



Konzeption einer Cache-Server-Infrastruktur auf dem Wissenschaftsnetz – Teil II

Abschlussbericht

Februar 2001, C. Grimm, J.-S. Vöckler

□ □ □ ■

R V S ■

□ ■ ■ ■

■ ● ■ ■

Lehrgebiet Rechnernetze und Verteilte Systeme (RVS)

Prof. Dr.-Ing. Helmut Pralle

Universität Hannover

Inhalt

| | | |
|---|---|----|
| 1 | Bezeichnung des Projektes | 2 |
| 2 | Vorbemerkung | 2 |
| 3 | Durchführung des Projektes | 3 |
| | 3.1 Personal | 3 |
| | 3.2 Investitionen | 3 |
| | 3.3 Vorträge, Veröffentlichungen | 4 |
| | 3.4 Weitere Aktivitäten | 4 |
| 4 | Vergleich neuer und alter DFN-Cache-Verbund | 5 |
| | 4.1 Hardware | 5 |
| | 4.2 System- und Betriebsparameter | 5 |
| | 4.3 Inter-Cache-Kommunikation | 6 |
| | 4.4 Partitionierung | 8 |
| | 4.5 Konzentration auf externe Objekte | 9 |
| | 4.6 Policy | 10 |
| | 4.7 Hilfen für Nutzer | 11 |
| | 4.8 Verkehrsbilanzen | 11 |
| 5 | Bewertung DFN-Cache-Verbund | 16 |
| | 5.1 Nutzen | 16 |
| | 5.2 Nutzung | 16 |
| | 5.3 Kosten für Hardware | 16 |
| | 5.4 Inter-Cache-Kommunikation | 17 |
| | 5.5 Vergleich mit europäischen Forschungsnetzen | 17 |
| | 5.6 Zusammenfassung | 17 |
| 6 | Ausblick – Content Distribution/Delivery Networks | 19 |
| | 6.1 Die Beteiligten | 19 |
| | 6.2 Der Ablauf | 20 |
| | 6.3 Mögliche Vor- und Nachteile | 23 |
| | 6.4 CDNs als Kontrollinstrumente | 24 |
| | 6.5 Der Content-Delivery-Markt | 25 |
| | 6.6 Fazit | 26 |
| 7 | Entwicklung des World Wide Web | 27 |

1 Bezeichnung des Projektes

| | |
|-------------------------|---|
| Projekttitel | Konzeption einer Cache-Server-Infrastruktur auf dem Wissenschaftsnetz Teil II |
| Kurztitel | DFN-Cache-Service Teil II |
| Ansprechpartner | Christian Grimm Jens-Sönke Vöckler Lehrgebiet Rechnernetze und Verteilte Systeme (RVS) an der Universität Hannover (Leiter: Prof. Dr.-Ing. Helmut Pralle) |
| Ausführende Einrichtung | RVS, Universität Hannover |
| Projektlaufzeit | 1. Mai 1998 bis Oktober 1999, 18 Monate |
| Verlängerung | 1. November 1999 bis 31. August 2000, 10 Monate |

2 Vorbemerkung

Das Projekt *Konzeption einer Cache-Server-Infrastruktur auf dem Wissenschaftsnetz Teil II* endete am 31. August 2000; der Betrieb des DFN-Cache-Verbundes wurde auf Empfehlung der Projektmitarbeiter zum 31. Dezember 2000 eingestellt. Mit dieser Empfehlung wurde eine der wesentlichen Aufgaben in dem Projekt, die Bewertung eines übergeordneten Cache-Service im Wissenschaftsnetz, beantwortet.

Der vorliegende Abschlussbericht dient weniger der Darstellung der im Projektverlauf durchgeführten Arbeiten, hier kann auf die Zwischenberichte verwiesen werden. Vielmehr sollen im folgenden die Erfahrungen aus annähernd 5 Jahren Planung, Aufbau und Betrieb eines zentralen Dienstes im Wissenschaftsnetz zusammengefasst werden. Die Schwerpunkte dieser Arbeiten werden in folgenden Abschnitten erläutert:

1. Planung und Aufbau von altem und neuem DFN-Cache-Verbund
2. Zusammenstellung der über den Projektzeitraum gewonnen Betriebserfahrungen
3. Begründung für die Empfehlung zur Beendigung des DFN-Cache-Service
4. Neue Dienste zur Verteilung von WWW-Inhalten
5. Aussagen über die Entwicklung und das Wachstum des World Wide Web

In den ersten beiden Abschnitten werden die Eigenschaften von altem und neuem DFN-Cache-Verbund gegenüber gestellt. Hierbei werden besonders die Erkenntnisse hervorgehoben, die in dem alten DFN-Cache-Verbund gesammelt wurden und wesentlich die Gestaltung des neuen DFN-Cache-Verbundes beeinflusst haben.

In der folgenden Bewertung werden die Gründe für die Empfehlung zur Abschaltung des DFN-Cache-Verbundes erläutert. Weiterhin werden Erfahrungen aus dem Projekt zusammengefasst, die für vergleichbare zukünftige Projekte oder Vorhaben des DFN – insbesondere im Bereich zentraler Dienste – von Nutzen sein können.

Diese möglichen Dienste werden in einem eigenen Abschnitt näher betrachtet. Mit den sog. *Content Distribution/Delivery Networks* stehen Verfahren zur Verfügung, die alternativ oder ergänzend zum Caching eingesetzt werden können, um die Verteilung von WWW-Inhalten im Wissenschaftsnetz zu verbessern.

Abschließend wird über ein bisher weniger beachtetes Ergebnis aus dem Projekt berichtet. Anhand von Cache-Statistiken konnten Langzeitanalysen über das Verkehrsaufkommen im World Wide Web erstellt werden. Diese Analysen bieten einen hervorragenden Überblick darüber, wie sich das World Wide Web in den vergangenen Jahren entwickelt und verändert hat.

3 Durchführung des Projektes

3.1 Personal

An dem Projekt waren folgende Personen beteiligt:

- Projektleitung
 - Prof. Dr.-Ing. Helmut Pralle
- Wissenschaftliche Mitarbeiter
 - Dipl.-Ing. Christian Grimm
 - Dipl.-Ing. Jens-Sönke Vöckler

3.2 Investitionen

Zur Durchführung des Projektes wurden folgende Investitionen getätigt:

Hardware aus Projektmitteln

- 2 Cache-Server für den Produktionseinsatz, jeweils ausgestattet mit:
 - Sun Server Enterprise 450
 - vier 400 MHz Prozessoren
 - 2 GByte Hauptspeicher
 - Plattenkapazität 144 GByte

Beide Server gingen nach Projektende an die DFN-Geschäftsstelle zurück.

Hardware aus Projektmitteln Teil I

- Zwei weitere Server, die bereits im Rahmen von Projekt Teil I beschafft wurden, konnten auch in Teil II für die Sammlung und Auswertung von Statistiken eingesetzt werden. Beide Server sind jeweils wie folgt ausgestattet:
 - Sun Ultra2 Enterprise
 - zwei 167 MHz Prozessoren
 - 256 MByte Hauptspeicher
 - Plattenkapazität 20 GByte

3.3 Vorträge, Veröffentlichungen

Ergebnisse aus dem Projekt wurden wie folgt präsentiert:

- Veröffentlichungen
 - Grimm, C.: *Was Administratoren über WWW-Caches wissen sollten*, Deutscher Internet Kongress, Mai 1998, Frankfurt/Main
 - Grimm, C., Pralle, H., Vöckler, J.: *Load and Traffic Balancing in Large Scale Cache Meshes*, 3rd Workshop on Caching, Juni 1998, Manchester
 - Grimm, C., Pralle, H., Vöckler, J.: *Request Routing in Cache Meshes*, 3rd Workshop on Caching, Juni 1998, Manchester
 - Grimm, C., Pralle, H., Vöckler, J.: *Load and Traffic Balancing in Large Scale Cache Meshes*, TERENA Networking Conference, Oktober 1998, Dresden
- Vorträge
 - Grimm, C.: *Cache-Service im WiN - es geht weiter !*, 29. DFN-Betriebstagung, September 1998, Berlin
 - Grimm, C.: *Verkehrsanalysen im Projekt DFN-Cache-Service*, DFN-Workshop Leistungsanalysen, Februar 1999, Berlin
 - Vöckler, J.: *Log File Analysis*, First DESIRE II Web Cache Managers Workshop, Februar 1999, Utrecht
 - Vöckler, J.: *Log Analysis and Cache Statistics*, Second DESIRE II Web Cache Managers Workshop, März 2000, Budapest
 - Grimm, C.: *Abschlussbericht Projekt DFN-Cache-Service*, 33. DFN-Betriebstagung, Oktober 2000, Berlin
- Fortlaufende Informationen für die Nutzer auf dem projekteigenen WWW-Server unter: <http://www.cache.dfn.de/>

3.4 Weitere Aktivitäten

- Mitglied im Programmkomitee des *3rd Workshop on Caching* in San Diego, USA, April 1999 (C. Grimm)
- Durchführung TERENA Projekt *Extended Cache Statistics*, Januar - Juni 2000, (J. Vöckler)
- Ständige aktive Teilnahme an der TERENA Caching Task Force (TF-Cache):
 - Oktober 1998, Dresden (C. Grimm, J. Vöckler)
 - Juni 1999, Lund (J. Vöckler)
 - Mai 2000, Lissabon (J. Vöckler)

4 Vergleich alter und neuer DFN-Cache-Verbund

Dieses Kapitel enthält eine Zusammenfassung der wichtigsten Merkmale und Betriebserfahrungen von altem und neuem DFN-Cache-Verbund. Der Schwerpunkt der Darstellungen liegt auf einer vergleichenden Betrachtung beider Cache-Verbünde. Detaillierte Informationen, wie z. B. Angaben zur Optimierung der Betriebssystemparameter, wurden bereits in den vergangenen Zwischenberichten aufgeführt.

4.1 Hardware

Tabelle 1 enthält eine Gegenüberstellung der Hardware von alten und neuen DFN-Caches.

| | alter Verbund | neuer Verbund |
|----------------|----------------------------|----------------------------|
| Anzahl Server | 10 | 2 |
| Bezeichnung | Sun Ultra Enterprise 2 | Sun Ultra Enterprise 450 |
| Prozessor | 2 x 167 MHz | 4 x 400 MHz |
| Hauptspeicher | 256 MByte | 2 GByte |
| Plattenplatz | 15 GByte | 180 GByte |
| Netzinterface | 1 x 100 Mbps Fast Ethernet | 2 x 100 Mbps Fast Ethernet |
| Betriebssystem | Solaris 2.5.1 / 2.7 | Solaris 2.7 |

Tabelle 1: Hardware DFN-Caches

Die Hardware im alten DFN-Cache-Verbund erschien zunächst ausreichend dimensioniert, um einen leistungsfähigen Cache-Verbund zu betreiben. Allerdings konnten die 10 Server nicht ausschließlich für Caching zur Verfügung gestellt werden. Vielmehr wurden sie für weitere Dienste mit hohem Bedarf an Ressourcen, wie die Sammlung von IP-Statistiken und Verteilung von News, eingesetzt. Die Hardware der alten Caches stellte somit keine optimale Plattform für einen Cache-Verbund dar. Im Gegensatz hierzu konnten die neuen DFN-Caches ausschließlich für Caching genutzt werden. Während die alten DFN-Caches dauerhaft an ihrer Lastgrenze betrieben wurden, waren auf den neuen DFN-Caches auch zu Spitzenzeiten Leistungsreserven vorhanden.

