbindungsverlust

- In Szenario1 gibt es nach einem Verbindungsverlust zum Zentralrechner keine Möglichkeit, Vermittlungs-Aktivitäten mit einem Rechner durchzuführen
- In den Szenarien 2,3,4 und 6 kann in dieser Situation nur auf lokale Daten zugegriffen werden. Replikate von Daten anderer Arbeitsämter existieren nicht.
- In Szenario5 kann neben dem Zugriff auf die lokalen Informationen auch nach Stellen im Pendelbereich gesucht werden, da sie als Replikate gespeichert sind.
- In Szenario7 ist eine fast vollständige Suche möglich, es fehlt nur der Zugriff auf die zentral verwalteten Daten.

Transaktionsraten

- In Szenario1 müssen alle Transaktionen zentral abgewickelt werden. Wie das Beispiel zeigt, kann man dabei leicht in Bereiche gelangen, die nur von einem Hochleistungszentralsystem bewältigt werden können (>100 TA/sec)
- In den Szenarien 2-7 liegen die Transaktionsraten auf den lokalen Rechnern um oder z.T. deutlich unter 1 TA/sec. Diese Transaktionsraten sind mit den heutigen Rechnern leicht zu bewältigen.

- Der Zentralrechner wird neben Szenario1 nur noch von den Szenarien 2 und 3 mit mehr als 10 TA/sec belastet. Dies kann von mittleren bis großen UNIX - Rechnern bearbeitet werden.
- In Szenario6 wird der Rechner des LAA sehr viel stärker belastet als die einzelnen Arbeitsamtsrechner. Die Rechner müßten daher entsprechend ausgebaut werden.

Netzbelastung

- Da in Szenario1 jeder Auftrag zentral bearbeitet werden muß , ergibt sich naturgemäß die höchste Netzbelastung, die weit über der der anderen Szenarien liegt.
- In den Szenarien 2 und 3 wird die Netzbelastung stufenweise reduziert, da ein Teil der Aufträge nur noch lokal abgehandelt wird.
- Szenario6 zeigt eine starke Netzbelastung zwischen den Arbeitsamtsrechnern und den LAA-Rechnern. Dies liegt daran, daß ein großer Teil der Aufträge über die LAA-Rechner abgewickelt werden.
- Ein Vergleich der Netzbelastung der Szenarien 4,5 und 7 zeigt starke Unterschiede bei der Kommunikation im Pendelbereich zwischen Szenario4 einerseits und den Szenarien5 und 7 andererseits. Dies liegt vor allem an der Tatsache, daß in Szenario4 die Suche nach Stellen in vielen Fällen über das Netz im Pendelbereich geschieht. Der Unterschied zwischen Szenario5 und 7 ist deswegen nicht so groß, weil in der Beispielrechnung sehr viel mehr Aktivitäten für die bewerberorientierte Stellenvermittlung angenommen wurde als für stellenorientierte Bewerbervermittlung.

7.5.4 Empfehlung

Aufgrund der gewonnenen Ergebnisse des mathematischen Modells können die Szenarien 5 oder 7 empfohlen werden. Sie

- ermöglichen eine fast vollständige Suche nach Arbeitssuchenden oder Stellen
 - erzeugen niedrige Transaktionsraten
 - zeigen die geringste Netzbelastung
- ermöglichen ein Arbeiten der Sachbearbeiter auch bei Netzausfall, wobei Szenario7 gegenüber Szenario5 den Vorteil hat, daß auch die stellenorientierte Bewerbervermittlung ohne große Einschränkung durchgeführt werden kann.

8 Glossar

Alphabetische Beschreibung der Variablen

- arbD : Gesamtanzahl der Sachbearbeiter aller Arbeitsämter im Bereich Arbeitsvermittlung
- AS : Gesamtanzahl der bundesweit registrierten Arbeitssuchenden
- AS+ : Gesamtanzahl von Arbeitssuchenden, die bundesweit jährlich neu registriert werden
- AT : Anzahl der Tage eines Jahres in denen die Arbeitsämter für den Publikumsverkehr geöffnet sind
- ATS : durchschnittliche Anzahl der jährlichen Arbeitstage eines coArb-Sachbearbeiters
- f : bei bewerberorientierter Stellenvermittlung: Gesamtanzahl der Folgeberatungen pro Jahr für einen Arbeitssuchenden
- f1 : bei bewerberorientierter Stellenvermittlung: jährliche Anzahl asynchroner Suchläufe zum routinemäßigen Suchen nach geeigneten Stellen.
- f2 : bei bewerberorientierter Stellenvermittlung: jährliche Anzahl der geplanten Folgetermine eines Arbeitssuchenden
- f3 : bei bewerberorientierter Stellenvermittlung durchschnittliche Anzahl der ungeplanten Beratungsbesuche eines Arbeitssuchenden
- fa : bei stellenorientierter Bewerbervermittlung: globaler, wählbarer Faktor zur Berücksichtigung der Folgeaktivitäten zur Besetzung von Stellen
- FS : Gesamtanzahl der täglich durchschnittlich bundesweit durchgeführten Folgeberatungen
- Fms : Anzahl der täglich in einem mittleren Arbeitsamt durchgeführten Folgeberatungen

- K : bei bewerberorientierter Stellenvermittlung: durchschnittliche Anzahl gefundener Stellen pro Bewerber
- K1 : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil lokal gefundener Stellen an der Gesamtanzahl K
- K2 : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil der im Pendelbereich gefundenen Stellen
- KN1 : bei bewerberorientierter Stellenvermittlung durchschnittliche Anzahl der gefundenen Stellen bei der asynchronen routinemäßigen Suche
- KN2 : bei bewerberorientierter Stellenvermittlung: durchschnittliche Anzahl gefundener Stellen beim Suchlauf vor dem geplanten Bewerberbesuch
- KN3 : bei bewerberorientierter Stellenvermittlung: durchschnittliche Anzahl gefundener Stellen beim Suchlauf vor dem ungeplanten Bewerbergespräch
- K3 : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil zentral gefundener Stellen
- KR : bei bewerberorientierter Stellenvermittlung für Szenario3: durchschnittliche Anzahl gefundener Stellen
- KR1 : bei bewerberorientierter Stellenvermittlung für Szenario3: prozentualer Anteil lokal gefundener Stellen
- I : bei stellenorientierter Bewerbervermittlung: prozentualer Anteil der von den Sachbearbeitern akzeptierten Vorschläge der gefundenen Bewerber
- I1 : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil der von den Sachbearbeitern akzeptierten Stellenvorschläge für den Fall, daß der Bewerber nicht anwesend ist
- I2 : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil der von den Sachbearbeitern akzeptierten Stellenvorschläge für den Fall, daß der Bewerber anwesend ist

- L** : durchschnittliche Anzahl der gefundenen Bewerber pro Stelle, wobei die Suche nach Erreichen der Obergrenze abgebrochen wird
- L1** : bei stellenorientierter Bewerbervermittlung: prozentualer Anteil der lokal gefundenen Bewerber
- L2** : bei stellenorientierter Bewerbervermittlung: prozentualer Anteil gefundener Bewerber aus anderen Arbeitsamtsbereichen
- lb** : bei stellenorientierter Bewerbervermittlung: prozentualer Anteil der Bewerber denen bei einer Folgeberatung eine Stelle angeboten werden kann
- LAA** : Anzahl der Landesarbeitsämter
- lk** : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil der von den Sachbearbeitern akzeptierten Stellenvorschläge unabhängig davon, ob der Bewerber anwesend ist oder nicht
- marbD** : Anzahl der coArb-Sachbearbeiter eines mittleren Arbeitsamtes
- mN** : Anzahl der Stellen, die jährlich neu in einem mittleren Arbeitsamt besetzt werden sollen
- N** : Anzahl der Stellen, die bundesweit jährlich neu besetzt werden sollen
- O** : maximale Anzahl von Bewerbem pro Stelle
- o*** : maximale Anzahl der Stellen pro Bewerber
- p** : durchschnittliche Anzahl der Arbeitsämter im Pendelbereich von Bewerbem
- q** : durchschnittliche Anzahl der Arbeitsämter im Pendelbereich von Stellen
- r1** : bei stellenorientierter Bewerbervermittlung: prozentualer Anteil der Fälle, in denen schon lokal genügend Bewerber gefunden werden

- r2** : bei stellenorientierter Bewerbervermittlung: prozentualer Anteil der Fälle, in denen auch in anderen Rechnern nach Bewerbem gesucht werden muß
- RAA** : Anzahl der Arbeitsämter mit einem eigenständigen UNIX-Rechner auf dem ein INFORMIX-Online-System zur Verfügung steht
- s1** : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil der Fälle, in denen nur lokal gesucht wird
- s2** : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil der Fälle, in denen auch im Pendelbereich gesucht wird
- s3** : bei bewerberorientierter Stellenvermittlung: prozentualer Anteil der Fälle, in denen für besondere Berufsgruppen auch zentral gesucht wird
- Teber** : durchschnittliche Anzahl der täglichen Erstkontakte eines Sachbearbeiters
- TmAS+** : Anzahl Arbeitssuchender, die in einem mittleren Arbeitsamt täglich neu registriert werden
- TmN** : durchschnittliche Anzahl neuer Stellen, die täglich in einem mittleren Arbeitsamt neu erfaßt werden
- TN** : durchschnittliche Anzahl neuer Stellen, die täglich bundesweit neu erfaßt werden
- transErfB** : durchschnittliche Anzahl der Transaktionen zur Erfassung eines neuen Arbeitssuchenden
- transLöN** : durchschnittliche Anzahl der Transaktionen zum Löschen einer Stelle
- transLöB** : durchschnittliche Anzahl der Transaktionen zum Löschen der Daten eines Bewerbers
- TransSachmB**: bei stellenorientierter Bewerbervermittlung: Gesamtanzahl der täglich in einem mittleren Arbeitsamt von den Sachbearbeitern initiierten Transaktionen zur Bearbeitung der Vorschläge der Team-Datei

- transSpBer** : durchschnittliche Anzahl der Transaktionen zur Speicherung der Berufsdaten eines neuen Arbeitssuchenden
- TransSpmN**: Gesamtanzahl der täglich anfallenden Transaktionen in einem mittleren Arbeitsamt zur Speicherung von Stellen
- transSpN** : durchschnittliche Anzahl der Transaktionen zur Erfassung einer Stelle
 $= \text{transSuFi} + \text{transSpN}^*$
- transSpN*** : durchschnittliche Anzahl der Transaktionen zur reinen Speicherung von Stellendaten
- transSpPers** : durchschnittliche Anzahl der Transaktionen zur Speicherung der Personaldaten eines neuen Arbeitssuchenden
- transSuAS** : durchschnittliche Anzahl der Transaktionen zur Suche nach einem Bewerber in den Datenbeständen
- transSuFi** : durchschnittliche Anzahl der Transaktionen zum Lesen von Firmendaten
- transSulokB** : durchschnittliche Anzahl der Suchvorgänge pro neuer Stelle inklusive Probesuchläufe
- transSulokV** : bei bewerberorientierter Stellenvermittlung : Anzahl der lokalen Suchvorgänge zur Ermittlung von geeigneten Stellen für einen neuen Arbeitssuchenden (einschließlich Probeversuche)
- TransSulokmB** : bei stellenorientierter Bewerbervermittlung: Gesamtanzahl der täglichen Transaktionen in einem mittleren Arbeitsamt, die durch lokales Suchen nach Bewerbern entsteht
- TransSumB** : Gesamtanzahl der täglichen Transaktionen in einem mittleren Arbeitsamt zum Suchen nach geeigneten Bewerbern für eine Stelle
- TransSuremmB** : bei stellenorientierter Bewerbervermittlung: Gesamtanzahl der von einem mittleren Arbeitsamt initiierten Transaktionen auf Fremden Rechnern zur Suche nach geeigneten Bewerbern

- transSuremV** : bei bewerberorientierter Stellenvermittlung: Anzahl der Suchtransaktionen nach geeigneten Stellen Bewerber auf fremden Arbeitsamtsrechnern
- TransSuvremmB** : bei stellenorientierter Bewerbervermittlung: Gesamtanzahl der täglichen Transaktionen, die von fremden Rechnern zur Suche nach Bewerbern auf dem lokalen Rechner gestartet werden
- transSuremB** : bei stellenorientierte Bewerbervermittlung: durchschnittliche Anzahl der Suchvorgänge auf anderen Rechnern pro Stelle
- transUpPers** : bei bewerberorientierter Stellenvermittlung: durchschnittliche Anzahl der Transaktionen zur Erweiterung der Personaldaten um in Frage kommende Stellen
- TransWrmTd** : bei stellenorientierter Bewerbervermittlung: Gesamtanzahl der täglich anfallenden Transaktionen in einem mittleren Arbeitsamt zum Schreiben in die Team-Datei.
- TZEber** : durchschnittliche Zeit, die ein Sachbearbeiter täglich für Erstkontakte benötigt
- TZFber** : durchschnittliche Zeit eines Sachbearbeiters für Folgeberatungen während der Öffnungszeit des Arbeitsamtes
- Tzopen** : durchschnittliche tägliche Zeit, in der ein Arbeitsamt für Publikumsverkehr geöffnet ist
- X** lokal : prozentualer Anteil der Stellen, der in Szenario2 nur angeboten werden
- X*** : prozentualer Anteil der Stellen der besonderen Berufsgruppe, die zentral verwaltet werden (Szenario 3-7)
- Zeber** : durchschnittliche Zeit für eine Erstberatung eines Arbeitssuchenden

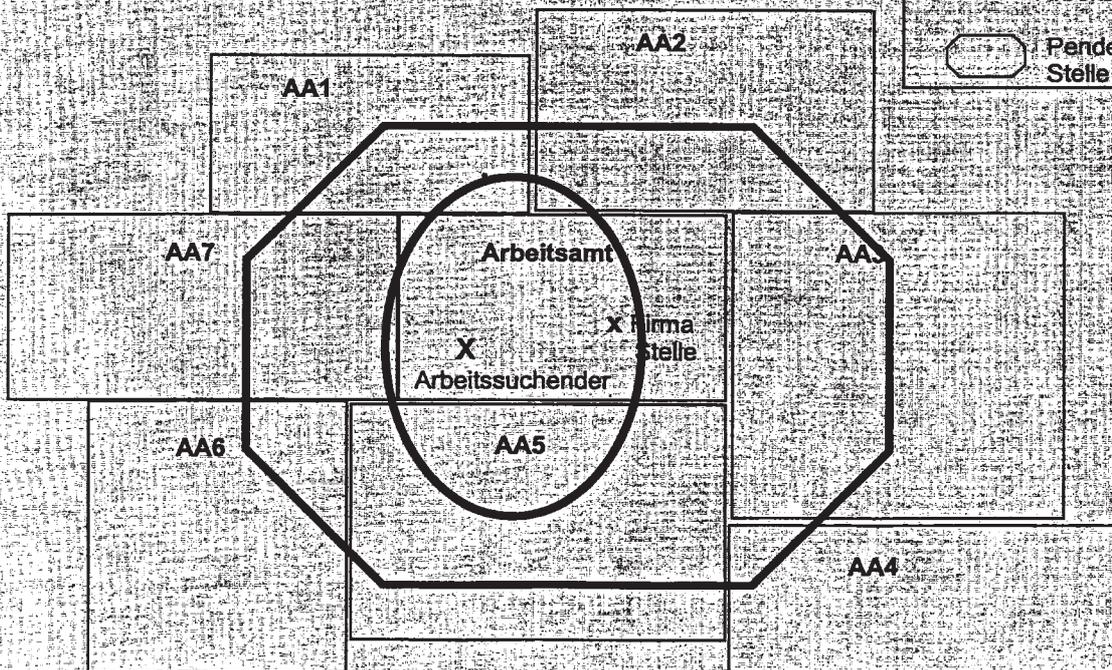
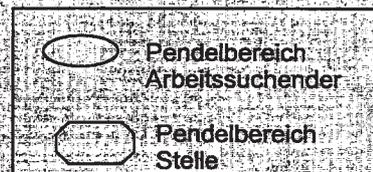
Studie zur verteilten Verarbeitung in Arbeitsämtern

Technische Einleitung

Arbeitsverwaltungen

SNI - 1/D

Pendelbereich



Beispiel der stellenorientierten Bewerbervermittlung im Pendelbereich

zu besetzende Stellen: 100
max. Anzahl Bewerber: 3

Ergebnis der lokalen Suche: 50%: 3 gefundene Bewerber = 150 Bewerber
20%: 2 " " " = 40 "
20%: 1 " " " = 20 "
10%: keine gef. Stelle
Zusammen: 210

in 50% aller Fälle muß im Pendelbereich gesucht werden
es werden aber 70% der gesuchten Bewerber lokal gefunden

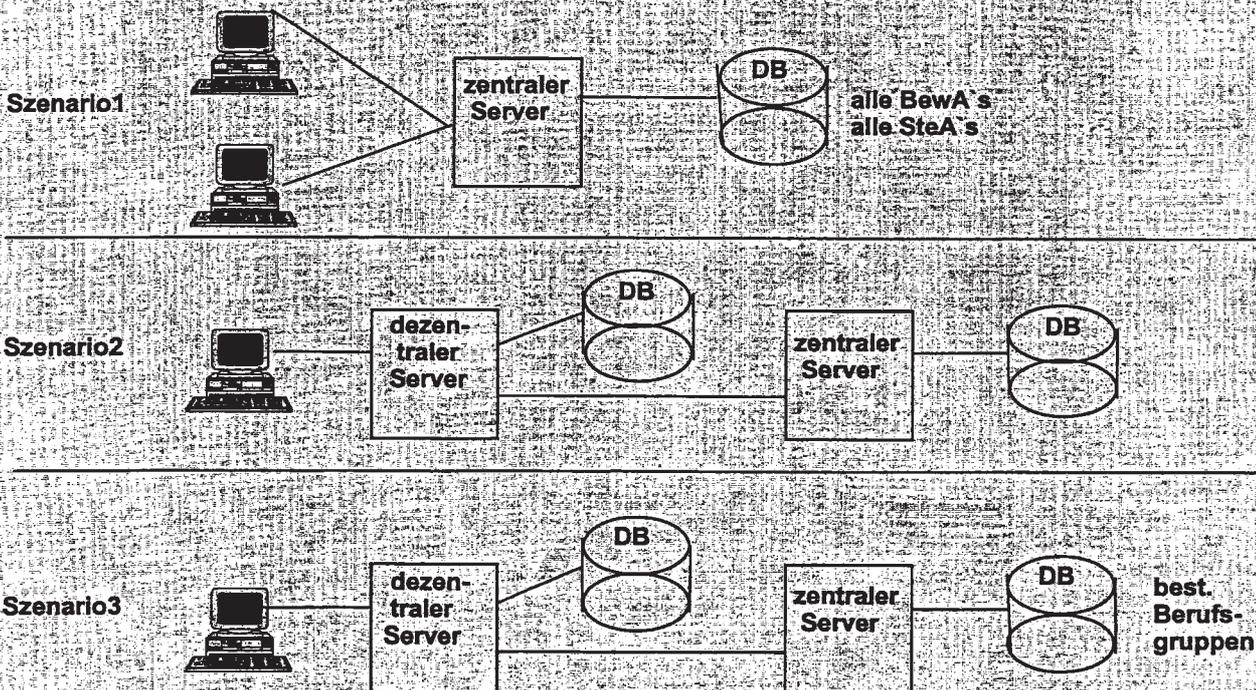
jedes AA erhält im Durchschnitt $p \cdot 50\%$ aller Suchfragen (p = Anzahl der AA im Pendelbereich)

jedes AA gibt im Durchschnitt so viel Bewerbervorschläge ab, wie es aus dem Pendelbereich als Vorschläge annimmt

