GWDG-Bericht Nr. 67

Christoph Gartmann, Jochen Jähnke
(Hrsg.)

21. DV-Treffen der Max-Planck-Institute

17. - 19. November 2004 in Göttingen

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Christoph Gartmann, Jochen Jähnke (Hrsg.)

21. DV-Treffen der Max-Planck-Institute

17. - 19. November 2004
in Göttingen

Christoph Gartmann, Jochen Jähnke (Hrsg.)

21. DV-Treffen der Max-Planck-Institute

17. - 19. November 2004
in Göttingen

GWDG-Bericht Nr. 67
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
Am Faßberg
D-37077 Göttingen
Telefon: 0551 201-1510
Telefax: 0551 201-2150
E-Mail: gwdg@gwdg.de

Satz: Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Druck: Offset- und Dissertationsdruck Kinzel, Göttingen-Weende

ISSN 0176-2516

Inhalt

Vorwort

Im vorliegenden Band werden einige Beiträge des 21. DV-Treffens der Max-Planck-Institute wiedergegeben. Das Treffen fand vom 17. bis 19. November 2004 bei der GWDG in Göttingen statt.

Eine vollständigere, elektronische Zusammenstellung der Vorträge findet sich im IT-Portal der MPG (https://it-portal.mpg.de) unter „DV-Treffen der Institute“.

Die Zahl der Vorträge war diesmal so groß, dass teilweise parallele Sitzungen abgehalten werden mussten. Diese hatten die Themen Content-Management-Systeme und Rechner- bzw. Speicher-Cluster sowie Grid-Computing als Schwerpunkte. Zu diesem Themenkomplex finden sich im vorliegenden Band dann auch drei Artikel.

Wissenschaftliche Datenverarbeitung bedeutet zumeist auch, insbesondere in der MPG, Dienstleistung für Wissenschaftler. Da liegt es nahe, auch die Verwaltung eines solchen Serviceangebotes durch den Einsatz der Datenverarbeitung effizienter zu gestalten. Deshalb finden sich dazu in diesem Band ebenfalls drei Vorträge.

Ein weiterer Schwerpunkt der Veranstaltung waren die Themen Sicherheit und Recht in der EDV, die aus dem Arbeitsumfeld nicht mehr wegzudenken sind. Leider ist von den vielen interessanten und aufschlussreichen Vorträgen hier nur der von Herrn Gercke vertreten.

Aus der breiten Themenvielfalt des 21. DV-Treffens seien beispielhaft noch der Einsatz von LDAP bei der GWDG erwähnt sowie die IP-Telefonie, die mittlerweile den Kinderschuhen entwachsen ist.

Schließlich möchten wir es als Veranstalter dieses DV-Treffens nicht versäumen, uns bei der GWDG für die Organisation vor Ort herzlich zu bedanken, insbesondere bei Herrn Otto, der die Hauptlast in Göttingen trug.

Freiburg, 04.10.2005 Christoph Gartmann, Jochen Jähnke

Ressourcenverwaltung in einer komplexen Anforderungs- und Verteilungsstruktur

Ulrich Schwardmann

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

1. Ressourcen-Management

Das Ressourcen-Management soll dazu dienen, die Nutzung der Ressourcen zu optimieren. Die Frage nach dem, was ein solches Optimum der Nutzung darstellt, ist nicht a priori offensichtlich, und ist insbesondere aus der Sicht der mit den Ressourcen beschäftigten Akteure sehr unterschiedlich:

Die Sicht der Benutzer

ist geprägt von den folgenden beiden Aspekten:

- Anwendung mit Bedarf an Ressourcen
- Anspruch auf Zuteilung von Ressourcen

Dabei besteht üblicherweise der Wunsch nach:

- kurzer Antwortzeit (Turn-Around-Zeit)
- evt. sogar sofortiger Antwortzeit (Echtzeitverarbeitung)
- hohem Grad der Automatisierung der Nutzung

- und nach Transparenz der Zuteilung
- sowie Beratungskompetenz bei Problemfällen

Die Bearbeitung der Anwendungen des Benutzers durch lokale Provider muss diesen Forderungen genügen. Die Anbindung des Benutzers an einen lokalen Provider leitet sich nur daraus ab.

Die Sicht der Betreiber

ist demgegenüber meist durch das Dilemma dominiert, zu viele Benutzer auf zu wenig Ressourcen abbilden zu müssen. Wegen dieser üblichen Ressourcenknappheit ist also ein maximaler Gesamtdurchsatz durch optimale Auslastung der Ressourcen zu gewährleisten.

Die Durchsatzmaximierung wird dabei durch die folgenden Punkte gestützt:

- Systemstabilität
- Zeitliche Optimierung (keine Verschnitte, keine Ausführungsverzögerungen)
- Nutzungsoptimierung (keine blockierenden Konkurrenzen, keine Mehrfachausführungen)

Es ergibt sich dabei für die Betreiber die zusätzliche Schwierigkeit, dass die Komplexität der Zuteilungsansprüche ständig zunimmt.

Diese zunehmende Zuteilungskomplexität entsteht durch:

- Verwendung einer sehr komplexen Ressourcenstruktur durch
  - hybride Parallelverarbeitung
  - I/O
  - Datenbanken
- und Anforderungsstruktur
  - Dialog
  - Grafik
  - Batch
  - Web
  - Datenbank
  - Echtzeitverarbeitung

- den Wunsch der Benutzer nach kurzfristiger starker Bevorzugung
- Server-Hosting evt. als Teil eines Parallelrechners
- überregionale Ressourcenzuteilungen (GRID)

2. Grid-Computing

Die zentrale Forderung des Grid-Computing kann auf den folgenden einfachen Grundgedanken gebracht werden:

„Irgendein Nutzer will mit irgendwelchen Programmen oder Daten irgendwo Ressourcen nutzen.“

Um die Auflösung all dieser hierin formulierten Unbekannten sowohl auf der Seite des Nutzers als auch der Seite der Betreiber bewerkstelligen zu können, muss ein beträchtlicher Aufwand an Infrastruktur aufgebaut werden.

Im Wesentlichen man braucht dafür:

- eine Anpassung an die Abstraktion von Ressourcen
- eine Instanz zur Zuteilung der Ressourcen
- eine Zertifizierungsauthorität (CA)
- einen Weg, Programme und Daten am gewünschten Ort zur Verfügung zu stellen
- eine Metrik, um den Aufwand dafür zu messen und
- ein Interface, um dies in seinem Verhältnis zum Nutzen darzustellen
- ein Zuteilungs- und Abrechnungsverfahren

3. Der Ressourcen-Manager

Erst wenn ein Mangel in der Verfügbarkeit von Ressourcen auftritt, muss dieser gerecht verwaltet werden. Dabei hängt der Aufwand, der für die Verwaltung betrieben wird, zuallererst davon ab, wie eng oder lose die Zusammenarbeit der beteiligten Akteure ist: Auf Abteilungsebene reicht als Ressourcen-Manager oft ein Wandkalender.

Erst bei der ersten Beschwerde, die nicht mehr im Rahmen gegenseitiger Übereinkunft geregelt werden kann, entsteht die Bedarf nach allgemein gültigen Regeln und nach einem Automatismus, eben einem Ressourcen-Manager, der diese Regeln durchsetzt.

- Dabei verlangt die Heterogenität der Ansprüche sehr bald vielfältige und transparente Steuerungsmechanismen für diesen Automatismus,
- und die Stabilität des Ressourcen-Manager, sowohl was die Robustheit als auch was die Jobkontrolle angeht, wird dabei vorausgesetzt.

4. Spezialfall Parallelverarbeitung

Parallelrechner stellen eine Ressource dar, die für die Ressourcenzuteilung einige zusätzliche Schwierigkeiten mit sich bringt. Die Parallelverarbeitung ist eine Überflussressource im folgenden Sinne:

- Sie verkürzt die Turn-Around-Zeit eines einzelnen Jobs,
- aber sie vermindert häufig den Gesamtdurchsatz des Systems,
- es sei denn, durch die Parallelität wird die Summe der Verarbeitungszeiten aller Threads kürzer: dies wird superlinearer Speedup.

Es wäre nun im Sinne der Gerechtigkeit des Ressourcen-Managements wünschenswert, über eine Kontrolle dieser parallelen Effizienz zu verfügen. Leider ist ein solcher Ansatz zur Zeit aber schwierig zu realisieren. In einigen Fällen lässt sich eine ineffiziente Nutzung durch die Auslastung der Prozessoren und das Kommunikationsprofil erkennen, aber im Allgemeinen gibt es keine sicheren Indikatoren für Möglichkeiten der Effizienzverbesserung. Insbesondere automatische Rückkopplungen derartiger Messgrößen sind noch nicht hinreichend entwickelt, dass sie sicher zum Einsatz kommen können.

Hinzu kommen für die Parallelverarbeitung die folgenden Merkmale:

- Parallele Aufgaben haben für gewöhnlich ein höheres Anforderungsvolumen,
- allein die parallele Ressourcenanforderung erhöht die Zuteilungskomplexität.
- Die sogenannte „Hybride Parallelität“, also die gleichzeitige Nutzung der Mehrprozessor-Eigenschaft vieler heutiger Rechner zusammen mit den Techniken des Message-Passing über Rechnergrenzen hinweg, erzeugt dann noch eine zusätzliche Komplexität in der Zuteilung.

5. Welche Prinzipien helfen bei der Ressourcenverwaltung?

Bei der Automatisierung eines Regelwerks zur Ressourcenverwaltung sind etliche Komponenten notwendig. Die Basis stellt üblicherweise ein Queuing-Mechanismus dar. An dieser Stelle sollen einige weitere prominente

Prinzipien behandelt werden, die über die grundlegenden Mechanismen des Queuing-Systems hinausgehen und die für ein komfortables Ressourcen-Management wertvoll sind:

- Jobkontrolle
- Backfill
- Fairshare
- Checkpoint/Restart
- Lastausgleich und Jobmigration
- Exklusivität bei der Steuerung der Ressourcennutzung

5.1 Jobkontrolle

Dies besagt zunächst, dass alle beteiligten Prozesse dem Ressourcen-Manager bekannt sein müssen. Diese scheinbare Selbstverständlichkeit ist im Umfeld der Parallelverarbeitung, insbesondere unter Beteiligung mehrerer Rechner, nicht einfach zu lösen. Die Forderung ist aber gleichzeitig von wesentlicher Bedeutung für die Stabilität des Gesamtsystems, da sonst die Verwaltung von Resten unkontrolliert beendeter Jobs zur aufwändigen Handarbeit der Systemadministratoren wird.

Ferner muss die Verwendung der Ressourcen für alle beteiligten Prozesse protokolliert werden: Accounting.

5.2 Backfill

Dies ist ein Mechanismus, der vor allem der höheren Auslastung eines Parallelrechners dient. Um einen parallelen Job mit vielen Prozessoren zur Ausführung zu bringen, ist deren vorzeitige Reservierung nötig.

Während der Zeit der Reservierung stehen diese Ressourcen normalerweise leer. Ist dem Batchsystem aber bekannt, wie lange auf die Zuteilung der restlichen Ressourcen für den parallelen Job noch zu warten ist, und ob ein anderer Job mit den reservierten Ressourcen für die fragliche Zeit bis zu deren geplanter Nutzung auskommt, kann dieser Job auf den reservierten Ressourcen rechnen, ohne den sonstigen Ablauf zu beeinträchtigen. Die für das Batchsystem benötigten Informationen lassen sich aus den Angaben der Benutzer über die geplante Rechenzeit der Jobs entnehmen.

Da die Gesamtauslastung des Systems damit steigt, lassen sich die Benutzer meist davon überzeugen, dass eine verlässliche Angabe dieser Information in ihrem eigenen Interesse ist.

Insgesamt wird dadurch die Turn-Around-Zeit der Aufgaben mit geringem Anforderungsvolumen vermindert.

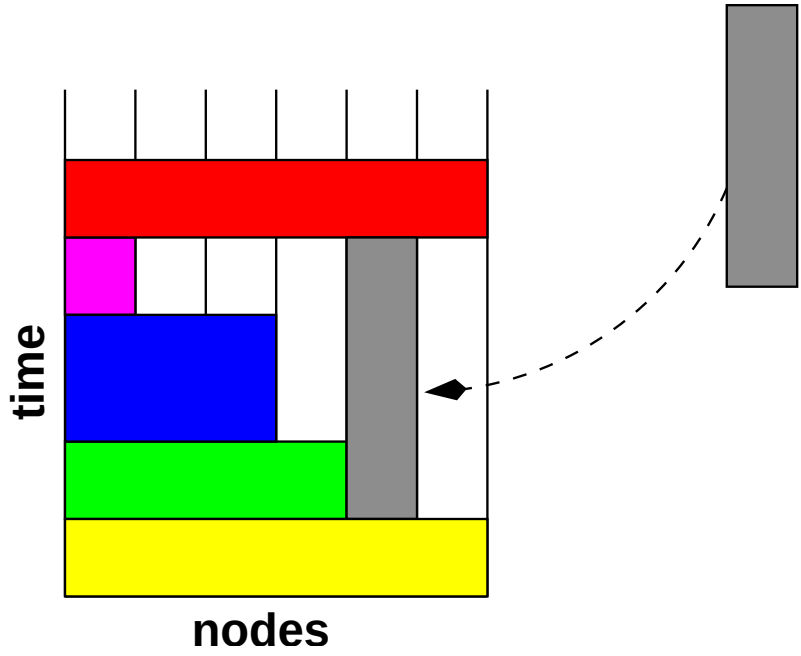


Abb. 1: Backfill - das ist wie Mogeln beim Tetris-Spielen

5.3 Fairshare

Fairshare ist ein Mechanismus zur Priorisierung von Jobs für Gruppen und Einzelpersonen, für Anforderungsprofile, die über die Queue-Definitionen gegeben werden, und für Nutzungsprofile. Über eine geeignete Formel wird hierbei die Priorität eines Jobs so geändert, dass das Gesamtsystem sich im Mittel dem gewünschten Verhalten anpasst. Auf diese Weise ist es zum Beispiel möglich, Benutzern oder Benutzergruppen im Durchschnitt eine Mindestzuteilung an Ressourcen (entsprechend zum Beispiel der finanziellen Beteiligung am Gesamtsystem) zukommen zu lassen oder zu Zeiten dringenden Bedarfs eine vorzügliche Behandlung zu gewähren, die später wieder ausgeglichen werden kann.

5.4 Checkpoint/Restart

Checkpoint/Restart beschreibt die Möglichkeit, einen unvollendeten Job kontrolliert an einem bestimmten Punkt der Ausführung unterbrechen zu

können, und ihn später und auf anderer Ressource wieder aufnehmen zu können. Dabei sind folgende Prinzipien wichtig:

- Checkpoint/Restart soll periodisch oder durch Signal erfolgen, und zwar auf Benutzer- oder auf Betriebssystemebene
- Betriebssystem-Checkpoint/Restart braucht Betriebssystemunterstützung (diese ist bei Linux zum Beispiel nicht gegeben)
- User-Checkpoint/Restart braucht die Implementation eines Signalhandlers
- das Ressourcen-Management-System braucht eine Schnittstelle dahin
- Checkpoint/Restart erhöht die Ausführungssicherheit und ermöglicht die zeitliche Zerlegung der Jobs
- es kann im Zusammenhang mit Backfill die Auslastung erhöhen oder es erhöht die Flexibilität durch Jobmigration (s. u.).

5.5 Lastausgleich und Jobmigration

Lastausgleich erhöht den Gesamtdurchsatz, indem die Last auf allen verfügbaren Ressourcen verteilt wird. Dies kann bei zu hoher Gesamtlast allerdings alle Ressourcen blockieren (Denial of Service). Daher muss für jede Ressource eine für den Durchsatz optimale Last angestrebt werden.

Zudem kann Jobmigration (Jobpriorisierung durch das Betriebssystem) Ressourcen für hochpriorisierte Anforderungen freistellen. Dies sind Anforderungen, die im Zusammenhang mit der Echtzeitverarbeitung und dem Dialogbetrieb nötig werden.

Die Jobmigration kann damit die Menge der zur Verfügung stehenden Ressourcen erhöhen. Insbesondere geschieht dies durch die Hinzunahme von Ressourcen mit anderweitigen Aufgaben, zum Beispiel Dialog- oder Arbeitsplatzrechner.

5.6 Exklusivität bei der Steuerung der Ressourcennutzung

Viele Ressourcen sind gemeinsam nutzbar: SMP-Proz, Speicher, Platten und NICs. Aber es gibt Prozesse, die für diese Ressourcen Exklusivität brauchen. Speziell bei der Parallelverarbeitung kann dies von besonderer Bedeutung sein, weil zum Beispiel das schnelle Netzwerk, die Speicherhierarchie oder der Plattenplatz nicht durch andere Jobs auf dem gleichen Rechner gleichzeitig beansprucht werden soll, um eine ausreichende Effizienz zu gewährleisten.

6. Ressourcen-Management-Systeme

Die folgenden Tabelle gibt eine Übersicht über die Verfügbarkeit der wichtigsten Prinzipien für das Ressourcen-Management, soweit dies zum Zeitpunkt des Vortrags bekannt war.

	Open-PBS	PBSPro	Grid-Engine	Load-leveler	LSF
Jobkontrolle	-	+	o	+	+
Backfill	-	?	-	+	+
Fairshare	-	?	-	+	+
Checkpoint	-	?	-	o	?
Jobmigration	-	?	-	-	?
Nutz.-Steuer.	o	?	o	o	+

7. Ein Spezialfall für das Ressourcen-Management: die Systempflege

Soweit die Systempflege nicht im Hintergrund des laufenden Betriebs durchgeführt werden kann, sondern die exklusive Nutzung eines Knotens voraussetzt, zum Beispiel um einen Neustart des Systems herbeizuführen, handelt es sich um eine Ressourcenzuweisung, die ebenfalls vom Ressourcen-Management-System behandelt werden kann und sollte.

Üblicherweise hat diese Systemaktivität maximale Priorität. Es ist ein sequentieller Prozess, der exklusiv auf (evt. genau) einem Knoten durchgeführt werden muss.

Leistungsvergleich verschiedener Global-Filesystem-Produkte in einer SAN-Umgebung

Reinhard Sippel

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

1. Einleitung

1.1 Vorteile globaler Filesysteme bei der Massenspeicherverwaltung

Globale Filesysteme ermöglichen in einer Storage Area Network (SAN) Umgebung mehreren Klienten heteregoner Plattformen, gemeinsam auf dieselben Massenspeicher-Ressourcen zuzugreifen.