4.2 System- und Betriebsparameter

Ein weiteres Problem zeigte sich im alten DFN-Cache-Verbund in der schlechten Netzanbindung, die aufgrund älterer Interfaces in den zugehörigen Routern nur mit 10 Mbps Ethernet anstelle der serverseitig möglichen 100 Mbps Fast Ethernet ausgeführt werden konnte.

In Tabelle 2 wird neben der Netzanbindung eine Reihe weiterer Betriebsparameter beider Cache-Verbünde gegenübergestellt. Die z. T. deutlichen Unterschiede ergeben sich zum einen aus der besseren Hardware, aber auch aus der unterschiedlichen Nutzungsform als gemeinsam genutzte bzw. im neuen DFN-Cache-Verbund dedizierte Server:

| | alter Verbund | neuer Verbund |
|---|----------------------|------------------------|
| max. CPUs pro Cache-Prozess | 1 | 2 |
| Netzanbindung an B-WiN (über Router) | 10 Mbps Ethernet | 100 Mbps Fast Ethernet |
| max. Plattenplatz pro Cache-Prozess | 4 GByte | 56 GByte |
| Volumen insgesamt auf allen Caches | 40 GByte | 224 GByte |
| Objekte insgesamt auf allen Caches | bis 2 Mio. | über 10 Mio. |
| max. Volumen je Objekt | 32 MByte | 96 MByte |
| max. Verweildauer der Objekte im Cache | unter 20 Stunden | 5 Tage |

Tabelle 2: Vergleich von altem und neuem DFN-Cache-Verbund

Aufgrund des wesentlich höheren Plattenvolumens wurde im neuen DFN-Cache-Verbund die maximale Größe der gespeicherten Objekte auf 96 MByte gesetzt. Somit konnten selbst größere Downloads, wie z. B. neue Versionen eines WWW-Browsers, auf den neuen DFN-Caches gespeichert und über einen längeren Zeitraum zur Verfügung gestellt werden.

Ein wichtiger Hinweis auf die bessere Ausstattung der neuen DFN-Caches war in der maximalen Verweildauer der Objekte zu sehen. Dieser Wert ist eine wesentliche Hilfe zur korrekten Dimensionierung des zur Verfügung gestellten Plattenplatzes. Ziel sollte eine Verweildauer von mehreren Tagen sein. Eine Verweildauer von unter einem Tag wie im alten DFN-Cache-Verbund deutete auf eine völlig unterdimensionierte Cache-Partition hin. Der Wert von ca. 5 Tagen für den neuen DFN-Cache-Verbund war akzeptabel.

4.3 Inter-Cache-Kommunikation

Wesentlicher Bestandteil eines Rechnerverbundes sind die Protokolle, über die die beteiligten Systeme kommunizieren. Ein Cache-Verbund muss über Protokolle verfügen, mit denen Kontrollinformationen über die Verfügbarkeit von Objekten und die Objekte selbst ausgetauscht werden. Für den Austausch von Kontrollinformationen in einem Cache-Verbund existieren mit dem *Inter Cache Protocol (ICP)* und *Cache-Digests* zwei Ansätze, die im DFN-Cache-Verbund eingesetzt und verglichen werden konnten. Tabelle 3 zeigt eine Übersicht der wichtigsten Merkmale beider Protokolle.

| | ICP | Cache-Digest |
|--------------------|---|---|
| Transportprotokoll | UDP | HTTP/TCP |
| Request enthält | Kontrollinformation, URL eines Objektes | HTTP-Request mit URL für den Cache-Digest |

Tabelle 3: Vergleich ICP und Cache-Digest

| | | |
|-----------------------------|--|---|
| Reply enthält | Kontrollinformation, URL eines Objektes | Verzeichnis aller momentan im Cache gespeicherten Objekte |
| Volumen Request | abhängig von Länge der URLs, typischer Mittelwert liegt bei ca. 75 Byte | typisch ca. 400 Byte |
| Volumen Reply | | abhängig von Anzahl der gespeicherten Objekte (bzw. Plattenplatz) des befragten Caches, codiert mit 5 Bit/Objekt |
| Klient sendet Request | unmittelbar nach Erhalt eines HTTP-Request, der nicht beantwortet werden kann (MISS) | zeitgesteuert, in der Regel stündlich |
| Häufigkeit Requests/Replies | proportional MISSes auf lokalen Caches | proportional Anzahl lokaler Caches |
| Verfügbarkeit | frei | <ul style="list-style-type: none"> – bisher nur in Squid – Patent pending by University of Wisconsin – keine kommerzielle Implementierung zu erwarten. |

Tabelle 3: Vergleich ICP und Cache-Digest

Der Tabelle können bereits die wichtigsten Kriterien für den Einsatz beider Protokolle entnommen werden. Hierbei wird deutlich, dass kein Protokoll für einen größeren Cache-Verbund mit hohem Verkehrsaufkommen geeignet ist. ICP skaliert schlecht, da für jeden MISS auf einem lokalen Cache ICP-Requests an alle konfigurierten DFN-Caches versendet werden. Somit hängt der auftretende ICP-Verkehr sowohl von der Anzahl der lokalen Caches als auch von deren Hitrate ab. Mit Abbildung 1 wird verdeutlicht, wie sehr das ICP-Verkehrsaufkommen auf den DFN-Caches von dem Verhalten der lokalen Caches bestimmt wurde.

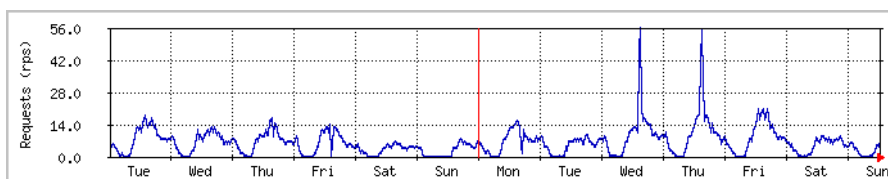


Bild 1: ICP-Requests im neuen DFN-Cache-Verbund

Die Lastspitzen am Mittwoch und Donnerstag der zweiten Woche wurden lediglich durch einen lokalen Cache verursacht, auf dem kurzzeitig Experimente mit transparentem Caching durchgeführt wurden. Eine Akkumulation oder Überlagerung ähnlicher Spitzen bei bis zu 80 lokalen Caches hätte umgehend zu einem Zusammenbruch des DFN-Caches-Service geführt.

Im Vergleich zu ICP zeigen Cache-Digests eine sehr gute Skalierbarkeit. Die Anzahl der auftretenden Requests ist lediglich von der Anzahl lokaler Caches abhängig. Die Höhe des übertragenen Volumens wird wiederum von den DFN-Caches bestimmt und richtet sich nach der

Anzahl der dort gespeicherten Objekte. Bei ca. 3 Millionen gespeicherten Objekten ergibt sich mit 5 Bit pro Objekt eine typische Größe der Cache-Digests von ca. 2 MByte. Entsprechend beträgt das tägliche Volumen zur Übertragung der Cache-Digests bei 4 DFN-Caches und 80 lokalen Caches $24 \cdot 4 \cdot 80 \cdot 2 \text{ MByte} = 15 \text{ GByte}$. Verglichen mit dem Volumen an übertragenen WWW-Objekten von annähernd 200 GByte (s. Verkehrsstatistiken Abschnitt 4.8) ist dieser Betrag durchaus signifikant. Der Vorteil der Cache-Digests gegenüber ICP liegt somit nicht in der Einsparung von übertragenem Datenvolumen zur Inter-Cache-Kommunikation, sondern

- in der geringen Anzahl Requests (genau 24 Requests pro lokalem Cache pro Tag) sowie
- in der Unabhängigkeit von dem Verkehrsaufkommen auf den lokalen Caches.

Die Planung eines Cache-Verbundes auf Basis von Cache-Digests ist somit wesentlich besser kalkulierbar als mit ICP.

Der entscheidende Nachteil von Cache-Digests ist die fehlende Verfügbarkeit in kommerziellen oder Hardware-basierten Implementierungen. Eine Inter-Cache-Kommunikation ausschließlich über Cache-Digests lässt sich nur in einer homogenen Squid-Umgebung realisieren.

4.4 Partitionierung

Als einzige durchführbare Möglichkeit einer Last- und Verkehrsverteilung der Anfragen lokaler Caches auf die DFN-Caches wurde eine sog. Domain-basierte Partitionierung vorgenommen. Im alten DFN-Cache-Verbund wurden hierfür zunächst zwei Caches der Domain *com* sowie ein Cache der Domain *de* zugeordnet, die restlichen Domains wurden auf die übrigen 7 DFN-Caches verteilt. Es zeigte sich bald, dass diese Verteilung nicht dem tatsächlichen Verkehrsaufkommen entsprach, z. B. entfielen annähernd 50 % aller ankommenden Requests auf die Domain *com*. Dieses Missverhältnis wurde bereinigt, indem zwei Gruppen zu je 5 DFN-Caches definiert wurden. Eine Gruppe wurde lediglich nach Objekten aus der Domain *com*, die andere Gruppe nach den übrigen Objekten (*!com* bzw. *nicht com*) befragt. Diese Verteilung wurde bis zum Betriebsende auf den alten DFN-Caches beibehalten.

Auf den neuen DFN-Caches wurde von Beginn an eine leicht modifizierte Variante betrieben. Hier wurden die vier Cache-Prozesse in zwei Gruppen zu je zwei Caches mit den virtuellen Domains *abroad* und *!abroad* unterteilt. Die virtuelle Domain *abroad* enthielt die Toplevel-Domains *com*, *edu*, *gov*, *org*, *int*, *mil*, *us* und *ca*. Der Name *abroad* orientierte sich an der Tatsache, dass Hosts innerhalb dieser Domains mit hoher Wahrscheinlichkeit jenseits einer Atlantikstrecke des Wissenschaftsnetzes zu finden sind. Die virtuelle Domain *!abroad* enthielt entsprechend alle Toplevel-Domains, die nicht von *abroad* abgedeckt wurden.

Tabelle 4 zeigt einige Betriebsparameter der neuen DFN-Caches. Es fällt zunächst auf, dass wesentlich mehr Objekte aus der Domain *!abroad* abgerufen werden. Im Gegensatz dazu ist das Volumen der aus der *abroad*-Domain abgerufenen Objekte deutlich größer. Dieser Sachverhalt erklärt sich durch die unterschiedliche mittlere Größe der Objekte beider Domains. Die mittlere Größe eines Objekts der Domain *!abroad* beträgt mit 19 KByte nur ca. 63 % der mittleren Größe eines Objektes der Domain *abroad*.

| | !abroad | abroad |
|---|----------------|---------------|
| auf Festplatten gespeicherte Objekte | 6,6 Millionen | 4,0 Millionen |
| an lokale Caches ausgelieferte Objekte | 8,2 Millionen | 4,6 Millionen |
| an lokale Caches ausgeliefertes Volumen | 82 GByte | 104 GByte |
| max. Request-Hitrate | 33 % | 33 % |
| max. Volumen-Hitrate | 25 % | 22 % |
| max. Verweildauer eines Objektes | 6 Tage | 4 Tage |
| Mittlere Größe / Objekt | 19 KByte | 30 KByte |

Tabelle 4: Betriebsdaten neuer DFN-Cache-Verbund (an Wochentagen)

An den Daten aus Tabelle 4 zeigt sich, dass bei einer Domain-basierten Partitionierung eine Lastverteilung (übertragene Objekte bzw. Requests) nicht mit einer Verkehrsverteilung (übertragenes Volumen) einhergehen muss. Daraus folgt, dass eine Lastverteilung in einem Cache-Verbund nach unterschiedlichen Gesichtspunkten durchgeführt werden kann. Die Einteilung nach Objekten verschiedener Domains ist dabei offenbar keine optimale Wahl. Erschwerend kommt hinzu, dass eine optimale Nutzung der Domain-basierten Partitionierung nur durch korrekt konfigurierte lokale Caches erreicht werden kann.