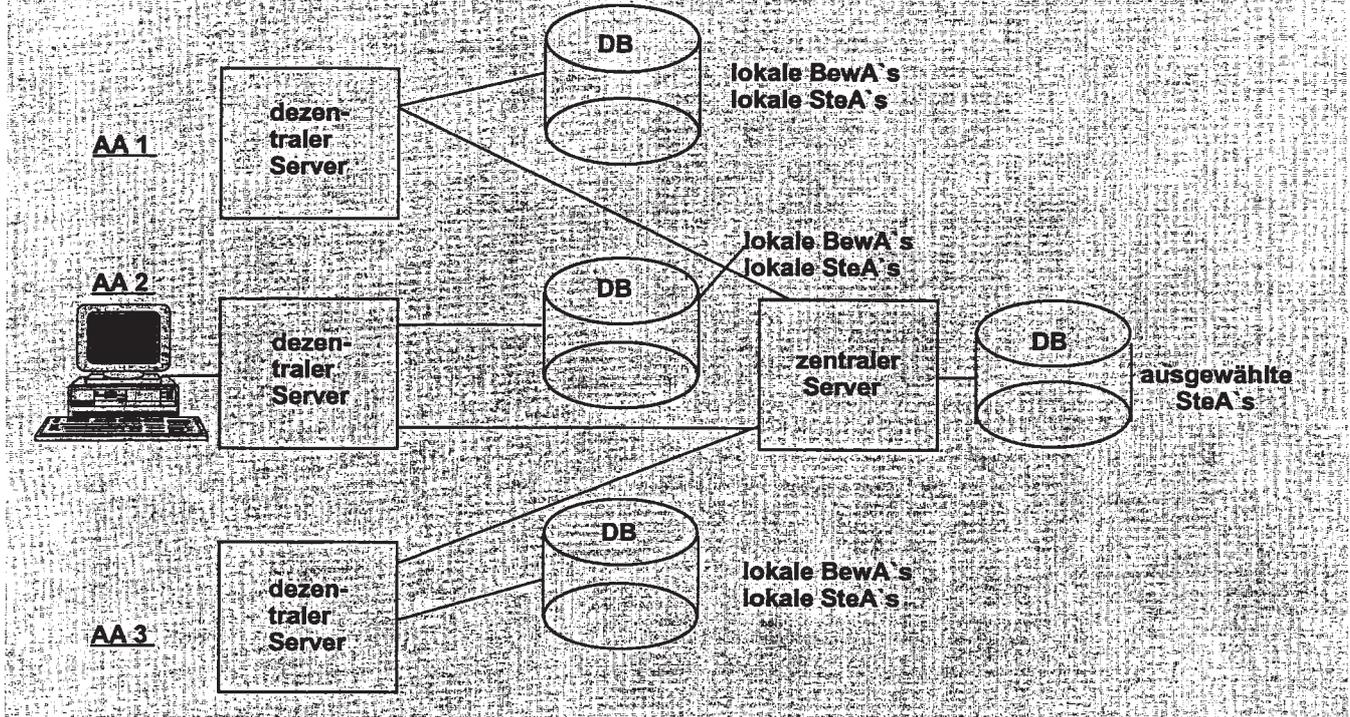


- AA1 sendet 50 Suchfragen je an AA2 u. AA3 und erhält auch je 50 Suchfragen
- AA1 erhält 90 Bewerbervorschläge von außen und gibt je 45 an AA2 u. AA3 ab

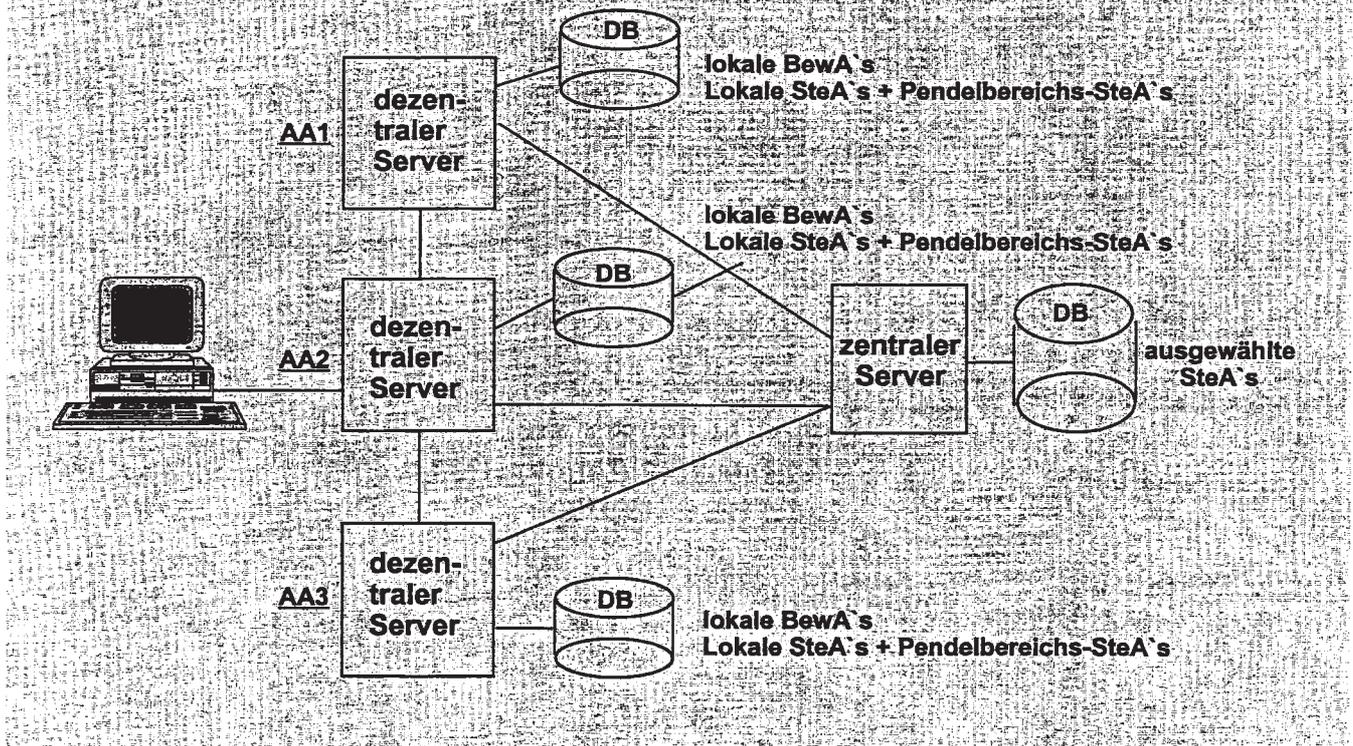
Szenarien



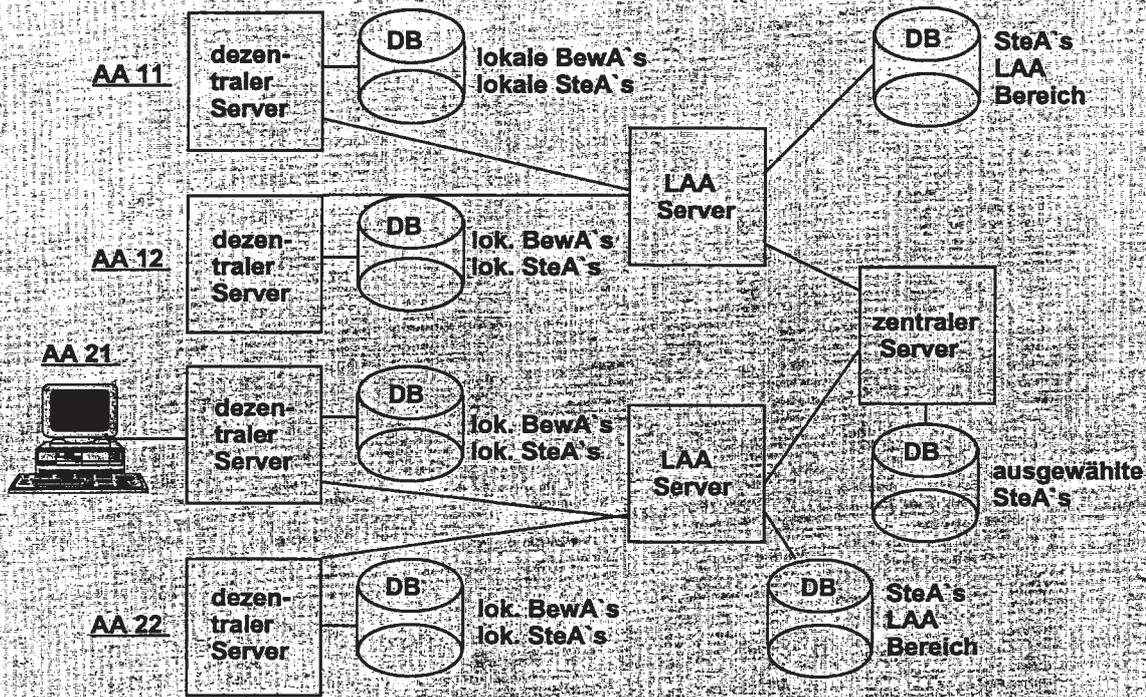
Szenario 4



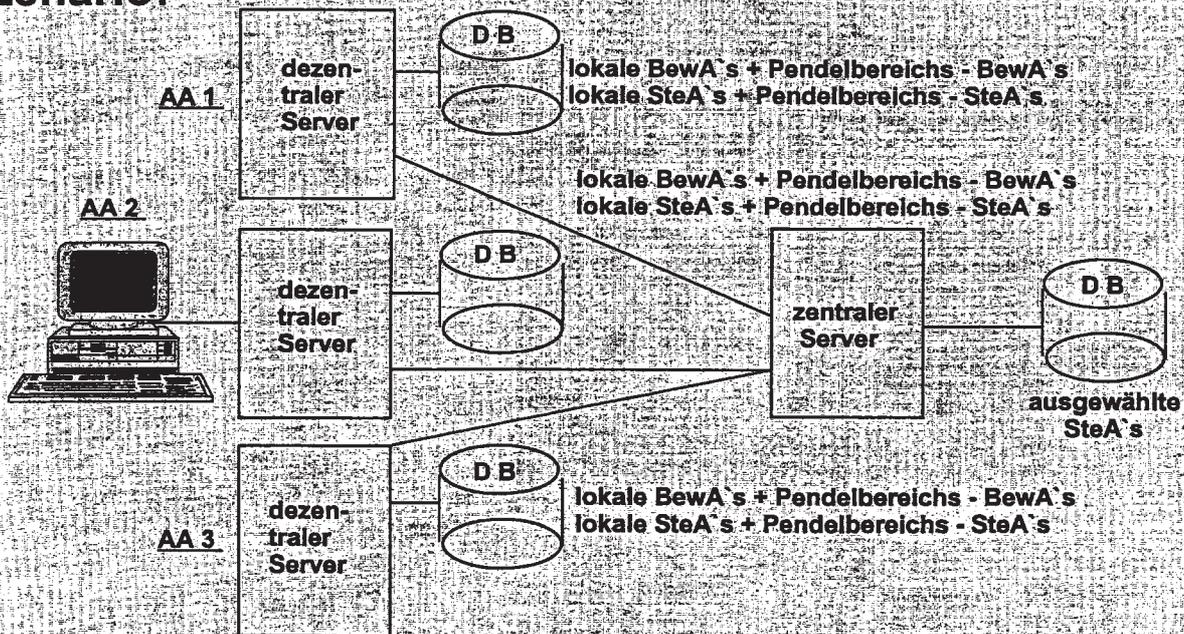
Szenario 5



Szenario 6



Szenario 7



Welche Informationen werden wann wo gespeichert

DATENART	Szenario 1	Szenario 2	Szenario 3	Szenario 4	Szenario 5	Szenario 6	Szenario 7	
1 Stellendaten								
1.1 Aufnahme / Löschen								1: Berufsgruppen die in Szenario 2 zentral gespeichert werden.
lokal	X	X	X	X	X	X	X	
Pendelbereich								
LAA								
zentral	X	X	1	2	2	2	2	
1.2 Update mit Bewerberdaten								2: ausgewählte Berufsgruppen, die zentral verwaltet werden.
lokal		X	X	X	X	X	X	
Pendelbereich								
LAA								
zentral	X	X	1	2	2	2	2	X
2 Bewerberdaten								
2.1 Aufnahme / Löschen								
lokal	X	X	X	X	X	X	X	
Pendelbereich								
LAA								
zentral	X							
2.2 Update Stellendaten								
lokal								
Pendelbereich		X	X	X	X	X	X	
LAA								
zentral	X							

Formeln für die stellenorientierte Bewerbervermittlung (1)

Aktion	Szenario 1	Szenario 2		Szenario 3	
	lokal	lokal	zentral	lokal	zentral
SpSt	$\text{transSpN} \cdot \text{TN}$	$\text{transSpN} \cdot \text{TmN}$	$\text{transSpN} \cdot \text{TN}$	$\text{transSpN} \cdot \text{TmN}$	$(100-X)/100 \cdot \text{transSpN} \cdot \text{TN}$
LöSt	$\text{transLöN} \cdot \text{TN}$	$\text{transLöN} \cdot \text{TmN}$	$\text{transLöN} \cdot \text{TN}$	$\text{transLöN} \cdot \text{TmN}$	$(100-X)/100 \cdot \text{transLöN} \cdot \text{TN}$
SuBw	$\text{transSulokB} \cdot \text{TN}$	$\text{transSulokB} \cdot \text{TmN}$	-	$\text{transSulokB} \cdot \text{TmN}$	-
STd	$L \cdot \text{TN}$	$L/100 \cdot L \cdot \text{TmN}$	-	$L/100 \cdot L \cdot \text{TmN}$	-
Sach	$(3+2 \cdot I/100) \cdot L \cdot \text{TN}$	$(1+2 \cdot (1+I/100)) \cdot L/100 \cdot L \cdot \text{TmN}$	-	$(1+2 \cdot (1+I/100)) \cdot L/100 \cdot L \cdot \text{TmN}$	-
FoAkt	$\text{fa} \cdot (\text{SuBw} + \text{STd} + \text{Sach})$	$\text{fa} \cdot (\text{SuBw} + \text{STd} + \text{Sach})$	-	$\text{fa} \cdot (\text{SuBw} + \text{STd} + \text{Sach})$	-

SpSt: Speichern einer Stelle LöSt: Löschen einer Stelle SuBw: Suche nach Bewerbern
STd: Schreiben in Team-Datei Sach: Aktivitäten der Sachbearbeiter FoAkt: Folgeberatung

Formeln für die stellenorientierte Bewerbervermittlung (2)

Aktion	Szenario4		Szenario5	
	lokal	zentral	lokal	zentral
SpSt	$\text{transSpN} \cdot \text{TmN}$	$(100 - X^*)/100 \cdot \text{transSpN} \cdot \text{TN}$	$(\text{transSpN} + q \cdot \text{transSpN}^*) \cdot \text{TmN}$	$100 - X^*/100 \cdot \text{transSpN} \cdot \text{TN}$
LöSt	$\text{transLöN} \cdot \text{TmN}$	$(100 - X^*)/100 \cdot \text{transLöN} \cdot \text{TN}$	$(\text{transLöN} + q \cdot \text{transLöN}^*) \cdot \text{TmN}$	$(100 - X^*)/100 \cdot \text{transLöN} \cdot \text{TN}$
SuBw	$(\text{transSulokB} + r/100 \cdot (1+q) \cdot \text{transSuremB}) \cdot \text{TmN}$	-	$(\text{transSulokB} + r/100 \cdot (1+q) \cdot \text{transSuremB}) \cdot \text{TmN}$	-
STd	$(1 + L/100) \cdot L \cdot \text{TmN}$	-	$(1 + L/100) \cdot L \cdot \text{TmN}$	-
Sach	$(1 + (1+l/100)(2+L/100)) \cdot L \cdot \text{TmN}$	-	$(3+l/100)(2+L/100) \cdot L \cdot \text{TmN}$	-
FoAkt	$fa \cdot (\text{SuBw} + \text{STd} + \text{Sach})$	-	$fa \cdot (\text{SuBw} + \text{STd} + \text{Sach})$	-

SpSt: Speichern einer Stelle LöSt: Löschen einer Stelle SuBw: Suche nach Bewerbern
 STd: Schreiben in Team- Datei Sach: Aktivitäten der Sachbearbeiter FoAkt: Folgebearbeitungen

Formeln für die stellenorientierte Bewerbervermittlung (3)

Aktion	Szenario6			Szenario7	
	lokal	LAA	zentral	lokal	zentral
SpSt	$\text{transSpN} \cdot \text{TmN}$	$\text{transSpN} \cdot \text{TN/LAA}$	$(100 - X^*)/100 \cdot \text{transSpN} \cdot \text{TN}$	$(\text{transSpN} + q \cdot \text{transSpN}^*) \cdot \text{TmN}$	$(100 - X^*)/100 \cdot \text{transSpN} \cdot \text{TN}$
LöSt	$\text{transLöN} \cdot \text{TmN}$	$\text{transLöN} \cdot \text{TN/LAA}$	$(100 - X^*)/100 \cdot \text{transLöN} \cdot \text{TN}$	$(\text{transLöN} + q \cdot \text{transLöN}^*) \cdot \text{TmN}$	$(100 - X^*)/100 \cdot \text{transLöN} \cdot \text{TN}$
SuBw	$\text{transSulokB} \cdot \text{TmN}$	-	-	$\text{transSulokB} \cdot \text{TmN}$	-
STd	$L/100 \cdot L \cdot \text{TmN}$	-	-	$(1 + L/100) \cdot L \cdot \text{TmN}$	-
Sach	$(1 + 2(1+l/100)) \cdot L \cdot \text{TmN}$	-	-	$(3+l/100)(2+L/100) \cdot L \cdot \text{TmN}$	-
FoAkt	$fa \cdot (\text{SuBw} + \text{STd} + \text{Sach})$	-	-	$fa \cdot (\text{SuBw} + \text{STd} + \text{Sach})$	-

SpSt: Speichern einer Stelle LöSt: Löschen einer Stelle SuBw: Suche nach Bewerbern
 STd: Schreiben in Team- Datei Sach: Aktivitäten der Sachbearbeiter FoAkt: Folgebearbeitungen

Beispielrechnungen für die stellenorientierte Bewerbervermittlung (1)

Annahmen: generell : $N=2.500.000$ $TN=10.000$ $TmN = 50$ $q=5$
 SpSt: Speichern: $transSuFl=1$ $transSpN^*=1$ $X=50\%$ $X^*=95\%$
 LöSt: Löschen : $transLöN^*=1$
 SuBw: Suchen : $r_1=50\%$ $transSuLokB=5$ $transSuremB=1$
 STd: Schreiben Team - Datei: $L1=70\%$ $L2=30\%$ $L=3$
 Sach: Sachbearbeiteraktivitäten : $I=60\%$
 FoAkt: Folgebearbeitungen : $fa = 2$

Beispielrechnungen für die Stellenorientierte Bewerbervermittlung (2)

Aktiv	Szenario 1		Szenario 2		Szenario 3		Szenario 4		Szenario 5		Szenario 6		Szenario 7	
	zentral	lokal	LAA	zentral	lokal	zentral								
SpSt	20.000	100	10.000	100	5.000	100	500	350	500	100	1.000	250	350	500
LöSt	20.000	100	10.000	100	5.000	100	500	350	500	100	1.000	250	350	500
SuBw	50.000	250		250		400		400		250			250	
STd	30.000	105		105		195		195		105			195	
Sach	126.000	441		441		702		552		630			552	
FoAkt	412.000	1592		1592		2594		2284		1990			1994	
Zus	658.000	2588	20.000	2588	10.000	4091	1000	4141	1000	3175	2000	1000	3691	1000

Beispielrechnung für die Aufteilung der Arbeitszeit der Sachbearbeiter

Anzahl neuer Arbeitssuchender : AS+ : 2.500.000 TAS+ : 10.000
 Anzahl der Sachbearbeiter : arbD : 7.500
 Öffnungszeit : AT : 250 Tage Arbeitszeit der Sachbearbeiter : ATS : 200 Tage
 tägliche Anzahl der Erstberatungen pro Sachbearbeiter : TEber : 1,66
 Annahme : durchschnittliche Zeit pro Erstkontakt : ZEber : 30 Min.
 tägliche Zeit für Erstkontakte : TZEber : 50 Min.
 Annahme : Zeit für Publikumsverkehr : TZopen : 4 Std. = 240 Min.
 Zeit für Folgeberatungen während der Öffnungszeit : TZFber : 190 Min.
 Annahme : tägliche Arbeitszeit eines Sachbearbeiters : TAZ : 7,5 Std.
 Zeit für Tätigkeiten außerhalb der Besuchszeit : TZaFber : 3,5 Std. = 210 Min.