Die Kombination von SAN mit einem globalen Filesystem bietet für den Betrieb von File-Servern folgende Vorteile:

- Erhöhung der Ausfallsicherheit

  Durch Verwendung eines globalen Filesystems kann der Fileservice redundant aufgebaut werden, so dass beim Ausfall eines Rechners ein anderer die entsprechenden Funktionalitäten übernehmen kann.

- Reduzierung der Ausfälle wegen Wartung

  Durch die redundante Auslegung der Umgebung können einzelne Systeme gewartet werden, ohne den Produktionsbetrieb zu unterbrechen.

- Verbesserung der Datenzugriffszeiten

  Durch den Zugriff auf die Massenspeicher-Ressourcen über das SAN werden Zugriffszeiten vergleichbar mit denen lokaler Platten erzielt.

- Verringerung des administrativen Aufwands

  Die SAN-Architektur, die Betriebsoberflächen der Storage Devices und die globalen Filesysteme stellen Administrationwerkzeuge zur Verfügung, mit deren Hilfe die Massenspeicher-Ressourcen zentral verwaltet werden können.

- Verbesserung der Skalierbarkeit des Massenspeichers

  Die globalen Filesysteme bieten die Möglichkeit, die Datenbereiche im laufenden Betrieb zu vergrößern.

1.2 Kriterien für die Evaluierung von Global-Filesystem-Produkten

Für die Evaluierung von Global-Filesystem-Produkten sind folgende Punkte relevant:

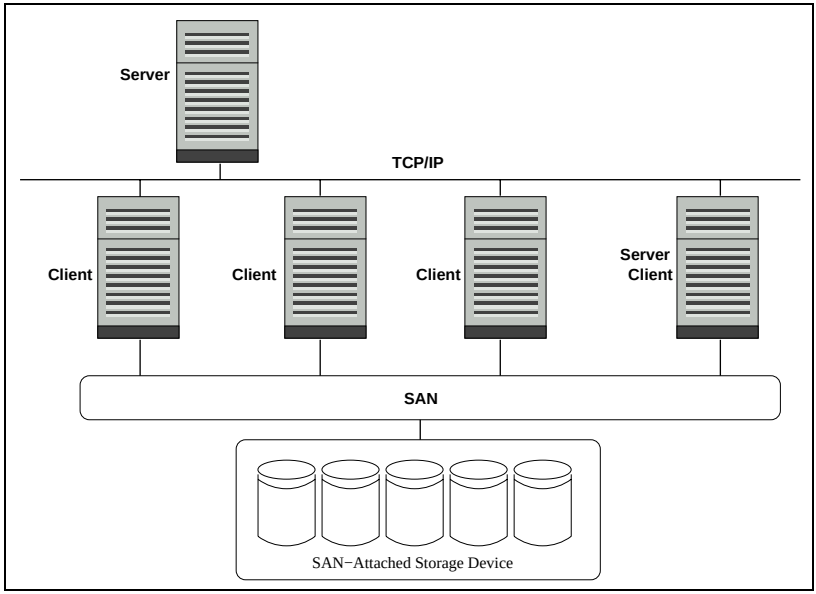
- unterstützte Plattformen
- Funktionalitäten wie Quota, NFS oder Exportmechanismen
- Stabilität
- Durchsatzleistung entsprechend den Benchmark-Tests
- Administrierbarkeit
- Hersteller-Support
- Kosten

2. Konfigurationsmöglichkeiten von Rechnern mit globalen Filesystemen

Die Global-Filesystem-Klienten greifen gemeinsam auf Massenspeicherbereiche im SAN zu. Auf den Klienten können Anwendungen laufen, wie beispielsweise Mail-Service, Datenbank-Service, Compute-Service oder NFS-Service

Um beim Zugriff mehrerer Rechner auf gemeinsame Bereiche Datenkorrumpierung zu vermeiden, müssen die Zugriffe durch Locking-Mechanismen oder durch den Einsatz von Metadatenservern gesteuert werden.

Die folgende Abbildung illustriert den allgemeinen Aufbau einer Global-Filesystem-Konfiguration. In dieser Umgebung wird Massenspeicher direkt über das SAN zur Verfügung gestellt (SAN-Attached Storage Devices). Mehrere Rechner sollen gemeinsam mit Hilfe eines globalen Filesystems auf die Storage Devices zugreifen. Wie in der Abbildung dargestellt, sind die Global-Filesystem-Klienten über Fiberchannel-Interface mit dem SAN verbunden. Die Nutzdaten werden über das SAN transferiert. Zur Synchronisation der Nutzdaten müssem Locking-Mechanismen oder Metadaten-Server bereitgestellt werden. In der Abbildung ist ein Medatatenserver über TCP/IP mit den Klienten verbunden. Der Metadaten-Server greift nicht auf die Nutzdaten im SAN zu. Der Abgleich der Metadaten erfolgt über TCP/IP. Es empfiehlt sich, für den Metadatentransfer ein eigenes Subnetz zur Verfügung zu stellen. (Es sind auch Konfigurationen möglich, bei denen der Metadaten-Server direkt auf einem Klienten implementiert ist; es muss allerdings darauf geachtet werden, dass die Last auf diesem Rechner nicht zu groß wird.)



3. Übersicht über die bei der GWDG evaluierten Produkte

Die GWDG hat die folgenden Global-Filesystem-Produkte getestet:

3.1 StorNext File System (SNFS)

Das StorNext File System (SNFS) der Firma adic unterstützt folgende Betriebssysteme:

- AIX
- RedHAT Enterprise Linux (ia64 ab AS 3.0)
- RedHAT Enterprise Linux (x86)
- Apple OS X (von Apple als Xsan angeboten)
- SGI IRIX
- Sun Solaris
- SuSE Linux (x86)
- Windows

Das Produkt umfasst folgende Funktionalitäten:

- Quota-Unterstützung
- NFS-Unterstützung
- Tuning-Möglichkeiten

3.2 Global File System (GFS)

Das Filesytem GFS 6.0 liegt komplett als Open-Source-Produkt vor. Mit Support wird GFS 6.0 im RedHAT-Enterprise-Linux-AS-3.0-Paket vertrieben.

Die unterstützten Betriebssysteme sind:

- RedHAT Linux (x86 und ia64)
- RedHAT Enterprise Linux AS 3.0 (x86 und ia64); (Support)
- SuSE Linux (x86)

Als Funktionalitäten stehen zur Verfügung:

- Quota-Unterstützung
- NFS-Unterstützung
- Tuning-Möglichkeiten

3.3 SAN Filesystem (SFS)

Das Filesystem SFS wird von der Firma DataPlow angeboten. Es unterstützt folgende Betriebssysteme:

- RedHAT Linux (x86) (ia64 ab Ende 2004)
- SGI IRIX
- Sun Solaris
- SuSE Linux (x86); (ia64 ab Ende 2004)
- Windows

Zu den Funktionalitäten gehören:

- Quota-Unterstützung (ab Ende 2004)
- NFS-Unterstützung
- Tuning-Möglichkeiten
- der Zugriff auf viele kleine Files ist optimiert

4. Kriterien zur Erstellung anwendungsnaher Leistungstests (Benchmarks)

In verteilten heterogenen DV-Umgebungen können auf verschiedenen Servern Anwendungen mit unterschiedlichen Anfordungen für den Massenspeicherzugriff vorliegen. Ein unter UNIX auf exim-Basis betriebener Mail-Server muss beispielsweise sehr schnell auf sehr viele kleine Files zugreifen können.

Demgegenüber liegt im UNIX-Cluster der GWDG eine Umgebung vor, in der kleine, aber insbesondere auch sehr große Files verarbeitet werden. Hier ist bei Benchmarks der Durchsatz für das Schreiben und Lesen großer Files die relevante Größe.

Zum Test einer Umgebung mit vielen kleinen Files wurde das Benchmark-Programm postmark von Network Appliance verwendet. Postmark führt File-Erzeugungen, Lese- und Schreiboperationen und File-Löschungen durch.

Für den postmark-Benchmark wurden folgende Parameter gesetzt:

- Anzahl Files: 10.000
- Anzahl Transaktionen: 10.000

- File-Größe: 100 Bytes bis 50 KBytes
- Block-Größe für Lese- und Schreibzugriffe: 512 Bytes

Für die Bestimmung des Durchsatzes in einer Umgebung mit großen Files wurde die Zeit für das

- Schreiben,
- Lesen und
- konkurrierendes Lesen und Schreiben

gemessen.

Als File-Größe wurde 2 GByte gewählt.

5. Übersicht über die bei der GWDG durchgeführten Benchmarks

Die I/O-Operationen wurden auf drei Klienten gleichzeitig durchgeführt. Bei der Auswertung des postmark-Benchmark gibt die Gesamtlaufzeit ein gutes Maß für die Qualität des Filesystems in Bezug auf die Verarbeitung vieler kleiner Files. Die folgende Tabelle zeigt das Ergebnis der Benchmarks:

FS	Postmark Gesamt-laufzeit	I/O mit großen Files			
		Lesen	Schreiben	Lesen bei Schreiben	Schreiben bei Lesen
SNFS	760 s	60 MB/s	30 MB/s	10 MB/s	27 MB/s
GFS	590 - 950 s	53 MB/s	28 MB/s	35 MB/s	15 MB/s
SFS	180 s	keine stabilen Werte			

6. Ergebnis der Evaluierung

6.1 Das bei der GWDG zukünftig eingesetzte Global-Filesystem-Produkt

Die GWDG wird als globales Filesystem das Produkt StorNext File System (SNFS) der Firma adic einsetzen. Gründe für die Entscheidung sind:

- Die Stabilität bei den Benchmarktests
- Die Liste der unterstützen Betriebssysteme

- Der Hersteller-Support

Insbesondere waren die Unterstüzung von AIX und Windows sowie die Lauffähigkeit auf einer ia64-Architektur bei der Auswahl ausschlaggebend.

6.2 Die zukünftige Konfiguration der UNIX-File-Server bei der GWDG

Die folgende Abbildung skizziert den Einsatz eines globalen Filesystems für den UNIX-File-Service der GWDG. Die Filesysteme werden auf drei Rechnern mit ia64-Architektur zur Verfügung gestellt; diese Rechner sind als SNFS-Klienten an das SAN angebunden. Zwei Rechner mit x86-Architektur werden als Metadaten-Server für diese Umgebung eingesetzt.

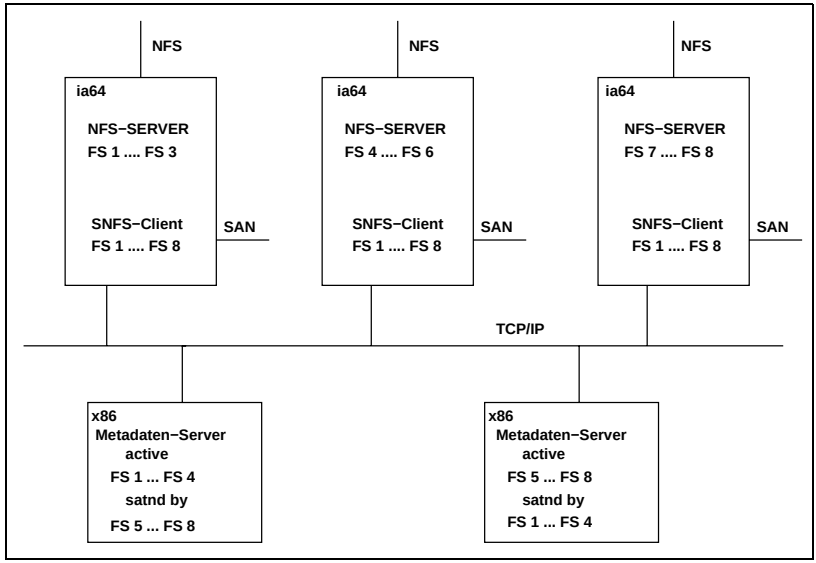
Um im Fehlerfall eine schnelle Rekonstruktion der Daten aus dem Backup sicherzustellen, wird der gesamte File-Service nicht in einem großen Filesystem abgebildet, sondern auf mehrere kleinere Filesysteme verteilt. In der Praxis werden die Daten nach ihren Funktionalitäten auf die einzelnen Filesysteme verteilt. So werden beispielsweise Anwendungs-, Benutzer- und temporäre Daten in unterschiedlichen Filesytemen abgelegt. Dabei werden die Benutzerdaten nach ihrer Institutionszugehörigkeit aufgeteilt.

In der Abbildung wird die Aufteilung der Nutzdaten abstrakt durch acht Filesysteme (FS 1 ... FS 8) dargestellt. Dabei sind alle Filesysteme FS 1 bis FS 8 auf allen drei SNFS-Klienten über das SAN im Zugriff. Bei dieser Konfiguration werden die Filesysteme über NFS an die Applikationsrechner exportiert. Das heißt, hier fungieren die SNFS-Klienten als NFS-Server.

Um die Last auf den Rechnern zu verteilen, stellen die NFS-Server nicht jeweils alle Filesysteme zur Verfügung, sondern nur eine Untermenge. So exportiert beispielsweise der erste Fileserver die Filesysteme FS 1 bis FS 3, der zweite FS 4 bis FS 6 und der dritte FS 7 bis FS 8. Im Fehlerfall oder für Wartungszwecke kann ein Fileserver die Filesysteme eines anderen mit übernehmen, da durch den Einsatz des globalen Filesystems auf jedem NFS-Server alle Filesysteme über SNFS im Zugriff sind. Somit ist ein unterbrechungsfreier Produktionsbetrieb gewährleistet.

Die SNFS-Metadaten-Server verwalten jeweils auch nur eine Untermenge von Filesystemen aktiv. Der erste Metadaten-Server ist aktiv für FS 1 bis FS 4, der zweite für FS 5 bis FS 8. Jeder Medatataten-Server steht für die Filesysteme des jeweils anderen als Standby zur Verfügung. Beim Ausfall eines Metadaten-Servers kann der andere Server die Filesysteme des ausgefallen Rechners mit übernehmen.

Das globale Filesystem soll bei zukünftigen Anwendungen nicht nur für den NFS-Service eingesetzt werden, sondern SNFS-Klienten werden auch auf Produktionssystemen, wie Parallelrechnern oder Backup-Servern, zur Verfügung gestellt.

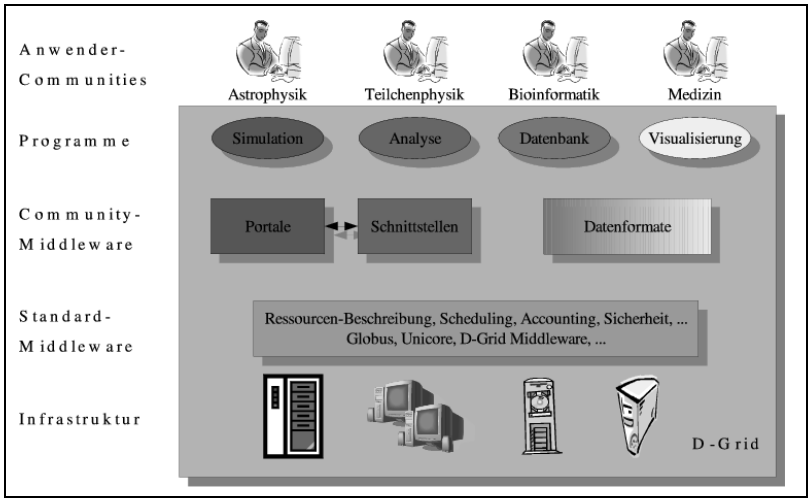


„Instant-Cluster“ für das D-Grid

Christian Boehme

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

1. Was ist das D-Grid?



Das D-Grid ist eine Initiative des Bundesministeriums für Bildung und Forschung (BMBF) zur Schaffung einer so genannten „e-Science“-Infrastruktur. Diese soll es Wissenschaftlern ermöglichen, IT-Ressourcen einfach durch Spezifikation der gewünschten Eigenschaften zu reservieren. In einem Schichtenmodell sind dazu zwischen den Anwendern (oben) und den vorhandenen Ressourcen vermittelnde (Software-)Schichten notwendig, die so genannte Middleware. Bei der Formierung des D-Grids kommt zwei Zielen besondere Bedeutung zu:

- **Ziel 1: Anwender in der Wissenschaft überzeugen**
  - Etablierung von „Community-Grids“, also Grids für bestimmte Forschungsrichtungen
  - Die „Gridifizierung“ von Anwendungen soll durch die D-Grid-Initiative begleitet werden
- **Ziel 2: Bereitstellung von Diensten, nicht Definitionen**
  - Das D-Grid soll von vorneherein eine breite Ressourcen-Basis haben
  - Dazu gehören nicht nur Ressourcen in Rechenzentren, sondern auch diejenigen in Forschungsinstituten

2. Motivation: Forschungsinstitute und D-Grid

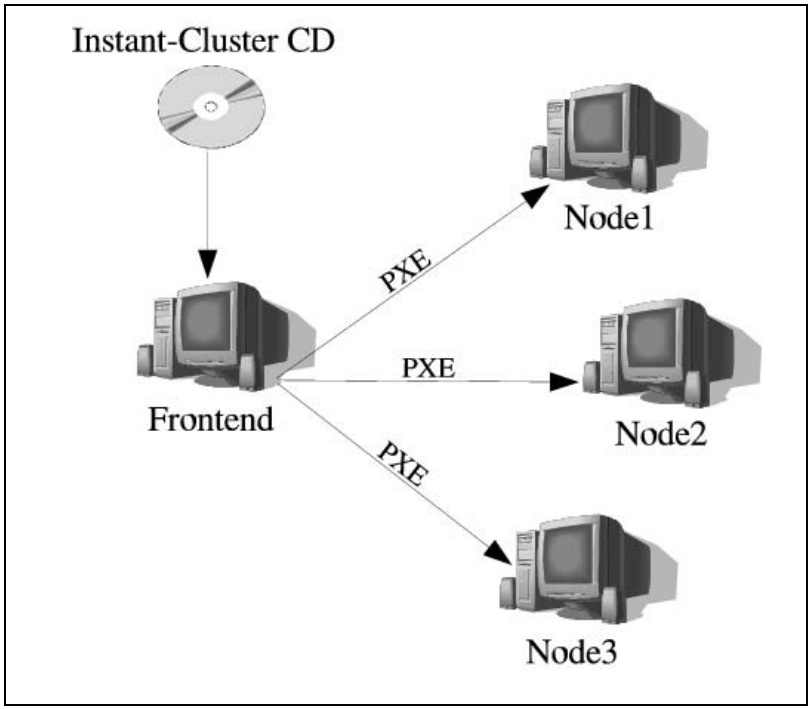
Betrachtet man die Situation an den Forschungsinstituten, so gibt es gute Gründe, sich am D-Grid zu beteiligen:

- Zeitweise freie Ressourcen
- Interesse
- Geeignete Software

Es spricht allerdings auch einiges gegen eine Beteiligung:

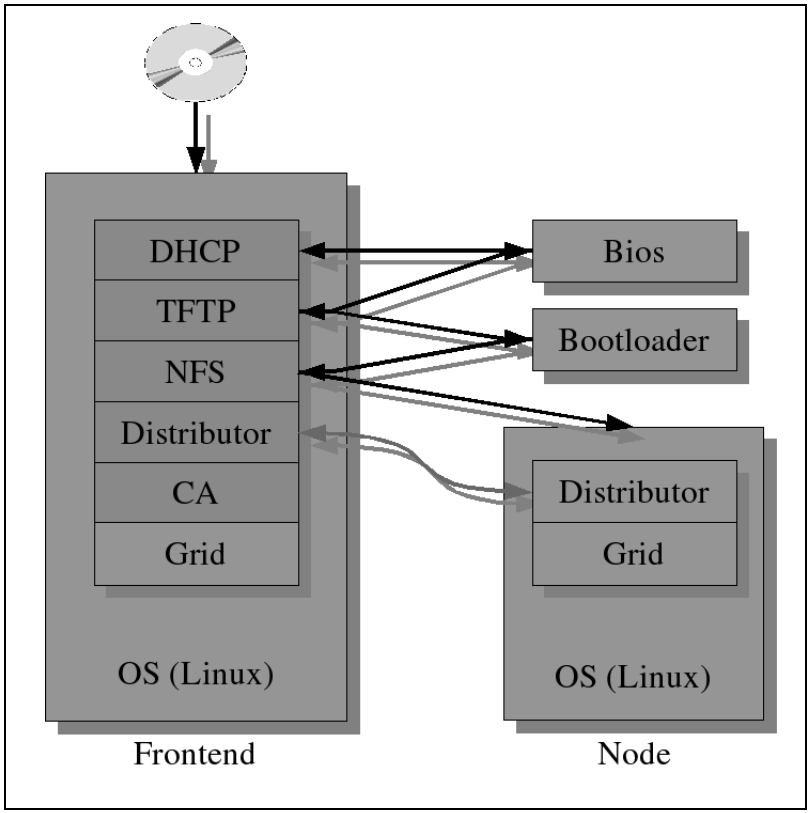
- Keine dauernd freien Ressourcen
- Hohe Einstiegshürde
- Keine geeigneten Test-, Entwicklungs- und Demonstrationsumgebungen

3. Das Konzept des „Instant-Clusters“



Der „Instant-Cluster“ soll die Hindernisse auf dem Weg zu einer Beteiligung am D-Grid wenigstens teilweise aus dem Weg räumen. Es handelt sich um eine mit minimalem Aufwand zu installierende, weitestgehend vorkonfigurierte Grid-Umgebung in Form eines PC-Clusters, der von einer CD-ROM bzw. per PXE(Pre-Execution Environment)-Boot gestartet wird. Die vorhandene Umgebung wird dabei nicht dauerhaft verändert, so dass beliebig zwischen Grid-fähigem Linux-Cluster und originaler Konfiguration gewechselt werden kann.

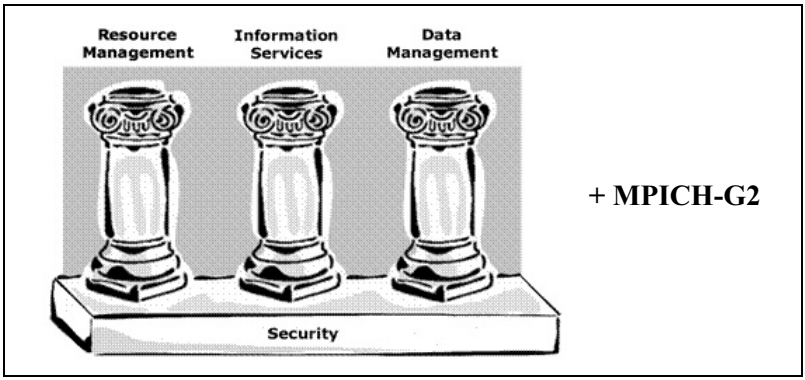
4. Der Startvorgang



Gestartet wird zunächst ein Terminalserver. Unter diesem Begriff sind verschiedene Server zusammengefasst die auf dem Cluster-Frontend vorkonfiguriert von der CD-ROM gestartet werden, um im Folgenden weiteren Rechnern (Clients) den Start und die Integration in den Cluster per PXE-Boot zu erlauben. Dazu gehören vor allem ein DHCP-Server für die Zuteilung der IP-Nummern und die Bekanntgabe der PXE-Boot-Informationen, ein TFTP-Server, auf dem ein Bootloader, Kernelimage und Miniroot für die Clients bereitgehalten, und ein NFS-Server, von dem die Clients das eigentliche Root-Verzeichnis (auf der CD-ROM enthalten) sowie eventuell Datenaustauschverzeichnisse einhängen können. Ferner läuft hier der Distributor-Server für das automatische Kopieren aktualisierter Konfigurationsdateien, für die automatische An- und Abmeldung neuer bzw. aus dem Cluster entfernter Clients und die automatische Bearbeitung von Host-Zertifizierungsanfragen.

Nach dem Terminalserver können die Clients gestartet werden. Entsprechend der vom DHCP-Server des Terminalservers erhaltenen Informationen laden die Clients den PXELINUX-Bootloader vom TFTP-Server. PXELINUX lädt dann Kernelimage und Miniroot vom TFTP-Server, die dann die CD-ROM per NFS als Rootfilesystem einhängen. Dann starten die Clients den Distributor-Client. Zu seinen Aufgaben gehört das Einhängen der vom Frontend exportierten Verzeichnisse und das Erstellen der für jeden Client notwendigen Host-Zertifizierungsanfrage.

5. Middleware: Globus-Toolkit



Entsprechend dem Stand der D-Grid-Planung Ende 2004 ist das Globus-Toolkit (Version 2.4) als Middleware im „Instant-Cluster“ integriert. Es besteht aus folgenden Teilen:

- **GRAM**: Ressourcen-Zuteilung
- **MDS**: Dynamische Grid-Informationen
- **GridFTP**: File-Transfer
- **GSI**: X509/SSL-basierte Transportsicherung

Eine Vorabversion der CD wurde auf der CeBIT 2005 verteilt.

6. Links und Danksagung

Links (weiterführende Literatur):

Knoppix	http://www.knoppix.org
Globus-Toolkit	http://www-unix.globus.org/toolkit/
D-Grid	http://www.d-grid.de

Danke an:

Jan Engelhardt	http://linux01.org:2222/

Leistungsmessung für Parallelrechner: Der HPC Challenge Benchmark

Oswald Haan

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

1. Einleitung

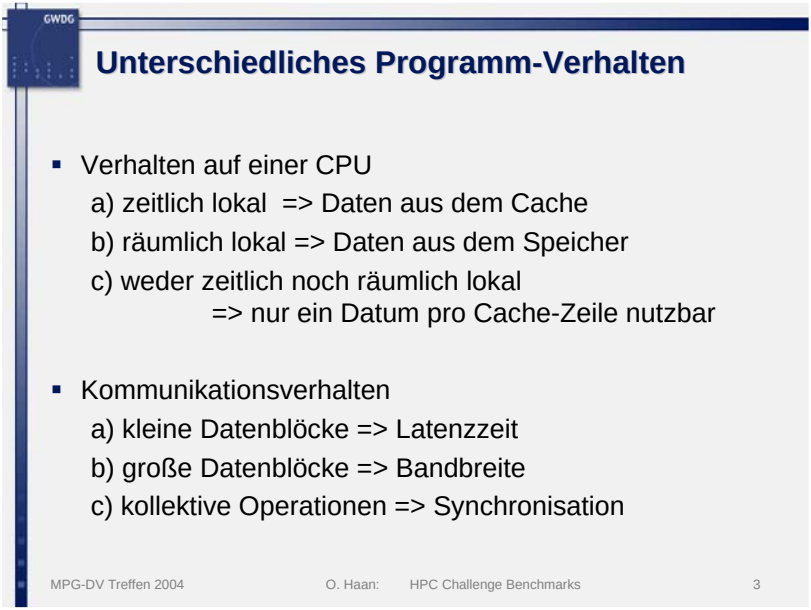
Bei der Beschaffung neuer Rechnerhardware für numerisch intensive Anwendungen steht als Auswahlkriterium die Frage nach der Leistung im Vordergrund. Bei vorgegebenen Investitionsmitteln wird die maximale Leistung gesucht, bei vorgegebenem Leistungsbedarf der minimale Preis. Leistung wird in dem Bereich des High Performance Computing (HPC) durch die Einheit Flop/s gemessen, das ist die Anzahl der Rechenoperationen mit Fließkommazahlen, die das System pro Sekunde bearbeiten kann. Entscheidend ist dabei nicht die maximal mögliche Rechenleistung, die sich aus der Zahl der Prozessoren, der Anzahl der pro Prozessor gleichzeitig arbeitenden Rechenpipelines und der für die Geschwindigkeit der Pipeline maßgebenden Taktrate ergibt als $r_{max} = n_{proc} * n_{pipe} * r_{cpu}$. Vielmehr hängt die vom System erzielbare Rechenleistung in erster Linie von den Eigenheiten der bearbeiteten Anwendungen ab, so dass ohne Berücksichtigung dieser Eigenheiten keine realistischen Aussagen über die Nutzleistung gemacht werden können. Diese Nutzleistung kann also nur über Benchmarks auf dem zu beschaffen-

den System festgestellt werden, wobei die Benchmarks die zu erwartenden Anwendungen repräsentieren müssen.

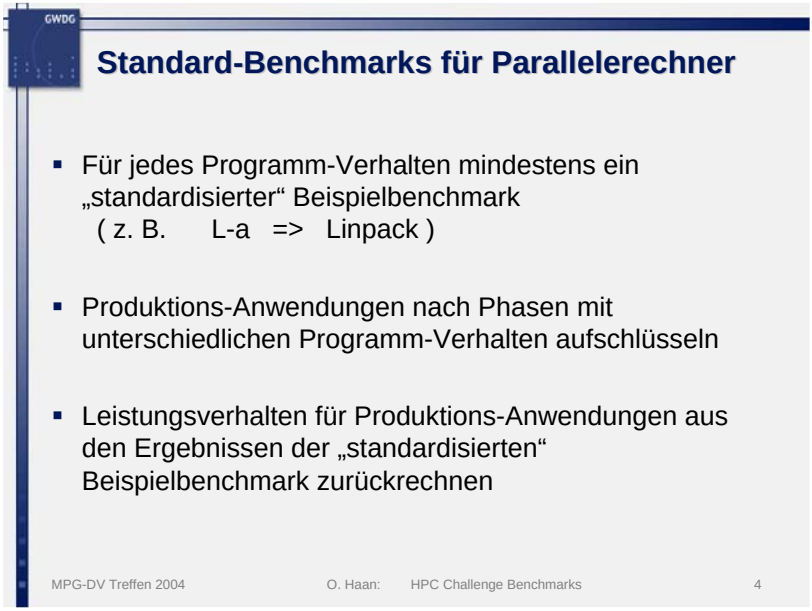
Im Folgenden wird der HPC Challenge Benchmark vorgestellt, der eine Reihe von Anwendungen unterschiedlicher Charakteristika enthält und dessen Ergebnisse durch geeignete Gewichtung Aussagekraft für verschiedenste Anwendungsprofile hat.

2. Der HPC Challenge Benchmark

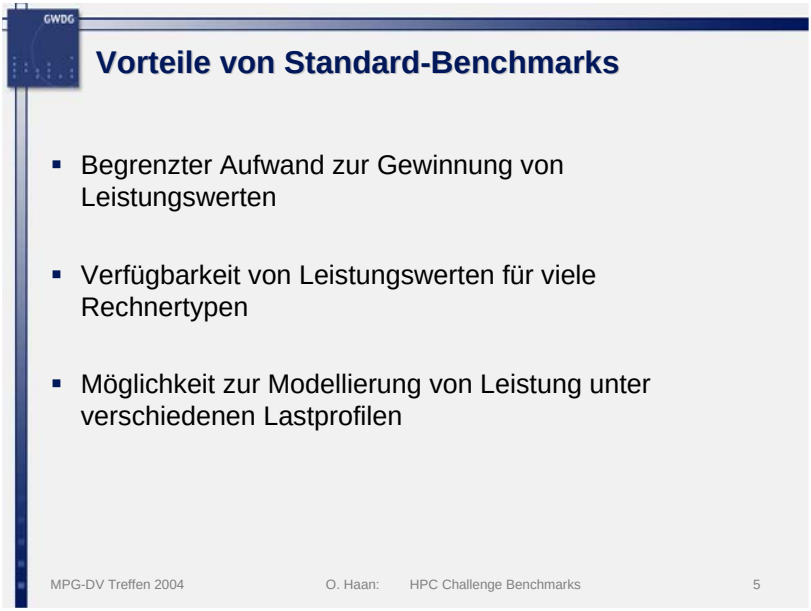
Anwendungsprogramme lassen sich durch einige typische Grundmuster von Operationsfolgen charakterisieren, die unterschiedliche Komponenten eines Rechnersystems beanspruchen und deren Bearbeitungsgeschwindigkeiten deshalb von der Leistung dieser Komponenten beeinflusst werden. Die sechs wichtigsten Grundmuster sind in der folgenden Abb. aufgeführt.



Die Leistungsfähigkeit eines Parallelrechnersystems für ein definiertes Anwendungsspektrum kann ermittelt werden, wenn die Verteilung der Grundmuster in dem Anwendungsspektrum bekannt ist und wenn für die Grundmuster Standard-Benchmarks vorliegen:



Die Vorteile der Nutzung solcher Standard-Benchmarks gegenüber speziellen Benchmarks für alle relevanten Anwendungen eines Anwendungsprofils sind in der folgenden Abb. aufgeführt:



Die Idee von Standard-Benchmarks für Grundmuster von Operationsfolgen wurde von einer Gruppe um Jack Dongara und Piotr Luszczek von der University of Tennesse aufgegriffen und in Form des HPC Challenge Benchmark realisiert. Im HPC Challenge Benchmark sind sieben Typen von Programmen enthalten, deren Zuordnung zu den verschiedenen Grundmustern in den folgenden Abbildungen wiedergegeben wird:

Der HPC Challenge Benchmark

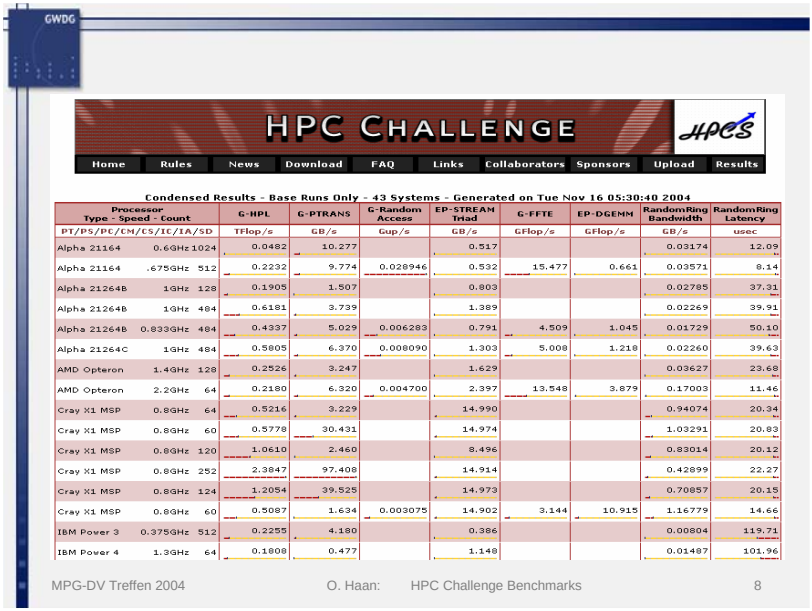
http://icl.cs.utk.edu/hpcc/index.html

- 7 verschiedene Benchmarks mit unterschiedlichem Programm-Verhalten
  - HPL (LINPACK) => L-a und K-b
  - DGEMM => L-a
  - STREAM => L-b
  - PTRANS (A <- A + B^+) => K-b
  - Random Access => L-c und K-a
  - FFT (1dim.) => L-b und K-b
  - b_eff => K-a und K-b
- Kein Benchmark testet K-c (kollektive Kommunikation)

Test-Varianten im HPCC Benchmark

HPL und PTRANS	G	Gesamtsystem
DGEMM und STREAM	SN	Single Node
	EP	Embarrassingly Parallel
Random Update	SN EP G	
FFT	EP G	
b_eff	paarweise sequentiell	
	Ringtausch parallel	

Ergebnisse der HPC Challenge Benchmarks sind auf den Webseiten des Projektes zu finden. Bisher haben Hersteller und Betreiber von Parallelrechner Daten von 40 Systemen auf diesen Seiten veröffentlicht. Die folgende Abb. zeigt einen Ausschnitt aus einer der möglichen Ansichten der Ergebnisse:



Processor Type - Speed - Count	G-HPL	G-PTRANS	G-Random Access	EP-STREAM Triad	G-FFTE	EP-DGEMM	RandomRing Bandwidth	RandomRing Latency
PT/PS/PC/CM/CS/IC/IA/SD	TFlop/s	GB/s	Gup/s	GB/s	GFlop/s	GFlop/s	GB/s	usec
Alpha 21164 0.6GHz 1024	0.0482	10.277		0.517			0.03174	12.09
Alpha 21164 .675GHz 512	0.2232	9.774	0.028946	0.532	15.477	0.661	0.03571	8.14
Alpha 21264B 1GHz 128	0.1905	1.507		0.803			0.02785	37.31
Alpha 21264B 1GHz 484	0.6181	3.739		1.389			0.02269	39.91
Alpha 21264B 0.833GHz 484	0.4337	5.029	0.006283	0.791	4.509	1.045	0.01729	50.10
Alpha 21264C 1GHz 484	0.5805	6.370	0.008090	1.303	5.008	1.218	0.02260	39.63
AMD Opteron 1.4GHz 128	0.2526	3.247		1.629			0.03627	23.68
AMD Opteron 2.2GHz 64	0.2180	6.320	0.004700	2.397	13.548	3.879	0.17003	11.46
Cray X1 MSP 0.8GHz 64	0.5216	3.229		14.990			0.94074	20.34
Cray X1 MSP 0.8GHz 60	0.5778	30.431		14.974			1.03291	20.83
Cray X1 MSP 0.8GHz 120	1.0610	2.460		8.496			0.83014	20.12
Cray X1 MSP 0.8GHz 252	2.3847	97.408		14.914			0.42899	22.27
Cray X1 MSP 0.8GHz 124	1.2054	39.525		14.973			0.70857	20.15
Cray X1 MSP 0.8GHz 60	0.5087	1.634	0.003075	14.902	3.144	10.915	1.16779	14.66
IBM Power 3 0.375GHz 512	0.2255	4.180		0.386			0.00804	119.71
IBM Power 4 1.3GHz 64	0.1808	0.477		1.148			0.01487	101.96