Bessere Methoden zur Last- und Verkehrsverteilung bieten aktive Netzkomponenten wie die sog. Layer 4-7 Switches. Diese Geräte verteilen dynamisch die Verkehrsströme in Abhängigkeit von der jeweiligen Lastsituation auf mehrere Cache-Server. Bisher existieren Lösungen jedoch nur für die Verteilung auf Cache-Server, die innerhalb eines lokalen Netzes installiert sind (sog. Cache-Cluster oder Cache-Farms). Ob und unter welchen Umständen Layer 4-7 Switches auch für die Verkehrsverteilung im Weitverkehrsbereich geeignet sind, ist bisher nicht eindeutig geklärt.

4.5 Konzentration auf externe Objekte

Tabelle 1 ist zu entnehmen, dass mit den neuen DFN-Caches wesentlich mehr Plattenplatz zur Verfügung gestellt wurde. Mit einer zusätzlichen Maßnahme wurde versucht, diesen Platz optimal zu nutzen. Durch entsprechende Konfiguration der Squid-Prozesse konnte die Speicherung im Cache nur auf Objekte von WWW-Servern außerhalb des Wissenschaftsnetzes (im folgenden externe Objekte genannt) beschränkt werden. Die zugrundeliegenden Verfahren zur Unterscheidung von internen und externen Objekten basieren auf AS-Informationen, die von sog. whois-Servern zur Verfügung gestellt werden (s. 2. Zwischenbericht).

Vorausgehende Überlegung für dieses Vorgehen war, dass

- externe Objekte Kosten auf den externen Anbindungen des Wissenschaftsnetzes verursachen,
- WWW-Server außerhalb des Wissenschaftsnetzes in der Regel schlechter zu erreichen sind als WWW-Server innerhalb.

So war es naheliegend, den zur Verfügung stehenden Plattenplatz lediglich für die "wertvolleren" externen Objekte bereitzustellen.

Für eine sinnvolle Umsetzung dieses Konzeptes musste in einem zweiten Schritt vermieden werden, dass Anfragen lokaler Caches nach internen Objekten überhaupt an die DFN-Caches gerichtet wurden. Ähnlich zur Domain-basierten Partitionierung mussten entsprechende Regeln zur Steuerung der Anfragen auf den lokalen Caches konfiguriert werden.

Auch wenn die in diesem Abschnitt vorgestellte Maßnahme sinnvoll erscheint, war der Nutzen gering. Eine Untersuchung der alten DFN-Caches, auf denen interne und externe Objekte gespeichert wurden, zeigte, dass weniger als 5 % des gesamten Volumens durch interne Objekte verursacht wurde.

4.6 Policy

Wie bereits erwähnt, konnten im neuen DFN-Cache-Verbund von Beginn an einige Schwächen des alten DFN-Cache-Verbundes vermieden werden. Ein wesentlicher Mangel des alten DFN-Cache-Verbundes war die offene Policy, die keine Zugangskontrolle oder Überwachung des Nutzungsverhaltens vorsah.

Diese offene Policy war noch aus den Anfangszeiten begründet, als die Teilnahme am DFN-Cache-Verbund möglichst einfach gestaltet werden sollte. Es zeigte sich jedoch, dass mit steigender Zahl lokaler Caches ein erhöhter Aufwand zur Überwachung und Administration der DFN-Caches notwendig wurde. Probleme mit – aufgrund fehlerhaft konfigurierter lokaler Caches – überlasteten DFN-Caches nahmen während des Projektverlaufs deutlich zu. Weiterhin wurden die DFN-Caches zunehmend von unbefugten Caches außerhalb des Wissenschaftsnetzes genutzt.

Im nachhinein stellte es sich als schwierig heraus, eine nachträgliche Nutzerbeschränkung einzuführen. Die hierfür notwendigen Informationen, wie z. B. die IP-Adressen der zur Nutzung befugten lokalen Caches, konnten nur mit hohem Aufwand eingeholt werden. Häufig waren die ehemals benannten Cache-Administratoren nicht mehr in ihrem Bereich tätig, zum Teil mußten die Namen und Adressen vermeintlicher Ansprechpartner mühevoll über die WWW-Seiten der jeweiligen Einrichtung ermittelt werden.

Für die Teilnahme am neuen DFN-Cache-Verbund mussten sich die nutzenden Einrichtungen registrieren lassen, bevor ihnen der Zugang explizit freigeschaltet wurde. Dieses Verfahren war zunächst mit höherem Verwaltungsaufwand für die Projektmitarbeiter verbunden, erwies sich im dauerhaften Betrieb jedoch als unumgänglich.

Verbunden mit der Teilnahme war auch die – allerdings nur in Grenzen verpflichtende – Anerkennung der für den neuen DFN-Cache-Verbund formulierten Policy (s. 2. Zwischenbericht). Ein wesentlicher Bestandteil der Policy war die Aufforderung zur Beachtung von Konfigurationshinweisen lokaler Caches. Anhand verschiedener, im Rahmen des Projektes entwickelter Monitoring-Tools konnte das Zugriffsverhalten der lokalen Caches auf den DFN-Caches umfassend überwacht werden. Konnten fehlerhaft konfigurierte lokale Caches beobachtet werden, erhielten die betroffenen Administratoren eine entsprechende Benachrichtigung. Dieses Vorgehen führte zu einem äußerst kontrollierten und dadurch problemfreien Betrieb der neuen DFN-Caches. Darüber hinaus waren die Reaktionen der Cache-Administratoren auf Hinweise auf mögliche Konfigurationsfehler durchgehend positiv.

4.7 Hilfen für Nutzer

Zu den weniger herausfordernden, jedoch unumgänglichen Arbeiten der Projektmitarbeiter, zählte die Betreuung der Nutzer. Aus Sicht des DFN-Cache-Verbundes konnten die Nutzer in zwei Gruppen unterteilt werden:

- die Administratoren der lokalen Caches,
- die tatsächlichen Endnutzer an ihrem Arbeitsplatz, die mit ihrem Browser über lokale Caches und DFN-Caches auf Inhalte im WWW zugreifen.

Der Aufwand zur Betreuung der Cache-Administratoren war offensichtlich notwendig, da zur korrekten Nutzung der DFN-Caches eine Reihe von Änderungen in der Squid-Konfiguration vorgenommen werden musste. Die hierfür notwendigen Anleitungen wurden über den gesamten Projektzeitraum gepflegt und waren nach Reaktionen der Cache-Administratoren stets eine große Hilfe. In einigen Fällen konnten sogar Konfigurationen gemeinsam per Mail oder Telefon kontrolliert und korrigiert werden.

Hilfen für die Endnutzer waren unter anderem immer dann notwendig, wenn auf den DFN-Caches eine Fehlermeldung generiert und zurückgesendet wurde. So werden z. B. Hostnamen in URLs nicht von den Browsern oder lokalen Caches auf Korrektheit überprüft. Da fehlerhafte URLs in der Regel gleichbedeutend mit MISSes in lokalen Caches sind, werden dieses Anfragen stets an die DFN-Caches geleitet. Erst hier wird eine DNS-Auflösung durchgeführt, die bei einem unkorrekten Hostnamen in einer entsprechenden Fehlermeldung endet. Diese Fehlermeldung wird von dem DFN-Cache auf einer WWW-Seite – in der Regel in englischer Sprache – dargestellt und an den Klienten übermittelt. Um den Nutzern einen größtmöglichen Komfort zu bieten, wurden die Fehlerseiten der DFN-Caches durch ausführliche deutschsprachige Texte ersetzt. Die neuen Texte umfassten nicht nur die Beschreibung des Fehlers, sondern auch Tipps zur Behebung.

Mit diesen erklärenden Fehlertexten konnten auch unnötige Anfragen der Endnutzer vermieden und somit Bearbeitungszeit für die Projektmitarbeiter eingespart werden. Die Nachfragen von Endnutzern nahmen zum Projektende deutlich ab, obwohl das Verkehrsaufkommen bzw. die Nutzung des Dienstes im gleichen Zeitraum anstieg.

Die Notwendigkeit der Nutzerfreundlichkeit eines DFN-Cache-Service oder vergleichbarer zukünftiger Dienste soll an dieser Stelle deutlich hervorgehoben werden. Neben einem technisch funktionierenden Service ist eine kompetente und informative Nutzerbetreuung unerlässlich. Die Bedeutung der Nutzerfreundlichkeit wird um so höher, wenn, wie in Fall der lokalen Caches oder des DFN-Cache-Verbundes, die Nutzung des Dienstes auf freiwilliger Basis geschieht.

4.8 Verkehrsbilanzen

Ergänzend zu den im 2. Zwischenbericht aufgeführten Verkehrsstatistiken werden hier die vollständigen Daten bis zum Projektende aufgeführt. Bis zur Entscheidung für die Abschaltung des DFN-Cache-Verbundes Ende November 2000 zeigte sich ein Anstieg der Nutzung, obwohl in den letzten Wochen kaum weitere lokalen Caches für die Nutzung gewonnen werden konnten. Als Ursache für den fortdauernden Anstieg könnte der Semesterbeginn im Oktober herangezogen werden. Entsprechende Anstiege im Datenaufkommen konnten auch im alten DFN-Cache-Verbund in den vergangenen Jahren beobachtet werden.

Zur Erläuterung der Verkehrsstatistiken in Abbildung 3 und 4 wird, wie bereits in vorangegangenen Berichten üblich, Abbildung 2 herangezogen. Folgende Datenströme werden in den Verkehrsstatistiken beobachtet:

- In* Verkehr zwischen DFN-Caches und WWW-Servern
- Out* Verkehr zwischen DFN-Caches und lokalen Caches
- Among* Verkehr zwischen DFN-Caches untereinander
- Savings* Out - In

Für den neuen DFN-Cache-Verbund ist die Beschriftung auf der Ebene mit den DFN-Caches in *abroad* und *!abroad* zu ändern.

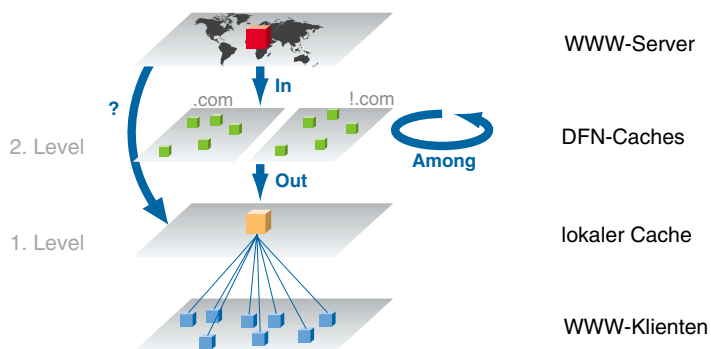


Bild 2: Datenströme in der Verkehrsbilanz.

Die vierte Größe in Bild 2, der mit ? markierte Verkehr zwischen WWW-Servern und lokalen Caches bzw. WWW-Klienten, lässt sich nur anhand von Statistiken der zentralen Router im Wissenschaftsnetz ermitteln. Die entsprechenden Messungen werden durch das DFN-NOC durchgeführt und auf eingeschränkt zugänglichen WWW-Seiten zur Verfügung gestellt.