Formeln für den Erstkontakt eines Bewerbers (1)

Aktivität	Szenario 1		Szenario 2	
	lokal	zentral	lokal	zentral
Erfassen Personaldata		$\text{transErfB} \cdot \text{TAS} +$	$\text{transErfB} \cdot \text{TmAS} +$	-
Stellensuche		$\text{transSulokV} \cdot \text{TAS} +$	$(\text{transSulokV} +$ $+ (\text{100} - \text{s1}) / \text{100} \cdot$ $\text{transSuremV}) \cdot \text{TmAS} +$	$(\text{100} - \text{s1}) / \text{100} \cdot$ $\text{transSuremV} + \text{TAS} +$ -
Update Personaldata		$\text{transUpPers} \cdot \text{TAS} +$	$\text{transUpPers} \cdot \text{TmAS} +$	-
Update Stellendata		$\text{transUpN} \cdot \text{K} \cdot \text{TAS} +$	$(\text{1} + (\text{100} - \text{K1}) / \text{100}) \cdot$ $\text{transUpN} \cdot \text{K} \cdot \text{TmAS} +$	$\text{transUpN} \cdot \text{K} \cdot \text{TAS} +$

Formeln für den Erstkontakt eines Bewerbers (2)

Aktivität	Szenario 3		Szenario 4	
	lokal	zentral	lokal	zentral
Erfassen Personaldaten	$\text{transERFB} \cdot \text{TmAS} +$		$\text{transERFB} \cdot \text{TmAS} +$	
Stellensuche	$(\text{transSulokV} +$ $+ (100 - X) / 100 \cdot$ $(100 - s1) / 100 \cdot$ $\text{transSuremV}) \cdot \text{TmAS} +$	$(100 - X) / 100 \cdot$ $(100 - s1) / 100 \cdot$ $\text{transSuremV} \cdot$ $\text{TAS} +$	$\text{transSulokV} +$ $(1 + p) \cdot s2 / 100 - s3 / 100 \cdot$ $\text{transSuremV} \cdot \text{TmAS} +$	$s3 / 100 \cdot$ transSuremV $+ \text{TAS} +$
Update Personaldaten	$\text{transUpPers} \cdot \text{TmAS} +$		$\text{transUpPers} \cdot \text{TmAS} +$	
Update Stellendaten	$\text{transUpN} \cdot \text{KR} \cdot$ $\cdot \text{TmAS} +$	$(100 - \text{KR}1) / 100 \cdot$ $\text{KR} \cdot \text{TAS} +$	$81 + (\text{K}2\text{K}3) / 100 \cdot \text{K} \cdot$ $\text{transUpN} \cdot \text{TmAS} +$	$\text{K}3 / 100 \cdot \text{K} \cdot \text{TAS}$ $+ \text{transUpN}$

Formeln für den Erstkontakt eines Bewerbers (3)

Aktivitäten	Szenario 5	
	lokal	zentral
Erfassen	$\text{transErfB} \cdot \text{TmAS} +$	
Stellensuche	$(\text{transSulokV} +$ $+ s3 / 100 \cdot$ $\text{transSuremV}) \cdot \text{TmAS} +$	$s3 / 100 \cdot$ transSuremV $+ \text{TAS} +$
Update Personaldaten	$\text{transUpPers} \cdot \text{TmAS} +$	
Update Stellendaten	$(1 + (\text{K}2 + \text{K}3) / 100) \cdot \text{K} \cdot$ $\text{transUpN} \cdot \text{TmAS} +$	$\text{K}3 / 100 \cdot \text{K} \cdot \text{TAS} +$ $+ \text{transUpN}$

Formeln für den Erstkontakt eines Bewerbers (4)

Aktivität	Szenario 6		
	lokal	LAA	zentral
Erfassen	$\text{transErfB} \cdot \text{TmAS} +$		
Stellensuche	$\text{transSutokV} \cdot \text{TmAS} +$ $+ (100 - s1) / 100 \cdot$ $\text{transSuremV} \cdot \text{TmAS} +$	$(100 - s1) / 100 \cdot$ $\text{transSuremV} \cdot$ $(1 / \text{LAA}) \cdot \text{TAS} +$	$s3 / 100 \cdot$ $\text{transSuremV} \cdot \text{TAS} +$
Update Personaldaten	$\text{transUpPers} \cdot \text{TmAS} +$		
Update Stellendaten	$(1 + (K2 + K3) / 100) \cdot K \cdot$ $\text{transUpN} \cdot \text{TmAS} +$	$(1 + K3 / 100) \cdot K \cdot$ $\text{transUpN} \cdot$ $(1 / \text{LAA}) \cdot \text{TAS} +$	$((100 - X) \cdot (K2 + K3) /$ $100 + K3 / 100) \cdot$ $\text{transUp} \cdot K \cdot$ $\text{TAS} +$

Formeln für den Erstkontakt eines Bewerbers (5)

Aktivität	Szenario 7	
	lokal	zentral
Erfassen	$(\text{transErfB} +$ $p \cdot \text{transErfB}) \cdot \text{TmAS} +$	
Stellen - Suche	$(\text{transSutokV} \cdot +$ $+ s3 / 100 \cdot$ $\text{transSuremV}) \cdot \text{TmAS} +$	$s3 / 100 \cdot$ transSuremV $\cdot \text{TAS} +$
Update Personaldaten	$\text{transUpPers} \cdot \text{TmAS} +$	
Update Stellendaten	$(1 + (K2 + K3) / 100) \cdot K \cdot$ $\text{transUpN} \cdot \text{TmAS} +$	$K3 / 100 \cdot K \cdot \text{TAS} +$ transUpN

Unterschied zwischen Szenario2 / Szenario 3 bei der Stellensuche

Anzahl der Bewerber : 100

max. Zahl angebotener Stellen pro Bewerber: $K = 3$
d.h. Es werden insgesamt 300 Stellen gesucht

Suchergebnis lokal : in 50% der Suchfragen werden 3 Stellen = 150 Stellen gefunden (GR1)
in 20% " " " 2 " " = 40 " " (GR2)
in 20% " " " 1 " " = 20 " " (GR3)
in 10% " " " 0 " " = 0 " " (GR4)

Annahme für Szenario3 : alle der ersten beiden Gruppen gehören zu den Berufsgruppen von X,
d.h. die Stellen dieser Gruppen werden nur lokal gespeichert

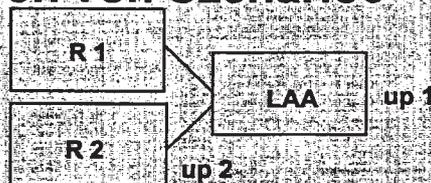
Ergebnis Szenario2: werden gefunden : 300 Stellen , davon 90 zentral; $K1=70\%$; $K3=30\%$

Ergebnis Szenario3: werden gefunden : 280 Stellen , davon 70 zentral; $KR1=75\%$; $KR3=25\%$

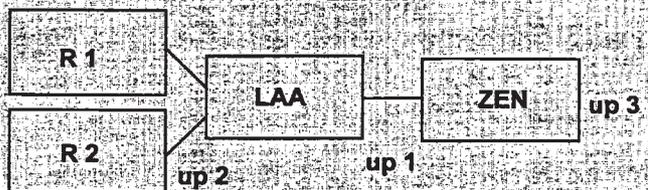
* Es fehlen die 20 Stellen von GR2 ; $KR=2.8$

Das Stellen - Update - Verfahren von Szenario6

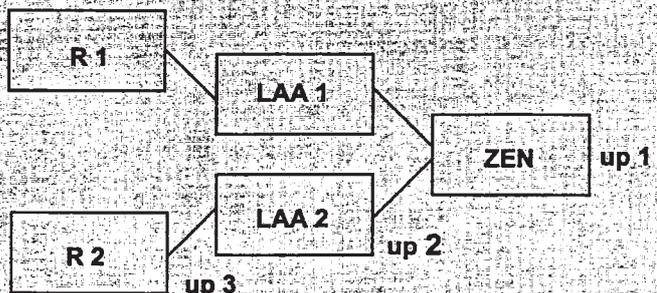
für nicht zentral
verw. Berufsgruppen



zentral verw.
Berufsgruppen



Stelle innerhalb
LAA -Bereich



Stelle außerhalb
LAA-Bereich

Beispielrechnung bewerberorientierte Stellenvermittlung (Erstkontakt 1)

Annahmen: generell: AS+ = 2.500.000 TAS+ = 10.000 TmAS+ = 50 P= 3

Er : Erfassen Personaldaten: transSuAS =1 transSpPers =1 transSpBer =1
transErfB = 3 transErfBP=1

SSu : Stellensuche: transSulokV =5 transSuremV =1
s1 =50% s2 = 45% s3 =5%
X = 50% X* = 95%

Up : Update Personaldaten: transUpPers =1

Us : Update Stellendate: transUpN = 1
K = KR = 3
K1 = 60% K2 =35 % K3 =5%
KR1 = 60%

Beispielrechnungen für die bewerberorientierte Stellenvermittlung (Erstkontakt)

Aktiv	Szenario 1		Szenario 2		Szenario 3		Szenario 4		Szenario 5		Szenario 6		Szenario 7	
	zentral	lokal	LAA	zentral	lokal	zentral								
ER	30.000	150	150	150	150	150	150	150	150	150			300	
SSU	50.000	275	5.000	262	2.500	343	500	253	500	275	500	500	253	500
Up	10.000	50	50	50	50	50	50	50	50	50			50	
Us	30.000	150	30.000	150	15.000	210	3.000	210	3.000	210	3300	4.350	210	3.000
Zus	120.000	625	35.000	612	17.500	753	3.500	663	3.500	685	3800	4.850	813	3.500

Formeln für die Folgeberatungen von Bewerbern (1)

Aktivität	Szenario1		Szenario2	
	lokal	zentral	lokal	zentral
asynchr. Suchen		transSufol * FS	transSufol * FmS + (100 - s1) / 100 *	(100 - s1) / 100 *
Schreiben Team - Daten		FS	FmS	
Lesen Team - Daten		FS	FmS	
Lesen Bew. - Daten		FS	FmS	
Lesen lokale Stellendaten		KNT * FS	K1 / 100 * KNT * FmS	
Lesen remote Stellendaten * FS			(100 - K1) / 100 * KNT * FmS	(100 - K1) / 100 * KNT
Update Bewerber		lb / 100 * FS	lb / 100 * FmS	
Update lokale Stellen		lk / 100 * KNT * FS	KNT * FmS	
Update in der Ferne / 100 *			lk / 100 * (100 - K1) / 100 *	lk / 100 * (100 - K1)
Update aus der Ferne			KNT * FmS lk / 100 * (100 - K1) / 100 KNT * FmS	KNT * FS

Formeln für die Folgeberatungen von Bewerbern (2)

Aktivitäten	Szenario 3	
	lokal	zentral
asynchr. Suchen	transSufol * FmS + (100 - s1) / 100 * (100 - X) / 100 * transSuremV * FmS	(100 - s1) / 100 * (100 - X) / 100 * transSuremV * FS
Schreiben Team - Daten	FmS	
Lesen Team - Daten	FmS	
Lesen Bew. Daten	FmS	
Lesen lokale Stellendaten	KR 1 / 100 * KRNT * FmS	
Lesen remote Stellendaten	(100 - KR 1) / 100 * KRNT * FmS	(100 - KR 1) / 100 * KRNT * FS
Update Bewerber	lb / 100 * FmS	
Update lokale Stellen	lk / 100 * KR 1 / 100 * KRNT * FmS	
Update in der Ferne	lk / 100 * (100 - KR 1) / 100 * KRNT * FmS	lk / 100 * (100 - KR 1) / 100 * KRNT * FS
Update aus der Ferne	lk / 100 * (100 - KR 1) / 100 * KRNT * FmS	

Formeln für die Folgeberatungen von Bewerbern (3)

Szenario 4

Aktivität	lokal	zentral
asynchr. Suchen	$\text{transSufol} \cdot \text{FmS} +$ $(1+p) \cdot s2/100 + s3/100$	$s3/100$
Schreiben Team-Datel	$\text{transSuremV} \cdot \text{FmS}$	$\text{transSuremV} \cdot \text{FS}$
Lesen Team-Datel	FmS	
Lesen Bew.Daten	FmS	
Lesen lokale Stellendaten	$K1/100 \cdot \text{KNT} \cdot \text{FmS}$	
Lesen remote Stellendaten	$(100 - K1)/100 \cdot \text{KNT} \cdot \text{FmS}$ $+ K2/100 \cdot \text{KNT}$	$K3/100 \cdot \text{KNT} \cdot \text{FS}$
Update Bewerber	$lb/100 \cdot \text{FmS}$	
Update lokale Stellen	$lk/100$ $\text{KNT} \cdot \text{FmS}$	
Update in der Ferne	$lk/100 \cdot (100 - K1)/100$ $\text{KNT} \cdot \text{FmS}$	$lk/100 \cdot K3/100$ $\text{KNT} \cdot \text{FmS}$
Update aus der Ferne	$lk/100 \cdot (100 - K1)/100$ $\text{KNT} \cdot \text{FmS}$	

Formeln für die Folgeberatungen von Bewerbern (4)

Szenario 5

Aktivität	lokal	zentral
asynchr. Suchen	$(\text{transSufol} +$ $s3/100) \cdot \text{FmS}$	$s3/100$
Schreiben Team - Datel	$\text{transSuremV} \cdot \text{FmS}$	$\text{transSuremV} \cdot \text{FS}$
Lesen Team-Datel	FmS	
Lesen Bew.Daten	FmS	
Lesen lokale Stellendaten	$(K1 + K2)/100 \cdot \text{KNT} \cdot \text{FmS}$	
Lesen remote Stellendaten	$K3/100 \cdot \text{KNT} \cdot \text{FmS}$	$K3/100 \cdot \text{KNT} \cdot \text{FS}$
Update Bewerber	$lb/100 \cdot \text{FmS}$	
Update lokale Stellen	$lk/100 \cdot K1/100$ $\text{KNT} \cdot \text{FmS}$	
Update in der Ferne	$lk/100 \cdot (100 - K1)/100$ $\text{KNT} \cdot \text{FmS}$	$lk/100 \cdot K3/100$ $\text{KNT} \cdot \text{FS}$
Update aus der Ferne	$lk/100 \cdot (100 - K1)/100$ $\text{KNT} \cdot \text{FmS}$	

Formeln für die Folgeberatungen von Bewerbern (5)

Szenario 6

Aktivität	lokal	LAA	zentral
asynchr. Suchen	$(\text{transSufol} + (-100 - s1)/100 * \text{transSuremV}) * \text{FmS}$	$(100 - s1)/100 * \text{transSuremV} * (1/LAA) * \text{FS}$	$s3/100 * \text{transSuremV} * \text{FS}$
Schreiben Team-Datfel	FmS		
Lesen Team-Datfel	FmS		
Lesen Bew.Daten	FmS		
Lesen lokaleStellendaten	$K1/100 * \text{KNT} * \text{FmS}$		
Lesen remote Stellendaten	$(K2 + K3)/100 * \text{KNT} * \text{FmS}$	$(K2 + K3)/100 * (1/LAA) * \text{KNT} * \text{FS}$	$K3/100 * \text{KNT} * \text{FS}$
UpdateBewerber	$lb/100 * \text{FmS}$		
Update lokale Stellen	$lk/100 * K1/100 * \text{KNT} * \text{FmS}$		
Update in der Ferne	$lk/100 * \text{KNT} * \text{FmS}$	$lk/100 * (1/LAA) * \text{KNT} * \text{FS}$	$lk/100 * K3/100 * \text{KNT} * \text{FmS}$
Update aus der Ferne	$lk/100 * (100 - K1)/100 * \text{KNT} * \text{FmS}$	$lk/100 * K3/100 * (1/LAA) * \text{KNT} * \text{FS}$	

Formeln für die Folgeberatungen von Bewerbern (6)

Szenario 7

Aktivität	lokal	zentral
asynchr. Suchen	$(\text{transSufol} + s3/100) * \text{transSuremV} * \text{FmS}$	$s3/100 * \text{transSuremV} * \text{FS}$
Schreiben Team-Datfel	FmS	
Lesen Team-Datfel	FmS	
Lesen Bew.Daten	FmS	
Lesen lokaleStellendaten	$(K1 + K2)/100 * \text{KNT} * \text{FmS}$	
Lesen remote Stellendaten	$K3/100 * \text{KNT} * \text{FmS}$	$K3/100 * \text{KNT} * \text{FS}$
UpdateBewerber	$lb/100 * \text{FmS}$	
Update lokale Stellen	$lk/100 * K1/100 * \text{KNT} * \text{FmS}$	
Update in der Ferne	$lk/100 * (100 - K1)/100 * \text{KNT} * \text{FmS}$	$lk/100 * K3/100 * \text{KNT} * \text{FS}$
Update aus der Ferne	$lk/100 * (100 - K1)/100 * \text{KNT} * \text{FmS}$	

Beispiel für Folgeberatungen von Bewerbern(1)

Annahmen:

Anzahl der Arbeitssuchenden AS: 5.000.000

jährliche Anzahl der asynchronen Suchläufe f1: 12
 jährliche Anzahl der Einbestellungen f2: 4
 jährliche Anzahl unangemeldeter Beratungen f3: 4 $f = f1 + f2 + f3 = 20$

Gesamtzahl der jährlichen Folgeberatungen: 100.000.000

tägliche Anzahl der Folgeberatungen FS: 400.000
 tägliche Anzahl der Folgeberatungen in einem Mittleren AA FmS: 2.000

Anzahl der Sachbearbeiter für Bewerberberatung: 7.500

tägliche Anzahl der Folgeberatungen pro Sachbearbeiter: 53,3
 davon 2/5 = 21,3 in der Öffnungszeit des Arbeitsamtes

Für Beispiel auf Folie 14 heißt das:

Zeit für 1 Folgeberatung während der Öffnungszeit TZFBerB: 8,9 Min.
 außerhalb TZaFberB: 6,5 Min.
 durchschnittliche Anzahl gefundener Stellen KNT=2 (KRNT=1,5)
 prozentualer Anteil gefundener Stellen lokal K1: 60% im Pendelbereich K2: 35%, zentral K3: 5%
 (für Szenario 3 KR1: 70%, KR3: 30%)
 prozentualer Anteil akzeptierter Stelle für f1 (nur Sachbearbeiter) l1: 50%
 für f2 u. f3 (Sachbearbeiter + Bewerber) l2: 50% also lk=50%
 prozentualer Anteil für Bewerber die eine Stelle erhalten lb: 55%