3. Anwendung von HPCC-Ergebnissen

Aus den veröffentlichten Ergebnissen der HPC Challenge Benchmarks sollen nun fünf verschiedene Rechnersysteme verglichen werden. Zunächst sollen die für die Einzelprozessorleistung relevanten Benchmarks zusammengestellt werden:

Einzelprozessor-Leistung für zeitlich lokales Programm-Verhalten

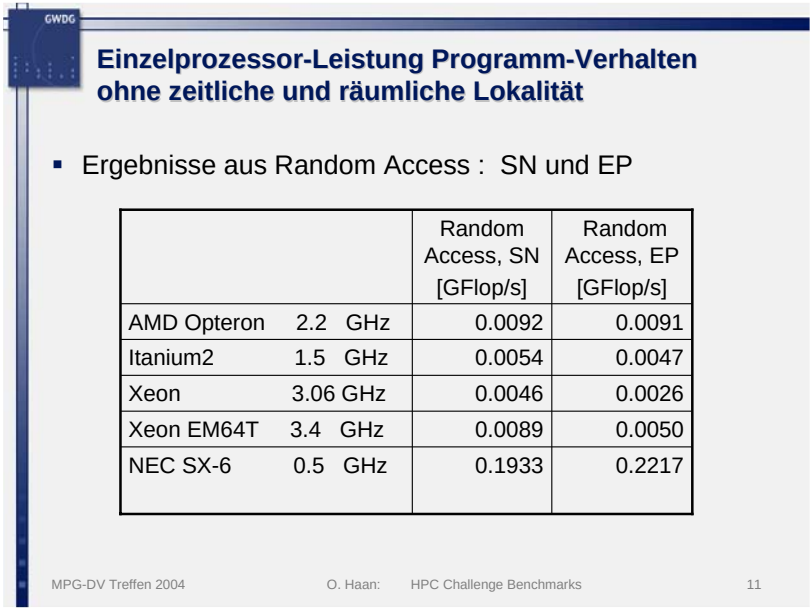
- Ergebnisse aus HPL pro Prozessor und DGEMM

	HPL / Proz. [GFlop/s]	DGEMM [GFlop/s]
AMD Opteron 2.2 GHz	3.4	3.9
Itanium2 1.5 GHz	5.0	5.8
Xeon 3.06 GHz	4.0	----
Xeon EM64T 3.4 GHz	5.5	6.1
NEC SX-6 0.5 GHz	7.1	7.9

Einzelprozessor-Leistung für räumlich lokales Programm-Verhalten

- Ergebnisse aus STREAM: SN und EP

	STREAM triad, SN [GFlop/s]	STREAM triad, EP [GFlop/s]
AMD Opteron 2.2 GHz	0.20	0.20
Itanium2 1.5 GHz	0.31	0.18
Xeon 3.06 GHz	0.14	0.06
Xeon EM64T 3.4 GHz	0.24	0.11
NEC SX-6 0.5 GHz	2.66	2.24



Eine Relevanz für bestimmte Anwendungen ergibt sich nun aus deren Lastprofil. Je nach Anteil der verschiedenen Grundmuster ergeben sich durch Berücksichtigung der gemessenen Leistung für diese Grundmuster unterschiedliche Einzelprozessorleistungen für die Gesamtanwendung.

Bewertung der Einzelprozessor-Leistung

- Angenommenes Lastprofil:
  - Bruchteil x mit Programmverhalten nach L-a
  - Bruchteil y mit Programmverhalten nach L-b
  - Bruchteil z mit Programmverhalten nach L-c
- Leistungsfähigkeit für dieses Lastprofil

$$r_E = \frac{1}{x \cdot r_a^{-1} + y \cdot r_b^{-1} + z \cdot r_c^{-1}}$$

Beispiele für Einzelprozessor-Leistung

- drei Last-Szenarien:

L1: x=0.3, y=0.6, z=0.1
L2: x=0.5, y=0.5, z=0.0
L3: x=1.0, y=0.0, z=0.0

		L1 [GFlop/s]	L2 [GFlop/s]	L3 [GFlop/s)]
AMD Opteron	2.2 GHz	0.07	0.38	3.9
Itanium2	1.5 GHz	0.04	0.35	5.8
Xeon	3.06 GHz	0.02	0.11	gesch. 4.6
Xeon 64T	3.4 GHz	0.04	0.22	6.1
NEC SX-6	0.5 GHz	1.89	3.50	7.9

Ebenso werden auch die für die Kommunikation zwischen Prozessoren aussagekräftigen Benchmarks dargestellt. Dabei ist natürlich die Leistung des die Prozessoren verbindenden Kommunikations-Netzwerkes der entscheidende Faktor.

Kommunikations-Leistung

- Ergebnisse aus b_eff

		Latenz PingPong [µs]	Latenz Ring [µs]	Bandbr. PingPong [GB/s]	Bandbr. Ring [GB/s]
AMD Opteron	QsNet II	1.55	6.24	0.873	0.364
Itanium2	Numalink	2.31	4.38	5.892	0.715
Xeon	Myrinet	7.77	21.99	0.242	0.054
Xeon 64T	Infiniband	5.96	6.21	0.654	0.201
NEC SX-6	Crossbar	6.17	9.01	8.575	6.774

MPG-DV Treffen 2004 O. Haan: HPC Challenge Benchmarks 14

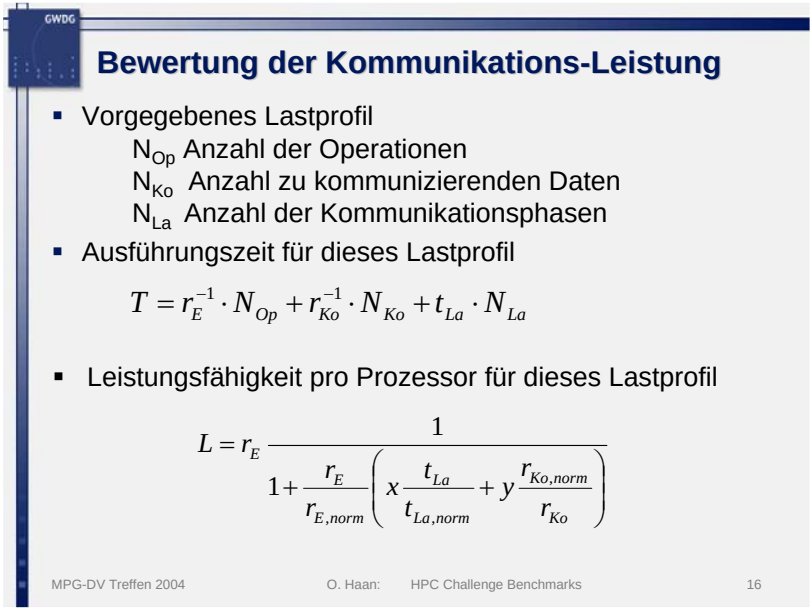
Kommunikations-Leistung

- Ergebnisse aus PTRANS, FFT, Random Access
  (In Klammern Einzelprozessor-Werte)

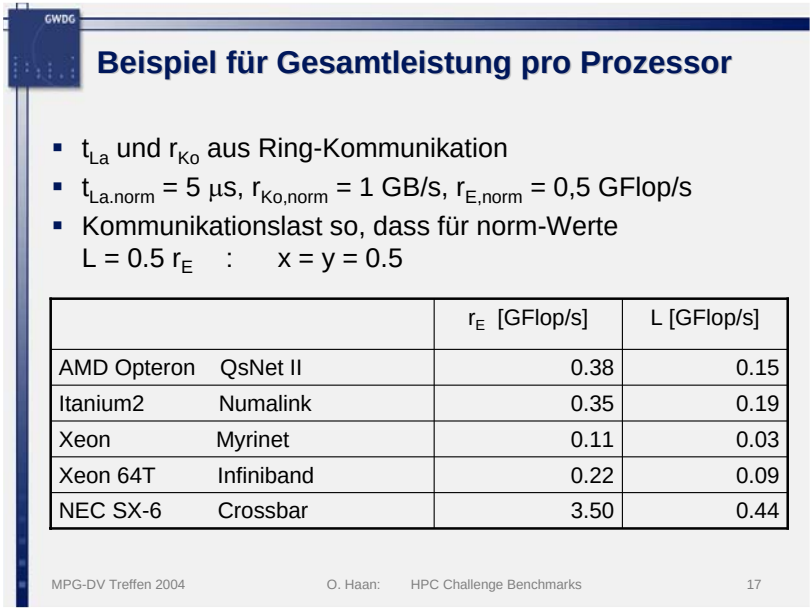
		PTRANS pro Proz. [GB/s]	FFT pro Proz. [GFlop/s]	R A pro Proz. [MFlop/s]
AMD Opteron	QsNet II	0.099	0.212 (0.57)	0.073 (9.1)
Itanium2	Numalink	0.059	0.110 (0.52)	0.088 (4.7)
Xeon	Myrinet	0.012	---	---
Xeon 64T	Infiniband	0.030	0.162 (0.57)	0.067 (5.0)
NEC SX-6	Crossbar	0.539	0.234 (0.58)	0.041 (211)

MPG-DV Treffen 2004 O. Haan: HPC Challenge Benchmarks 15

Auch hier wird die Relevanz der verschiedenen Leistungszahlen erst durch die Angabe eines Lastprofils deutlich, das nun auch die Kommunikationsmuster beschreiben muss. Ein Beispiel für die Parametrisierung der Kommunikationslast gibt die nächste Abb.:



Aus den beiden Komponenten Einzelprozessorleistung und Kommunikationsleistung kann nun die Leistung des Gesamtsystems bezüglich einer definierten Zusammensetzung der Last ermittelt werden. Die folgende Abb. gibt ein Beispiel:



		r_E [GFlop/s]	L [GFlop/s]
AMD Opteron	QsNet II	0.38	0.15
Itanium2	Numalink	0.35	0.19
Xeon	Myrinet	0.11	0.03
Xeon 64T	Infiniband	0.22	0.09
NEC SX-6	Crossbar	3.50	0.44

4. Schlussbemerkung

Die Vorteile der Leistungsbestimmung eines Parallelrechnersystems durch standardisierte Benchmarks für typische Grundmuster von Programmverhalten sind von der GWDG bei zwei Beschaffungsmaßnahmen genutzt worden. Dabei wurden, da zur Zeit der ersten Ausschreibung die HPC Challenge Benchmarks noch nicht existierten, von der GWDG ausgewählte Benchmarks eingesetzt.

Erfahrungen bei der GWDG

- Beschaffung PC-Cluster 4Q. 2003
  99 Dual Xeon 3.06 GHz, SCI
- Beschaffung 64bit-Cluster 4Q. 2004
  32 Dual Opteron 2.2 GHz, Infiniband

- Verwendete Benchmarks:

1000 x 1000 LINPACK	(anstatt DGEMM)
DGEMV	(anstatt STREAM)
PingPong und MPI_ALLTOALL	(anstatt b_eff)

Trouble-Ticket-Systeme – Kriterien, Auswahl, Erfahrungen

Wilfried Grieger

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Einleitung

Die GWDG hat sich dazu entschlossen, eingehende Probleme, Anfragen, Aufträge und Fehlermeldungen weitestgehend elektronisch zu verwalten und innerhalb dieses Systems auch zu bearbeiten. Die Gründe dafür sind vielfältig. Zum einen wurde von Benutzerseite beklagt, dass einige Anfragen bei der GWDG bisher doch verloren gehen können. Zum anderen wurde die Dauer der Abarbeitung bemängelt. Ein elektronisches System soll also auf jeden Fall das Verlorengehen verhindern und zumindest statistisch ausweisen können, wie lang die Bearbeitungszeit gedauert hat.

Als elektronisches System bietet sich ein Trouble-Ticket-System an, das in einem Helpdesk eingesetzt wird. Was ein Trouble-Ticket-System leisten kann oder auch leisten muss, ist im RFC 1297[1] beschrieben. Ein Trouble Ticket wird dabei mit einem Krankenblatt im Krankenhaus verglichen, das einen Patienten von der Aufnahme bis zur Entlassung begleitet. Ähnlich soll

1. http://www.faq.org/rfcs/rfc1297.html

es mit einem gemeldeten Problem sein: Im Trouble Ticket wird all das vermerkt, was von der Meldung bis zur Lösung an Aktivitäten erfolgt ist.

1. Anforderungen der GWDG an ein Trouble-Ticket-System

Die GWDG hat einen eigenen Anforderungskatalog an ein Trouble-Ticket-System erstellt:

1. Eingehende Anfragen, Aufträge und Fehlermeldungen sollen erfasst, bearbeitet und weiter geleitet werden können. Nach einer festzulegenden Zeit, in der das Trouble Ticket nicht bearbeitet wird, soll automatisch ein Eskalationsmechanismus in Gang gesetzt werden.
2. Abgeschlossene Probleme sollen nachverfolgt werden können, um eventuelle Lösungssackgassen oder auch Musterlösungen zu erkennen.
3. Über die Bearbeitung von Trouble Tickets sollen Statistiken erzeugt werden können, um die Güte der Bearbeitung nachzuweisen.
4. Die Kunden- und Bearbeiter-Schnittstellen sollen per WWW-Browser oder per E-Mail bereitgestellt werden. Ein eigener Client-Zugang ist nicht erwünscht, weil in der heterogenen Wissenschaftsumgebung nur über einen WWW-Browser allen der Zugang eröffnet wird.
5. Eine Wissensdatenbank soll erstellt werden, um Probleme schon auf Grund bereits vorhandener Einträge in der Wissensdatenbank lösen zu können.
6. Das Trouble-Ticket-System soll leicht durch automatische Fehlermeldung von Servern oder anderen Überwachungssystemen erweitert werden können.

Die Überwachung von Service Level Agreements (SLAs) soll erst später in ein Trouble-Ticket-System integriert werden.

2. ITIL-Konformität

Heutzutage wird häufig bei neu einzuführenden Geschäftsprozessen, wie beispielsweise bei der Bearbeitung eines Trouble Tickets, die ITIL-Konformität gefordert. ITIL ist dabei die Abkürzung für **I**nformation **T**echnology **I**nfrastructure **L**ibrary. Diese vom Office of Government Commerce (OGC) zusammen gestellte Bibliothek beschreibt IT-Prozesse, die einen effizienten und effektiven Betrieb der gesamten IT-Infrastruktur ermöglichen, um die zwischen der IT-Organisation und ihren Kunden vereinbarten Service Levels einhalten zu können.

Die von der GWDG untersuchten Trouble-Ticket-Systeme beanspruchten alle für sich, ITIL-konform zu sein.

3. Help-Desk-Forum

Auf dem Markt wird zur Zeit eine schier unübersehbare Menge von Trouble-Ticket-Systemen angeboten, so dass eine Auswahl extrem schwerfällt. Glücklicherweise findet alljährlich in Mainz eine Konferenz des Help-Desk-Forums[2] statt, auf der insbesondere über Trouble-Ticket-Systeme diskutiert und vorgetragen wird. Eine mehr als 50 Anbieter umfassende Liste der führenden Systeme ist auf den dortigen WWW-Seiten abgelegt, aus der eine Auswahl getroffen werden kann.

4. Auswahl

Aus der Liste des Help-Desk-Forums wurden die folgenden Systeme ausgewählt und genauer untersucht:

1. **Touchpaper/VEGA**
   Die Firma Touchpaper hat mit ihrem Produkt VEGA ein eigenständiges System entwickelt, das sich jedoch nur mit Hilfe der Firma anpassen und erweitern lässt. Auch Schnittstellen zu anderen Systemen sind nicht offen, sondern müssen durch Touchpaper programmiert werden. Als Datenbanken lassen sich Oracle, SQL-Server und Sybase anschließen. Benachrichtigungsmechanismen und Eskalationen werden ausschließlich innerhalb des Systems abgewickelt. Die Software ist rein client-basiert und verfügt nur ansatzweise über einen WWW-Zugang. Zur Erzeugung von Statistiken ist der mitgelieferte Report Generator unübertrefflich.

2. **IntraWare/Gedys**
   Das von den Firmen IntraWare und Gedys angebotene System basiert auf Lotus Software und verwendet dementsprechend auch eine Lotus-Datenbank. Angepasst und erweitert werden kann es dann nur durch Programmiereingriffe der Firmen. Als WWW-Client müsste ein eigenes Modul beschafft werden, das auf dem Internet Explorer basiert. Reports müssten ebenfalls durch die Firmen programmiert werden. Eskalationen können sowohl innerhalb des Systems als auch über E-Mail oder SMS abgewikkelt werden.

3. **HP ServiceDesk**
   Das von der Firma Hewlett Packard angebotene ServiceDesk integriert in

2. http://www.help-desk-forum.de

idealer Weise alle von HP-Produkten erzeugten Meldungen in ein Trouble-Ticket-System. Leider lässt sich dieses nur innerhalb der HP-Produkte anpassen und nur schwer mit Meldungen fremder Produkte erweitern. Als Datenbanken lassen sich sowohl Oracle als auch SQL-Server anschließen. Reports können nur über HP-Produkte erzeugt werden. Ein WWW-Client wird erst ab dem ersten Quartal 2005 verfügbar sein. Eine Authentifizierung über einen externen LDAP-Server ist nicht möglich, da ServiceDesk eine eigene Benutzerverwaltung integriert hat.

4. **CA Unicenter**
   Das Helpdesk-System der Firma Computer Associates ist nur ein Teil des Management-Systems CA Unicenter. Es integriert demzufolge auch ausschließlich die Module dieses Systems. Es existieren zwar Schnittstellen zu Fremdprodukten; diese müssten allerdings vollständig neu programmiert werden. Reports sind lediglich konfigurierbar, nicht neu erstellbar. Als Datenbanken lassen sich sowohl Oracle als auch SQL-Server anbinden.

5. **Remedy/ARS**
   Das Action Request System (ARS) der Firma Remedy besteht aus einem beliebig erweiterbaren Baukasten, der an bestehende Konfigurationen leicht angepasst werden kann. Schnittstellen zu anderen Systemen stehen über Perl und C zur Verfügung. Nahezu beliebige Daten können importiert werden. Als Datenbank sind Oracle und viele andere Systeme anschließbar. Über ein eigenes Modul können alle gängigen WWW-Browser als Client eingesetzt werden. Ein Report-Generator wird mitgeliefert.