In Tabelle 5 ist der eingehende Datenverkehr auf den wichtigsten externen Anbindungen des Wissenschaftsnetzes (Stand Dezember 2000) aufgeführt. Die angegebenen Datenraten wurden den grafischen Darstellungen des DFN-NOC entnommen.

| | Aufpunkt im WiN | Datenrate eingehend (Tagesmittel Wochentag) | entspricht Volumen/Tag |
|------------|-----------------|---|------------------------|
| USA | Frankfurt | 160 Mbps | 1.688 GByte |
| USA | Hannover | 140 Mbps | 1.477 GByte |
| TEN 155 | Frankfurt | 35 Mbps | 369 GByte |
| DE-CIX | Frankfurt | 100 Mbps | 1.055 GByte |
| T-Internet | Hannover | 20 Mbps | 211 GByte |
| T-Internet | Stuttgart | 15 Mbps | 158 GByte |
| Summe | | 470 Mbps | 4.957 GByte |

Tabelle 5: Übersicht externe Anbindungen im Wissenschaftsnetz (Dezember 2000)

Mit der Annahme, dass der Anteil an WWW-Verkehr ca. 75 % des gesamten Verkehrs beträgt, ergibt sich für den täglich in das Wissenschaftsnetz eingehenden WWW-Verkehr ein Volumen von 3,7 TByte. Diese Volumen entspricht einer über den Tag gemittelten Leitungskapazität von 350 Mbps.

Die durch den neuen DFN-Cache-Verbund gewonnenen Einsparungen von bis zu 50 GByte pro Wochentag (Abbildung 4) fallen im Vergleich zu den 3,7 TByte an tatsächlich empfangenem Datenvolumen gering aus. Es ist festzustellen, dass auch der neue DFN-Cache-Verbund in der dargestellten Konfiguration und Leistungsfähigkeit keinen feststellbaren Nutzen bzgl. der Einsparung von Ressourcen auf den externen Anbindungen erzielte.

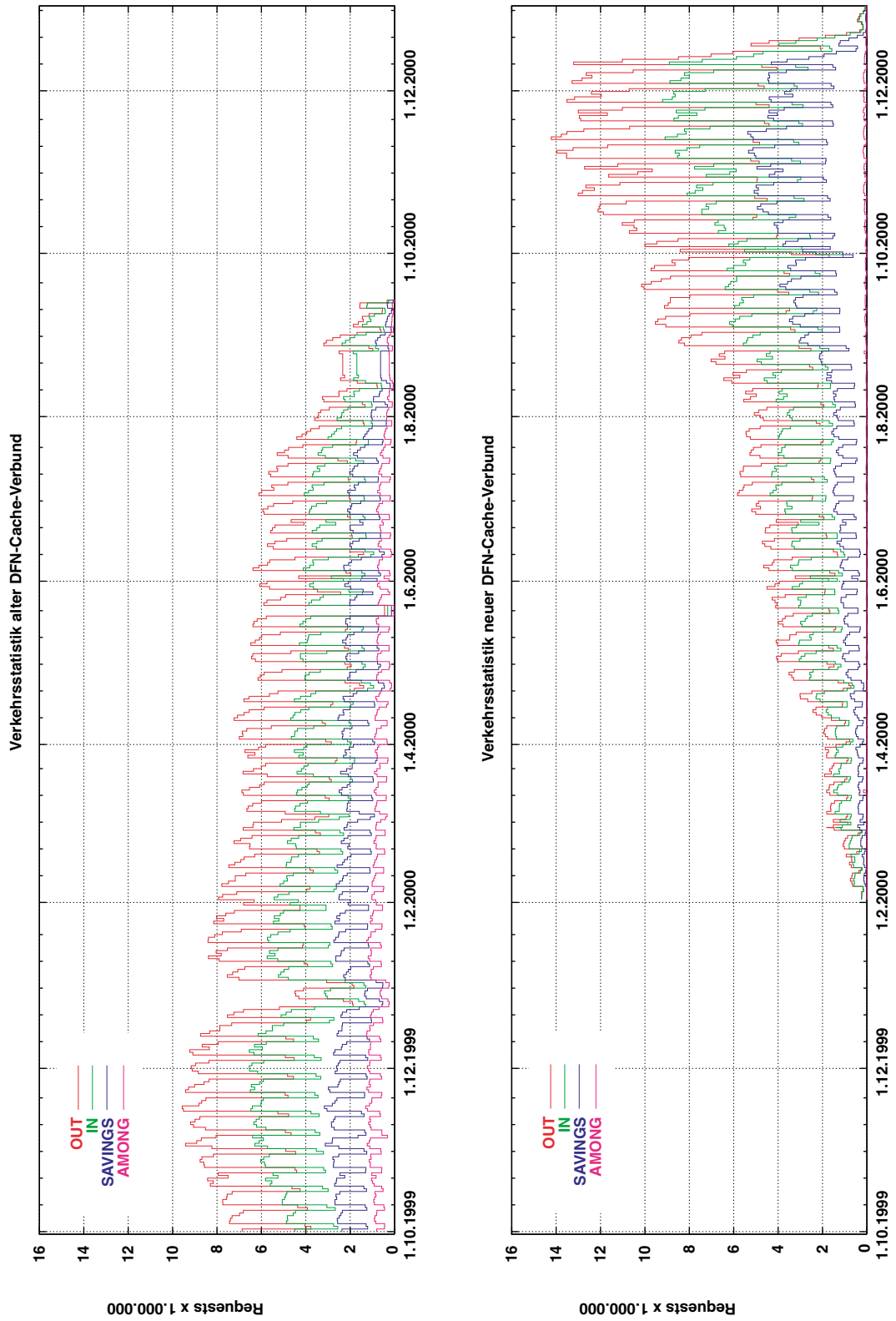


Bild 3: Verkehrsbilanz Requests alter und neuer DFN-Cache-Verbund

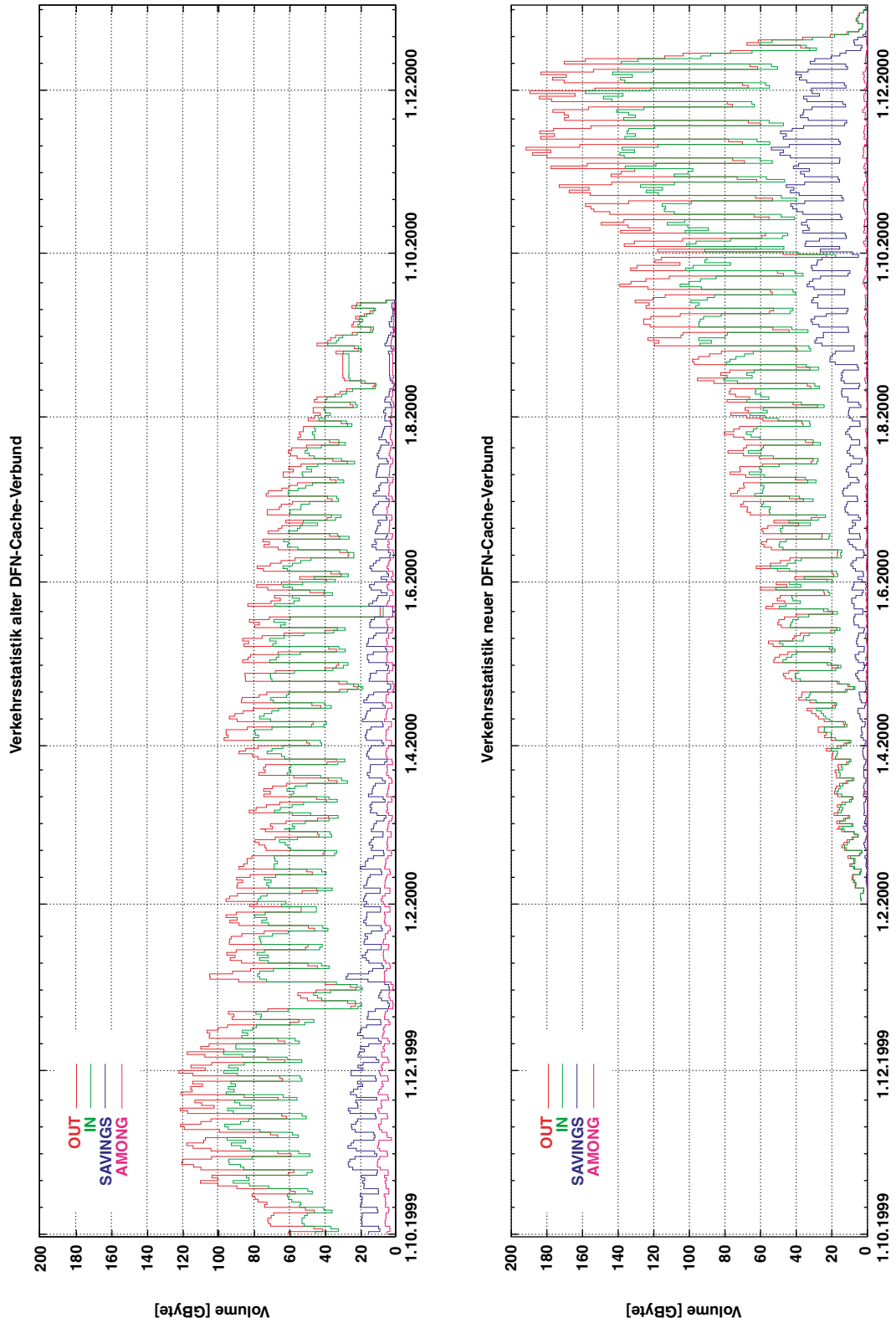


Bild 4: Verkehrsbilanz Volumen alter und neuer DFN-Cache-Verbund

5 Bewertung DFN-Cache-Verbund

In Vorträgen auf der 33. DFN-Betriebstagung und vor dem DFN-Betriebsausschuss im Oktober 2000 wurde von den Projektmitarbeitern empfohlen, eine Fortführung des DFN-Cache-Service zunächst nicht anzustreben. Diesem Vorschlag wurde seitens der DFN-Geschäftsleitung zugestimmt und die Beendigung des Dienstes zum 31. Dezember 2000 bekanntgegeben. Im folgenden werden die Gründe erläutert, die zur Einstellung des DFN-Cache-Verbundes führten. Weiterhin wird dargelegt, warum keine Versuche zur Etablierung eines neuen DFN-Cache-Verbundes unternommen wurden.

5.1 Nutzen

Vorangiges Ziel des DFN-Cache-Verbundes war die Einsparung von Leitungskapazität auf den externen Anbindungen des Wissenschaftsnetzes. Bereits im alten DFN-Cache-Verbund musste festgestellt werden, dass die erzielten Einsparungen keine signifikante Verringerung des WWW-Verkehrs bewirkten. Als wesentlicher Grund hierfür wurde die unzureichende Hardware der 10 alten DFN-Caches angeführt.

Im Projekt Teil II konnte auch mit verbesserter Hardware keine Veränderung der Situation beobachtet werden. Das eingesparte Volumen konnte zwar mehr als verdoppelt werden, in demselben Zeitraum stieg jedoch der gesamte WWW-Verkehr um den Faktor 4 an. Somit hat sich das Verhältnis von eingespartem WWW-Verkehr zu gesamtem WWW-Verkehr im Wissenschaftsnetz sogar verschlechtert.

5.2 Nutzung

Die Einsparungen durch einen Cache hängen von dessen Nutzung ab. Unter Beibehaltung der freiwilligen Teilnahme am DFN-Cache-Verbund kann jedoch kein Grad der Nutzung garantiert werden. Hieraus ergibt sich eine Planungsunsicherheit sowohl bei der Dimensionierung der Caches als auch bei einer Kosten/Nutzenprognose.