Beispielrechnungen Gesamtanzahl der Transaktionen

Aktiv	Szenario 1		Szenario 2		Szenario 3		Szenario 4		Szenario 5		Szenario 6		Szenario 7	
	zentral	lokal	LAA	zentral	lokal	zentral								
Stor BewV	658.000	2.588	20.000	2.588	10.000	4.091	1.000	4.141	1.000	3.175	2.000	1.000	3.691	1.000
Bew StVE	120.000	625	35.000	612	17.500	753	3.500	663	3.500	658	3.800	4.850	813	3.500
Bew StVF	3.020.000	16.900	680.000	13.550	370.000	21.000	140.000	6.600	40.000	18.100	100.000	40.000	16.600	140.000
L6 Bew	20.000	100		100		100		100		100			200	
Zus	3.818.000	20.213	735.000	16.850	397.500	25.944	144.500	21.504	144.500	22.060	105.800	45.850	1.304	144.500

Beispielrechnungen für die Bewerberorientierte Stellenvermittlung (Folgeberatung)

Aktiv	Szenario 1	Szenario 2		Szenario 3		Szenario 4		Szenario 5		Szenario 6		Szenario 7		
	lokal	lokal	zentral	lokal	zentral	lokal	zentral	lokal	zentral	lokal	LAA	zentral	lokal	zentral
AsSu	400.000	3.000	200.000	2.500	100.000	5.700	20.000	2.100	20.000	3.000	20.000	20.000	2.100	20.000
STeD	400.000	2.000		2.000		2.400		2.000		2.000			2.000	
LTeD	400.000	2.000		2.000		2.000		2.000		2.000			2.000	
LBew	400.000	2.000		2.000		2.000		2.000		2.000			2.000	
LISt	800.000	2.400		2.100		2.100		3.800		2.400			3.800	
LrSt		1.600	320.000	900	180.000	2.800	80.000	400	80.000	1.600	36.000	80.000	400	
UBew	220.000	1.100		1.100		1.100		1.100		1.100			1.100	
UISt	400.000	1.200		1.050		2.000		1.200		1.200			1.200	
UrSt		800	160.000	450	90.000	800	40.000	800	40.000	2.000	40.000	40.000	800	40.000
UaSt		800		450		800		800		800	4.000		800	
Zus	3.020.000	16.900	680.000	13.550	370.000	21.000	140.000	16.600	140.000	18.100	100.000	140.000	6.600	140.000

Transaktionsraten und Netzbelastung aufgeteilt auf die Tageszeiten (1)

Szenario 1

Aktivität	Zeit für Publikumsverkehr (4 Std.)			Arbeitszeit außerhalb Publik. Verkehr (4Std.)			Restzeit ohne Sachbearbeiter			Netzbel. gesamt	
	TA/sec	Netzbel.	Anz. GB*km	TA/sec	Netzbel.	Anz. GB*km	TA/sec	Netzbel.	Anz. GB*km	Anz. GB*km	Anz. GB*km
	lokal/zentral	lokal/zentral		lokal/zentral	lokal/zentral						
Stellenorientierte Bewerbervermittlung Stellenbearbeitung		4.17	60.000	45.0	4.17	60.000	45.0	2.98		120.000	90.0
Bewerberbearbeitung				26.25	378.000	283.5				378.000	283.5
bewerberorientierte Stellenvermittlung											
Erstkontakte		8.33	120.000	90.0						120.000	90.0
Folgeberatung		41.11	520.000	333.0	123.36	1.776.000	1332	11.85		2.300.000	17.25
Gesamt											

Transaktionsraten und Netzbelastung aufgeteilt auf die Tageszeiten (2)

Szenario 2

Aktivität	Zeit für Publikumsverkehr (4 Std.)				Arbeitszeit außerhalb Publik. Verkehr (4Std.)				Restzeit ohne Sachbearbeiter				Netzbel. gesamt	
	TA/sec		Netzbel.		TA/sec		Netzbel.		TA/sec		Netzbel.		Anz. GB*km	
	lokal	zentral	Anz.	GB*km	lokal	zentral	Anz.	GB*km	lokal	zentral	Anz.	GB*km		
<u>Stellenorientierte Bewerbervermittlung</u> Stellenbearbeitung	0.016	0.69	10.000	7.5	0.018	0.69	10.000	7.5	0.013				20.000	15.0
Bewerberbearbeitung					0.091									
<u>bewerberorientierte Stellenvermittlung</u> Erstkontakte	0.043	2.43	35.000	26.25									35.000	26.25
Folgeberatung	0.253	9.43	168.000	126	0.864	26.67	512.000	384.0	0.07	3.7	160.000	120.0	340.000	630.0
Gesamt	0.312	12.55	213.000	159.75	0.851	27.36	522.000	391.5	0.083	3.7	160.000	120.0	895.000	671.25

Transaktionsraten und Netzbelastung aufgeteilt auf die Tageszeiten (3)

Szenario 3

Aktivität	Zeit für Publikumsverkehr (4 Std.)				Arbeitszeit außerhalb Publik. Verkehr (4Std.)				Restzeit ohne Sachbearbeiter				Netzbel. gesamt	
	TA/sec		Netzbel.		TA/sec		Netzbel.		TA/sec		Netzbel.		Anz. GB*km	
	lokal	zentral	Anz.	GB*km	lokal	zentral	Anz.	GB*km	lokal	zentral	Anz.	GB*km		
<u>Stellenorientierte Bewerbervermittlung</u> Stellenbearbeitung	0.016	0.345	5.000	3.75	0.018	0.345	5.000	3.75	0.013				10.000	7.5
Bewerberbearbeitung					0.091									
<u>bewerberorientierte Stellenvermittlung</u> Erstkontakte	0.043	1.215	17.500	13.12									17.500	13.12
Folgeberatung	0.203	5.05	92.000	69.0	0.56	15.0	288.000	216.0	0.066	1.85	60.000	60.0	460.000	345.0
Gesamt														

Transaktionsraten und Netzbelastung aufgeteilt auf die Tageszeiten (4)

Szenario 4

Aktivität	Zeit für Publikumsverkehr (4 Std.)				Arbeitszeit außerhalb Publik. Verkehr (4Std.)				Restzeit ohne Sachbearbeiter				Netzbel. gesamt	
	TA/sec		Netzbel.		TA/sec		Netzbel.		TA/sec		Netzbel.		Anz.GB*km	
	lokal	zentral	Anz. GB*km		lokal	zentral	Anz. GB*km		lokal	zentral	Anz. GB*km			
Stellenorientierte Bewerbervermittlung														
Stellenbearbeitung	0.100	0.035	Z:500 P:85.5	0.375 1.283	0.100	0.375	Z:500 P:85.5	0.375 1.283	0.022		Z:0 P:215	0 3.24	Z:1000 P:385	0.75 5.806
Bewerberbearbeitung					0.146		Z:0 P:215	0 3.24						
bewerberorientierte Stellenvermittlung														
Erstkontakte	0.052	0.139	Z:2.000 P:120	1.500 1.800									Z:2.000 P:120	1.50 1.80
Folgeberatung	0.25	1.11	Z:20.000 P:960	15.00 14.40	0.572	3.33	Z:64.000 P:1680	48.00 29.2	0.114	0.3	Z:16.000 P:2.160	12.0 32.4	Z:100.000 P:4.800	63.00 72.60
Gesamt	0.402	2.28	Z:22.500 P:1165.5	16.875 17.083	0.818	3.705	Z:64.500 P:1980.5	48.375 29.72	0.136	0.3	Z:16.000 P:2375	12.0 36.64	Z:103.000 P:5520	77.25 82.8

Transaktionsraten und Netzbelastung aufgeteilt auf die Tageszeiten (5)

Szenario 5

Aktivität	Zeit für Publikumsverkehr (4 Std.)				Arbeitszeit außerhalb Publik. Verkehr (4Std.)				Restzeit ohne Sachbearbeiter				Netzbel. gesamt	
	TA/sec		Netzbel.		TA/sec		Netzbel.		TA/sec		Netzbel.		Anz. GB*km	
	lokal	zentral	Anz. GB*km		lokal	zentral	Anz. GB*km		lokal	zentral	Anz. GB*km			
Stellenorientierte Bewerbervermittlung														
Stellenbearbeitung	0.100	0.035	Z:500 P:335.5	0.375 5.033	0.100	0.375	Z:500 P:335.5	0.375 5.033	0.022		Z:0 P:215	0 3.24	Z:1000 P:886	0.75 13.29
Bewerberbearbeitung					0.115		Z:0 P:81	0 1.215					P:81	1.215
bewerberorientierte Stellenvermittlung														
Erstkontakte	0.052	0.139	Z:2.000 P:52.5	1.500 0.788									Z:2.000 P:52.5	1.50 0.788
Folgeberatung	0.22	1.11	Z:20.000 P:140	15.00 2.10	0.66	3.33	Z:64.000 P:560	48.00 8.4	0.06	0.3	Z:16.000	12.0	Z:100.000 P:700	75.00 10.50
Gesamt	0.327	2.28	Z:22.500 P:527.0	16.875 7.91	0.857	3.705	Z:64.500 P:976.5	48.375 14.648	0.082	0.3	Z:16.000 P:215	12.0 3.24	Z:103.000 P:1719.5	77.25 25.793

Transaktionsraten und Netzbelastung aufgeteilt auf die Tageszeiten (6/1)

Szenario 6

Aktivität	Zeit für Publikumsverkehr (4 Std.)					Arbeitszeit außerhalb Publik. Verkehr (4Std.)				
	lokal	LAA	zentral	Anz.	GB*km	lokal	LAA	zentral	Anz.	GB*km
stellenorientierte Bewerbermittlung										
Stellenbearbeitung	0.019	0.069	0.017	AL: 10.000 LZ: 250	4.50 0.013	0.019	0.069	0.017	AL: 10.000 LZ: 250	4.50 0.013
Bewerberbearbeitung						0.084	0.26	0.131	AL: 37.800 LZ: 1890	17.01 0.85
bewerberorientierte Stellenvermittlung										
Erstkontakte	0.048	0.253	0.018	AL: 36.500 LZ: 2600	16.43 0.78					
Folgeberatungen	0.069	0.277		AL: 40.000 LZ: 4000	18.00 1.80	0.91	5.0	4.31	AL: 1.080.000 LZ: 78.000	486.0 35.1
Gesamt	0.136	0.599	0.035	AL: 86.500 LZ: 6850	38.93 2.89	1.013	5.329	4.458	AL: 1.117.800 LZ: 80.340	507.51 36.06

Transaktionsraten und Netzbelastung aufgeteilt auf die Tageszeiten (6/2)

Szenario 6

Aktivität	Restzeit ohne Sachbearbeiter (15Std.)					Netzbel. gesamt	
	lokal	LAA	zentral	Anz.	GB*km	Anz.	GB*km
stellenorientierte Bewerbermittlung							
Stellenbearbeitung	0.049					AL: 20.000 LZ: 500	9.0 0.226
Bewerberbearbeitung						AL: 37.800 LZ: 1.900	17.01 0.85
bewerberorientierte Stellenvermittlung							
Erstkontakte						AL: 36.500 LZ: 2.600	16.43 0.78
Folgeberatungen	0.074	0.296	0.296	AL: 160.000 LZ: 16.000	72.0 7.2	AL: 1.280.000 LZ: 98.000	576.0 44.1
Gesamt							

(identisch zur Abb. S. 40, auf die im ersten Teil der Studie verwiesen wird)

SIEMENS
NIXDORF

Bewertung der einzelnen Szenarien							
Kriterium	Szenario1	Szenario2	Szenario3	Szenario4	Szenario5	Szenario6	Szenario7
Vollständigkeit bzgl. der Suche nach - Stellen - Bewerbern	++	++	+	(++)	(++)	+	(++)
Vollständigkeit der Suche nach Netzausfall bzgl. - Stellen - Bewerbern				(-)	(++)	(-)	(++)
Transaktions- raten - lokal - zentral - LAA	entf.	+	+	+	+	+	+
Netzbelastung - zur Zentrale - im Pendelber. - zum LAA- Rechner	entf.	entf.	(-)	(+)	(++)	entf.	+

2 Kurzfassung Arbeitsamt 2000

(als Anlage genehmigt durch Herrn Funke, vgl. Literaturangabe [AA96])



Kurzfassung

zur Grobbeschreibung des Organisationskonzepts „Arbeitsamt 2000“

Vorbemerkung

- Der Vorstand der BA hat am 18. Dezember 1996 - nach Vorberatung im Vorstandsausschuß für Personalwesen, Organisation und Informationsverarbeitung am 17. Dezember 1996 - von der Kurzfassung zur Grobbeschreibung des Organisationskonzepts „Arbeitsamt 2000“ sowie der Absicht, eine Erprobung in Modellämtern durchzuführen, Kenntnis genommen.

Von einzelnen Mitgliedern des Vorstands wurden dabei folgende Erwartungen bzw. Wünsche geäußert:

- Der Option, die Aufgaben der stellenorientierten Arbeitsvermittlung zu konzentrieren und den Mitarbeiter-Teams für arbeitgeberbezogene Aufgaben zuzuordnen, soll bei der Modell-erprobung entsprechende Aufmerksamkeit gewidmet werden (s. Abschnitt C, Ziffer 1, Nr. 1.2 der Kurzfassung).
- Während der Modell-erprobung sollen Lösungsansätze zu der Frage entwickelt und erprobt werden, wie die Kontakte zu den Betrieben bei der Akquisition und Entgegennahme von Arbeits- und Ausbildungsstellenangeboten zwischen den Mitarbeiter-Teams für die einzelnen Kundengruppen koordiniert werden können, um die Anzahl der Ansprechpartner für die Betriebe zu reduzieren. Eine Veränderung der in der Konzeption vorgesehenen Aufbauorganisation ist damit nicht verbunden.

Mit der Beratung im Vorstand haben die Ergebnisse der Projektarbeit zum „Arbeitsamt 2000“ einen Verbindlichkeitsgrad erreicht, der es zuläßt, diese zur Verfügung zu stellen.

- Gegenstand der Erprobung werden vorrangig die Handlungsfelder sein, von denen unmittelbare aufbauorganisatorische Veränderungen in den Arbeitsämtern ausgehen, wie die Handlungsfelder „Einrichtung kundenorientierter Mitarbeiter-Teams“ und „Anpassung der Organisationsstrukturen in den Arbeitsämtern“ sowie die Qualifizierung der Mitarbeiter und Führungskräfte entsprechend den diesem Handlungsfeld zugrundeliegenden Ergebnissen der Projektarbeit. Die Ergebnisse zu den übrigen Handlungsfeldern werden sukzessive ausgewertet und - soweit wie möglich - zu gegebener Zeit berücksichtigt.

Rückfragen beantworten:
Referatsleiter: Herr Pielenz, Tel. 0911 / 179 3715
Bearbeiter: Herr Munker, Tel. 0911 / 179 2269
Herr Funke, Tel. 0911 / 179 2395

Arbeitsamt 2000

A Ausgangslage

Die gegenwärtige organisatorische Gliederung der Dienststellen der BA in Abteilungen (Sparten) richtet sich an einer fachlich-funktionalen Aufgliederung der übertragbaren Aufgabenkomplexe aus. So sind die Arbeitsvermittlung und Arbeitsberatung, die Berufsberatung und die Gewährung von Leistungen in verschiedenen Abteilungen angesiedelt. Zwar ist es mit dieser Organisationsform gelungen, die Komplexität der Aufgaben und auch besondere Herausforderungen zu beherrschen, wie zum Beispiel Aufbau und Integration der Arbeitsverwaltung in den neuen Bundesländern zeigen. Sie hat aber zwangsläufig eine differenzierte arbeitsteilige Gliederung der Arbeitsprozesse und damit eine typisch geprägte Struktur zur Folge - häufig mit mehreren Ansprechpartnern in verschiedenen Abteilungen für zusammenhängende Kundenanliegen. Bei der Suche nach verbesserten Organisationsformen ist verstärkt zu berücksichtigen, daß die Bürger eine möglichst ganzheitliche Erledigung ihrer Anliegen erwarten, die von der bisherigen Organisationsform so nicht geleistet werden kann.

Deshalb wird der Übergang von der Spartenorganisation zur kundenorientierten Teamorganisation angestrebt, die verschiedenen Anliegen der Kunden sollen in überschaubaren Mitarbeiter-Teams möglichst vollständig, d.h. ganzheitlich erledigt werden. Kernpunkte dieses Lösungsansatzes sind Dezentralisierung und Integration der Aufgaben. Ergänzend sind die Möglichkeiten zur Selbstinformation für die Kunden zu verbessern.

B Konzept „Arbeitsamt 2000“

Unter der Bezeichnung „Arbeitsamt 2000“ hat die BA ein Gesamtkonzept über die zukünftige Gestaltung ihrer Organisation entwickelt, dem der Vorstand am 30. März 1995 zugestimmt hat. Dabei ist sie von den vier grundlegenden Organisationszielen

- Kundenorientierung,
- Wirksamkeit,
- Wirtschaftlichkeit und
- Mitarbeiterorientierung

ausgegangen, die möglichst gleichmäßig ohne eine bestimmte Rangfolge anzustreben sind. Das Konzept enthält 13 nach Sachzusammenhängen gegliederte Handlungsfelder, zu denen im Rahmen einer Projektorganisation die - im folgenden beschriebenen - Lösungsansätze entwickelt wurden (vgl. Anhang II, Abb. 1).

C Handlungsfelder

I. Aufbau- und ablauforganisatorische Maßnahmen zur Ausrichtung der Organisation der BA auf ihre Kunden

1. Einrichtung kundenorientierter Mitarbeiter-Teams

Diesem Handlungsfeld kommt im Projekt Arbeitsamt 2000 eine Schlüsselrolle zu, da von der Einrichtung kundenorientierter Mitarbeiter-Teams ein wesentlicher Einfluß auf die übrigen Handlungsfelder ausgeht.

Bei der Bildung von Mitarbeiter-Teams gilt einerseits die Forderung, zusammengehörende Kundenanliegen zur Bearbeitung zu integrieren. Andererseits sind die Aufgaben auch künftig mit hoher fachlicher Kompetenz wahrzunehmen. Hieraus ergibt sich ein Spannungsfeld, weil eine ganzheitliche Aufgabenerledigung in der Tendenz eine teilweise Abkehr von der Spezialisierung bedeutet, die ein hohes Maß an Fachlichkeit gewährleistet. Es kommt darauf an, Dienstleistungen so zusammenzufassen, daß keine qualitativen Defizite eintreten. Da die gesamte Bandbreite der Dienstleistungen für die Kunden des Arbeitsamtes von den Mitarbeitern¹ eines Teamtyps ganzheitlich nicht erbracht werden kann, werden bestimmte Kundenanliegen, zwischen denen ein Zusammenhang besteht, gebündelt und zusammengeführt.

Für folgende Kundengruppen ist es möglich, funktionsfähige - das heißt transparente und steuerbare sowie mit fachlich kompetenten Mitarbeitern besetzte - Mitarbeiter-Teams einzurichten (vgl. Anhang II, Abb. 2):

- Ausbildungsmarktpartner²,
- Arbeitsmarktpartner¹,
- Behinderte und
- Kindergeldberechtigte.

Diese kundenorientierten Mitarbeiter-Teams sind in ihrer Grundstruktur möglichst gleichartig gehalten; erfordert es jedoch die Aufgabenvielfalt bzw. sind Optionen freigegeben, können auch unterschiedliche Teamtypen zugelassen werden. Übereinstimmende Merkmale sind Verantwortung innerhalb der Teams für die Aufgabenerledigung, klarer Auftrag sowie eine die Kompetenzen der Teams akzeptierende zielorientierte Führung.

Neben den Teams für die genannten Kundengruppen sind Mitarbeiter-Teams für ordnungspolitische Aufgaben zu bilden.

1.1 Ausbildungsmarktpartner

In den Mitarbeiter-Teams für die Kundengruppe Ausbildungsmarktpartner erfolgt die ganzheitliche Betreuung der Anliegen von Berufsanfängern - einschließlich der Abiturienten und Studierenden - von ungelerten jungen Erwachsenen, die eine Berufsausbildung anstreben, sowie von Ausbildungsplatzanbietern. In den Mitarbeiter-Teams werden auch die leistungsrechtlichen Anliegen für diese Kundengruppe bearbeitet.