6. **OTRS**
   Das Open source Ticket Request System (OTRS) basiert auf Open-Source-Software, lässt sich damit nahezu beliebig anpassen und erweitern, wobei die Programmierarbeit selber oder von einer Fremdfirma durchgeführt werden muss. Die Benachrichtungsmechanismen basieren ausschließlich auf E-Mail. Als Client gibt es nur den WWW-Browser, da die Oberfläche durch reines HTML dargestellt wird. Es stehen allerdings keine vorgefertigten Reports zur Verfügung.

5. Engere Wahl

Auf Grund der oben zusammen gestellten Haupteigenschaften der untersuchten Systeme sind Remedy/ARS und OTRS in die engere Wahl gekommen.

Die Vorteile von Remedy/ARS gegenüber den anderen Systemen sind: Es handelt sich um den Marktführer im Helpdesk-Bereich. Es handelt sich um ein Baukastenprinzip. Es gibt sehr unterschiedliche Eskalationsmechanismen. Ein Report-Generator wird mitgeliefert. Die Nachteile sind: Sowohl die Anschaffung als auch der Betrieb des Systems sind sehr teuer.

Die Vorteile von OTRS gegenüber den anderen Systemen sind: Die Software ist Open Source. Die Oberflächen sind rein HTML-basiert. Beliebige Eskalationsmechanismen lassen sich programmieren. Die Anschaffung ist extrem billig. Die Akzeptanz bei den Mitarbeitern ist wegen der offenen Strukturen sehr hoch. Die Nachteile sind: Es gibt keine fertigen Reports. Die Dokumentation ist schlecht.

Nach dem Abwägen der Vor- und Nachteile fiel die Entscheidung zu Gunsten von OTRS.

6. OTRS im Einsatz

Die Installation des Systems ist bereits erfolgt. Zur Zeit (November 2004) wird das System konfiguriert und dabei auch innerbetrieblich getestet.

Ab dem ersten Quartal 2005 soll das Trouble-Ticket-System dann von allen Mitarbeiterinnen und Mitarbeitern der GWDG genutzt werden, wobei gleichzeitig intensive Schulungsmaßnahmen durchgeführt werden. Ab dem zweiten Quartal 2005 soll es dann in den Benutzerbetrieb überführt werden.

Als besonders nachteilig hat sich die schlechte Dokumentation erwiesen, trotzdem scheint die Entscheidung für OTRS richtig zu sein.

Überblick über die Cybercrime Konvention des Europarates und ihre Implikation für die Tätigkeit von Systemadministratoren

Marco Gercke

Ludwig-Maximilians-Universität München

Die „Convention on Cybercrime" (im folgenden Cybercrime Konvention) des Europarates (Treaty 185) zählt neben dem Rahmenbeschluss der EU zum Schutz vor Angriffen auf Informationssysteme[1] zu den bedeutensten interna-

1. Rahmenbeschluss 2005/222/JI des Rates. Das mit zahlreichen Initiativen auf EU-Ebene verbundene Ziel, die Sicherheit innerhalb der für die moderne Informations- und Dienstleistungsgesellschaft so bedeutenden Kommunikationsinfrastruktur sicher zu stellen, wurde vom Rat im Rahmenbeschluss im Wesentlichen auf die Harmonisierung des materiellen Strafrechts beschränkt. Darin kommt die Überzeugung zum Ausdruck, dass die Angleichung der Strafnormen die einerseits notwendige und andererseits ausreichende Voraussetzung für eine effektive Bekämpfung der Computerkriminalität schaffe. Ausweislich der Begründung zum ersten Vorschlag des Rahmenbeschlusses sah der Rat „die rechtlichen Unterschiede und Diskrepanzen" als ein Hindernis für die wirksame polizeiliche und justizielle Zusammenarbeit" an.

tionalen Übereinkommen im Bereich des Informationsstrafrechts. Neben der Harmonisierung des materiellen Strafrechts beinhaltet die Konvention eine Reihe von strafprozessualen Ermittlungsinstrumenten, von denen insbesondere Art. 19, der die Durchsuchung und Beschlagnahme gespeicherter Computerdaten thematisiert, für die Tätigkeit der Systemadministratoren von zentraler Bedeutung sein wird.

1. Einleitung

Am 23.11.2001 haben 30 Staaten, unter ihnen auch Deutschland, in Budapest die Cybercrime Konvention des Europarates unterzeichnet.[2] Bei der Konvention handelt es sich um einen völkerrechtlichen Vertrag, der in Übereinstimmung mit den meisten nationalen Regelungen gemäß Art. 36 Abs.2 der Konvention der Ratifikation, bzw. eines vergleichbaren Aktes der Umsetzung in nationales Recht bedarf. [3] Bisher haben zehn Unterzeichner die Konvention ratifiziert.[4] Anders als bei dem im Hinblick auf die Zielrichtung vergleichbaren Rahmenbeschluss der Europäischen Union über Angriffe auf Informationssysteme, besteht keine Möglichkeit, die Unterzeichner der Konvention nach Ablauf einer Frist zur Umsetzung der Regelungen in nationales Recht zu zwingen.[5]

---

2. Neben Deutschland haben am 23.11.2001 die Europaratsmitglieder Albanien, Armenien, Österreich, Belgien, Bulgarien, Estland, Finnland, Frankreich, Griechenland, Großbritannien, Holland, Italien, Kroatien, Moldawien, Norwegen, Polen, Portugal, Rumänien, Spanien, Schweden, Schweiz, Ehemalige Jugoslawische Republik Mazedonien, Ukraine, Ungarn und Zypern, sowie die Nichtmitgliedsstaaten Kanada, Japan, Südafrika und USA die Konvention unterzeichnet. Später folgten Bosnien und Herzegowina (2005), Dänemark (2003), Irland (2002), Island (2001), Lettland (2004), Litauen (2003), Luxemburg (2003), Malta (2002), Slovakei (2005), Slowenien (2002), Tschechien (2005) - (Stand: April 2005). Der aktuellen Stand der Unterzeichnungen kann auf den Internetseiten des Europarates abgerufen werden.

3. Zu den allgemeinen Regeln des Völkerrechts, die ohne Transformationsakt Teil des Bundesrechts sind, zählen nur solche Normen, die von der weit überwiegenden Mehrheit der Staaten als verpflichtend anerkannt sind. Vgl. BverfG 15, 34.; Stern, Band 1, § 14.

4. Albanien, Bulgarien, Ehemalige jugoslawische Republik Mazedonien, Estland, Kroatien, Litauen, Rumänien, Slovenien, Ungarn und Zypern.

Obwohl die Konvention wichtige Instrumente enthält, die sowohl für die nationale Strafverfolgung, als auch für die Koordinierung internationaler Ermittlungsmaßnahmen unerlässlich sind, fehlt es in Deutschland bislang an einer Umsetzung der Konvention in nationales Recht.

2. Aufbau und Inhalt der Konvention

Die Konvention besteht aus vier Kapiteln. Im ersten Kapitel werden zentrale Begriffe legaldefiniert. Das zweite Kapitel enthält einen Katalog von Strafvorschriften, die als internationaler Mindeststandard die grenzüberschreitende Strafverfolgung erleichtern sollen. Ohne eine entsprechende Harmonisierung der nationalen Strafvorschriften würden sich transnationale Ermittlungen schwierig gestalten. Zu den im zweiten Kapitel aufgeführten Strafvorschriften zählen u.a. der Schutz von Computersystem und Daten vor unberechtigten Zugriffen und die Strafbarkeit der Verbreitung von Kinderpornographie über das Internet. Weite Teile der in der Konvention enthalten Strafvorschriften sind bereits im geltenden deutschen Strafrecht enthalten, so dass sich der Anpassungsbedarf insoweit auf wenige Strafvorschriften beschränkt. Darüber hinaus umfasst das zweite Kapitel einen Katalog grundlegender strafprozessualer Ermittlungsinstrumente. Neben einer Reihe von Eingriffsbefugnissen, die bereits in der geltenden Strafprozessordnung enthalten sind, enthält die Konvention auch neue Instrumente. Für die Tätigkeit der Systemadministratoren von besonderer Bedeutung sind dabei die Anordnung der beschleunigten Sicherung in Art. 16 und die Erweiterung der bestehenden Durchsuchungsvorschriften durch Art. 19 der Konvention, die im Folgenden näher dargestellt werden. Das dritte Kapitel enthält Regelungen zur internationalen Kooperation und Rechtshilfegesuchen, während das letzte Kapitel die Schlussbestimmungen enthält.

Die Konvention wurde im ersten Zusatzprotokoll vom 7.11.2002 um Strafvorschriften im Hinblick auf rassistische und fremdenfeindliche Inhalte ergänzt.[6] Eine Integration dieser Vorschriften in die Konvention selbst erwies sich als problematisch, da sie Ländern, wie den USA, mit einer extensiven, verfassungsrechtlich garantierten Meinungs- und Kommunikationsfreiheit, die Unterzeichnung der Konvention unmöglich gemacht hätte.[7]

---

5. Zum rechtspolitischen Mehrwert der EU-Initiative vgl. auch Sanchez-Hermosilla, CR 2003, 778.

6. Vgl. dazu Kugelmann, DuD 2003, 345ff.; Gengler, Computer Fraud & Security, Volume 2002, Issue 4, Page 4ff.

3. Anordnung der beschleunigten Speicherung („quick freeze“)

Gemäß Art. 16 der Konvention sind die Vertragsparteien verpflichtet, im nationalen Recht die Möglichkeit einer Anordnung der beschleunigten Sicherung von Computerdaten, insbesondere Verbindungsdaten, vorzusehen. Die Vorschrift trägt dem Umstand Rechnung, dass die für die Rückverfolgung von Übertragungsvorgängen notwendigen Verbindungsdaten meist nur für sehr kurze Zeit bei den beteiligten Diensteanbietern vorliegen.[8] Die Anordnung der beschleunigten Speicherung soll den Strafverfolgungsbehörden insoweit die Möglichkeit geben, die Löschung von beweisrelevanten Computerdaten zeitnah durch eine entsprechende Sicherungsanordnung verhindern zu können. Sie umfasst nur die Sicherung der Daten, aber berechtigt die Strafverfolgungsbehörden nicht, die Herausgabe der Daten anzuordnen. Die Herausgabe der Daten erfolgt über eine gesonderte Anordnung, die in Art. 18 der Konvention im Detail geregelt ist.

Für die Systemadministratoren ist in soweit von Bedeutung, dass die Maßnahme im Unterschied zur klassischen Durchsuchung und Beschlagnahme, die nur mit einer Duldungspflicht einhergeht, eine aktive Mitwirkungspflicht der Betroffenen – im Regelfall der Systemadministratoren – begründet. Bislang existiert ein entsprechendes Instrument im deutschen Recht. Die bestehenden Ermittlungsinstrumente, wie beispielsweise die Beschlagnahme von Datenträgern mit Verbindungsdaten greifen erheblich weiter in die Rechtssphäre der Betroffenen ein als die Anordnung der Sicherung, weshalb sie unter einem Richtervorbehalt stehen. Die durch die gerichtliche Prüfung und den physischen Zugriff bedingte Verfahrensdauer hat sich im Zusammenhang mit Ermittlungen im Bereich der Internetkriminalität als problematisch erwiesen. Unter dem Druck der praktischen Notwendigkeit, aber weitgehend ungeklärten rechtlichen Rahmenbedingungen hat sich in den letzten Jahren in vielen Strafverfahren ein der Regelung des Art. 16 vergleichbares Procedere durchgesetzt.[9]

---

7. Schwarzenegger in Festschrift für Trechsel, S. 309; Rechtsvergleichend zum Spannungsverhältnis von Meinungsfreiheit und strafrechtlicher Verantwortlichkeit am Beispiel illegaler Inhalte Holznagel, ZUM 2000, 1007.

8. Zur Definition der Verbindungsdaten vgl. Art. 1 d) der Konvention. Zur Funktion der Verbindungsdaten bei der Strafverfolgung im Internet vgl. Gercke, DuD 2002, 477ff.

Für die kooperationsbereiten Unternehmen und insbesondere Systemadministratoren beinhaltet die in Art. 16 enthaltene Regelung zunächst eine wesentliche Voraussetzung für die Zusammenarbeit mit den Ermittlungsbehörden, da die Diensteanbieter durch die Anordnung sowohl von häufig bestehenden vertraglichen Verpflichtungen gegenüber ihren Nutzern zur Geheimhaltung der Verbindungsdaten, als auch von datenschutzrechtlicher Bestimmungen im geltenden Recht befreit werden. Die Umsetzung der Regelung in das nationale Recht schafft damit erstmals die Grundlage für die Kooperation zwischen den Diensteanbietern und den Ermittlungsbehörden.

Darüber geht mit der Implementierung eines entsprechenden Ermittlungsinstrumentes aber auch die Notwendigkeit einher, die technischen und personellen Voraussetzungen dafür zu schaffen, zeitnah auf die Anordnung reagieren zu können. Nur so kann das Ziel des Ermittlungsinstrumentes erreicht werden. Dabei ist von besonderer Bedeutung, dass die Anordnung nur solche Daten umfasst, die zum Zeitpunkt der Anordnung in gespeicherter Form vorliegen; sie verpflichtet die Diensteanbieter weder zur Erhebung bestimmter Daten, noch begründet sie eine Pflicht zur Vorratsdatenspeicherung.[10]

Im Vergleich zu den Auswirkungen und Risiken der Einführung einer derzeit intensiv diskutierten Vorratsdatenspeicherpflicht[11] zeigt sich der Vorteil der Anordnung der sofortigen Speicherung, die als begrenztes Ermittlungsinstrument einer pauschalen Speicherung vorzuziehen ist.

Während grundsätzlich, wie oben dargestellt, die Speicheranordnung die Systemadministratoren nur zur Sicherung der Daten, nicht aber zu deren Herausgabe an die Ermittlungsbehörden verpflichtet, findet sich in Art. 17 der Konvention eine Sonderregelung für Verbindungsdaten, welche die Spei-

9. Auf die informelle Anfrage und Anregung von Strafverfolgungsbehörden werden sensible Daten von den Providern häufig vorläufig gesichert, um die Chancen für einen Zugriff in einem späteren formellen Verfahren zu wahren.

10. Explanatory Report Nr. 152.

11. Die vom Bundesrat, basierend auf der Empfehlung des Rechtsausschusses (BR-Drs 755/03), angeregte Vorratsspeicherung von Verkehrsdaten hat in der Novelle des TKG (BR-Drs. 15/2329) keine Mehrheit gefunden. Die Bundestagsfraktionen sprachen sich geschlossen gegen eine pauschale Kriminalisierung der Nutzer aus. Zur Vorratsspeicherung von Nutzungs- und Verbindungsdaten vgl. Dix, DuD 2003, 234; Heidrich, DuD 2003, 237; Breyer, DuD 2003, 491.

cheranordnung mit einer begrenzten Herausgabeaufforderung kombiniert. Art. 17 der Konvention verpflichtet die Vertragsparteien im Hinblick auf Verbindungsdaten, die umfangreiche Sicherung auch dann zu ermöglichen, wenn mehrere Diensteanbieter an Datenübertragungsprozessen beteiligt sind. Der Grund für die Sonderbehandlung der Verbindungsdaten ist der Umstand, dass diese, anders als Inhaltsdaten, meist keinen direkten Rückschluss auf die unmittelbaren Beteiligten zulassen. Aufgrund der Beteiligung unterschiedlicher Diensteanbieter muss die Rückverfolgung vielmehr schrittweise erfolgen.[12] Bereits im Rahmen der beschleunigten Speicheranordnung muss daher erstmals die Kommunikationsroute rekonstruiert werden, um die Diensteanbieter, die an der Übertragung beteiligt waren, zu identifizieren und zur Speicherung der Verbindungsdaten aufzufordern. Um die Ermittlungsbehörden in die Lage zu versetzen, eine solche Rückverfolgung vorzunehmen, sieht Art. 17 neben der Speicheranordnung des Art. 16 einen Herausgabeanspruch im Hinblick auf den Teil der gesicherten Daten vor, der zur Rekonstruktion der Kommunikationsroute erforderlich ist.[13]

4. Durchsuchung und Beschlagnahme

Art. 19 Abs.1 der Konvention verpflichtet die Vertragsparteien, die gesetzlichen Voraussetzungen für die Durchsuchung von Computersystemen, Datenträgern und den gespeicherten Daten zu schaffen. Das Ziel der Regelung ist die Harmonisierung und Modernisierung der teilweise sehr unterschiedlichen nationalen Regelungen.[14] Neben der Erweiterung der Durchsuchungsobjekte und der Befugnisse der Ermittlungsbehörden im Rahmen der Beschlagnahme liegt der bedeutendste Unterschied zum geltenden Recht in Art. 19 Abs.4 der Konvention. Dieser erweitert die rein passive Duldungspflicht hinsichtlich der Durchsuchung und Beschlagnahme dahingehend, dass Personen, die Kenntnisse über die Funktionsweise der von den vorgenannten Ermittlungsmaßnahmen betroffenen Computer oder vorhandener Maßnahmen zum Schutz der darin enthaltenen Daten haben, also insbesondere Systemadministratoren, aktive Mitwirkungspflichten auferlegt werden können. Die Konvention sieht eine Verpflichtung dieser Personen

12. Zu den technischen und juristischen Anforderungen an eine Rückverfolgung von Straftätern im Internet vgl. Gercke, DuD 2002, 477ff.

13. Im Gegensatz zum zentralen Herausgabeanspruch des Art. 19 stellt der Teilherausgabeanspruch des Art. 17 damit ebenso wie Art. 16 eine Sofortmaßnahme dar.

14. Explanatory Report Nr. 184.

vor, zur Ermöglichung der Ermittlungsmaßnahmen in „vernünftigem Maße“ Auskunft zu erteilen.[15] Die Aufnahme des Auskunftsanspruchs basiert auf der Erkenntnis, dass für die Ermittlungsbehörden die Durchsetzung eines Durchsuchungsbeschlusses, insbesondere bei komplexen oder besonders gesicherten Computersystemen, mit erheblichen Schwierigkeiten verbunden ist. Zugleich sollen kooperationsbereite Diensteanbieter, ähnlich wie im Rahmen von Art. 16 der Konvention durch die gesetzliche Verpflichtung zur Auskunftserteilung von entgegenstehenden, vertraglichen oder sonstigen Geheimhaltungspflichten entbunden werden.