Der Einsatz lokaler Caches muss ebenfalls vor dem Hintergrund der freiwilligen Nutzung bewertet werden. Es ist derzeit nicht abzusehen, ob und wie weit der Einsatz lokaler Caches im Wissenschaftsnetz vorangetrieben wird. Da ein direkter Zugriff der Endnutzer auf die DFN-Caches aus technischer und administratorischer Sicht nicht realisierbar ist, würden mit fehlenden lokalen Caches die Nutzer eines DFN-Caches-Verbund fehlen. Bei mangelnder Bereitschaft, Caches auf lokaler Ebene einzusetzen, fehlt die wichtigste Voraussetzung zum Betrieb eines übergeordneten Cache-Verbundes.

5.3 Kosten für Hardware

Für eine deutliche Leistungssteigerung ist die vorhandene Hardware unterdimensioniert. Die Ergänzung des neuen DFN-Cache-Verbundes um weitere gleichartige Server würde keine Abhilfe schaffen. Als einzige Möglichkeit einer signifikanten Leistungssteigerung käme die Installation von Hardware-basierten Caches, z. B. von Herstellern wie Network Appliance oder CacheFlow, im DFN-Cache-Verbund in Betracht. Aus unabhängigen Untersuchungen (und mittlerweile durch eigene Erfahrungen bestätigt) ist bekannt, dass zumindest die größeren Geräte dieser Hersteller die Leistungsfähigkeit Squid-basierter Caches deutlich übertreffen.

Die Kosten für einen DFN-Cache-Verbund auf Basis Hardware-basierter Caches würden jedoch im siebenstelligen Bereich liegen. Unter Berücksichtigung der im vorigen Abschnitt erläuterten Unsicherheit für die Nutzung eines solchen Dienstes erscheint die Rechtfertigung derartiger Investitionen fraglich.

5.4 Inter-Cache-Kommunikation

Auch wenn eine Entscheidung für die Etablierung eines neuen DFN-Cache-Verbundes gefallen wäre, tritt mit dem Problem der Inter-Cache-Kommunikation eine weitere Schwierigkeit auf. Da der Einsatz von Hardware-basierten Caches vorausgesetzt werden muss, steht als Protokoll zur Inter-Cache-Kommunikation lediglich ICP zur Verfügung. Die Eigenschaften von ICP, insbesondere die schlechte Skalierbarkeit bei hohem Verkehrsaufkommen, wurde bereits in Abschnitt 4.3 erläutert. Somit erscheint eine effektive Kommunikation zwischen lokalen und DFN-Caches unmöglich.

Alternativ bieten sich zwei Lösungen an:

- Bildung einzelner Inseln mit Cache-Clustern, auf die jeweils nur ein beschränkter Nutzerkreis zugreifen darf,
- Verzicht auf jegliche Inter-Cache-Kommunikation durch Einführung von transparentem Caching auf Ebene des G-WiN-Backbones.

Beide Lösungen werfen jedoch neben den finanziellen Fragen auch konzeptionelle Probleme auf. Die Installation mehrerer Cache-Cluster bringt wieder den hohen administrativen Aufwand; vor dem Einsatz von transparentem Caching im Backbone wäre zunächst eine Reihe technischer und organisatorischer Fragen zu klären.

5.5 Vergleich mit anderen europäischen Forschungsnetzen

Vergleichbare Cache-Projekte in anderen europäischen Forschungsnetzen wurden mit ähnliche Begründungen eingestellt bzw. werden derzeit "schleichend" aufgelöst. Lediglich in Ländern mit schwach ausgebauter Infrastruktur bieten übergeordnete Cache-Verbünde noch einen deutlichen Nutzen.

5.6 Zusammenfassung

Die Einstellung des DFN-Cache-Verbundes war sowohl in den technische Schwierigkeiten (Hardware-Ausstattung, Inter-Cache-Kommunikation) als auch in der hohen Planungsunsicherheit (Zukunft lokaler Caches) begründet. Beide Aspekte führten auch bei einer kritischen Kosten/Nutzenanalyse zu einem negativen Ergebnis. Für den dauerhaften Betrieb eines DFN-Cache-Service muss ein klarer, durch signifikante Einsparungen erkennbarer Nutzen erkennbar sein. Diese Einsparungen können jedoch nur durch eine erhöhte Nutzung erreicht werden. Einer erhöhten Nutzung steht die Notwendigkeit leistungsfähiger, teurer Cache-Hardware gegenüber.

Weiterhin zeigte die Erfahrung, dass bei Projekten im Bereich zentraler Netzdienste eine Reihe von Serviceleistungen (Administration, Hilfen für Nutzer) erbracht werden muss. Diese Leistungen sind als Daueraufgaben im Rahmen von Forschungsprojekten unattraktiv. Eine alternative Übernahme der Betreuung der DFN-Caches z. B. durch Mitarbeiter von DFN-GS oder DFN-NOC erscheint nicht ratsam, da der Zeitaufwand für die Einarbeitung und Betreuung eines Cache-Verbundes hoch ist. Die notwendige Optimierung der UNIX-Ser-

ver ist sehr aufwändig und kann nur von erfahrenen Mitarbeitern durchgeführt werden. Die Betreuung und Konfiguration der Cache-Prozesse erfordern spezielle Kenntnisse und ständige Kontrolle.

Mit den dargestellten Folgerungen werden die Überlegungen zu einem übergeordneten Cache-Verbundes im Wissenschaftsnetz abgeschlossen. Aus Sicht der Projektmitarbeiter wird es in absehbarer Zeit keine Gründe für den Betrieb eines solchen Dienstes geben.

6 Ausblick – Content Delivery/Distribution Networks

Firmen, die bisher im *web caching* tätig waren, schwenken derzeit mit vielen anderen der *caching* Gemeinde auf *streaming* um: Der wahlfreie Zugriff auf größere Datenmengen (Audio, Video, 3D) ohne die Notwendigkeit einer kompletten und bei populären Objekten mehrfachen Übertragung. Idealerweise ist nur noch die Kanalkapazität zwischen Provider und Surfer der limitierende Faktor. Daneben spielen großvolumige statische Objekte ebenfalls eine Rolle.

In den folgenden Abschnitten¹ werden die Beteiligten an einem *content distribution network* (CDN) vorgestellt, eine mögliche Funktionsweise dargestellt, einige Vor- und Nachteile der Funktionsweise und Geschäftsmodelle aufgezeigt sowie ein Blick auf die am Markt vertretenen Mitspieler geworfen.

6.1 Die Beteiligten

In dem CDN-Szenario gibt es mindestens vier Akteure, die miteinander in Verbindung stehen und deren logische Beziehungen zueinander in Abbildung 5 dargestellt wird. Der Konsument (*content consumer*) ist der Nutzer eines Angebotes. Üblicherweise wird das der Browser eines Surfers sein, ist aber nicht darauf beschränkt.

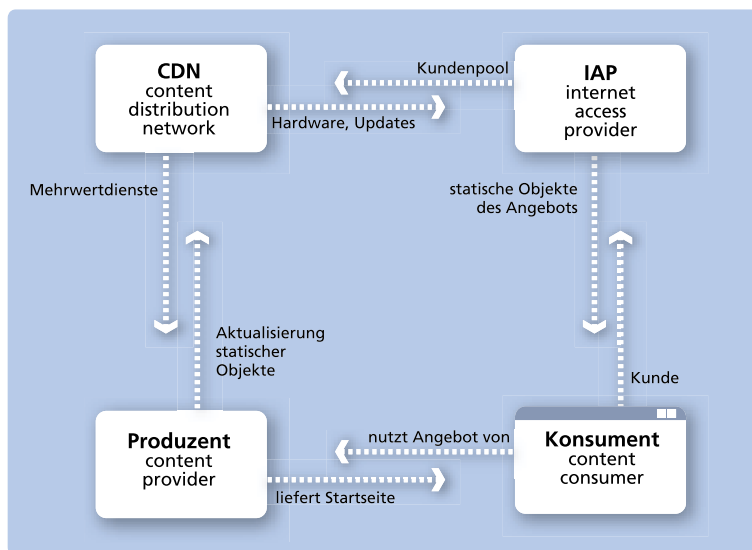


Bild 5: Logische Beziehung zwischen den Beteiligten

Der Produzent (*content provider*) stellt ein Angebot zur Verfügung. Mit Hilfe des *content distribution network* werden Teile des Angebotes des Produzenten an verschiedenen Stellen im Netz bereitgestellt. Dafür werden Rechner (*content nodes*) bei verschiedenen *Internet access provider* (IAP) aufgestellt. Der IAP enthält im Gegenzug zu seinem Kundenstamm und für die Unterbringung der Rechnerknoten die notwendige Hardware und eine entlastete Außenanbindung.

1. Auszug aus einem Artikel in *ix* 04/2001. Ko-Autor Bert Ungerer, Verlag Heinz Heise.

Der Konsument ist ein Kunde eines IAP. Von einem Rechnerknoten beim IAP bezieht der Konsument die statischen und *streaming* Objekte eines Angebotes. Die Startseite des Angebotes – das Dokument, in dem die Objekte eingebettet sind – bezieht der Konsument immer noch direkt vom Produzenten.

Der Produzent übergibt statische Objekte dem CDN, welches für die Verteilung an die Kanten seines virtuellen Netzes sorgt - bis zum IAP. Im Gegenzug ermöglicht es ein CDN dem Produzenten durch Mehrwertdienste die Kontrolle über die verteilten Objekte zu behalten.

Nähe und Nachbarschaft bezieht sich immer auf Netzwerkdimensionen, nicht auf geographische Gegebenheiten. Die round trip time (RTT) ist ein mögliches Maß für Nähe. Alternative Metriken können auf Paketverlustrate, Hops oder BGP-AS-Pfaden aufbauen.

6.2 Der Ablauf

Die CDN Betreiber lassen sich nur ungern in ihre Karten schauen. Mit Hilfe einiger Beobachtungen und neuerer Dokumente¹ läßt sich jedoch beschreiben, wie ein CDN funktionieren kann. Zunächst liefert ein Server oder ein nahegelegener Caching-Proxy des Anbieters die Startseite eines Angebotes. Das ausgelieferte Dokument enthält Referenzen auf große statische Objekte oder Streams (Abbildung 6). Alle Zugriffe auf das Internet erfolgen über den gewohnten Provider.

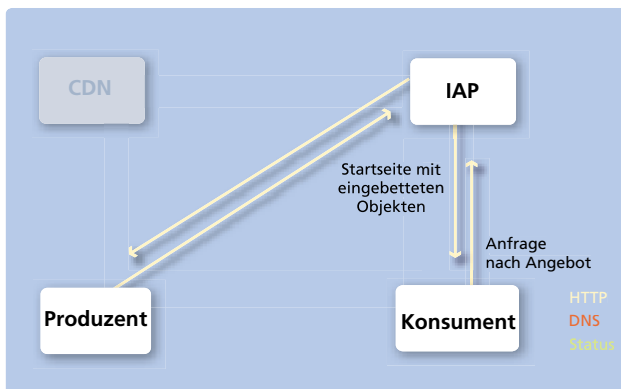


Bild 6: Anforderung der Startseite

Die oft multimedialen Referenzen verweisen auf Adressen im Einflussgebiet des CDN. Die Startseite kann durchaus statisch sein. Dynamisches HTML, beispielsweise um Referenzen anzupassen, ist in diesem Schritt nicht notwendig.