Die Mitarbeiter-Teams bestehen jeweils aus einem Beraterbereich und einem Kundenbüro. Im Beraterbereich werden die Fachaufgaben - insbesondere Berufsorientierung, berufliche Beratung und Ausbildungsvermittlung - wahrgenommen. Das Kundenbüro ist für die Erledigung von Aufgaben zuständig, die in der Regel anlässlich des ersten Kontaktes mit dem Kunden oder bei der Behandlung von einfachen Anliegen anfallen. Ob bestimmte Aufgaben, beispielsweise Leistungsbearbeitung und maßnahmeträgerbezogene Aufgaben, unter Effizienzaspekten in einem erweiterten Kundenbüro oder in einem Mitarbeiter-Team für zentrale Aufgaben erledigt werden müssen, wird nach im Modellversuch erprobt.

1) Aus Gründen der besseren Lesbarkeit wurde in der Regel die männliche Bezeichnung gewählt. Selbstverständlich beziehen sich alle Ausführungen auch auf die weibliche Form.

2) Bei den Begriffen Ausbildungsmarktpartner und Arbeitsmarktpartner handelt es sich noch um Arbeitstitel.

Für die Betreuung der Studierenden/Hochschulabsolventen werden an Hochschulstandorten mit mehr als 20.000 Studierenden entsprechend dem Beschluß des Vorstands Hochschulteams gebildet, die aus Mitarbeitern der jeweiligen Teams für die Kundengruppen „Ausbildungsmarktpartner“ und „Arbeitsmarktpartner“ bestehen.

1.2 Arbeitsmarktpartner

In den Mitarbeiter-Teams für diese Kundengruppe erfolgt die ganzheitliche Bearbeitung der Anliegen von Arbeitnehmern, soweit diese nicht zu den Kundengruppen Ausbildungsmarktpartner oder Behinderte gehören, und Arbeitgebern in ihrer Eigenschaft als Arbeitsplatzanbieter.

Zur Bewältigung der Aufgabenvielfalt und -komplexität sind für Arbeitsmarktpartner grundsätzlich zwei verschiedene Arten von Mitarbeiter-Teams vorgesehen, nämlich

- Mitarbeiter-Teams für arbeitnehmerbezogene Aufgaben und
- Mitarbeiter-Teams für arbeitgeberbezogene Aufgaben.

Die stellenorientierte Arbeitsvermittlung - einschließlich der Akquisition von Arbeitsstellen - ist im Regelfall den Mitarbeiter-Teams für arbeitnehmerbezogene Aufgaben zugeordnet, weil der Vermittlungsvorgang grundsätzlich als eine Einheit anzusehen ist und eine Aufspaltung der Maklerfunktion der BA auf dem Arbeitsmarkt vermieden werden sollte. Die Arbeitsämter erhalten jedoch die Option, die stellenorientierte Arbeitsvermittlung - einschließlich der Akquisition von Arbeitsstellen - den arbeitgeberbezogenen Mitarbeiter-Teams zuzuordnen, wenn dies zur Intensivierung der Vermittlungsarbeit erforderlich ist.

Die arbeitnehmerbezogenen Mitarbeiter-Teams bestehen aus Mitarbeitern mit jeweils im Schwerpunkt

- beratertischen und vermittelnden Aufgaben sowie
 - leistungrechtlichen und sachbearbeitenden Aufgaben.
- Für die Erledigung von Aufgaben, die in der Regel anlässlich des ersten Kontaktes mit dem Kunden bei dessen jeweiliger Vorsprache oder bei der Behandlung von einfachen Anliegen anfallen, wird in den Mitarbeiter-Teams jeweils ein Kundenbüro eingerichtet.
- Es ist zu erproben, ob innerhalb der Teams jeweils eine Fachkraft für die beratertischen und vermittelnden sowie die leistungrechtlichen und sachbearbeitenden Aufgaben oder eine gemeinsame „Erste Fachkraft“ anzusetzen ist.

Wesentliche Aufgabenschwerpunkte in den Mitarbeiter-Teams für arbeitgeberbezogene Aufgaben sind insbesondere die Information, Auskunft und Beratung über Geldleistungen an Arbeitgeber sowie die Gewährung dieser Leistungen.

Zur Bildung von Hochschulteams siehe Ausführungen unter Ziffer 1.1.

1.3 Behinderte

In den Mitarbeiter-Teams für diese Kundengruppe erfolgt die ganzheitliche Betreuung der Behinderten, bei denen besondere Hilfen zur beruflichen Rehabilitation erforderlich sind, sowie der schwerbehinderten und Gleichgestellten im Sinne des Schwerbehindertengesetzes. Auch behinderte Jugendliche in spezifischen Einrichtungen, beispielsweise in Förderschulen für Lernbehinderte, bei denen üblicherweise bzw. überwiegend im Laufe der Integration besondere Hilfen im o. a. Sinne erforderlich werden, sind diesen Mitarbeiter-Teams zugeordnet.

Es sind grundsätzlich Mitarbeiter-Teams vorgesehen, in denen die Anliegen sowohl der jugendlichen als auch der erwachsenen Behinderten behandelt werden. Unter dem Gesichtspunkt einer effizienten Organisationsgestaltung erhalten die sehr großen Arbeitsämter als Option die Möglichkeit, jeweils gesonderte Mitarbeiter-Teams für jugendliche und erwachsene Behinderte einzurichten.

Auch die Teams für Behinderte bestehen aus Mitarbeitern mit im Schwerpunkt

- beratertischen und vermittelnden Aufgaben sowie
 - leistungrechtlichen und sachbearbeitenden Aufgaben.
- Darüber hinaus ist auch in diesen Teams jeweils ein Kundenbüro vorgesehen.

1.4 Kindergeldberechtigte

Es wurde nicht als sinnvoll angesehen, neue Organisationsstrukturen für das Aufgabengebiet Kindergeld bereits vor der Klärung der Frage der personellen Ausstattung zu entwickeln. Als Folge der Rechtsänderungen durch das Jahressteuergesetz 1998 muß im Kindergeldbereich vom Jahre 1997 an Personal abgebaut werden. Die organisatorischen Strukturen im Aufgabengebiet Kindergeld werden deshalb im Zusammenhang mit der Anpassung der Personalkapazität an den geänderten Personalbedarf neu gestaltet.

1.5 Mitarbeiter-Teams für ordnungspolitische Aufgaben

Unabhängig von der Zuordnung zu einer Kundengruppe ist vorgesehen, auch für die ordnungspolitischen Aufgaben, insbesondere zur Bekämpfung des Leistungsmissbrauchs sowie der illegalen Beschäftigung Mitarbeiter-Teams einzurichten. Wegen des besonderen Charakters dieser Aufgaben bietet es sich nicht an, sie im Verbund mit den Dienstleistungen für die einzelnen Kundengruppen wahrzunehmen.

1.6 Zusätzliche Anmerkungen

Weitergehende Erläuterungen

- zur Bildung jeweils gesonderter Kundengruppen für Ausbildungsmarktpartner und Arbeitsmarktpartner,
 - zur Bearbeitung der Anliegen der Arbeitgeber sowie
 - zur Bildung der Kundengruppe Behinderte
- ergeben sich aus den Ausführungen im Anhang I.

2. Räumliche Dezentralisierung des Dienstleistungsangebots

Abgeleitet aus dem Organisationsziel „Dienstleistungen am Kunden orientieren“ wird angestrebt, die ortsnahe Versorgung der Kunden mit den Dienstleistungen der BA von der ländlichen Region bis in die Großstadt zu verbessern. Deshalb müssen Aufgaben stärker als bisher in der Fläche wahrgenommen werden. Die Grenzen der damit verbundenen Dezentralisierung sind jedoch erreicht, wenn Effizienzverluste bei der Aufgabenerfüllung eintreten oder die Qualität der Dienstleistungen beeinträchtigt wird. Es gilt daher der Leitgedanke

„soviel Kundennähe wie möglich, soviel zentrale Dienstleistung wie nötig“.

Den Kunden können die Dienstleistungen, bezogen auf die Ortsnähe, in drei Abstufungen angeboten werden: Durch

- unmittelbare Präsenz beim Kunden (z.B. Schule, Betrieb, Träger),
- dezentrales Angebot in Nebenstellen, in sonstigen ausgelagerten Stellen und mit Hilfe elektronischer Medien,
- zentrales Angebot im Hauptamt.

Ziel ist es, in den Nebenstellen soweit wie möglich kundenorientierte Mitarbeiter-Teams einzurichten. Auch hier sollen dem Kunden die wesentlichen Dienstleistungen möglichst vollständig angeboten werden. In Abhängigkeit von der Größe der Nebenstellen ist vorgesehen:

- Ein Standard-Dienstleistungsangebot im Regelfall,
- ein erweitertes (Standard-)Dienstleistungsangebot in großen Nebenstellen, und
- ein eingeschränktes Dienstleistungsangebot in kleinen Nebenstellen und
- ein Mindest-Dienstleistungsangebot in Kleinstnebenstellen.

Das Standard-Dienstleistungsangebot umfasst die wesentlichen kundenorientierten Dienstleistungen in der Nebenstelle, vorrangig Information, Beratung, Vermittlung und Gewährung der wichtigsten Geldleistungen.

Aus Gründen der verstärkten Kundenorientierung, der Effizienz der Aufgabenerfüllung und der höheren Arbeitszufriedenheit der Mitarbeiter sollte die Dezentralisierung großer Arbeitsämter durch die Bildung von Nebenstellen insbesondere in Großstädten fortgeführt werden.

Es wird auch künftig nicht möglich sein, alle Dienstleistungen flächendeckend über Nebenstellen effizient anzubieten. Deshalb werden Alternativen (z.B. mobile Einrichtungen) in die Überlegungen zur Dezentralisierung des Dienstleistungsangebots einbezogen.

3. Vereinfachung der Arbeitsabläufe durch vorgangorientierte Bearbeitung

Gleichzeitig mit der Bildung kundenorientierter Mitarbeiter-Teams wird eine Orientierung der Ablauforganisation an den Objekten - das sind die Anliegen der Kunden, z.B. Beratung, Anträge auf Leistungen, Stellen- bzw. Bewerberangebote - angestrebt. Die grundsätzlichen Geschäftsvorgänge wurden daraufhin mit dem Ergebnis untersucht, daß beim Großteil aller Geschäftsvorfälle eine objektorientierte Ablauforganisation erfolgen kann. Damit ist in der Regel eine ganzheitliche Bearbeitung innerhalb der Mitarbeiterteams einhergehend mit einer deutlichen Reduzierung von Schnittstellen möglich.

Die Detailergebnisse stehen in der Dokumentation der Projektarbeit zur Verfügung; sie dienen u.a. als Grundlage für die Vorbereitung und Durchführung von Modellversuchen.

4. Anpassung der Organisationsstrukturen in den Arbeitsämtern

Die künftige objektorientierte Aufgabenerfüllung des Arbeitsamts 2000 durch kundenorientierte Mitarbeiter-Teams - auch in den Nebenstellen - erfordert neue, ganzheitliche Organisationsstrukturen beim Leitungsgeschehen. Gebildet werden ein Kundenbereich und ein Verwaltungsbereich. Das Erfordernis flexibler organisatorischer Lösungen ist zu berücksichtigen.

Im Kundenbereich sind sämtliche kundenorientierten und ordnungspolitischen Aufgaben zusammengefaßt, die bisher in den Abteilungen AVUAB, Berufsberatung und in der Leistungsabteilung wahrgenommen werden.

Der Verwaltungsbereich stellt die personellen und infrastrukturellen Ressourcen zur Erledigung der Fachaufgaben bereit und entspricht von der Aufgabenstellung im Prinzip der heutigen Abteilung Verwaltung. Lösungsansätze für Struktur und Aufgaben dieses Bereichs sind noch zu entwickeln.

Zwischen Direktor und Mitarbeiter-Teams werden im Kundenbereich zwei Führungsebenen - Kundenbereichsleiter³⁾ und Teamleiter³⁾ - eingerichtet. Die neue Leitungsstruktur basiert auf dem Einliniensystem. Die Zuordnung der Mitarbeiter-Teams zu Teamleitern und Kundenbereichsleitern erfolgt nach dem Prinzip der Mengenteilung, d.h. die Zuordnung der Mitarbeiter-Teams wird nach dem Aufgabenvolumen und nicht primär nach der Art der Aufgaben (= Sparten) vorgenommen (vgl. Anhang II, Abb. 3).

In der Regel werden zwei Kundenbereichsleiter angesetzt; Richtwerte zum zahlenmäßigen Ansatz von Kundenbereichsleitern müssen ebenso wie konkrete Aufgabenbeschreibungen noch erarbeitet werden.

Um die erforderliche Fachkompetenz zu gewährleisten, ist auf der Ebene der Kundenbereichsleiter zusätzlich eine fachliche Schwerpunktsetzung vorgesehen. Ein Kundenbereichsleiter wird im Schwerpunkt für besondere Fragen aus dem Bereich „Markt“, ein anderer für ebenso wichtige Fragen aus dem Bereich „Recht“ zuständig sein. Sie beraten und unterstützen den Direktor und die ggf. anderen Kundenbereichsleiter entsprechend ihrem fachlichen Schwerpunkt. Daraus darf jedoch kein Mehrlinienensystem abgeleitet werden. Unmittelbare Eingriffe in den Verantwortungsbereich des jeweils anderen sind nicht statthaft. Klarheit und Eindeutigkeit der Zuständigkeiten im Einliniensystem bleiben unantastbar.

Die Kundenbereichsleiter werden bei der Wahrnehmung ihrer Führungsaufgaben von Teamleitern unterstützt, denen Mitarbeiter-Teams gesamtverantwortlich zugeordnet werden. Aufgabenbeschreibungen und Richtwerte zur Anzahl der unterstellten Teams müssen noch ausgearbeitet werden.

Die Nebenstellenleiter - zugleich Teamleiter - sind einem Kundenbereichsleiter zugeordnet. Sie sind gesamtverantwortlich für alle in ihrer Dienststelle angebotenen Dienstleistungen.

Die Leiter besonders großer Nebenstellen sind als Kundenbereichsleiter unmittelbar dem Direktor unterstellt. Eine Ausweitung der Leitungsspanne tritt dabei nicht ein. Kleine Nebenstellen, in denen kein vollständiges Mitarbeiter-Team angesetzt werden kann, sind einem Teamleiter des Hauptamtes oder einer anderen Nebenstelle zugeordnet. Neben der fachlichen Verantwortung für die Aufgabenerfüllung in der Nebenstelle obliegt dem Leiter kleiner Nebenstellen in diesem Falle insbesondere die Vertretung der Dienststelle nach außen.

Die organisatorische Zuordnung der Aufgabengebiete Ärztlicher Dienst, Psychologischer Dienst, Technischer Beratungsdienst und Bearbeitungsstelle für Angelegenheiten nach dem SGG sowie der Selbstinformationseinrichtungen ist noch nicht festgelegt.

5. Anpassung der Organisationsstrukturen in den Landesarbeitsämtern

Im Kontext des Projekts „Arbeitsamt 2000“ gilt es, auch die Organisation der Landesarbeitsämter (LAA) weiterzuentwickeln. Grundsätzlich erscheint eine Neustrukturierung erst dann sinnvoll, wenn sich die künftigen konkreten Organisationsstrukturen der Arbeitsämter im Rahmen der Modellprüfung abzeichnen. Deshalb wird die Neuorganisation der LAA im Detail solange zurückgestellt. Im folgenden wird auf den gegenwärtigen Diskussionsstand eingegangen.

³⁾ Bei den Begriffen Kundenbereichsleiter und Teamleiter handelt es sich noch um Arbeitstitel.



Die LAÄ konzentrieren sich künftig auf die Wahrnehmung von Kernfunktionen im Sinne von Kernaufgaben. Darunter sind insbesondere zu verstehen

- Controlling mit seinen Instrumenten zur Abstimmung der Planung (Zielsetzung, Zielvereinbarung) zwischen der Hauptstelle und den Landesarbeitsämtern auf der einen wie den Landesarbeitsämtern und Arbeitsämtern auf der anderen Seite, zur gemeinsamen Überprüfung des Grades der Zielerreichung, zur Steuerung sowie zur Durchführung der Fachaufsicht,
- allgemeine Servicefunktionen gegenüber den Arbeitsämtern - vor allem in den Bereichen Information, Beratung, Unterstützung - sowie
- Zusammenarbeit mit Landesregierungen und überregional auf Landesebene zuständigen Behörden und Verbänden.

Ausführungsaufgaben nehmen die LAÄ nur noch wahr, wenn dies unter Berücksichtigung der vier Organisationsziele, insbesondere Wirksamkeit und Wirtschaftlichkeit, geboten ist. Ansonsten erfolgt je nach Aufgabe eine Delegation auf Arbeitsämter, Sitzpunkt-Arbeitsämter oder Stützpunkt-LAA. Zu prüfen ist, welche Aufgaben wegen zentraler Erledigung in der Hauptstelle nicht mehr von den LAÄ wahrgenommen werden müssen.

Als Grundlage für organisatorische Weiterentwicklungen ist eine neue Beschreibung der Aufgaben der LAÄ noch zu erarbeiten. Diese soll insbesondere die Brückenfunktion der LAÄ zwischen Arbeitsämtern und Hauptstelle deutlich machen.

Zur Organisation der LAÄ wurden alternative Lösungsansätze erarbeitet. So sollte nach einem Vorschlag die Organisation der LAÄ die objektorientierte Struktur der Arbeitsämter widerspiegeln. Der andere Lösungsansatz sieht vor, die Fachlichkeit durch eine Mischform von objekt- und verrichtungsorientierter Organisation sicherzustellen: Es kommt darauf an, eine ausgewogene Lösung zu finden, die beiden Anliegen Rechnung trägt. Für alle LAÄ wird eine einheitliche Grundstruktur angestrebt. Die Anzahl der Abteilungen und der Referate ist im Zuge einer Straffung der Organisation zu reduzieren. Gebildet werden sollen neben der Geschäftsführung durch den Präsidenten und Vizepräsidenten zwei Bereiche, in denen die kundenorientierten und ordnungspolitischen Aufgaben wahrgenommen werden sollen sowie ein Bereich Verwaltung.

II. Ergänzende und unterstützende organisatorische Maßnahmen

1. Ausbau der Selbstinformationseinrichtungen

Die Selbstinformationseinrichtungen (SIE) sollen erweitert werden, um dem Anliegen der Kunden nach mehr Selbstbedienung zu entsprechen. Aufzubauen ist dabei auf bestehende Einrichtungen und Dienste, beispielsweise BIZ (Berufs-Informations-Zentrum) und SIS (Stellen-Informations-Service). Individueller Beratungsbedarf läßt sich durch den Ausbau der SIE nicht ersetzen, jedoch werden sich die Ansprüche der Kunden an die Mitarbeiter inhaltlich und qualitativ verändern. Das BIZ neuer Prägung soll den Kern und gleichzeitig die Klammer aller Selbstinformationseinrichtungen und -verfahren im Arbeitsamt bilden und ist daher - mehr noch als in der Vergangenheit - als kundengruppenübergreifende Einrichtung konzipiert.

Neben den zentralen Selbstinformationseinrichtungen mit ihrem notwendigen breiten Angebot vom Printmedium bis zum dialogfähigen EDV-System soll künftig die Versorgung in der Fläche optimiert werden. So ist an die Einrichtung von Selbstinformationsszonen im Wartebereich der kundenorientierten Mitarbeiter-Teams sowie an BIZ-Dependancen in größeren Nebenstellen gedacht.