Für die Systemadministratoren wird dabei in praktischer Hinsicht von entscheidender Bedeutung sein, ob und ggf. wie der Gesetzgeber die wenig präzise Vorgabe der Konvention hinsichtlich des Maßes der erforderlichen Kooperation konkretisiert. In Anbetracht der zu erwartenden Konfliktsituationen, in denen sich die Systemadministratoren befinden, ist eine präzise Anordnung erforderlich. Die Begründung der Konvention ist insoweit nicht unproblematisch, da sie Tendenzen erkennen lässt die Prüfungspflicht, ob eine Auskunftspflicht besteht, faktisch auf die Systemadministratoren zu verlagern. Weigern diese sich unberechtigt, kann dies zu einer unzulässigen Behinderung der Ermittlungen führen – kooperieren sie, ohne dass eine Verpflichtung besteht, kann dies eine Vertragsverletzung gegenüber den Nutzern darstellen und möglicherweise zivilrechtliche Konsequenzen nach sich ziehen.

5. Zusammenfassung

Die Konvention enthält zahlreiche Vorschriften, die für die Strafverfolgungsbehörden von zentraler Bedeutung sind und ohne deren weltweite Implementierung die Identifikation von Straftätern im Internet erheblich behindert würde. Im Rahmen der Umsetzung der Konvention werden die Systemadministratoren sowohl im Rahmen der Schaffung der technischen und organisatorischen Maßnahmen zur Umsetzung von Speicheranordnungen, als auch im Rahmen von unmittelbar an sie gerichtete Auskunftsersuchen während Durchsuchungsmaßnahmen in die Pflicht genommen. Die Umsetzung der Konvention in das nationale Recht wird zeigen, ob es dem Gesetzgeber gelingt, die für die Praxis so wichtigen Vorgaben der Konvention umzusetzen, ohne die Prüfungspflichten hinsichtlich der Voraussetzungen für die

15. Mögliche Mitwirkungshandlungen könnten beispielsweise das Offenbaren von Passwörtern oder die Lokalisierung von Daten sein. Vgl. Explanatory Report 201ff.

Mitwirkungspflichten in unbilliger Weise auf die Systemadministratoren zu verlagern.

LDAP in der GWDG – Einsatzspektrum

Konrad Heuer, Andreas Ißleiber

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Einleitung

In diesem Vortrag wird dargestellt, wie sich

- das Einsatzspektrum von OpenLDAP bei der GWDG zurzeit darstellt,
- wie die aktuelle Konfiguration der Server aussieht,
- wie sich die Verzeichnisdaten darstellen,
- welche Zielvorgaben es gibt,
- und wie die Anbindung der RADIUS-Server vorgesehen ist.

1. Einsatzzwecke

Insbesondere viele der bei der GWDG im Einsatz befindlichen Linux-Systeme, an erster Stelle zu nennen die Knoten der als Parallelrechner betriebenen Compute-Cluster, verwenden OpenLDAP zur Benutzeranmeldung und -verwaltung. Hinzu kommen einige Spezialdienste wie Lotus- oder Apache-Web-Server sowie Rechner mit FreeBSD oder Mac OS X als Betriebssystem.

2. Server-Konfiguration

OpenLDAP ist auf unterschiedlichen UNIX-Varianten sowie Linux einsetzbar. Die GWDG betreibt OpenLDAP-Server unter FreeBSD und Linux.

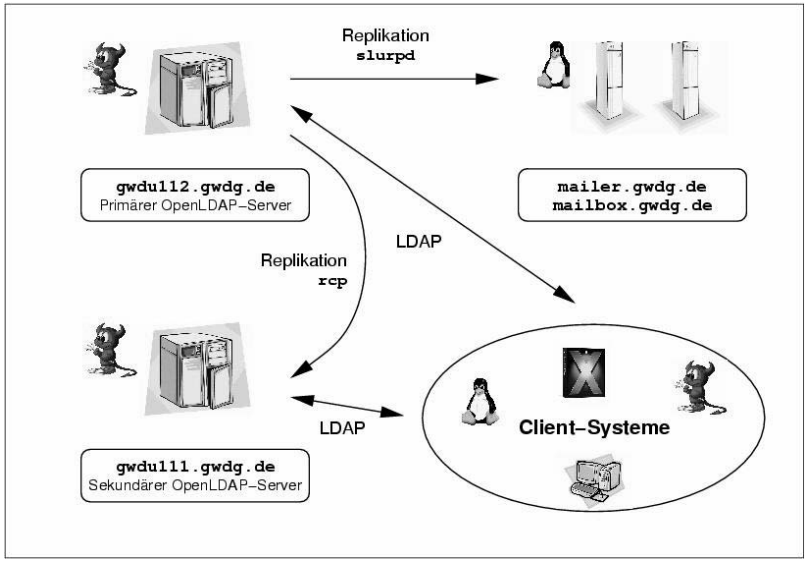


Abb. 1: OpenLDAP-Server der GWDG

Abb. 1 zeigt die Zusammenhänge zwischen den von der GWDG betriebenen OpenLDAP-Servern und -Klienten, wobei das jeweilige Betriebssystem durch Symbole angedeutet wird (Daemon – FreeBSD, Pinguin – Linux, X – Mac OS X).

Der primäre Server **`gwdu112`** ist die Quelle aller Verzeichnisdaten; der Abgleich mit der traditionellen UNIX-Benutzerdatenbank der GWDG NIS (Network Information System von SUN, früher *yellow pages* genannt) erfolgt zurzeit mit Hilfe kleiner Programme.

Zur Erhöhung der Ausfallsicherheit und zur Verteilung der Last werden zusätzliche OpenLDAP-Server verwendet, welche die Verzeichnisdaten per Replikation übermittelt bekommen.

OpenLDAP bietet zwei Mechanismen zur Replikation an: Zum einen kann der primäre Server über einen als UNIX-Daemon laufendes Programm **`slurpd`** alle Änderungen zeitnah an sekundäre Server weitergeben, oder diese holen periodisch Änderungen beim primären Server ab (sog. *sync*

replication). Das zweite Verfahren ist relativ neu und gilt als noch nicht stabil genug.

Die **slurpd**−Replikation wird verwendet, um die OpenLDAP-Server auf den Linux-Mailsystemen der GWDG zu aktualisieren. Die Mailserver sind selbst OpenLDAP-Klienten mit einem hohen Lastpotenzial und wenden sich an ihre eigenen exklusiven Replikationsserver, um die Last nicht anderen Servern aufzubürden und auch bei Ausfall der anderen OpenLDAP-Server oder der Netzwerkkommunikation dorthin autonom weiterarbeiten zu können.

Da auch die **slurpd**-Replikation in seltenen Fällen zu Inkonsistenzen führen kann, kopiert der sekundäre OpenLDAP-Server **gwdu111** periodisch die kompletten Datenbanken, auf die OpenLDAP sich stützt, per **rcp**-Befehl von dem Rechner **gwdu112**. Beide Rechner sind u. a. zu diesem Zweck unmittelbar über separate Netzwerkadapter und ein gekreuztes Netzwerkkabel verbunden.

Abgesehen von den Linux-Mailsystemen wenden sich alle OpenLDAP-Klienten an die Server **gwdu112** und **gwdu111**.

Der Zugang zu den Verzeichnisdaten ist bis auf einzelne Ausnahmen auf Rechner im IP-Adressbereich des GÖNET beschränkt.

3. Verzeichnisdaten

In Abb. 2 ist ein Beispieleintrag für eine GWDG-Kursbenutzerkennung zu erkennen. Bislang beinhaltet das von OpenLDAP verwaltete Verzeichnis hauptsächlich Daten, die für eine Benutzeranmeldung unter UNIX (z. Z. nur FreeBSD oder Mac OS X) oder Linux benötigt werden.

```
dn: cn=4kurst00,ou=GKRS,ou=gwdgadm,dc=gwdg,dc=de
objectClass: inetOrgPerson
objectClass: person
objectClass: posixAccount
objectClass: shadowAccount
objectClass: inetLocalMailRecipient
objectClass: GWDGuser
objectClass: top
cn: 4kurst00
cn: 4kurst00 Kursleiter(in)
gidNumber: 5050
ou: GKRS
gecos: 4kurst00 Kursleiter(in) ,,,
givenName: 4kurst00
mail: 4kurst00@gwdg.de
mailHost: -
mailPrefs: -
homeDirectory: /usr/users/4kurst00
loginShell: /bin/ksh
sn: Kursleiter(in)
uid: 4kurst00
uidNumber: 5265
userPassword:{crypt}4lq/F9j/UbOWg
```

Abb. 2: Beispiel für einen Verzeichniseintrag

Zu erkennen ist, dass ein Nutzer durch mehrere Objektklassen beschrieben wird, beispielsweise durch **`posixAccount`** und **`shadowAccount`**. Diese Objektklassen beinhalten ihrerseits Attribute wie **`homeDirectory`** und **`loginShell`**, die notwendig sind, damit OpenLDAP die Benutzerverwaltung für UNIX-Systeme übernehmen kann.

Ein großer Sicherheitsvorteil gegenüber der Benutzerverwaltung per NIS ist, dass das verschlüsselte Benutzerkennwort nur bei einer autorisierten Anmeldung am OpenLDAP-Server durch den Nutzer selbst oder einen besonders privilegierten Administrator sichtbar wird.

Alle Nutzer der existierenden NIS-Datenbasis sind automatisch in das Verzeichnis übernommen worden, und die Einträge werden kontinuierlich gepflegt.

Ein weiterer Vorteil von OpenLDAP gegenüber NIS ist die Möglichkeit der hierarchischen Strukturierung der Benutzereinträge. Abb. 2 zeigt dies

anhand des *distinguished name* (**dn**-Zeile ganz oben): Jedes Institut stellt im Verzeichnis eine eigene Organisationseinheit (*organizational unit* – **ou,** im Beispiel **GKRS**) dar, wodurch sich bei Bedarf Verantwortung für die jeweiligen Benutzereinträge leichter delegieren lässt. Auch kann der Zugriff von Klienten so auf ein einzelnes Institut begrenzt werden.

Neben der Organisationseinheit **gwdgadm** für die GWDG-Benutzerverwaltung (ebenfalls in Abb. 2 zu erkennen) existieren andere Organisationseinheiten wie **gwdgovid** für die OVID-Benutzer der Max-Planck-Gesellschaft oder auch schon institutsspezifische Organisationseinheiten, die Instituten eine Absicherungsmöglichkeit für selbstbetriebene OpenLDAP-Server bieten können.

4. Zielvorgaben

Für das Jahr 2005 ist geplant, das OpenLDAP-Verzeichnis als primäre Datenquelle für die GWDG-Benutzerverwaltung einzusetzen und auch einen gewissen Abgleich mit dem *Active Directory* der Windows-Welt vorzunehmen. Dazu ist die Einführung einer ganzen Reihe weiterer Attribute erforderlich, die in Abb. 2 noch nicht zu erkennen sind. Sie sind in einem GWDG-eigenen Schema definiert und der Objektklasse **GWDGuser** zugeordnet.

Über ein Kennwort-Portal sollen Verschlüsselungen für alle Betriebssysteme gesetzt werden können, um ein einheitliches Kennwort zu ermöglichen; in diesem Zusammenhang wird auch eine Anbindung der Samba-Server vorgenommen werden, die bislang mit eigenen Kennwortdateien arbeiten.

Bei dieser Neustrukturierung des OpenLDAP-Verzeichnisses im Zuge der Übernahme der primären Benutzerdatenbankfunktion wird es sogar möglich sein, unterhalb der Institutsebene Abteilungen als zusätzliche Organisationseinheiten einzuführen, was für große Institute wie z. B. das Max-Planck-Institut für biophysikalische Chemie wichtig sein kann. Die gemeinsame Mitgliedschaft in einer UNIX-Gruppe (hier dann **MBPC**) bleibt davon unberührt.

5. Anbindung der RADIUS-Server

RADIUS ist ein sehr verbreitetes Authentifizierungsverfahren, welches mittlerweile von nahezu jeder Hardware unterstützt wird. RADIUS benötigt einen Zugriff entweder auf eine eigene lokale Datenbank oder auf externe Datenbanken, um Benutzer authentifizieren zu können. Bei der GWDG dient RADIUS als Schnittstelle zwischen den Systemen, die über das RADIUS-Protokoll die Benutzer authentifizieren, und der zentralen Benutzerdaten-

bank in Form eines LDAP-Servers. Der RADIUS-Server benutzt den LDAP-Server hierbei als reine Authentifizierungsinstanz. Weitergehende Attribute, welche RADIUS den anfragenden Systemen (z. B. Einwahlsysteme oder VPN-Gateways) zur Verfügung stellt, definiert der RADIUS-Server selbst. Das kann z. B. die Vergabe von bestimmten IP-Adressbereichen bei der Einwahl in Abhängigkeit von der entsprechenden Gruppe eines Benutzers sein.

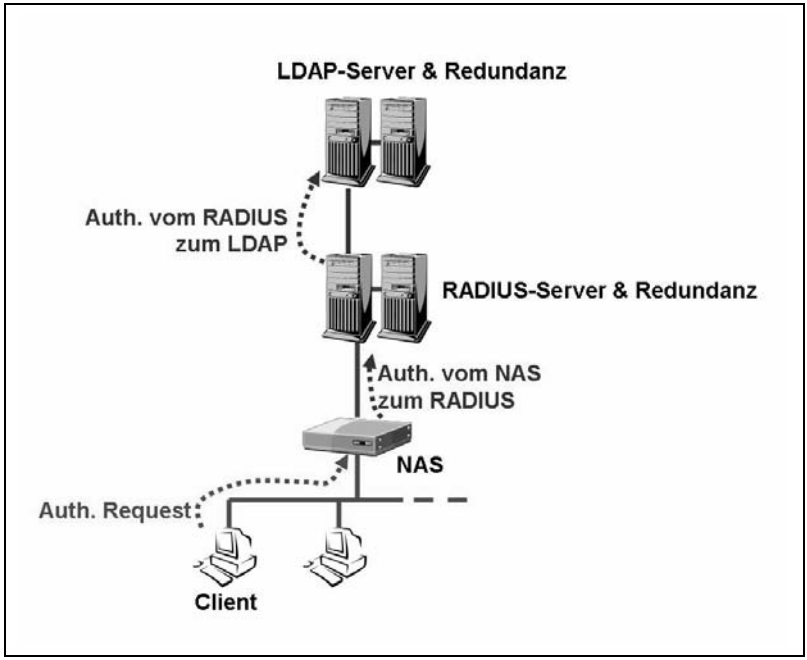


Abb. 3: Kommunikation zwischen RADIUS- und LDAP-Server

PKI-Leistungen der GWDG

Sebastian Rieger

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

1. Einleitung

Das Thema IT-Sicherheit ist in den letzten Jahren zu einem Schlüsselthema innerhalb der Informationstechnologie geworden. Viele Institute und Einrichtungen arbeiten an Sicherheitskonzepten, die u. a. auch die Integration von Zertifikaten (signierte öffentliche Schlüssel bzw. public keys) beinhalten. Die GWDG bietet für die Umsetzung dieser Anforderung praxisnahe Lösungsmöglichkeiten an.

Für die Verwaltung von Zertifikaten und deren Anwendung sind dabei Public-Key-Infrastrukturen (PKI) erforderlich, die die Basis für das Vertrauen in digitale Signaturen und die Verschlüsselung von Daten liefern. Zertifikate nach dem X.509-Standard basieren auf einem hierarchischen Vertrauensmodell, im Vergleich zum Web-of-Trust, bei dem einzelne Teilnehmer gegenseitig Ihre Schlüssel signieren (vgl. PGP). An der Spitze des hierarchischen Modells steht hierbei eine Wurzelzertifizierungsstelle, die durch einen gemeinsamen vertrauenswürdigen Dritten betrieben wird, und daher auch Vertrauensstellungen ohne persönlichen Kontakt der Kommunikationspartner oder die Notwendigkeit von gemeinsamen Bekannten ermög-

licht. Wurzelzertifizierungsstellen vieler Unternehmen werden bereits standardmäßig z. B. mit gängigen Web-Browsern oder Betriebssystemen ausgeliefert.

X.509-Zertifikate finden daher in vielen Anwendungen, z. B. der Authentifizierung von Web-Servern und Clients über HTTPS oder der Signatur und Verschlüsselung von E-Mails mittels S/MIME, Verwendung. In den meisten Anwendungen lassen sich Zertifikate ohne zusätzliche Erweiterungen direkt für die Erhöhung der IT-Sicherheit verwenden. Zertifikate ermöglichen dabei die Umsetzung der Grundlagen der IT-Sicherheit durch die vertrauliche Übertragung von Daten, die Wahrung ihrer Integrität und die verbindliche Zuordnung eines Absenders sowie eines Adressaten innerhalb der Kommunikation. Durch den Einsatz von Zertifikaten für die Authentifizierung von Servern, Clients und Benutzern und die Abwehr von unberechtigen Dritten kann überdies die Verfügbarkeit von Systemen als weiteres Grundelement der IT-Sicherheit unterstützt werden.

2. Struktur der GWDG-CA

Die PKI der GWDG stellt eine dreistufige Hierarchie, wie in der folgenden Abbildung gezeigt, dar:

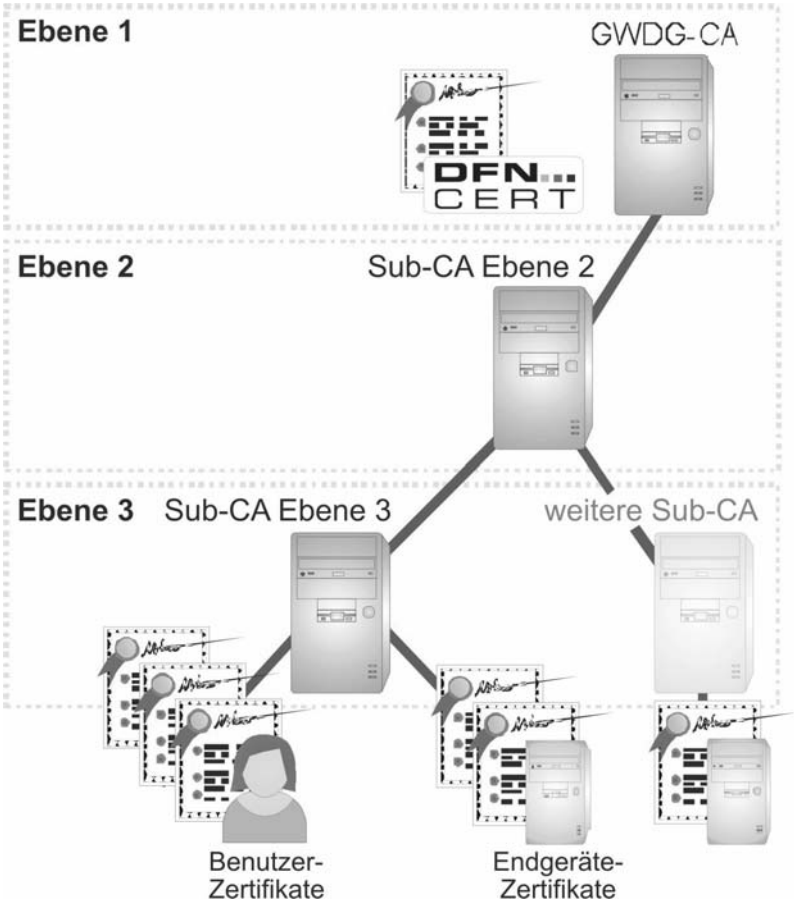


Abb. 1: Zertifizierungshierarchie der GWDG

Auf der **Ebene 1** der Hierarchie befindet sich dabei die oberste Zertifizierungsstelle der GWDG (GWDG-CA; CA steht für Certification Authority), deren Zertifikat von der Wurzel-Zertifizierungsstelle des DFN-Vereins (DFN-PCA) signiert wurde. Vertrauensgrundlage für die von der GWDG-CA ausgestellten Zertifikate stellen daher die Vergaben und die Struktur der PKI des DFN-Vereins dar.