Im nächsten Schritt löst der Browser die symbolische Adresse mit Hilfe des Domain Name System (DNS) in eine numerische IP-Adresse auf. Dazu befragt er den DNS-Server seines

1. <http://www.ietf.org/internet-drafts/draft-cain-cdn-known-request-routing-01.txt> (Feb 2001)

Zugangsproviders (Schritt 1 in Abbildung 7). Dieses Beispiel setzt voraus, dass kein Eintrag im DNS-Cache des Providers vorliegt (Schritt 2).

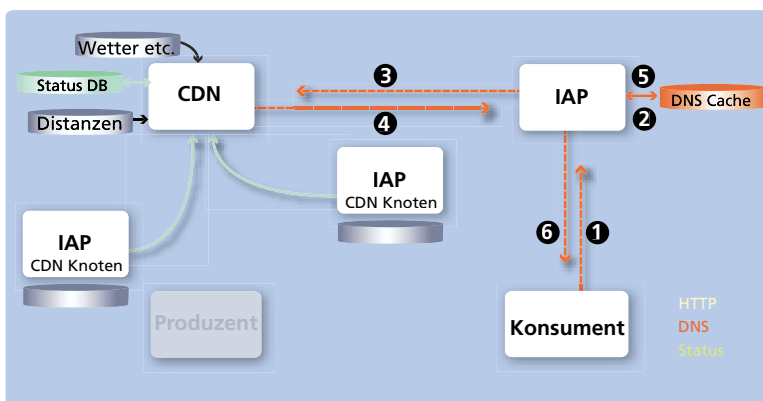


Bild 7: DNS Anfrage für eine Stream-Verknüpfung

Mit Hilfe der DNS-Hierarchie kontaktiert der Nameserver des Providers denjenigen des CDN (Schritt 3). Anders als bei einer herkömmlichen DNS-Antwort macht letzterer seine Antwort unter anderem von der Adresse des Anfragenden – hier des Provider-Nameservers - abhängig.

Beispielsweise kann die Absendeadresse dazu dienen, die Autonomous System Number (ASN) des Senders zu ermitteln. Mit Hilfe einer ASN-"Landkarte" des Internet lässt sich eine Menge der CDN-Server (content node) ermitteln, die die kleinste Netzdistanz zum Anfragenden aufweist. Am Beispiel: Für Tabelle 6 wurde ein Akamai Server zugrundegelegt, wie er von den Apple Startseiten referenziert wird. Dasselbe Ziel ist im G-WiN anders erreichbar als beispielsweise bei COLT.

Tabelle 6: Zu a772.g.akamai.net bei COLT (unten, zu 212.31.251.13 bei Colt Italy) und im G-WiN (rechts, zu 195.176.255.137 bei Switch)

| Nr | Netzwerk | ASN | RTT [ms] | | |
|----|------------------|-------|----------|----|----|
| 1 | 213.61.0.0/16 | 9126 | 1 | 1 | 1 |
| 2 | 213.61.0.0/16 | 9126 | 2 | 2 | 2 |
| 3 | 213.61.0.0/16 | 9126 | 3 | 2 | 2 |
| 4 | 212.121.128.0/19 | 9126 | 7 | 8 | 7 |
| 5 | 212.74.65.0/24 | 8220 | 8 | 8 | 8 |
| 6 | 212.74.65.0/24 | 8220 | 18 | 17 | 17 |
| 7 | 212.74.71.0/24 | 8220 | 17 | 18 | 17 |
| 8 | 212.31.224.0/19 | 12761 | 18 | 17 | 18 |
| 9 | 212.31.224.0/19 | 12761 | 19 | 17 | 18 |
| 10 | 212.31.224.0/19 | 12761 | 18 | 18 | 20 |

| Nr | Netzwerk | ASN | RTT [ms] | | |
|----|------------------|-------|----------|----|----|
| 1 | 130.75.0.0/16 | 680 | 1 | 1 | 1 |
| 2 | 130.75.0.0/16 | 680 | 4 | 1 | 1 |
| 3 | 130.75.0.0/16 | 680 | 1 | 1 | 1 |
| 4 | 188.1.46.0/24 | (680) | 1 | 2 | 2 |
| 5 | 188.1.29.0/24 | (680) | 1 | 1 | 2 |
| 6 | 188.1.18.0/24 | (680) | 6 | 6 | 6 |
| 7 | 188.1.80.0/24 | (680) | 7 | 6 | 6 |
| 8 | 212.1.192.0/21 | 8933 | 7 | 6 | 7 |
| 9 | 212.1.192.0/21 | 8933 | 7 | 7 | 7 |
| 10 | 212.1.192.0/21 | 8933 | 19 | 19 | 20 |
| 11 | 212.1.192.0/21 | 8933 | 22 | 22 | 22 |
| 12 | 130.59.0.0/16 | 559 | 27 | 29 | 27 |
| 13 | 195.176.224.0/19 | 559 | 29 | 28 | 27 |

Von den in Frage kommenden Rechnerknoten ermittelt das CDN den günstigsten Zielknoten und gibt ihn in der DNS-Antwort an (Schritt 4 in Abbildung 7). Zustandsinformationen – etwa über die Verfügbarkeit bestimmter Kundenobjekte – über alle Zielknoten stehen dem CDN ständig aktualisiert und unabhängig von DNS-Anfragen zur Verfügung. Besonders intelligente Verfahren können die aktuelle "Internet-Wetterlage" in die Ermittlung der Antwort einbeziehen. Der DNS-Cache des Providers speichert die Antwort des CDN (Schritt 5) und liefert sie an den Kunden aus (Schritt 6).

Der Konsument nutzt die vom DNS zurückgegebene Adresse, um auf den Inhalt zuzugreifen. So wird etwa aus `www.msnbc.com` die speziell für diese Übertragung ermittelte Adresse `a799.ms.akamai.net/7/799/388/3f585017485515/www.msnbc.com`. Statische oder Streaming-Objekte liefert ein Server in der Nähe des Konsumenten, der direkt beim Provider oder in dessen Nachbarschaft zur Verfügung steht (Abbildung 8).

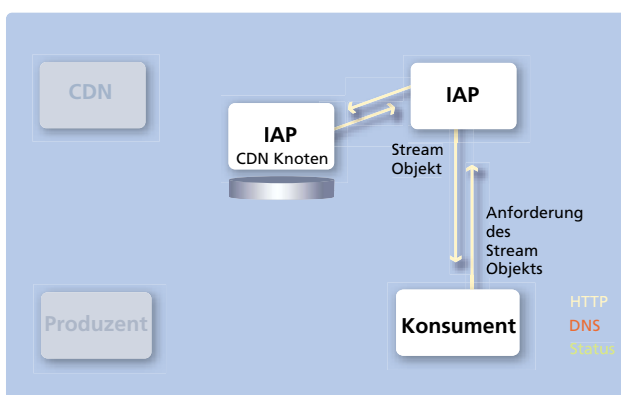


Bild 8: Abruf des Multimedia Objektes

Trotz der von CDNs aufgestellten Serverfarmen ist es unwahrscheinlich, dass dort alle vom Konsumenten gewünschten Objekte vorliegen. Das CDN besorgt fehlende Objekte entweder direkt vom Originalanbieter ("Caching Surrogate"¹) oder von einem anderen Server des CDN (Knoten arbeitet als Caching Proxy) – siehe auch Abbildung 9.

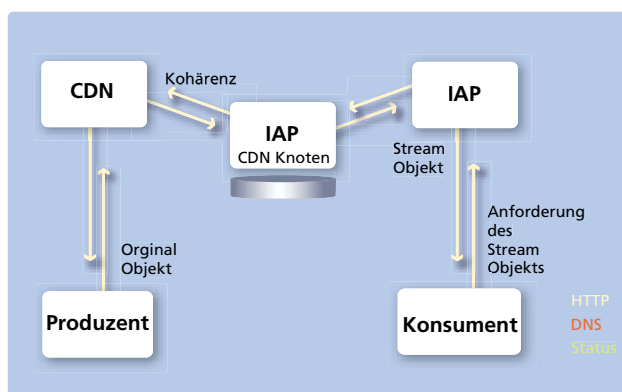


Bild 9: Aktualisierung von Objekten

1. RFC 3040.

Die von "Content Replication" und Caching bekannten Verfahren kommen auch innerhalb eines CDN zum Einsatz, mit all ihren Vor- und Nachteilen. Die Aktualisierung muss nicht notwendigerweise auf Grund einer Anfrage geschehen. Anders als ein regulärer Caching Proxy muss ein CDN-Knoten nur Objekte von CDN-Kunden vorhalten, was sich günstig auf die Hit-Rate auswirkt. Mit Hilfe proprietärer Protokolle kann jeder Produzent seine Objekte aktiv aktualisieren oder invalidieren.

Es ist nicht notwendig, dass alle CDN-Rechnerknoten alle Objekte eines Kunden vorhalten. Eine Partitionierung, bei der mindestens zwei Knoten je "Region" Objekte bereithalten, ermöglicht es, mehr Objekte bereitzustellen.

Alternative Verfahren können an Stelle einer DNS-Umleitung das initiale Segment der Transportschicht oder auch weitergehende Protokollinformationen auf Applikationsebene auswerten, um daraufhin die Anfrage an den nächstgelegenen Rechnerknoten des CDN weiterzuleiten. Dabei kommt in der Regel die echte Adresse des Klienten zum Einsatz. Ein solch hoher Aufwand ist aber nur für langlebige Datenflüsse zu rechtfertigen.

6.3 Mögliche Vor- und Nachteile

Wenn die DNS Anfrage nicht in unmittelbarer Nähe des Servers generiert wird, verweist die Antwort des CDN unter Umständen auf einen Server, der nicht optimal zum Konsumenten plaziert ist. Dies kann beispielsweise beim Einsatz von DNS-Server-Hierarchien (DNS-Forwarder und -Resolver-Caches) geschehen, mit deren Hilfe einige Zugangsprovider und Backbone ISPs das aus Sicht der Kunden wichtige DNS Antwortverhalten optimieren. Unter Umständen kontaktiert daher ein weit vom ursprünglichen Konsumenten entfernter ISP-DNS den CDN-DNS.

Sollten die Nameserver eines Providers temporär nicht erreichbar sein, kann die ermittelte Antwort über den Fallback bei einem entfernten Provider ebenfalls nicht optimal sein. Bei globaler Betrachtung ist sie in vielen Fällen aber oft noch besser, als den Content-Produzenten am anderen Ende einer Transatlantikleitung direkt zu befragen. Außerdem ist beim Ausfall des Provider-DNS die Latenzzeit so schlecht, daß erhöhte Übertragungszeiten an anderer Stelle kaum noch ins Gewicht fallen.

Für Web-Caches können analoge Betrachtungen herangezogen werden. Hier kann ebenfalls ein weit vom Nutzer entfernter Cache die DNS Anfrage durchführen. Da die Auslieferung des Objektes jedoch an den anfragenden Cache geschieht, und die endgültige Auslieferung des Objektes Aufgabe der Caching-Hierarchie ist, sind Cache-Hierarchien verträglich mit dem CDN Konzept.

Viele Verfahren zur Ermittlung von Netz-Nachbarschaften betrachten nur eine Richtung, entweder den Hin- oder Rückweg. Da viele Routen im Internet jedoch asymmetrisch sind, ist das Ergebnis oft wenig aussagekräftig für die tatsächliche Entfernung und die vom Anwender wahrgenommene Verzögerung. Außerdem lassen sich aktive Verfahren nur periodisch starten und scheitern oft an Firewalls.