Eine deutlich größere Rolle werden künftig die neuen interaktiven Medien spielen; AA online befindet sich bereits in der Erprobung. Das Informationsangebot, exemplarisch seien hier SIS und AIS (Arbeitgeber-Informations-Service) genannt, reicht nicht mehr nur in die Arbeitsamtsdienststellen, sondern bis in Unternehmen und die Wohnungen der Nutzer. Zugleich werden damit Teilbereiche des Selbstinformationsangebotes von den Öffnungszeiten der Dienststellen entkoppelt und ein Datenzugriff auch in den Abendstunden oder an Wochenenden ermöglicht.

Beim Ausbau der SIE ist ein Stufenkonzept angedacht, neue interaktive Medien sollen danach nur schrittweise entwickelt und eingeführt werden.

Umfangreichere **Baumaßnahmen** sollten nur dann vorgenommen werden, wenn sie im Zuge von grundlegenden Aus- und Neubauten **kostengünstig zu verwirklichen sind. Eine Ausnahme** davon sollten BIZ-Dependancen in größeren Nebenstellen von Flächenbezirken mit schlecht erreichbarem Hauptamt bilden.

2. Verlagerung von Organisations- und Budgetkompetenzen auf die Ebene der Arbeitsämter

Ausgangspunkt der Überlegungen war, daß sich die Hauptstelle und die Landesarbeitsämter auf ihre **Kernaufgaben** beschränken. Aufgaben, die in der jeweils nachgeordneten Instanz erledigt werden können, sollen auch dort bearbeitet und entschieden werden. Erforderlich ist daher eine möglichst weitgehende Delegation unter anderem von Organisations- und Budgetkompetenzen jeweils innerhalb des rechtlichen und haushaltsmäßigen Rahmens auf die Arbeitsämter, diese muß einhergehen mit der Übertragung von Verantwortung für lokale Organisationsmaßnahmen und den Ressourceneinsatz.

Organisationskompetenzen können beispielsweise durch die **Wahlmöglichkeit alternativer organisatorischer Lösungsansätze** - entsprechend den Bedürfnissen des jeweiligen Arbeitsamtes im Sinne von Optionen - verlagert werden. Auch **Rahmenvorgaben** anstelle detaillierter Organisationsregelungen sind ein geeignetes Mittel, um künftig mehr Flexibilität bei der Gestaltung der Organisation in den Arbeitsämtern zu erreichen.

Der Spielraum für die Verlagerung von Budgetkompetenzen stellt sich vergleichsweise geringer dar, da den Arbeitsämtern im Zuge gesetzlicher Pflichtleistungen bereits heute die Entscheidungskompetenz für mehr als die Hälfte ihres Haushaltsvolumens obliegt. Es bleiben somit die Bereiche, in denen Haushaltsmittel den Arbeitsämtern zugeteilt werden. Das sind insbesondere arbeitsmarktpolitische Ermessensleistungen sowie Sächliche Verwaltungsausgaben und Investitionen.

3. Einführung von Controlling

Um die Aufgaben einer großen Organisation wie der BA auch künftig erfolgreich bewältigen zu können, ist eine systematische, prozessorientierte Betrachtung ihres Verwaltungshandelns und seiner Rahmenbedingungen sowie dessen Optimierung erforderlich. Primär von Interesse ist hierbei der Grad der Zielerreichung, insbesondere beim Organisationsziel Effizienz, aber auch bei den weiteren Zielen Effektivität, Kunden- und Mitarbeiterorientierung mit ihren jeweiligen Unterzielen. Aufgabe von Controlling ist es, durch die Bereitstellung von entscheidungsrelevanten Informationen sowie die Übernahme von Analyseaufgaben durch Controlling-Instrumente alle Führungs- und Ausführungskräfte beim Regelkreis aus Planung, Kontrolle und Steuerung zu unterstützen und zu entlasten.

Die Planung orientiert sich an den Vorgaben aus dem politischen Raum und an der Ressourcenzuteilung; sie schließt Zielvereinbarungsprozesse und - wo dies nicht möglich ist - Zielvorgaben ein.

Wesentlichster Teil der Kontrolle ist die Abweichungsanalyse (Wo stehen wir im Augenblick? Wo wollten wir stehen?) mit ihrer zielorientierten Bewertung der Ursachen.

Eine Steuerung ist erforderlich, wenn aus der Abweichungsanalyse hervorgeht, daß Korrekturen erfolgen müssen; Steuerung ist Planung mit neuen Informationen. Aufbauend auf den „Produkten“ der BA - darunter werden hier Dienstleistungen verstanden, die von den Kunden nachgefragt werden und für deren Erbringung in der Regel Aufwendungen entstehen (d.h. „ein Preis bezahlt wird“) - wird ein Kennzahlensystem zur Messung des qualitativen und quantitativen Grades der Zielerreichung entwickelt.

Controlling soll in bestehenden Verfahren harmonisiert und weiterentwickelt und ggf. durch neue Instrumente ergänzt werden. Dabei sollen einzelne Instrumente, beispielsweise eine Kosten- und Leistungsrechnung und ein Planungsverfahren für das Arbeitsamt, mit Priorität entwickelt und in der Praxis erprobt werden.

Die so gewonnenen Informationen werden für die Führungskräfte über ein Führungsinformationssystem (FIS) und für die übrigen Mitarbeiter über einen geeigneten Service bereitgestellt. Dabei werden die im Rahmen des Projekts „Leistungsorientierte Führung“ gewonnenen Erkenntnisse nach erfolgreicher Erprobung in das FIS integriert.

Ein Fach- und Zentralcontrolling soll fest im Aufbau der BA verankert werden.

4. Ausbau der Informationsverarbeitung

Die Umsetzung des Konzepts „Arbeitsamt 2000“, das eine ganzheitliche objekt- bzw. prozessorientierte Vorgangsbearbeitung vorsieht, erfordert die Weiterentwicklung der Informations-Technik (IT). Lösungsansätze hierzu wurden aufgrund der Komplexität der Materie und des Erfordernisses, verstärkt informationstechnischen Sachverstand einzusetzen, nicht im Rahmen der Projektorganisation „Arbeitsamt 2000“ entwickelt. Die Weiterentwicklung der Informations-Technik ist jedoch mit dem Konzept „Arbeitsamt 2000“ abzustimmen und im Grundsatz „organisationsneutral“ zu gestalten.

Grundlegende Defizite der derzeitigen Anwendungssoftware - insbesondere uneinheitliche Anwendungs- und Datenstrukturen - können nur im Wege einer grundsätzlichen Neugestaltung behoben werden. Dies soll mit dem „IT-Konzept 2000“ erreicht werden. Eine gemeinsame BA-Datenbasis mit einheitlichen Benutzer- (Bediener-) Oberflächen ist dabei ein wesentliches Ziel. Unter anderem soll ein BA-Netz alle wesentlichen Dienste der BA unterstützen und Anwender sowie Daten miteinander verbinden. Zur Umsetzung ist ein Stufenkonzept vorgesehen.

III. Maßnahmen zur Schaffung der Rahmenbedingungen für die Realisierung der organisatorischen Veränderungen

1. Qualifizierung der Mitarbeiter und Führungskräfte

Der Erfolg grundlegender organisatorischer Veränderungen hängt nicht zuletzt von der Vorbereitung der Mitarbeiter auf ihre neue Aufgabe ab. Daher muß die Realisierung der Handlungsfelder des Konzepts „Arbeitsamt 2000“ von Qualifizierungsmaßnahmen begleitet werden.

Mit der Einrichtung kundenorientierter Mitarbeiter-Teams werden Veränderungen in der fachlichen, personalen, sozialen und organisatorischen Kompetenz der Mitarbeiter einhergehen. Dies ist bei der Anpassungsfortbildung, der Ausbildung und Aufstiegsfortbildung sowie bei der Personalauswahl zu berücksichtigen.

Für die Anpassungsfortbildung wurden verschiedene Maßnahmen konzipiert, die entsprechend einem Zeitraster sechs Monate vor Beginn der Implementierung des Konzepts „Arbeitsamt 2000“ in die Modellarbeitsämter beginnen. Bei der Umsetzung der Maßnahmen sind die bei den Mitarbeitern und Führungskräften bereits vorhandenen Kompetenzen zu berücksichtigen; dies bedeutet, daß Anzahl und Dauer der Maßnahmen jeweils an die individuellen Bedürfnisse und amtspezifischen Problemstellungen anzupassen sind. Die Maßnahmen werden entsprechend der jeweiligen Situation und der Entwicklung in den Arbeitsämtern zeitnah durchgeführt.

Die Rahmenpläne für die Aus- und Fortbildungswege im mittleren und gehobenen Dienst müssen an die veränderten Kompetenzanforderungen angepaßt werden; für den mittleren Dienst ist außerdem eine Anpassung der Ausbildungsordnung notwendig, die nicht von der BA allein erfolgen kann.

Die Einweisung der Nachwuchskräfte für die allgemeine Verwendung im höheren Dienst ist bereits in den letzten Jahren grundlegend reformiert worden, ein aktueller Handlungsbedarf besteht daher nicht. Bei der Personalauswahl sollen die veränderten Kompetenzanforderungen berücksichtigt werden.

2. Strategie der Personalentwicklung

Personalentwicklung ist die Integration aller Maßnahmen der Auswahl, Entwicklung und Förderung von Mitarbeitern, um den aktuellen und zukünftigen qualitativen und quantitativen Personalbedarf zu decken. Grundlegende Vorarbeiten wurden mit dem vom Vorstand gebilligten Personalentwicklungskonzept für den höheren Dienst (RdErI 49/96) bereits geleistet.

Das Personalentwicklungskonzept soll mit den Organisationszielen und den in den Leitlinien für Führung und Zusammenarbeit verankerten Grundsätzen in Einklang stehen. Die Personalführung in ihrer hierauf gerichteten Arbeit soll unterstützt werden, denn Fordern und Fördern von Mitarbeitern ist eine originäre Führungsaufgabe. Somit stellt die Personalentwicklungskonzeption eine weitere Säule zur Weiterentwicklung eines innovativen Personalmanagements dar.

Die folgenden Leitsätze werden der Personalentwicklung (PE) zugrundegelegt:

- Jeder Mitarbeiter sollte wissen, wo er steht, wie er gesehen wird und ob bzw. welche Entwicklungsmöglichkeiten er hat (vertikale PE; seine Eigeninitiative hierzu wird erwartet und gefördert).
- Förderung bedeutet nicht immer auch Beförderung. Viele nicht ausgeschöpfte Entwicklungsmöglichkeiten bietet so oft schon der derzeitige oder ein gleichwertiger Ansatz (horizontale PE).
- Jede erkannte und realisierte Entwicklungsmöglichkeit fördert die Zufriedenheit und die Motivation der Mitarbeiter.

Die bestehenden Instrumente der Personalplanung und Personalführung werden neu geordnet, durch zusätzliche Elemente ergänzt und zu einem ganzheitlichen System oder Verbund zusammengefaßt.

Die Schwerpunkte werden u. a. auf die folgenden Elemente gelegt:

- Institutionalisiertes Mitarbeitergespräch für alle Mitarbeiter
- Modernes Beurteilungssystem
- Personalentwicklungskonferenzen in den Arbeitsämtern und Landesarbeitsämtern
- Zielorientierte Verwendungsplanung und Verknüpfung von Personalplanung und Fortbildungsplanung
- Stufenweise Entwicklung eines Personalentwicklungscontrolling.

Das PE-Konzept als Teil der Personalpolitik soll dynamischen Organisationsprozessen entsprechend fortentwickelt werden, so daß es sowohl unter Berücksichtigung gegenwärtiger als auch zukünftiger Strukturen umgesetzt werden kann.

3. Gestaltung von Weisungen

Streben nach Einzelfallgerechtigkeit und Perfektionsdenken haben in der Vergangenheit zu sehr detaillierten Weisungen und zu einer hohen Vorschriftendichte geführt. Das behindert auch die angestrebte Delegation von Kompetenzen. Um ein wirksames und wirtschaftliches Verwaltungshandeln zu ermöglichen, sind soweit wie möglich detaillierte Handlungsanweisungen durch Zielvereinbarungen bzw. Rahmenregelungen zu ersetzen. Nur soweit es die zu regelnde Materie zwingend erfordert, insbesondere um eine einheitliche Rechtsanwendung sicherzustellen, sind detaillierte Handlungsanweisungen herauszugeben.

Die Praxis soll in der Regel bei der Erstellung zentraler Weisungen der Hauptstelle bzw. regionaler Weisungen der LAA beteiligt werden.

Die Möglichkeiten der EDV bei Distribution und Zugriff auf Weisungen sollen genutzt werden; dadurch sind Einsparungen möglich.

Eine Reduzierung des Gesamt-Weisungsvolumens in der BA wird angestrebt.

4. Anpassung der Infrastruktur, insbesondere Bau- und Liegenschaftswesen

Die im Projekt „Arbeitsamt 2000“ entwickelten organisatorischen Veränderungen haben auch Auswirkungen auf die Infrastruktur, vor allem der Arbeitsämter mit ihren Nebenstellen. Da das Organisationskonzept in erster Linie in bereits vorhandenen Gebäuden umzusetzen ist, soll im Grundsatz eine enge Orientierung an den gegebenen Baulichkeiten erfolgen.

Die Mitarbeiter-Teams für die einzelnen Kundengruppen bzw. für den Verwaltungsbereich sollen jeweils zusammenhängend und auf einer Ebene untergebracht werden. Besonderheiten der Kundengruppe Behinderte (möglichst ebenerdige Unterbringung der Mitarbeiter-Teams) und der zentralen Selbstinformationseinrichtungen (separater Zugang erforderlich) sind zu berücksichtigen.

Um die neue Organisationsform in den Dienststellen realisieren zu können, werden hauptsächlich Umbaumaßnahmen in bestehenden Gebäuden erforderlich sein.

Bei Neubauten muß noch stärker als bisher auf eine flexible Nutzungskonzeption geachtet werden, um künftigen Entwicklungen ohne großen Aufwand entsprechen zu können.

D Wirtschaftlichkeitsbetrachtung (WiBe) für das Projekt „Arbeitsamt 2000“

Zum Projekt „Arbeitsamt 2000“ erarbeitet die Verwaltung als Orientierungshilfe eine Wirtschaftlichkeitsbetrachtung (WiBe). Dabei werden die Empfehlungen des Bundesministers des Innern für die Durchführung von Wirtschaftlichkeitsbetrachtungen im Bereich der Informationstechnik (IT-WiBe) in der öffentlichen Verwaltung zur Orientierung herangezogen.

Die Wirtschaftlichkeit wird in der WiBe aus dem Verhältnis der durch die neue Organisationsform entstehenden Kosten und dem eintretenden Nutzen abgeleitet.

- Monetär quantifizierbare Kosten und Nutzen Größen bilden die Wirtschaftlichkeit im engeren Sinne. Dabei ist zwischen haushaltswirksamen Mehrbelastungen bzw. Einsparungen und kalkulatorischen Positionen, die sich haushaltsmäßig nicht (zusätzlich) auswirken, zu unterscheiden. Die zeitliche Verteilung der Kosten- und der Nutzenseite fließt in den Bewertungsprozeß ein.
- In der WiBe im weiteren Sinne werden - entsprechend den Organisationszielen Wirksamkeit, Mitarbeiter- und Kundenorientierung des Projektes „Arbeitsamt 2000“ - die nicht monetär quantifizierbaren Nutzeffekte durch eine Nutzwertanalyse berücksichtigt.

Die Ergebnisse einer WiBe, insbesondere bei der Berücksichtigung monetärer Nutzen-Kriterien, sind abhängig von der Genauigkeit der zugrunde gelegten Daten. Schon nach den ersten Ansätzen wird deutlich, wie schwierig überhaupt eine Erhebung der Kosten wäre. Bei einer Grobbeschreibung können diese zwangsläufig nur erste Anhaltswerte darstellen, die zum großen Teil geschätzt werden müßten, oder überhaupt zur Zeit noch nicht erfaßt werden (z.B. Kosten für Baumaßnahmen). Nahezu unüberwindlich sind gegenwärtig die Probleme, den künftigen Nutzen nachvollziehbar und planbar zu quantifizieren.

Aus diesen Gründen ist es derzeit nicht möglich, eine WiBe unter Berücksichtigung aller relevanten Faktoren zu erstellen.

Bildung von Kundengruppen für Ausbildungsmarktpartner und Arbeitsmarktpartner

Wie mit der Bezeichnung Ausbildungsmarktpartner und Arbeitsmarktpartner bereits zum Ausdruck kommt, gehören zu diesen Kundengruppen neben den Berufsanfängern und Arbeitnehmern jeweils auch die Arbeitgeber/Betriebe. Damit werden die beiden Seiten des Ausbildungsmarkts und des Arbeitsmarkts jeweils zu gemeinsamen Kundengruppen zusammengefaßt.

Es bestanden Überlegungen, die Kundengruppen Ausbildungsmarktpartner und Arbeitsmarktpartner in eine Gruppe zu integrieren, da beide Kundengruppen das Arbeitsamt mit ähnlichen Anliegen in Anspruch nehmen - zum Beispiel bei der Vermittlung von Ausbildungsstellen und Arbeitsstellen. Von einer Zusammenlegung wurde jedoch insbesondere abgesehen, weil die mit einer integrierten Betreuung von Ausbildungsmarktpartnern und Arbeitsmarktpartnern verbundene fachliche Komplexität der Aufgabenstellung (einschließlich Leistungsgewährung) gegenwärtig nicht beherrschbar erscheint.

Bearbeitung der Anliegen der Arbeitgeber

Die entwickelten Lösungsansätze sehen vor, daß für die unterschiedlichen Anliegen der Arbeitgeber - zum Beispiel die Besetzung von Ausbildungsstellen einerseits und von Arbeitsstellen andererseits - auch künftig verschiedene Stellen im Arbeitsamt zuständig sind. Um die Zahl der Ansprechpartner für Arbeitgeber/Betriebe im Arbeitsamt zu reduzieren, wurde eingehend darüber diskutiert, ob Arbeitgeber/Betriebe bei der Entwicklung organisatorischer Lösungen als eine gesonderte Kundengruppe angesehen werden sollten. Hiervon wurde jedoch aus folgenden Gründen abgesehen:

- Nach dem Ergebnis einer Studie der Gesellschaft für Sozialforschung und Marktforschung mbH (SINUS) hat für Arbeitgeber/Betriebe die Erwartung, im Arbeitsamt „kompetente“ Gesprächspartner vorzufinden, einen höheren Stellenwert als der Wunsch nach möglichst wenigen Ansprechpartnern. So erwartet man vom Arbeitsamt nicht, daß es in den Betrieb immer nur ein und denselben Mitarbeiter entsendet; insbesondere in Großbetrieben möchte man je nach Problemlage möglichst von Spezialisten besucht werden.

Damit sich die Fachkräfte weitgehend auf die besonderen Fragestellungen des Ausbildungsmarktes oder des Arbeitsmarktes spezialisieren und ihre Fachkompetenz für den jeweiligen Bereich erhalten und ausbauen können, erschien es insbesondere wegen der umfangreichen und differenzierten Gesamthematik zweckmäßig, auch bei den Arbeitgeberanliegen nach den beiden Marktsegmenten Arbeitsmarkt und Ausbildungsmarkt zu differenzieren.

Darüber hinaus sollen die Beratungs- und Vermittlungsfachkräfte sowohl über intensive Kenntnisse der Bewerber- als auch der Stellenseite des Ausbildungs- oder Arbeitsmarktes verfügen. Deshalb wurde bei der Bildung von Kundengruppen eine Trennung der Anliegen von Berufsanfängern bzw. Arbeitnehmern auf der einen und Arbeitgebern auf der anderen Seite vermieden.