Die zweite Hierarchieebene (**Ebene 2**) ermöglicht die Integration unterschiedlicher Organisationen in die dargestellte PKI. Im Ausnahmefall können auf dieser Ebene bereits Zertifikate für Endbenutzer bzw. Endgeräte ausgestellt werden, z. B. wenn diese für den Betrieb untergeordneter CAs erforderlich sind.

Zertifizierungsstellen ab der Ebene 2 können über einen Netzanschluss verfügen. Eine Absicherung der Systeme erfolgt nach der jeweils aktuellen IT-Sicherheitsleitlinie der GWDG sowie den zugehörigen spezifischen Richtlinien.

Auf der **Ebene 3** innerhalb der in Abb. 1 gezeigten Hierarchie können Zertifizierungsstellen eingerichtet werden, die ihrerseits Zertifikate für Endbenutzer bzw. Endgeräte ausstellen. Die Signierung von Zertifikaten für Endbenutzer und Endgeräte ist, abgesehen von der genannten Ausnahme für die Ebene 2, erst ab dieser Ebene möglich.

Die PKI-Leistungen der GWDG sind integraler Bestandteil des Gesamtkonzepts für eine einheitliche Authentifizierung und Identity Management.

Weitere Informationen zur GWDG-CA bietet die Web-Seite

`https://ca.gwdg.de`

Hier kann auch das Wurzel-Zertifikat des DFN über eine gesicherte Verbindung auf dem eigenen Rechner installiert werden. Die Web-Seite verwendet hierfür ein von allen gängigen Browsern bereits akzeptiertes Zertifikat der Fa. Trust Center.

2.1 Aufgabe

Die PKI der GWDG bietet eine digitale Vertrauensstruktur innerhalb der GWDG und des GÖ*-Projekts. PKI-Leistungen der GWDG-CA werden daher derzeit von verschiedenen Max-Planck-Instituten, dem Rechenzentrum Garching, der Universität Göttingen sowie der Stadt Göttingen genutzt.

Innerhalb der PKI der GWDG werden seit Juni 2004 Zertifikate auf den Ebenen 2 und 3 vergeben. Seit September 2004 werden zusätzlich Zertifikate für die MPG ausgestellt.

2.2 Technische und organisatorische Realisierung

Die oberste Zertifizierungsstelle der GWDG-CA verfügt über keinerlei Netzanbindung. Die GWDG-CA wird ausschließlich bei Bedarf in Betrieb genommen. Zugriff auf die GWDG-CA besitzen vier Administratoren sowie

die Geschäftsleitung der GWDG. Backup und Datenaustausch erfolgt über separate Datenträger.

Die Daten der GWDG-CA werden ausschließlich verschlüsselt gehalten. Die GWDG-CA ist gemäß den aktuellen Vorgaben der innerhalb der GWDG verwendeten IT-Sicherheitsleitlinie sowie den zugehörigen Richtlinien realisiert.

Für die technische Realisierung der GWDG-CA werden die Microsoft-Windows-Zertifizierungsdienste sowie die OpenCA (**`http://www.openca.org`**) unter Linux verwendet. Durch die Verwendung verschiedener PKI-Software können effiziente Lösungen umsetzt werden, die Schwächen einer Einzellösung kompensieren, ohne kostenintensive Speziallösungen zu erfordern. Die Interoperabilität der PKI-Software ist dabei durch den X.509-Standard gewährleistet. Zertifikate werden in gemeinsamen Verzeichnisstrukturen (basierend auf Active Directory und OpenLDAP) sowie im Web veröffentlicht und verwaltet.

Innerhalb der beschriebenen PKI-Struktur existieren ebenfalls keine expliziten Hard- und Software-Vorgaben. Für Sub-CAs können verschiedene Betriebssysteme und Plattformen eingesetzt werden, sofern diese der jeweils aktuellen IT-Sicherheitsleitlinie der GWDG und ihren zugehörigen spezifischen Richtlinien entsprechen.

Zertifikate können direkt über die Web-Seite

`https://ca.gwdg.de`

beantragt werden. Bei der Beantragung des ersten, persönlichen Zertifikats ist eine persönliche Identifizierung des Zertifikatnehmers erforderlich. Folgezertifikate für den gleichen Zertifikatnehmer können direkt digital signiert eingereicht werden.

Zukünftig wird neben der Beantragung über die Web-Seite auch die Erstellung eines Schlüsselpaares und Zertifikats durch die GWDG angeboten. Dem Zertifikatnehmer wird anschließend auf gesichertem Weg sein Schlüsselpaar (PKCS 12) verschlüsselt ausgehändigt. Nach der Übergabe des Schlüsselpaares wird das Schlüsselpaar auf Seiten der GWDG vollständig und unwiederbringlich gelöscht.

2.3 Pilot-Anwendung der neuen Policy des DFN-Vereins

Seit dem 1. Quartal 2005 nimmt die GWDG als einer von vier „Piloten“ am Pilotbetrieb der neuen PKI des DFN-Vereins teil. Der DFN strebt mit seiner neuen PKI den Ausbau und die Integration von weiteren europaweiten und

internationalen Zertifizierungsstrukturen an, um die Akzeptanz des Zertifikats des DFN-Vereins zusätzlich und nachhaltig zu stärken.

Zertifikatnehmer können innerhalb der neuen Struktur optional auf eine Archivierung der verwendeten Schlüssel (gesicherte Hinterlegung des privaten Schlüssels für Verschlüsselungs-Zertifikate) sowie die Identifizierung durch externe Dienstleister (z. B. PostIdent) zurückgreifen. Durch letzteres wird beispielsweise die Reise nach Göttingen für die persönliche Identifizierung der Zertifikatnehmer entfallen.

3. Nutzung der PKI-Leistungen der GWDG

Die GWDG bietet Zertifikate z. B. für Endbenutzer, Endgeräte, aber auch für eigene Zertifizierungs- und Registrierungsstellen interessierter Institute und Anwender an. Institute können so eigene PKI-Leistungen realisieren und dabei Aufgaben für deren Betrieb an die GWDG flexibel auslagern. Beispielsweise kann durch eine im Institut realisierte Registrierungsstelle die Identifizierung der Zertifikatnehmer (Anwender) im Institut erfolgen, während Betrieb und Wartung der zugehörigen Zertifizierungsstelle innerhalb der GWDG erfolgen. Hierbei kann auch der technische Betrieb der Registrierungsstelle selbst von der GWDG angeboten werden.

Zertifikate der GWDG-CA können z. B. für die Signatur und Verschlüsselung von E-Mails, die Authentifizierung von Benutzern, Clients und Servern (auch mittels Smart Card und Token) sowie für Dokument- und Code-Signaturen verwendet werden.

4. Realisierte Anwendungen

Durch den X.509-Standard und seine Unterstützung in den meisten gängigen Anwendungen können Zertifikate plattformübergreifend interoperabel eingesetzt werden. Die GWDG hat hierfür folgende Anwendungen und Betriebssysteme erfolgreich getestet:

Anwendungen	
E-Mail (Signatur, Verschlüsselung)	KMail, Mail.app, Mozilla, Netscape, mutt, Outlook, pine, PC-Pine, Thunderbird, (ListProc) …
Authentifizierung Server	Apache, IIS, Notes, LDAP, RADIUS, …

Authentifizierung Client	802.1X (EAP), Web, LDAP, IPsec, …
Verschlüsselung u. Sig. Daten	Dateiverschlüsselung EFS, PDF-Dokumente, Word bzw. MS-Office-Dokumente, …
SUB-CA	OpenCA, openssl, Windows-CA
SCEP	Cisco im Test
Smart Card / Token	
Verwendung in Applikation	Firefox, Mozilla, Netscape, Windows, openssl
Login	Windows (Aladdin, Microsoft), PAM / Linux
Betriebssysteme	
Unix / Linux	openssl, Firefox, Mozilla, KDE, GNOME
Mac OSX	Schlüsselbund, Safari, Firefox, Camino
Windows	„Windows“, Firefox, Mozilla, Netscape

Tab. 1: Innerhalb der GWDG realisierte Anwendungen mit Zertifikat-Verwendung

5. Zukünftige Planung

Für die sichere Speicherung der Schlüsselpaare und Zertifikate werden zukünftig für Endanwender auch praxisnahe Token- und Smart-Card-Lösungen optional angeboten. Hierbei werden zusätzlich die Datei-Verschlüsselung anhand von Zertifikaten sowie die optionale Archivierung von zugehörigen Schlüsseln geboten.

Zertifikate werden von Benutzern hierbei selbständig beantragt und können über ein Self-Service-Web-Interface verwaltet werden. Der Umgang mit Zertifikaten wird dadurch schrittweise vereinfacht und pragmatische

Lösungsansätze realisiert. Dies ermöglicht die Steigerung der Akzeptanz der PKI sowie die Reduzierung ihrer Komplexität aus Sicht der Anwender bei gleich bleibend hoher Gewährleistung der IT-Sicherheit.

6. Verfügbarkeit der MPG-CA

Seit dem zweiten Quartal 2005 steht neben der GWDG-CA auch die offizielle Zertifizierungsstelle der MPG zur Verfügung. PKI-Leistungen der MPG können daher direkt unter

`https://ca.mpg`

und

`https://user-ca.mpg.de`

bezogen werden.

Weitere Informationen zur GWDG- sowie MPG-CA können unter

`https://ca.gwdg.de`

sowie

`https://ca.mpg.de`

bezogen werden. Weitere Fragen können an

`gwdg-ca@gwdg.de`

gerichtet werden.

VoIP und Videolösungen bei der GWDG sowie im MPI-Umfeld

Andreas Ißleiber

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

Die GWDG betreibt und bewertet die Einsatzfähigkeit von VoIP-Systemen bereits seit dem Jahr 2000. Insbesondere VoIP-Systeme der beiden weit verbreiteten Anbieter Siemens und Cisco stehen bis heute bei der GWDG in einem direkten Vergleich.

Die ersten VoIP-Systeme bei der GWDG waren allerdings Spectralink-Funktelefone, die die Funk-LAN Infrastruktur als Übertragungsmedium nutzen konnten. Diese Telefone stellen in der VoIP-Umgebung einen Sonderfall dar, weil damit ausschließlich via Funk-LAN telefoniert wird und das System nicht für große kabelgebundene Telefoniesysteme geeignet ist. Die sechs Spectralink-VoIP-Telefone sind bis heute bei der GWDG im Einsatz.

1. VoIP von Siemens (HiPath)

In den GWDG-Nachrichten März 2003 wurde bereits über den Stand der damaligen VoIP-Installation berichtet. In der Zwischenzeit hat sich bezüglich der verfügbaren Geräte sowie der Software- und Hardwareversionen Einiges getan. Die GWDG beschaffte im Mai 2003 eine neues Siemens-

VoIP-System, „HiPath 5000“ genannt. Dieses bestand aus einem Gatekeeper, sechs VoIP-Telefonen sowie einem Gateway, um damit den Anschluss an die hauseigene Telefonanlage zu realisieren. Überdies wurden zehn Soft-Clients beschafft, die VoIP-Kommunikation als reine Softwarelösung von einem beliebigem PC oder Laptop aus erlauben.

Anfang Januar 2005 ist die VoIP-Anlage von Siemens (HiPath 3500) durch eine deutlich modernere, die HiPath 4000, ausgetauscht worden.

Der Umfang der an dieser Anlage angebundenen Telefone ist mit sechs Hardware- und 20 Softwaretelefonen zwar relativ klein, genügt aber vollkommen, um eine Aussage über die Integrationsfähigkeit in bestehenden Netzwerkstrukturen treffen zu können.

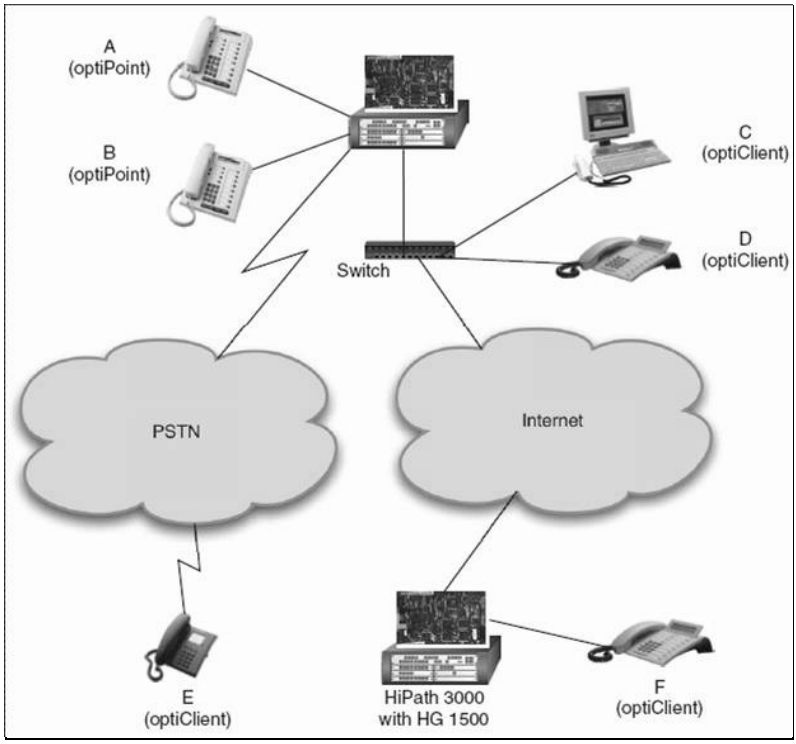


Abb. 1: Siemens-HiPath-System

Wesentliche Eckdaten der Siemens HiPath sind:

- Erlaubt den Anschluss von Systemtelefonen (non-VoIP)
- Gleichzeitige Anbindung von Siemens-VoIP-Telefonen
- 2 x S0 Anschluss (erweiterbar)
- Gatekeeper & Gateway in einem Gerät
- Protokoll:
  - H.323
  - proprietäre CorNet-IP
  - SIP (zukünftig)

Das HiPath-System vereint gleichzeitig den Gatekeeper sowie das Gateway, welches für die externe Kommunikation erforderlich ist. Ein weiterer getrennter Server als Gatekeeper ist nicht mehr erforderlich.

Das Siemens-System beherrscht natürlich H.323 als grundlegendes VoIP-Protokoll. Die Kommunikation mit den systemeigenen Telefonen geschieht allerdings über das proprietäre, Siemens-eigene CorNet-IP-Protokoll. Erweiterungen in Richtung SIP sind bei Siemens bereits abzusehen, so dass mit zukünftigen Geräten auch SIP das zentrale Protokoll sein wird.

2. VoIP von Cisco (CCM)

Parallel zur Siemens-Anlage betreibt die GWDG auch ein VoIP-System von Cisco. Es ist der CCM (Cisco Call Manager 3.5), welcher bei der GWDG bereits im April 2003 installiert wurde. Das Cisco-System besteht im Wesentlichen aus einer reinen Softwarelösung auf einem Server, der unter Windows 2000 (Server) eingerichtet ist. Zukünftige CCM-Versionen (4.1 -> 5.0) werden unter Windows 2003 sowie Linux als Betriebssystem laufen. Cisco plant, den CCM als Gesamtsystem zu vertreiben, so dass eine Trennung zwischen Betriebssystem und der eigentlichen Serversoftware (CCM) weniger deutlich wird und das System dadurch eine Einheit bildet.

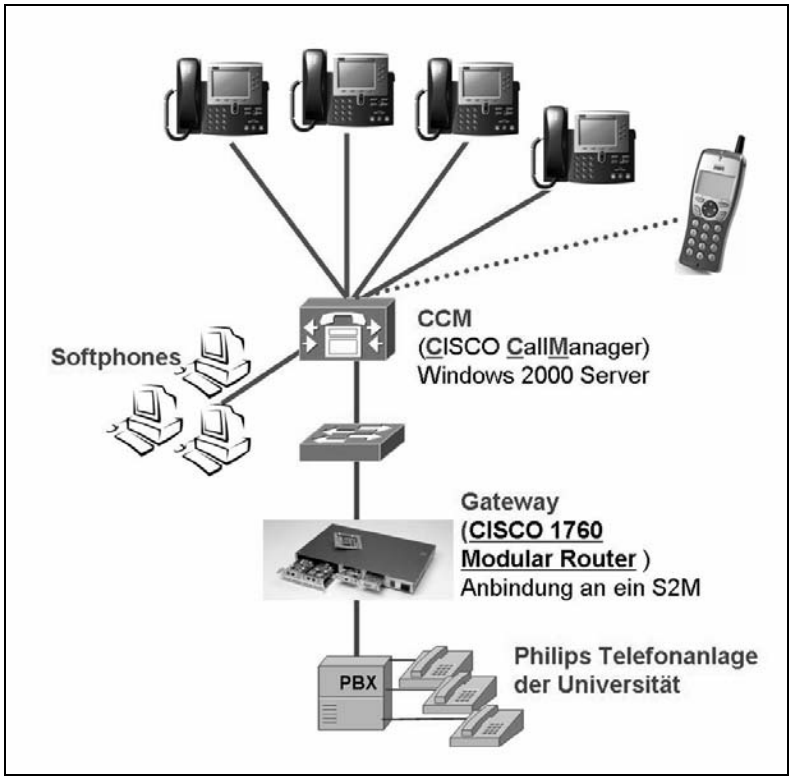


Abb. 2: Derzeitiges VoIP-System von Cisco bei der GWDG

3. Komponenten der Cisco-VoIP-Anlage bei der GWDG

3.1 Gatekeeper und Gateway

Das Cisco-System besteht aus der Software CCM, welche die Vermittlung der Gespräche übernimmt (Callmanager). Ein getrenntes Gateway (Hardware) ist bei der GWDG ein Cisco-Router (Cisco 1760), welcher durch entsprechende Interfaces in der Lage ist, als VoIP-Gateway zu fungieren. Das Gateway besitzt ein S2M-Interface, welches direkt an der Telefonanlage der Universität Göttingen angeschlossen ist. Darüber lassen sich gleichzeitig bis zu 30 Telefonate via ISDN führen. Der Cisco 1760 verbindet die VoIP-Welt mit der klassischen Telefonie.