CDN-Anbieter stellen ihre Knoten gerne hinter den Dialup-Pools eines Zugangsproviders auf. Der Zeitbedarf für die Auslieferung von Inhalten ist dann bedingt durch die hohe Lokalität lediglich von der Anbindung des Konsumenten abhängig. Selbst wenn für die Inhalte ein Knoten bei einem anderen ISP dient, ist dies oft einer interkontinentalen Übertragung vorzuziehen. Ein CDN-Anbieter kann eine um so bessere Qualität zur Verfügung stellen, bei je mehr ISPs ein Knotenrechner seines virtuellen Netzes untergebracht ist.

ISPs bekommen als "Gegenleistung" für das Einbringen ihres Kundenstamms, der die Teilnehmerzahl des CDN-Anbieters erweitert, einen verringerten Bedarf externer Kanalkapazität durch den direkten Zugang zu allen vom CDN vertriebenen Inhalten. Ein CDN benutzt zwar den Kundenstamm eines Zugangsproviders und verkauft genau diesen den Inhaltsproduzenten, aber auf den aufgestellten Rechnerknoten hat der Provider keinen Einfluß. Oftmals kann er als "Host Service Provider" (HSP) tätig werden, indem er als Reseller des CDN gegenüber Produzenten im eigenen Netzbereich auftritt.

Zu den Vorteilen für Content-Anbieter gehört, dass die weltweite Verteilung der Daten eine Form von Hochverfügbarkeit und Entlastung des eigenen Server für die verbreiteten Inhalte gewährt. Produzenten können mit Hilfe der bislang proprietären Protokolle, die ein CDN zur Verfügung stellt, eine bessere Kontrolle über die eigenen Inhalte ausüben, als dies mit den derzeit verfügbaren Caching Proxies möglich ist. So kann der Produzent die Kopien seiner Inhalte selbst auf dem aktuellen Stand halten.

6.4 CDNs als Kontrollinstrumente

Genau an diesem Punkt setzt die Kritik an CDNs ein. Als auf IP aufbauende, von einer Instanz betriebene virtuelle Netze stellen sie Instrumente dar, Inhalte und Vertriebskanäle zu kontrollieren. Sie erlauben nur solche Innovationen, die mit den Zielen des Netzbetreibers in Einklang stehen. Blicke das virtuelle Netz neutral bezüglich der Anwendungen, könnten sich innovativere Produkte durchsetzen, etwa wie das Web gegenüber *gopher*.

Proprietäre Protokolle sind nur eine Facette dieses Problemkreises, die Ablösung der Ende-zu-Ende-Strukturen durch ein CDN jedoch von elementarer Auswirkung¹. Noch gibt es eine Ungleichheit zwischen den zwei Welten, die Inhalte verteilen:

- Rein kommerzielle Inhalte wie Filme, Bücher und Musik unterliegen dem Copyright. Deren Verteilung ist genauestens kontrolliert; wenige Produzenten haben das Sagen über Herstellung, Bereitstellung und Verteilung.
- Eine Kontrolle der Verbreitung freier Inhalte, die viele Urheber haben, verteilt angelegt sind und frei miteinander geteilt werden dürfen, ist dagegen kaum möglich.

Während das ursprüngliche Internet den zuletzt genannten Inhalten entgegenkommt, kann eine Instanz der kommerziellen Seite mit Hilfe eines CDN die Kontrolle über seine Inhalte auch im Internet behalten. Der Zusammenschluß von AOL mit Time Warner und die Mitgliedschaft von AOL im Content Bridge Konsortium (siehe unten) zeigen, dass Hollywood durchaus weiß, wie die Kontrolle von der Herstellung bis zur Auslieferung trotz Internet aufrecht zu erhalten ist.

Das Content-Bridge-Konsortium vertritt sogar die Meinung, dass heutige CDNs noch nicht weit genug gehen: Das Konsortium sieht durchaus die Probleme, dass proprietäre Protokolle ein geschlossenes System erschaffen, das das Internet in kontrollierte, weltumspannende Inseln fragmentiert.

Es gibt im CDN-Ansatz viele Mitspieler, die sich einig sein müssen. Durch Reselling beispielsweise verliert ein CDN die direkte Kontrolle über seine Inhalte. Andererseits wird der Zugangsprovider für die Bereitstellung seines Kundenstamms oft nur ungenügend vergütet.

1. Lawrence Lessig; *Cyberspace's Architectural Constitution*; 1999; Lecture given at WWW9; <http://cyber.law.harvard.edu/works/lessig/www9.pdf>

Nach den Vorstellungen des Content Bridge Konsortiums besteht die Notwendigkeit eines zentralen und anerkannten Operators zu Abrechnungszwecken. Es geht ums Geld: Die Vergütungen werden in der Kette vom Anbieter der Inhalte bis zum IAP gemäß Aufwand und Nutzung abgerechnet. Ein Rückkanal für statistische Informationen soll das Angebot dem Anbieter der Inhalte helfen, sein Angebot zu verbessern. Neue, Stream-basierte Protokolle finden neben HTTP Beachtung.

Der Ansatz hat zwar einen gewissen Charme, aber es gibt darin Haken und Ösen. Ein zentraler Operator ist der "Single Point of Failure" und mit den bei verteilten Systemen dieser Größenordnung bekannten Problemen behaftet. Der Rückkanal für die statistischen Informationen hat in der Vergangenheit zu oft für negative Schlagzeilen gesorgt (Microsoft, Real, Intel), als dass ein Vertrauensvorschuss angeraten wäre.

6.5 Der Content-Delivery-Markt

Der CDN-Markt ist im Wesentlichen dreigeteilt:

- Zu den großen Namen gehört zweifelsohne der Marktführer Akamai Technologies.
- Cisco Systems hat mit Cable&Wireless, Digital Island, Genuity, GlobalCenter, Mirror Image Internet, NaviSite, Network Appliance, PSINet und ServInt die "Content Alliance" gegründet¹, die einen Standard für die Internet Content Distribution schaffen soll.
- In einer weiteren Allianz namens "Content Bridge" haben sich unter anderem Inktomi, AOL, Adero, Apogee, Mirror Image, Digital Island, madge web, Alteon Websystems, Net Rail, Portal und Storage Networks zusammengetan². Einige der Firmen sind mit Web Caching groß geworden, andere gehören zu den Content-Distributoren.

Viele Firmen und Persönlichkeiten entstammen dem Umfeld von Forschung und professionellen Web-Caching-Verfahren: Peter Danzig, vom Harvest Information Discovery System kommend, später chief architect bei Network Appliance für die Caching Produktlinie zuständig, ist jetzt VP of technology bei Akamai. Pei Cao, ehemals im Caching-Bereich an der University of Wisconsin tätig, Gründerin von Tasmania, lässt sich mit ihrer Kollegin Dan Li als eine Kraft hinter Ciscos Standardisierungsvorschlägen vermuten. Bei den Standardisierungsverfahren ist der Einfluss von John Martin, einst als CTO bei Terena für Caching und Indexing zuständig, derzeit CTA bei Network Appliance, nicht zu unterschätzen. Die Akquisition kleinerer Firmen aus dem Umfeld des Caching hat Cisco um wichtiges Know-how bereichert: Tasmania für Caching-Verfahren, Netiverse für Content Acceleration, Arrowpoint für Content Switching und SightPath für Content Delivery.

Die Vorschläge zu neuen Standards befinden sich bereits in Diskussion bei der IETF. Mehrere Gruppen beschäftigen sich mit dem Themenbereich CDN und Caching. Web Caching and Replication (WREC³) hat fast die selbst gesteckten Ziele erreicht, beispielsweise die Schaffung einer einheitlich nutzbaren Benennung oder das Aufzeigen bekannter Probleme. Web Intermediaries (WEBI) ist als Nachfolger geplant. Eine selbstständige Content Peering Working Group ohne IETF-Segen unter Leitung der Content Alliance hat im November 2000 etliche Vorschläge als Internet Draft eingebracht, beispielsweise zum Peering beziehungsweise Transit zwischen verschiedenen CDNs.

1. www.content-peering.org

2. www.content-bridge.org

3. www.wrec.org

6.6 Fazit

Primäres Ziel eines CDN ist die Minimierung der Verzögerungszeit, indem ein Angebot so nahe wie möglich am Abrufenden bereitgestellt wird. Einsparungen auf externen Anbindungen sind kein unmittelbares Ziel eines CDN.

Sollte sich der DFN Verein dafür entscheiden, die genannten Bewerber näher zu betrachten, so dürfte die Endkundenbasis von über einer Million G-WiN Nutzern (Schätzung) sicherlich eine starke Verhandlungsbasis bieten. Der Aufwand von Aufstellung und Betrieb eines CDN-Knotens, der Nutzen eines solchen Knotens im eigenen Netz und insbesondere der Mit-Nutzen durch andere, benachbarte Netzwerke ist dabei genau abzuwägen. Hierbei stellt sich die Frage, ob die Einführung solcher Dienste zunächst im Rahmen von Projekten geprüft werden muss oder als unmittelbare Betriebsaufgabe des DFN durch DFN-GS oder DFN-NOC durchgeführt werden kann.

7 Entwicklung des World Wide Web

Anhand der auf den DFN-Caches gesammelten Verkehrsstatistiken konnten über ein Zeitraum von mehr als 4 Jahren Aussagen über Entwicklung und Wachstum des World Wide Web gewonnen werden. Während der Projektlaufzeit waren diese Langzeitbeobachtung notwendig, um das Betriebsverhalten des Cache-Verbundes beurteilen zu können. Weiterhin lieferten sie unverzichtbare Anhaltspunkte für die Planung des neuen DFN-Cache-Verbundes. Unabhängig von dem DFN-Cache-Projekt liefern die im folgenden aufgeführten Ergebnisse wertvolle Informationen für Projekte im Bereich von Verkehrsanalysen im Internet, insbesondere zur Modellierung von WWW-Verkehr.

Für die Darstellung des gesamten Zeitraumes von Anfang 1997 bis Ende 2000 wurden fünf repräsentative Intervalle zu je 28 Tagen ausgewählt: für den alten Cache-Verbund jeweils der Monat Februar in den Jahren 1997, 1998, 1999 und 2000; für den neuen DFN-Cache-Verbund die Tage vom 13. November bis 10. Dezember 2000. Die Auswahl der Intervalle orientierte sich an zwei Forderungen:

- in den betrachteten Zeiträumen wurde der Cache-Verbund im eingeschwungenen Zustand betrieben, d. h. Änderungen an der Konfiguration oder gravierende Hardware-Ausfälle traten nicht auf,
- insgesamt sollte ein möglichst großer Zeitraum betrachtet werden.

In allen Darstellungen wird ausschließlich der Verkehr zwischen lokalen Caches und den DFN-Caches betrachtet (Out-Verkehr, vgl. Abbildung 2, Abschnitt 4.8). Weiterhin wurden nur HTTP-Replies mit den Statuscodes 20x berücksichtigt. Abbildung 10 zeigt zunächst die absoluten Zahlen von übertragenen Objekten und dem resultierenden Datenvolumen. Deutlich wird der stetige, zwischen 1999 und 2000 aber schwächere Anstieg im alten DFN-Cache-Verbund sowie der deutlich höhere Durchsatz im neuen DFN-Cache-Verbund (Spalte 2000.11).