- In großen und mittleren Betrieben sind für die unterschiedlichen Anliegen, wie zum Beispiel für Personaleinstellungen oder Nachwuchs- und Ausbildungsfragen, in der Regel verschiedene Stellen zuständig, die im Arbeitsamt - bei Aufteilung der Anliegen - jeweils einen „eigenen Ansprechpartner“ haben. Das Problem mehrerer Ansprechpartner stellt sich - wenn überhaupt - somit in erster Linie nur für kleinere Betriebe bei unterschiedlichen Anliegen. Gegenüber der

noch Anhang I

genannten höheren Fachkompetenz der Ansprechpartner ist dieser Aspekt eher als nachrangig anzusehen.

- Im übrigen ist der Umstand, sich zur Erledigung seiner Anliegen an verschiedene Stellen wenden zu müssen, für Arbeitgeber/Betriebe im Regelfall weitaus unproblematischer als für Berufsanfänger, Arbeitnehmer und Behinderte. Durch die meist langfristige Zusammenarbeit mit dem Arbeitsamt sind den Arbeitgebern/Betrieben vielfach die verschiedenen Ansprechpartner für ihre jeweiligen Anliegen bekannt.

Unabhängig vom Verzicht auf eine gesonderte Kundengruppe Arbeitgeber ist für die Kundengruppe Arbeitsmarktpartner im Regelfall vorgesehen, die arbeitnehmerbezogenen und arbeitgeberbezogenen Aufgaben zu trennen und jeweils gesonderten Mitarbeiter-Teams zuzuordnen. Damit erfolgt - mit Ausnahme der stellenorientierten Arbeitsvermittlung, die in den Mitarbeiter-Teams für arbeitnehmerbezogene Aufgaben angesiedelt ist - eine Zusammenfassung der Arbeitgeberanliegen. Im Sinne einer Option ist darüber hinaus vorgesehen, den Arbeitsämtern die Möglichkeit einzuräumen, auch die stellenorientierte Arbeitsvermittlung (einschließlich der Akquisition von Arbeitsstellen) bei den arbeitgeberbezogenen Mitarbeiter-Teams anzusiedeln.

Bildung der Kundengruppe Behinderte

Mit der Einrichtung einer besonderen Kundengruppe Behinderte werden die bisher in drei verschiedenen Abteilungen im Arbeitsamt angesiedelten Aufgaben für jugendliche und erwachsene Behinderte, Schwerbehinderte und Rehabilitanden zusammengefaßt. Dadurch sollen den Behinderten die Leistungen ganzheitlich an einer Stelle mit hoher fachlicher Kompetenz angeboten und die Arbeitsabläufe rationaler und effektiver gestaltet werden.

Eine im Rahmen der Projektarbeit vorgeschlagene Integration der Behinderten in die Kundengruppen Ausbildungsmarktpartner und Arbeitsmarktpartner wurde nicht weiterverfolgt. Maßgebend dafür war, eine fachliche Überforderung der Mitarbeiter zu vermeiden.

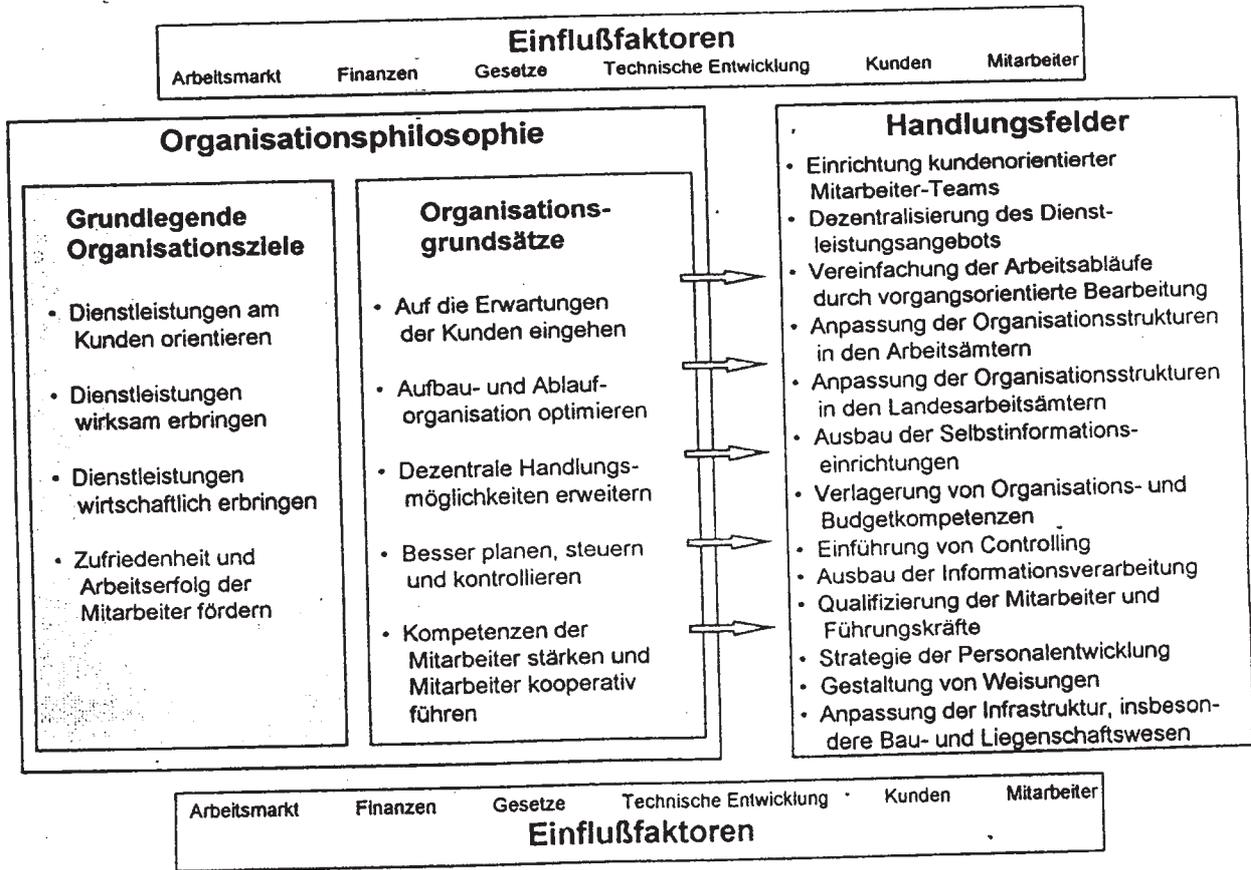
Es wird nicht verkannt, daß die Aufgabenstellung der Mitarbeiter-Teams sehr komplex ist und entsprechend hohe Anforderungen an die Mitarbeiter stellt. Deshalb wurde geprüft, ob es zweckmäßig ist, den Mitarbeiter-Teams für Behinderte im Sinne einer Minimalübung lediglich Aufgaben zuzuordnen, die im unmittelbaren Zusammenhang mit der Gewährung von berufsfördernden und ergänzenden Leistungen zur Rehabilitation und der Durchführung des Schwerbehindertengesetzes stehen.

Die Prüfung hat ergeben, daß mit einer Minimallösung den Organisationszielen der BA in nur geringem Maße Rechnung getragen werden kann. Bei einer Realisierung dieser Lösung wären im Vergleich zu der o. a. umfassenden Aufgabenzuordnung keine ganzheitliche Behandlung der Anliegen der Behinderten und eine deutliche Zunahme von Schnittstellen zu verzeichnen. Viele Schnittstellen beeinträchtigen das Erreichen der Organisationsziele, insbesondere das Ziel der Kundenorientierung. Schnittstellen ergeben sich - soweit diese Aufgaben nicht in den Teams für Behinderte bearbeitet werden - vor allem bei der

- Vermittlung von Rehabilitanden,
 - Gewährung von Lohnersatzleistungen und
 - Gewährung von nicht behindertenspezifischen Leistungen,
- weil für die Anliegen der Behinderten zusätzlich Teams für die Kundengruppe Arbeitsmarktpartner zuständig wären.

"Arbeitsamt 2000"

Kurzfassung „Arbeitsamt 2000“



Anhang II
Abbildung 1

noch Anhang I

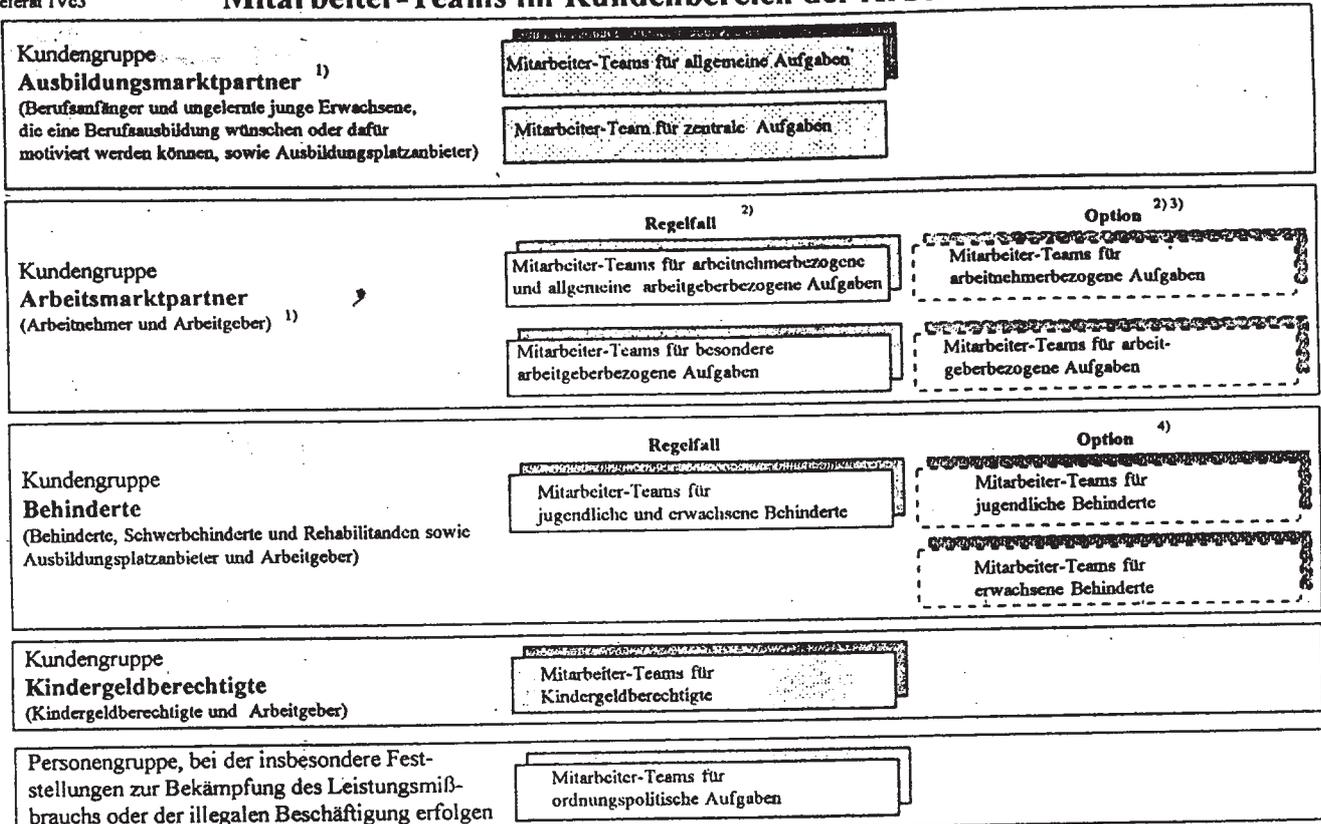
Zusätzliche Schnittstellen würden weiterhin auftreten bei der Durchführung der Reha-Maßnahmen, insbesondere bei der Anpassung des Gesamtplans, bei der Betreuung der Rehabilitanden, bei Abbrüchen von Maßnahmen, bei Abruf von Teilnehmern, bei der Modifizierung des Reha-Ziels, bei der Gewährung von zusätzlichen Hilfen, bei der Gewährung von Lohnersatzleistungen betreffenden Änderungen sowie bei der Eingliederung der Behinderten auf dem Arbeitsmarkt.

Bei Realisierung der o.a. umfassenden Aufgabenzuordnung und damit einer weitgehenden gänzlichen Aufgabenerledigung ist es hingegen möglich, im Vergleich zum derzeitigen IST-Zustand die Schnittstellen deutlich zu reduzieren.

Im Falle einer Minimallösung wäre es darüber hinaus nicht möglich, in kleinen Arbeitsämtern (bis etwa 250 Plankräften) funktionsfähige Teams, bei denen auch eine gegenseitige Vertretung der Mitarbeiter sichergestellt ist, einzurichten.

Im Interesse einer wirtschaftlichen Aufgabenerledigung ist bis zum Beginn von Modellversuchen jedoch gesondert zu untersuchen, ob die Aufgaben „Vermittlung in Arbeit“ und „Gewährung von Lohnersatzleistungen bei Arbeitslosigkeit“ zusätzlich auch in den Mitarbeiter-Teams für Behinderte vorgehalten werden müssen oder diese beiden Bereiche ohne Nachteile für die Behinderten von den Mitarbeiter-Teams der Kundengruppe Arbeitsmarktpartner mit wahrgenommen werden können. Dies gilt nicht für die Vermittlung in Ausbildung. Ein Beurteilungskriterium wird dabei sein, in welcher Häufigkeit diese Aufgaben im Falle ihrer Zuordnung in den Teams für Behinderte anfallen werden.

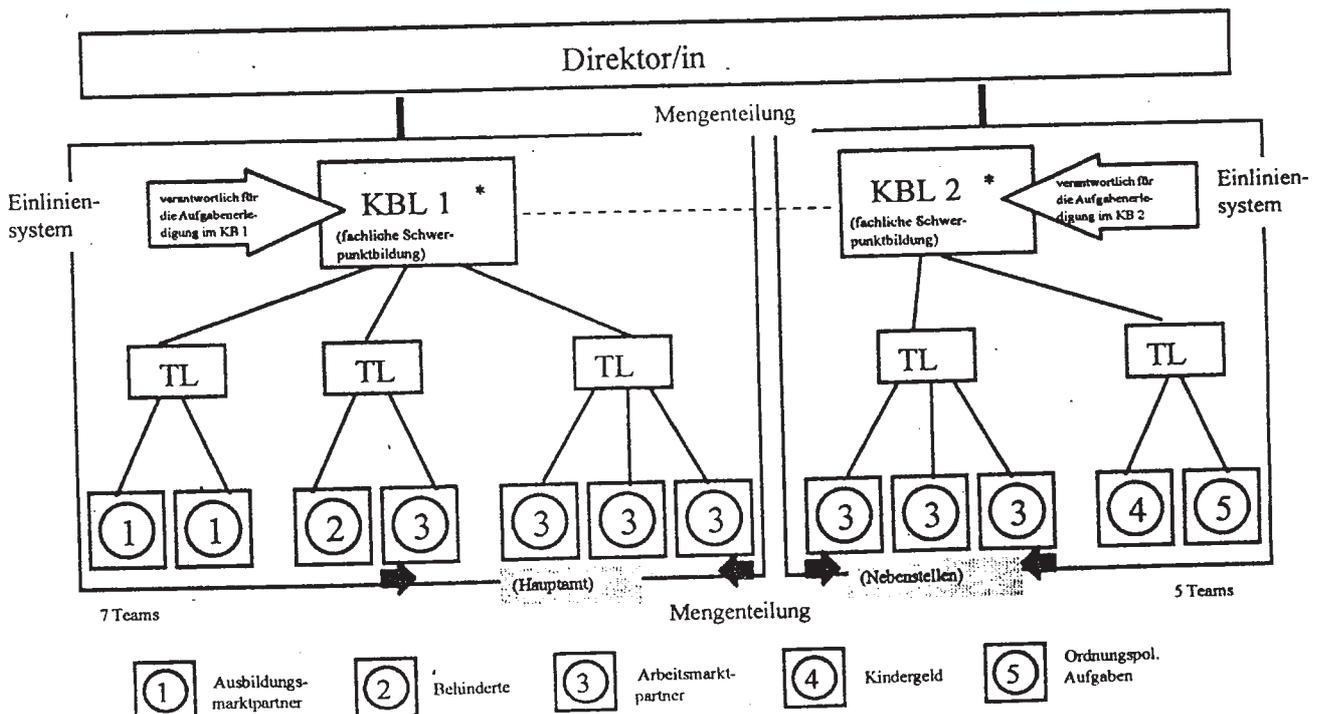
Kurzfassung Arbeitsamt 2000



- 1) Für die Betreuung der Studierenden und Hochschulabsolventen werden in Arbeitsamtsbezirken mit mehr als 20.000 Studierenden Hochschulteams eingerichtet, die aus Mitarbeitern der jeweiligen Teams für die Kundengruppe "Ausbildungsmarktpartner" und "Arbeitsmarktpartner" bestehen. In Bezirken mit weniger als 20.000 Studierenden bleibt es den Direktoren der Arbeitsämter überlassen, ein Hochschulteams zu bilden (entsprechend RdErl 46/96 Abschnitt III).
- 2) Im Rahmen der Modellversuche ist die Zuordnung sonstiger Aufgaben zu prüfen und nach der Modellierung darüber endgültig zu entscheiden.
- 3) im Sinne des Lösungsansatzes A (RdErl 24/94)
- 4) möglich unter dem Gesichtspunkt einer effizienten Organisationsgestaltung nur in sehr großen Arbeitsämtern

Abbildung 2
noch Anhang II

Gesamtübersicht der Führungsstrukturen im Arbeitsamt (beispielhaft)



* Je nach Größe des Arbeitsamtes sollen im Regelfall zwei Kundenbereichsleiter angesetzt werden. In jedem Arbeitsamt mit mindestens zwei Kundenbereichsleitern ist zur Gewährleistung der notwendigen Fachlichkeit eine ergänzende Schwerpunktbildung für die Bereiche "Markt" und "Recht" vorzusehen ("Rucksack"). Das bedeutet aber keine unmittelbare Einflussnahme der Kundenbereichsleiter auf die Teamleiter bzw. Teams der jeweils anderen Kundenbereiche. Die fachliche Schwerpunktbildung ist bei der Personalplanung und -auswahl zu berücksichtigen.