Mit dem Cisco-System sind bei der GWDG

- sechs Hardwaretelefone,
- ein Funktelefon und
- acht Softphones

installiert.

Für eine Bewertung der Leistungsfähigkeit ist die Anzahl an Telefonen ausreichend. Eine weitere Planung bei der GWDG sieht jedoch eine vollständige Versorgung der GWDG durch Cisco-Telefone vor.

3.2 Merkmale des Cisco-VoIP-Telefon

- LDAP-Integration
- Ticker Internetintegration)
- Content Transformation
- Radio
- SMS-Integration
- Zentrale Excel-Liste
- Raumsteuerung
- Datenbankanbindung ohne CTI
- Kalender
- Kantinenplan
- Zeiterfassung

Abb. 3: Cisco-VoIP-Telefon

Die Cisco-Telefone bieten eine ganze Reihe von Merkmalen, die bei klassischen Telefonen in der Regel nicht existieren. Ein wesentlicher Vorteil ist die Möglichkeit der direkten Anbindung an Datenbanken und Verzeichnisdienste wie LDAP. Zentrale Namens- und Telefonlisten sind so für jedes VoIP-Telefon zugänglich. Vergleicht man jedoch die Leistungsmerkmale, die in bisherigen klassischen Telefonanlagen zur Verfügung stehen, mit den Merkmalen von VoIP-Telefonen, so erkennt man deutlich, dass VoIP-Systeme meistens weniger Merkmale aufweisen als klassische Telefone. Jedoch ist fraglich, ob Benutzer tatsächlich die nicht selten in den dreistelligen Bereich reichende Anzahl an Merkmalen überhaupt nutzt (oder jemals nutzen kann). Durch Einzug der VoIP-Telefone ist die Diskussion über die Merkmale erneut entbrannt und dadurch eine Konsolidierung von Merkmalen bei der Entwicklung neuer Telefone entstanden. Der direkte Vergleich der Merkmale zwischen VoIP und klassischer Telefone geht meist zu ungunsten von VoIP aus. Allerdings ist der Vergleich meist so sinnlos, wie viele der ungenutzten Merkmale selbst. VoIP-Systeme nutzen die Stärken in anderen Bereichen.

3.3 Protokolle

Auch das Cisco-System benutzt ein eigenes, proprietäres Protokoll für die Sprachkommunikation (Skinny, SCCP=*Skinny Client Control Protocol).* Aber auch bei Cisco ist ein deutlicher Wechsel in Richtung SIP als primäres Kommunikationsprotokoll zu erkennen. Neuere System von Cisco werden SIP verstehen.

4. VoIP-System der Wahl

Die GWDG hat sich aufgrund der Ergebnisse der Tests und der Integrationsfähigkeit des VoIP-Systems in die bestehende Netzwerkinfrastruktur für das System von Cisco entschieden. Im Vergleich zur Siemens-Anlage besitzt der CCM eine größere Flexibilität hinsichtlich der Netzwerkeinstellungen. Überdies ist der CCM eine reine VoIP-Anlage und kein Hybridsystem wie bei Siemens, an dem auch Systemtelefone Anschluss finden.

Dennoch bleibt der Blick der GWDG auf die VoIP-Landschaft offen, so dass ggf. auf Neuerungen anderer Hersteller geeignet reagiert werden kann.

5. Ist das Netzwerk „VoIP ready“?

Diese Frage ist ganz entscheidend und unbedingt vor einer größeren Installation von VoIP zu klären. Bei der Frage geht es nicht immer nur um die zur Verfügung stehende Bandbreite. Eine ganze Reihe weiterer Faktoren bestimmt, ob ein lokales Netzwerk „VoIP ready“ ist.

Das folgende Bild zeigt die ganz unterschiedliche Übertragungweise bei Daten und Voice-Daten und deren Merkmale.

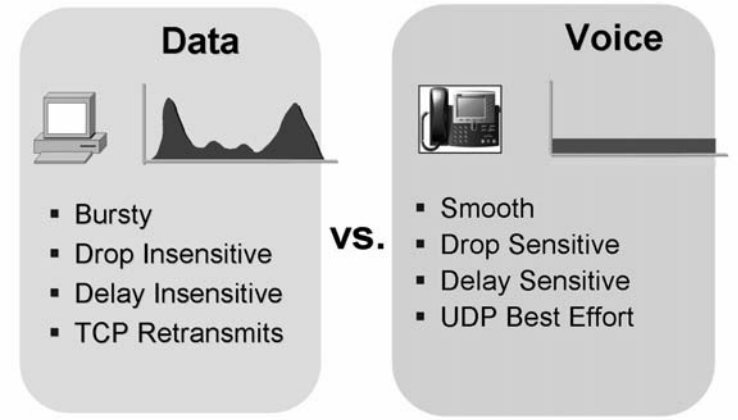


Abb. 4: Vergleich von Daten- und Voice-Daten-Übertragung

Bei VoIP treten seltener Lastsspitzen (Bursts) auf. Die Übertragung und die dazu erforderliche Bandbreite ist relativ konstant und berechenbar. Allerdings ist allein schon aufgrund des verwendeten IP-Protokolls (UDP) eine Überprüfung, ob Pakete auch beim Empfänger angekommen sind, nicht möglich. Dieses reduziert allerdings auch ganz deutlich den „Overhead“ und verringert die Latenz, da keine Antwortpakete geschickt und beim Absender wiederum verarbeitet werden müssen.

6. Bandbreite und Verzögerung (Latenz)

6.1 Bandbreite

Für eine Planung der Integration von IP-Telefonie im lokalen Netz ist vor allem auch das zu erwartende Telefonieaufkommen zu berücksichtigen. Wenn bekannt ist, wie viele Telefonate gleichzeitig in welchem Netzwerkbereich geführt werden sollen, so kann die dafür erforderliche Bandbreite bereits im Voraus berechnet und mit einem entsprechenden „Sicherheitsaufschlag“ in die Planung einbezogen werden.

Die erforderliche Bandbreite ist entscheidend vom verwendeten Codec abhängig, den die VoIP-Anlage verwendet. Oft werden verschiedene Codes von einer Anlage unterstützt und auch benutzt. Die folgende Tabelle zeigt den Codec und die dazugehörige Bandbreite an. Hieraus und aus der Anzahl der gleichzeitigen Telefonate ermittelt man einen ungefähren Bandbreitenbedarf.

Die Tabelle „Call Control“ zeigt den Bandbreitenbedarf für die Kontrollpakete während des Telefonates und des Verbindungsaufbaus und verläuft etwa linear mit der Anzahl der Telefone.

Bandbreiten

CODEC	Packets per sec.	Bandwidth
G.711	50	80 kbps
G.729	50	24 kbps
G.711 with cRTP	50	65.4 kbps
G.729 with cRTP	50	9.6 kbps

Bandbreite “Call Control”

Centralized Call Processing	
# of IP Phones, Gateways	Bandwidth
1 to 30	8 kbps
50	11 kbps
100	23 kbps
150	34 kbps

Abb. 5: Bandbreitentabellen

6.2 Latenz

Natürlich muß für die IP-Telefonie ausreichend Bandbreite zur Verfügung stehen, jedoch sind Verzögerungen innerhalb des Netzwerkes ein weiteres, sehr entscheidendes Merkmal.

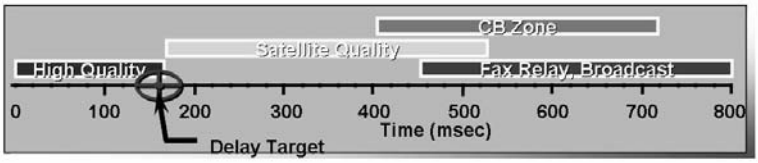


Abb. 6: Verzögerungen und Qualitätsdefinition

Zu große Verzögerungen im Netzwerk führen sehr schnell zu inakzeptabler Verständigungsqualität. Bei VoIP werden Verzögerungen von bis zu 150 ms als noch gerade akzeptabel angesehen.

Jitter, also Schwankungen in den Verzögerungszeiten, sind auch ein Merkmal eines Netzwerkes. Ein für VoIP akzeptabler Wert von 30 ms hat sich als obere Grenze bei der Sprachkommunikation herausgestellt.

6.3 Paketverluste

Natürlich sind auch die Paketverluste in einem Netzwerk eine Eigenschaft, die sich auf die Sprachkommunikation außerordentlich negativ auswirken kann.

7. Das Netzwerk für den Einsatz von VoIP vorbereiten

Folgende Punkte sind für ein VoIP-Netzwerk entscheidend.:

- QoS bei den eingesetzten Netzkomponenten (Router und! Switches)
- Alternativ große Bandbreiten zur Verfügung stellen (kein QoS)
- Latenzen reduzieren
- Abtrennen des VoIP-Netzes vom Datennetz entweder durch VLAN oder physikalisch
- Absichern der VoIP-Umgebung durch Firewalls
- Bei Verbindungen zwischen VoIP-Partnern über die lokalen Netzwerkgrenzen hinweg verschlüsselte Verbindung ermöglichen (ggf. VPN)

7.1 „VoIP ready“ testen … aber wie?

Eine erste Einschätzung über die VoIP-Fähigkeit des lokalen Netzwerkes bekommt man über folgende Verfahren:

- Bandbreiten- und Latenztest zu unterschiedlichen Uhrzeiten messen durch Sniffer oder durch Auslesen der Statistiken (SNMP) der Netzkomponenten.
- Feststellen der Latenzen im einfachsten Fall durch PING; manche Sniffer können eine Aussage über Latenzen während des „Mitschnüffelns“ einer VoIP-Verbindung geben.
- Erzeugung von Lasten: Für die Lastsimulation eignen sich „Trafficgeneratoren“, die gelegentlich auch in Sniffer (Software) mit eingebaut sind. Alle Latenz- und Jittermessungen müssten dann unter Last erneut durchgeführt werden. Hierfür gibt es auch spezielle Software wie z. B. „NetAlly“.

8. Sicherheit

8.1 Sicherheit bei klassischen Telefonanlagen

Auch bei klassischen Telefonanlagen ist Sicherheit ein Thema, welches oft zuwenig betrachtet wird. Auch klassische Telefonanlagen sind attackierbar. Die folgende Auflistung zeigt einen Ausschnitt von Angriffspunkten bei klassischen TK-Anlagen.

- Mithören von Raumgesprächen
  - Analog (Frequency Flooding)
  - Digital
- Mithören von Telefongesprächen
  - Konferenzschaltung
  - Zeugenschaltung (Silent Monitoring)
  - Aufschaltung (Seamless Intrusion)
- Physikalischer Zugriff
  - z. B. Geräte manipulieren

- Service Terminal
    - z. B. über Internet Liste von Hersteller-Standardpasswörtern; nach Login freie Konfiguration
- Fernwartungszugang
    - oft ab Werk aktiviert und nur Einfachsicherung,
    - Standard-Passwortliste im Internet
- Gebührenbetrug
- Verlust der Verfügbarkeit
- z. B. DoS via Massenanruf, Aktivierung Prüfschleife etc.
- Zugriff auf Signalisierungs- und Accounting-Daten (Communication Profiling)

8.2 Sicherheit bei VoIP

Einige der Attacken sind auch bei VoIP-Systemen durchführbar, jedoch unterscheiden diese sich in den zu ergreifenden Maßnahmen zur Erhöhung der Sicherheit.

Im Folgenden sind wesentliche Punkte für die Erhöhung der Sicherheit aufgeführt:

- Physikalische Sicherheit
  (Zugriff auf die Telefone)
- Sicherung der Netzwerkkomponenten
  Physischer und netztechnischer Zugriff auf Switches und Router
- Netzwerk-Design (VLANs etc.)
  Abtrennung der Netzbereiche, Layer 3 und Layer 2
- Zugriffsschutz (Firewall, NAT)
- Verschlüsselung der Übertragung
- Absichern zentraler (ggf. personenbezogener) Daten

9. Video over IP

Das Cisco-VoIP-System, bietet eine einfache Einbindung von Videoarbeitsplätzen in eine bestehende IP-Telefonieumgebung des CCM. Hiermit lassen sich auf dem Rechner des Anwenders leicht Videoverbindungen (Punkt zu

Punkt) zu Kollegen aufbauen, ohne eine größere Installation in Gang zu setzen.

In Kürze wird die GWDG die Erweiterungen der Cisco-VoIP-Lösung im Rahmen des neueren CallManagers (CCM 4.1) installieren und neben der Telefonie auch die Videoübertragung testen. Kombinationslösungen wie Voice und Video in Verbindung mit Applicationsharing sind unverkennbarer Trend im Bereich des VoIP. Sie sollen neben der klassischen Telefonie einen spürbaren Mehrwert für das wissenschaftliche Umfeld bringen. Über das Ergebnis dieses Tests wird zu gegebener Zeit in den GWDG-Nachrichten berichtet.

In der Reihe GWDG-Berichte sind zuletzt erschienen:

Nähere Informationen finden Sie im Internet unter
`http://www.gwdg.de/forschung/publikationen/gwdg-berichte`

Nr. 40 *Plesser, Theo und Peter Wittenburg* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 1994
1995

Nr. 41 *Brinkmeier, Fritz* (Hrsg.):
Rechner, Netze, Spezialisten. Vom Maschinenzentrum zum Kompetenzzentrum - Vorträge des Kolloquiums zum 25jährigen Bestehen der GWDG
1996

Nr. 42 *Plesser, Theo und Peter Wittenburg* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 1995
1996

Nr. 43 *Wall, Dieter* (Hrsg.):
Kostenrechnung im wissenschaftlichen Rechenzentrum - Das Göttinger Modell
1996

Nr. 44 *Plesser, Theo und Peter Wittenburg* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 1996
1997

Nr. 45 *Koke, Hartmut und Engelbert Ziegler* (Hrsg.):
13. DV-Treffen der Max-Planck-Institute - 21.-22. November 1996 in Göttingen
1997

Nr. 46 **Jahresberichte 1994 bis 1996**
1997

Nr. 47 *Heuer, Konrad, Eberhard Mönkeberg und Ulrich Schwardmann*:
Server-Betrieb mit Standard-PC-Hardware unter freien UNIX-Betriebssystemen
1998

Nr. 48 *Haan, Oswald* (Hrsg.):
Göttinger Informatik Kolloquium - Vorträge aus den Jahren 1996/97
1998

Nr. 49 *Koke, Hartmut und Engelbert Ziegler* (Hrsg.):
IT-Infrastruktur im wissenschaftlichen Umfeld - 14. DV-Treffen der Max-Planck-Institute, 20. - 21. November 1997 in Göttingen
1998

Nr. 50 *Gerling, Rainer W.* (Hrsg.):
Datenschutz und neue Medien - Datenschutzschulung am 25./26. Mai 1998
1998

Nr. 51 *Plesser, Theo und Peter Wittenburg* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 1997
1998

Nr. 52 *Heinzel, Stefan und Theo Plesser* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 1998
1999

Nr. 53 *Kaspar, Friedbert und Hans-Ulrich Zimmermann* (Hrsg.):
Internet- und Intranet-Technologien in der wissenschaftlichen Datenverarbeitung - 15. DV-Treffen der Max-Planck-Institute, 18. - 20. November 1998 in Göttingen
1999

Nr. 54 *Plesser, Theo und Helmut Hayd* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 1999
2000

Nr. 55 *Kaspar, Friedbert und Hans-Ulrich Zimmermann* (Hrsg.):
Neue Technologien zur Nutzung von Netzdiensten - 16. DV-Treffen der Max-Planck-Institute, 17. - 19. November 1999 in Göttingen
2000

Nr. 56 *Plesser, Theo und Helmut Hayd* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 2000
2001

Nr. 57 *Hayd, Helmut und Rainer Kleinrensing* (Hrsg.):
17. und 18. DV-Treffen der Max-Planck-Institute
22. - 24. November 2000 in Göttingen
21. - 23. November 2001 in Göttingen
2002

Nr. 58 *Plesser, Theo und Volker Macho* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 2001
2003

Nr. 59 *Suchodoletz, Dirk von*:
Effizienter Betrieb großer Rechnerpools - Implementierung am Beispiel des Studierendennetzes an der Universität Göttingen
2003

Nr. 60 *Haan, Oswald* (Hrsg.):
Erfahrungen mit den IBM-Parallelrechnersystemen RS/6000 SP und pSeries690
2003

Nr. 61 *Rieger, Sebastian:*
Streaming-Media und Multicasting in drahtlosen Netzwerken - Untersuchung von Realisierungs- und Anwendungsmöglichkeiten
2003

Nr. 62 *Kremer, Kurt und Volker Macho* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 2002
2003

Nr. 63 *Kremer, Kurt und Volker Macho* (Hrsg.):
Forschung und wissenschaftliches Rechnen - Beiträge zum Heinz-Billing-Preis 2003
2004

Nr. 64 *Koke, Hartmut* (Hrsg.):
GÖ* – Integriertes Informationsmanagement im heterogenen eScience-Umfeld: GÖ*-Vorantrag für die DFG-Förderinitiative „Leistungszentren für Forschungsinformation“
2004

Nr. 65 *Koke, Hartmut* (Hrsg.):
GÖ* – Integriertes Informationsmanagement im heterogenen eScience-Umfeld: GÖ*-Hauptantrag für die DFG-Förderinitiative „Leistungszentren für Forschungsinformation“
2004

Nr. 66 *Bussmann, Dietmar und Andreas Oberreuter* (Hrsg.):
19. und 20. DV-Treffen der Max-Planck-Institute
20. - 22. November 2002 in Göttingen
19. - 21. November 2003 in Göttingen
2004

Nr. 67 *Gartmann, Christoph und Jochen Jähnke* (Hrsg.):
21. DV-Treffen der Max-Planck-Institute
17. - 19. November 2004 in Göttingen
2005