Kurzfassung Arbeitsamt 2000

Abbildung 3
noch Anhang II

3 Herausnehmbare Übersichten über alle in den Kapiteln 3 und 4 verwandten Variablenbezeichner⁵⁶

⁵⁶ in dieser elektronischen Form stattdessen auf den folgenden Seiten eingebracht

arbD	Anzahl der PCs, an denen Arbeitsvermittlung durchgeführt wird. Sie entspricht der Anzahl der Sachbearbeiter für die Arbeitsvermittlung (bundesweit).
AS	Gesamtzahl der bundesweit registrierten und in Datenbanken gespeicherten Arbeitsuchenden
AS+	Gesamtanzahl Arbeitsuchender, die bundesweit jährlich neu registriert werden
AT	Anzahl der Tage, in denen ein Arbeitsamt im Jahr für Publikum geöffnet ist
ATS	Anzahl der Arbeitstage eines Sachbearbeiters, an denen er seine Aufgabe durchführt
f	durchschnittliche jährliche Anzahl der bewerberorientierten Folge-Suchläufe falls ein Bewerber ein komplettes Jahr arbeitsuchend bleibt
fa	pauschalisiert wählbarer Faktor zur Beschreibung des Aufwands aus stellenorientierten Bewerbervermittlungen, die nicht mit Neuerfassung zusammenhängen
fc	durchschnittliche jährliche Anzahl der bewerberorientierten Folge-Suchläufe in Fall c, $c \in \{1, 2, 3\}$ wobei Fall 1 alle periodisch stattfindenden Suchläufe umfaßt und die Fälle 2 und 3 alle sonstigen (entweder durch den Arbeitsvermittler oder Arbeitsuchenden ausgelöst), sofern ein Bewerber ein komplettes Jahr arbeitsuchend bleibt
FmS	durchschnittlich täglich in einem „mittleren“ AA stattfindende bewerberorientierte Folgesuchläufe
FS	durchschnittliche Gesamtanzahl an täglich bundesweit stattfindenden bewerberorientierten Folgesuchläufen
I_{SZ}	$I_{SZ} = 1$ für alle Szenarien, deren Ordnungszahl in der Menge SZ , $SZ \subseteq \{2, \dots, 7\}$ aufgeführt ist, ansonsten 0. Verkürzt werden die Mengenklammern von SZ weggelassen.
K	durchschnittliche Anzahl gefundener Stellen je Bewerber, wobei angenommen wird, daß K für die Szenarien 1, 2 und 4 - 7 (bis auf zu vernachlässigende Abweichungen durch die eingeschränkten Pendelbereiche an den LAA-Grenzen) identisch ist
K1	durchschnittlicher prozentualer Anteil lokal gefundener Stellen je Bewerber von K
K2	durchschnittlicher prozentualer Anteil der im Pendelbereich gefundenen Stellen je Bewerber von K
K3	durchschnittlicher prozentualer Anteil zentral gefundener Stellen je Bewerber von K
KNc	mit $c \in \{1, 2, 3\}$, durchschnittlich jährlich gefundene Anzahl an Stellen bei bewerberorientierten Folge-Suchläufen im Fall c, falls ein Bewerber ein komplettes Jahr arbeitsuchend bleibt. Fall 1 umfaßt alle periodisch stattfindenden Suchläufe, die Fälle 2 und 3 alle sonstigen, entweder durch den Arbeitsvermittler oder Arbeitsuchenden ausgelösten Fälle; KNc gilt nicht für $SZ 3$.
KNT	durchschnittliche Anzahl gefundener Stellen je bewerberorientiertem Folge-Suchlauf, außer bei $SZ 3$
KR	durchschnittliche Anzahl gefundener Stellen je Bewerber im Szenario 3
KR1	durchschnittlicher prozentualer Anteil lokal gefundener Stellen in $SZ 3$ von KR
KR2	durchschnittlicher prozentualer Anteil der im Pendelbereich gefundenen Stellen in $SZ 3$ von KR
KR3	prozentualer Anteil zentral gefundener Stellen in $SZ 3$ von KR
KRNc	mit $c \in \{1, 2, 3\}$, durchschnittlich jährlich gefundene Anzahl an Stellen bei bewerberorientierten Folge-Suchläufen im Falle c, falls ein Bewerber ein komplettes Jahr arbeitsuchend bleibt. Fall 1 umfaßt alle periodisch stattfindenden Suchläufe, die Fälle 2 und 3 alle sonstigen, entweder durch den Arbeitsvermittler oder Arbeitsuchenden ausgelösten Fälle; gilt nur für $SZ 3$
KRNT	durchschnittliche Anzahl gefundener Stellen je bewerberorientiertem Folge-Suchlauf bei $SZ 3$
I	durchschnittlicher prozentualer Anteil an den gefundenen Bewerbern L, der von den Sachbearbeitern als geeignet akzeptiert wird
L	durchschnittliche Anzahl gefundener Bewerber je Stelle* [* s. Endnote]. In den Modellrechnungen wird angenommen, daß stets gilt: $L=O$ (O Obergrenze)
I1	durchschnittlicher prozentualer Anteil der nur durch den Sachbearbeiter für einen Vermittlungsvorschlag akzeptierten Stellenangebote in bewerberorientierten Folgesuchläufen
L1	prozentualer Anteil lokal gefundener Bewerber (vgl. Endnote) von L
I2	durchschnittlicher prozentualer Anteil der durch Sachbearbeiter und Arbeitsuchenden gemeinsam akzeptierten Stellenangebote in bewerberorientierten Folgesuchläufen
L2	prozentualer Anteil der durch Anfragen im Pendelbereich in $SZ 4$ und $SZ 5$ gefundenen Bewerber von L (vgl. Endnote), $L2 = 100 - L1$

LAA	Anzahl der Rechner auf Landesarbeitsamtsebene in Modell 6. Diese ist identisch zur bundesweiten Anzahl an LAÄ.
lb	durchschnittlicher prozentualer Anteil der gefundenen Bewerber, denen Stellen aufgrund von bewerberorientierten Folgesuchläufen angeboten werden
lk	durchschnittlicher prozentualer Anteil für einen Vermittlungsvorschlag akzeptierter Stellenangebote in bewerberorientierten Folgesuchläufen
mEntfAA	mittlere Entfernung zwischen den AA-Rechnern
mEntfAL	mittlere Entfernung des AA-Rechners zum LAA-Rechner
mEntfAZ	mittlere Entfernung eines AA-Rechners zum Zentralrechner
mEntfLZ	mittlere Entfernung eines LAA-Rechners zum Zentralrechner
mNRL1	mittlere Nachrichtenlänge eines Auftrages
mNRL2	mittlere Nachrichtenlänge des Ergebnisses eines Auftrages
mNRLg	mittlere Nachrichtenlänge aus Auftrag inklusive Ergebnis, $mNRLg = mNRL1 + mNRL2$
N	Anzahl der Stellen, die bundesweit jährlich neu besetzt werden sollen. Diese Stellen werden gespeichert.
O	durchschnittliche Obergrenze für die Anzahl der zu suchenden Bewerber
O*	Obergrenze O* , begrenzt die Anzahl der Stellen, die als Suchergebnis angezeigt werden
p	durchschnittliche Anzahl an AA-Rechnern im Pendelbereich eines BewAs
q	durchschnittliche Anzahl der Arbeitsämter im Pendelbereich von Stellen
r1	prozentualer Anteil der stellenorientierten Bewerbungen, die lokal vollständig bis zur Obergrenze gelöst werden können
r2	prozentualer Anteil der stellenorientierten Bewerbungen, die entfernt angestoßen werden müssen um (ggf.) die Obergrenze noch zu erreichen, $r2 = 100 - r1$
RAA	Anzahl der Rechner auf Arbeitsamtsebene. Diese ist nicht identisch zur bundesweiten Anzahl an Arbeitsämtern, da manche Ämter kooperieren.
s1	gibt den prozentualer Anteil der Fälle wieder, bei denen in der bewerberorientierten Stellensuche nur lokal gesucht wird
s2	gibt den prozentualen Anteil der Fälle wieder, bei denen in der bewerberorientierten Stellensuche auch im Pendelbereich gesucht wird
s3	gibt den prozentualer Anteil der Fälle wieder, bei denen in der bewerberorientierten Stellensuche für besondere Berufsgruppen auch zentral gesucht wird
TAS+	durchschnittlich täglich bundesweit hinzukommende Zahl an Arbeitssuchenden, $TAS+ = AS+ / AT$
TEber	durchschnittliche tägliche Anzahl an Erstkontakten je Sachbearbeiter
TmAS+	durchschnittlich täglich in einem „mittleren Arbeitsamt“ hinzukommende Zahl an Arbeitssuchenden, $TmAS+ = AS+ / AT / RAA$
TmN	durchschnittlich täglich neu zu erfassende Stellenangebote in einem mittleren Arbeitsamt, $TmN = TN / RAA = N / AT / RAA$
TN	durchschnittliche Anzahl der täglich im Bundesgebiet neu zu erfassenden SteAs, $TN = N/AT$
transErfB	durchschnittliche Anzahl an Transaktionen zur Ersterfassung von Bewerberdaten, $transErfB = transSuAS + transSpPers + transSpBer$
transErfBP	durchschnittliche Anzahl an Transaktionen zum Übersenden der Bewerberdaten in den Pendelbereich
<i>TransErfmB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zur Erfassung der Bewerberdaten in einem „mittleren“ AA
<i>TransErfzenB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zur Erfassung neuer Bewerberdaten im zentralen Szenario 1
<i>TransFolmN</i>	durchschnittliche Anzahl an täglich in einem mittleren Arbeitsamt entstehenden Transaktionen aufgrund von Folgebearbeitungen aus stellenorientierten Bewerbervermittlungen
<i>TransFolzenN</i>	durchschnittliche Anzahl an täglich entstehenden Transaktionen im zentralen Fall aufgrund von Folgebearbeitungen aus stellenorientierten Bewerbervermittlungen
transLöB	durchschnittliche Anzahl an Transaktionen zum Löschen von Bewerberdaten bzw. zu deren „gelöscht“-Kennzeichnung
<i>TransLöLAAN*</i>	durchschnittliche Anzahl täglicher Transaktionen auf einem LAA-Rechner zum Löschen replizierter SteAs (SZ 6)

<i>TransLömB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zum Löschen von Bewerberdaten bzw. zu deren „gelöscht“-Kennzeichnung in einem „mittleren“ AA
<i>TransLömN</i>	durchschnittliche Anzahl täglicher Transaktionen zum Löschen von Stellendaten in einem „mittleren“ Arbeitsamt
transLöN	durchschnittliche Anzahl der Transaktionen zur Kennzeichnung eines SteAs als „vermittelt“
transLöN*	durchschnittliche Anzahl der Transaktionen zum wirklichen Löschen eines SteAs (wird nur für Replikate vorgenommen)
<i>TransLözenB</i>	durchschnittlich erforderliche tägliche Anzahl an Transaktionen zum Löschen von Bewerberdaten bzw. zu deren „gelöscht“-Kennzeichnung im zentralen Szenario 1
<i>TransLözenN</i>	durchschnittliche Anzahl täglich anfallender Transaktionen zum Löschen von zentral gelagerten Originalstellendaten in SZ 1
<i>TransLözenN*</i>	durchschnittliche Anzahl täglicher Transaktionen zum Löschen zentral replizierter SteAs
<i>TransRlmS</i>	durchschnittliche Anzahl der täglich lokal in einem „mittleren“ Amt entstehenden Transaktionen durch das Lesen von SteAs des eigenen Datenbestandes nach bewerberorientierten Folgesuchläufen
<i>TransRrLAAS</i>	durchschnittliche Anzahl der täglich auf LAA-Rechnern entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen
<i>TransRrmS</i>	durchschnittliche Anzahl der täglich lokal in einem „mittleren“ Amt entstehenden Transaktionen durch Leseanforderungen auf SteAs von entfernten Rechnern
<i>TransRrzenS</i>	durchschnittliche Anzahl der täglich zentral entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen
<i>TransSachmB</i>	durchschnittliche Anzahl an täglich in einem mittleren Arbeitsamt entstehenden Transaktionen, die durch das Abarbeiten der Teamdateien durch die Sachbearbeiter aufgrund der stellenorientierten Bewerbervermittlung entstehen. Angegebene Indizes weisen darauf hin, daß die Formel nur für das diesem Index entsprechende Szenario gilt.
<i>TransSachzenB</i>	durchschnittliche Anzahl an täglich entstehenden Transaktionen, die durch das Abarbeiten der Teamdateien durch die Sachbearbeiter im zentralen Fall aufgrund der stellenorientierten Bewerbervermittlung entstehen
transSpBer	durchschnittliche Anzahl an Transaktionen zum Speichern der Berufsdaten
<i>TransSpLAAN*</i>	durchschnittliche Anzahl täglicher Transaktionen zur Speicherung von Replikaten neu erfaßter Stellenangebote auf einem LAA-Rechner (SZ 6)
<i>TransSpmN</i>	durchschnittliche Anzahl täglicher Transaktionen zur Speicherung neuer Stellendaten in einem „mittleren“ Arbeitsamt
<i>transSpN</i>	durchschnittliche Anzahl der Transaktionen zur Erfassung eines SteAs
transSpN*	durchschnittliche Anzahl der Transaktionen zur reinen Speicherung eines SteAs inklusive des Verweises auf die zugehörigen Betriebsdaten
transSpPers	durchschnittliche Anzahl an Transaktionen zum Speichern von Personendaten
<i>TransSpzenN</i>	durchschnittliche Anzahl täglich anfallender Transaktionen zur zentralen Speicherung von SteAs
transSuAS	durchschnittliche Anzahl an Transaktionen zum Suchen von Personendaten
transSuFi	durchschnittliche Anzahl der Transaktionen zum Lesen der Betriebsdaten
transSufol	durchschnittliche Anzahl der zum lokalen Suchen bei bewerberorientierten Folgesuchläufen benötigten Transaktionen
<i>TransSufolLAA</i>	durchschnittliche Anzahl der täglich auf einem LAA-Rechner entstehenden Transaktionen durch bewerberorientierte Folgesuchläufe
<i>TransSufolm</i>	durchschnittliche Anzahl der täglich auf einem Rechner eines „mittleren“ Arbeitsamtes entstehenden Transaktionen durch bewerberorientierte Folgesuchläufe
<i>TransSufolzen</i>	durchschnittliche Anzahl der täglich zentral entstehenden Transaktionen durch bewerberorientierte Folgesuchläufe
<i>TransSuLAAV</i>	durchschnittliche Anzahl an Transaktionen, die täglich zur bewerberorientierten Stellensuche in SZ 6 auf einem LAA-Rechner anfallen
transSulokB	durchschnittliche Anzahl an Transaktionen für eine stellenorientierte Bewerbersuche
transSulokV	durchschnittliche Anzahl der lokalen Suchtransaktionen bei einer mit dem Bewerber abgestimmten bewerberorientierten Stellensuche
<i>TransSumB</i>	durchschnittliche Anzahl täglicher Transaktionen für stellenorientierte Bewerbersuchen auf

	einem Rechner eines "mittleren" Arbeitsamtes
<i>TransSumV</i>	durchschnittliche Anzahl an Transaktionen, die täglich zur bewerberorientierten Stellensuche auf einem Rechner eines „mittleren“ Arbeitsamtes anfallen
transSuremB	durchschnittliche Anzahl an Transaktionen, die für eine entfernte Suche oder durch eine Suche von einem entfernten Rechner im Pendelbereich im Rahmen der stellenorientierten Bewerbersuche erforderlich werden. Die Studie nimmt an: transSuremB = 1.
transSuremV	durchschnittliche Anzahl an Suchtransaktionen nach geeigneten Stellen in der Ferne für eine bewerberorientierte Stellensuche
<i>TransSuzenB</i>	durchschnittliche Anzahl täglich anfallender Transaktionen für stellenorientierte Bewerbersuchen auf dem Zentralrechner in SZ 1
<i>TransSuzenV</i>	durchschnittliche Anzahl an Transaktionen, die täglich zur bewerberorientierten Stellensuche zentral anfallen
<i>TransUpamS:</i>	durchschnittliche Anzahl der täglich lokal entstehenden Transaktionen durch Updateanforderungen auf SteAs von entfernten Rechnern nach bewerberorientierten Folgesuchläufen auf einem Rechner in einem „mittleren“ AA
<i>TransUpemS:</i>	durchschnittliche Anzahl der täglich lokal entstehenden Transaktionen durch Updateanforderungen auf SteAs auf entfernten Rechnern nach bewerberorientierten Folgesuchläufen auf einem Rechner in einem „mittleren“ AA
<i>TransUpLAAN</i>	durchschnittliche Anzahl täglich auf einem LAA-Rechner entstehender Transaktionen für Stellenupdates in Folge von bewerberorientierten Erstsuchen
<i>TransUpLAAS</i>	durchschnittliche Anzahl der täglich auf LAA-Rechnern entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen
<i>TransUplmS</i>	durchschnittliche Anzahl der täglich lokal entstehenden Transaktionen durch die Aktualisierung von SteAs des lokalen Datenbestandes nach bewerberorientierten Folgesuchläufen auf einem Rechner in einem „mittleren“ AA
<i>TransUpmN</i>	durchschnittliche Anzahl täglich in einem „mittleren“ Arbeitsamt erforderlicher Transaktionen für Stellenupdates in Folge von bewerberorientierten Erstsuchen
<i>TransUpmPers</i>	durchschnittliche tägliche Anzahl an Transaktionen für das Update der Personendaten nach einem akzeptierten Suchvorschlag in einem „mittleren“ Arbeitsamt
transUpN	durchschnittliche Anzahl an Transaktionen zur Ausführung eines Stellen-Updates
transUpPers	durchschnittliche Anzahl an Transaktionen zur Speicherung der für Vermittlungsvorschläge akzeptierten Stellendaten und der abgestimmten Suchanforderung
<i>TransUpzenN</i>	durchschnittliche Anzahl täglich zentral erforderlicher Transaktionen für Stellenupdates in SZ 1 in Folge von bewerberorientierten Erstsuchen
<i>TransUpzenPers</i>	durchschnittliche tägliche Anzahl an Transaktionen für das Update der Personendaten nach einem akzeptierten Suchvorschlag in SZ 1
<i>TransUpzenS</i>	durchschnittliche Anzahl der täglich zentral entstehenden Transaktionen durch das Lesen entfernter SteAs nach bewerberorientierten Folgesuchläufen
<i>TransWrmTD</i>	durchschnittliche Anzahl an täglich in einem „mittleren“ Arbeitsamt entstehenden Transaktionen für Einträge in die Teamdateien, die aus stellenorientierten Bewerbersuchen nach Stellen-Ersterfassung resultieren
<i>TransWrzenTD</i>	durchschnittliche Anzahl täglicher Transaktionen für Einträge in die Teamdateien, die in SZ 1 aus stellenorientierten Bewerbersuchen nach Stellen-Ersterfassung entstehen
<i>TZEber</i>	durchschnittlicher Tageszeitbedarf für Erstkontakte eines Vermittlers
<i>TZFber</i>	durchschnittliche Zeit für Folgeberatungen mit Publikumsverkehr
TZopen	durchschnittliche tägliche Öffnungszeitdauer mit Publikumsverkehr
TZTB2	durchschnittlich tägliche Zeit, die für Aktivitäten des Tagesbereiches 2 (Arbeiten ohne Publikum) zur Verfügung steht
TZTB3	durchschnittlich tägliche Zeit, die für Aktivitäten des Tagesbereiches 3 (ohne Sachbearbeiteranwesenheit) zur Verfügung steht
X	prozentualer Anteil der in SZ 3 nur lokal angebotenen Stellen
X	prozentualer Anteil in SZ 2 - 7 nur lokal angebotener Stellen, $X=0$ für SZ 2, $X=X$ für SZ 3, $X=X^*$ für SZ 4 - 7
X*	prozentualer Anteil in SZ 4 - 7 nur lokal angebotener Stellen
ZEber	durchschnittliche Zeitdauer für einen Erstkontakt

zentraleKz	durchschnittliche Anzahl an Transaktionen zur Kennzeichnung der zentral gehaltenen Stellenreplika im Rahmen der stellenorientierten Bewerbersuche
zentraleKzLAA	durchschnittliche Anzahl an Transaktionen zur Kennzeichnung der in einem LAA gehaltenen Stellenreplika im Rahmen der stellenorientierten Bewerbersuche

* alternativ:

<i>L</i>	durchschnittliche Anzahl gefundener Bewerber je Stelle auf zentralem Bewerberdatenbestand, Annahme $L \leq O$ (Obergrenze für Lösungen)
<i>L1</i>	durchschnittlicher prozentualer Anteil von <i>L</i> , der aus lokalen Arbeitsamtsrechnerdaten hervorgeht
<i>L2</i>	durchschnittlicher prozentualer Anteil von <i>L</i> , der aus den Daten von Rechnern des Pendelbereiches oder Replikaten dieser Daten hervorgeht. In SZ 2, 3, 6 gilt: $L2 = 0$, denn diese Daten sind dort bei bewerberorientierten Stellensuchen unsichtbar.

4 Diskette mit Kalkulationsbeispielen zur Studie⁵⁷

Diese Anlagen sind der Arbeit lose beigelegt.

⁵⁷ Diese Dateien sind gesondert unter dem Namen Anlage4 als Excel-Dateien abzurufen