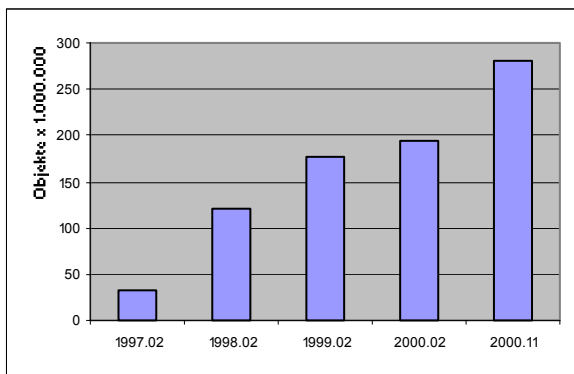


Bild 10a: Objekte DFN-Caches - lokale Caches

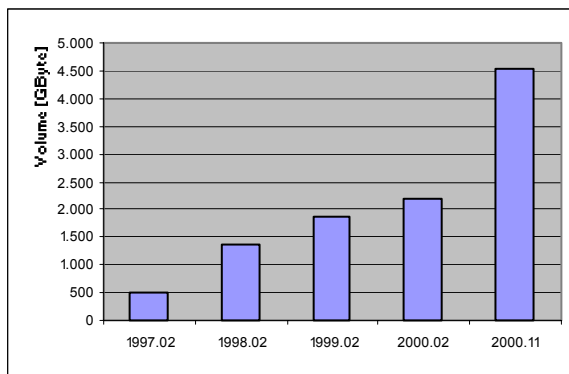


Bild 10b: Volumen DFN-Caches - lokale Caches

Die absoluten Werte dienen nur der Einführung und sollen hier nicht näher untersucht werden. In den folgenden Darstellungen liegt der Schwerpunkt in einer qualitativen Analyse der Datenströme zwischen DFN-Caches und lokalen Caches. Daher werden ausschließlich vergleichende Betrachtungen, basierend auf den Zahlen aus Abbildung 10, erläutert.

Abbildung 11 zeigt die Verteilung von Objekten und Volumen auf die verschiedenen Dienste, die von den DFN-Caches übertragen wurden. Abgesehen von 1997 ist in der Verteilung der

Objekte kein Dienst außer HTTP erkennbar. Dem gegenüber nimmt das Volumen an übertragenem FTP-Verkehr in allen Zeiträumen einen deutlich Anteil ein. Hierdurch wird die allgemein angenommene Tatsache bestätigt, dass FTP-Transfers häufig zur Übertragung größerer Dateien (ab einigen 100 KByte) eingesetzt werden. Weiterhin fällt das Absinken des übertragenen FTP-Volumens von über 20 % auf unter 10 % auf. Hier bestätigt sich die Vermutung, dass Dateitransfers zunehmend direkt über HTTP durchgeführt werden.

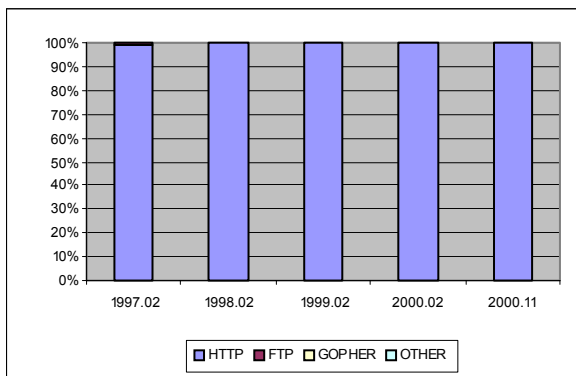


Bild 11a: Objekte nach Service

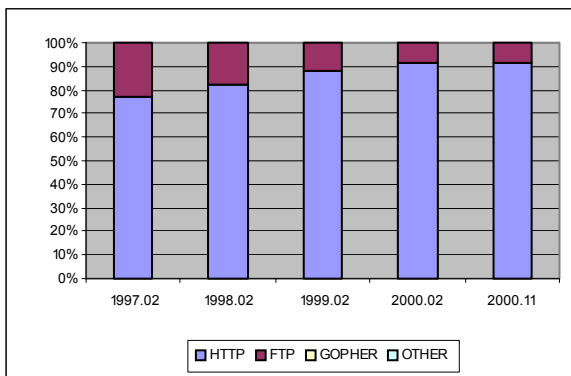


Bild 11b: Volumen nach Service

Abbildung 12 zeigt zwei für die Konzeption des DFN-Cache-Verbundes wichtige Diagramme. Wie bereits in Abschnitt 4.4 dargestellt, spielte die Zuordnung der DFN-Caches auf verschiedene Toplevel-Domains eine wichtige Rolle. Während des Projektverlaufs wurden drei verschiedene Partitionierungen untersucht, deren Eignung anhand Abbildung 12 beurteilt werden kann.

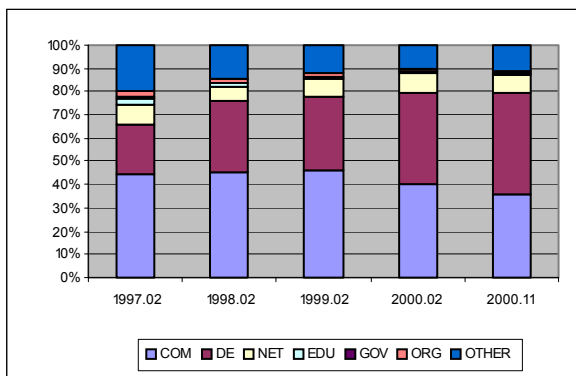


Bild 12a: Objekte nach Toplevel Domain

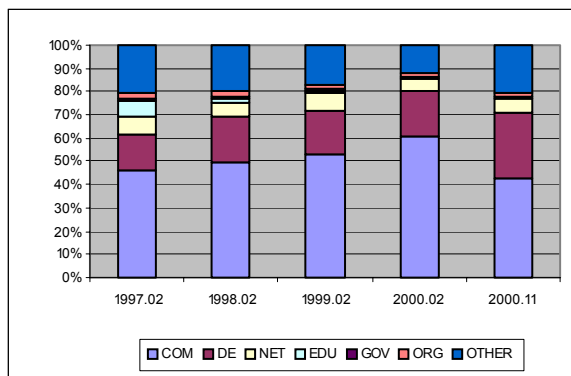


Bild 12b: Volumen nach Toplevel Domain

Wie aus beiden Abbildung zu ersehen ist, war die erste Aufteilung wenig effizient: mit zwei DFN-Caches für die Toplevel-Domain *com* und einem für *de* fielen über 60 % der Objekte und des Volumens auf drei der zehn DFN-Caches. Die veränderte Partitionierung nach *com* und *!com* ergab eine wesentliche Verbesserung, in der die Aufteilung sowohl der Objekte als auch des Volumens an den optimalen Wert von 50 % heranreichte.

Im neuen DFN-Cache-Verbund wurde eine leicht veränderte Partitionierung gewählt. Ausschlaggebend war das Absinken des *com*-Verkehrs gegenüber dem *de*-Verkehr. Wie weit die neuen virtuellen Domains *abroad* und *!abroad* eine brauchbare Verteilung zwischen den DFN-Caches bewirkte, kann aus Abbildung 12 (Spalte 2000.11) nur ungefähr entnommen

werden. Hier fällt auf, dass der Anteil von Objekten aus der Domain *de* zwar 45 %, das entsprechende Datenvolumen jedoch nur 25 % beträgt. Ein nahezu umgekehrtes Verhältnis zeigen die Objekte aus der Domain *com*. Hier wird das ungleiche Verhältnis von übertragenen Objekten zu resultierendem Datenvolumen in den Partitionen *abroad* und *!abroad* bestätigt (vgl. Tabelle 4, Abschnitt 4.4).

Eine weitere Feststellung betrifft die WWW-Server aus den Domains *edu*, *gov* und *org*. Bereits zu Beginn der Beobachtungen waren die Anteile von diesen Servern gering, heute tragen sie keinen relevanten Anteil mehr zum WWW-Verkehr bei.

Eine der häufigsten Untersuchungen des WWW-Verkehrs richtet sich auf die Verteilung der Mimetypes. Abbildung 13 zeigt die entsprechenden Diagramme für den DFN-Cache-Verbund.

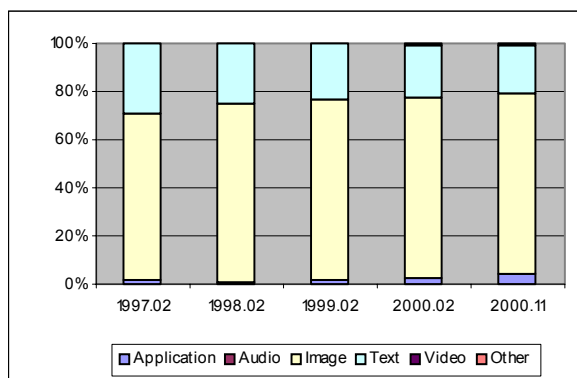


Bild 13a: Requested objects by mimetype

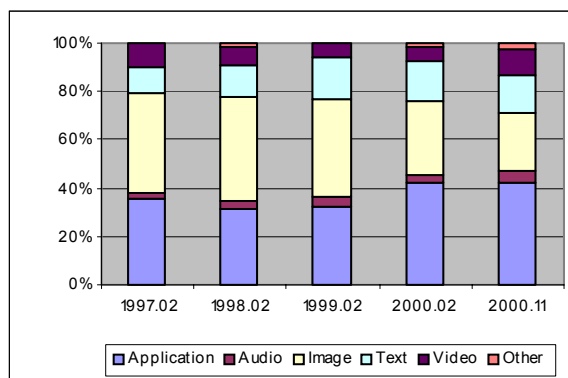


Bild 13b: Requested volume by mimetype

Zunächst fällt auf, dass bei der Verteilung der Objekte keine gravierenden Veränderungen auftraten. Der Anteil an Objekten der Mimetypes Image und Application hat sich über die Jahre leicht erhöht, wogegen weniger Objekte des Mimetypes Text übertragen wurden.

Eine deutlichere Änderungen ist bei den Datenvolumen der einzelnen Mimetypes zu erkennen. Das Volumen von Objekten des Mimetypes Image ist von über 40 % auf weniger als 20 % gesunken. Deutliche Anstiege bei anderen Mimetypes sind jedoch nicht zu erkennen. Der größte Anteil von 40 % wird durch Objekte des Mimetypes Application verursacht. Objekte des Mimetype Text übertreffen mit einem Anteil 15 % auch heute noch das Volumen der Mimetypes Audio und Video zusammen.

Neben den bisherigen Betrachtungen des übertragenen WWW-Verkehrs liefert Abbildung 14 eine weitere Herleitung, mit der sich das Wachstum des WWW belegen lässt. In Abbildung 14a wird die Anzahl der WWW-Server dargestellt, von denen Objekte über den DFN-Cache-Verbund abgerufen wurden. Im Vergleich dazu sind die Messungen der Netcraft Surveys¹ aufgetragen. Ziel der monatlichen Netcraft Surveys ist es, Anzahl, Betriebssystem und eingesetzte Software aller WWW-Server im Internet zu bestimmen.

1. <http://www.netcraft.com/>

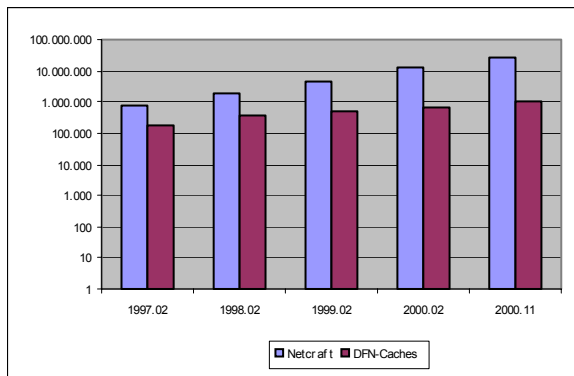


Bild 14a: WWW-Server im Internet

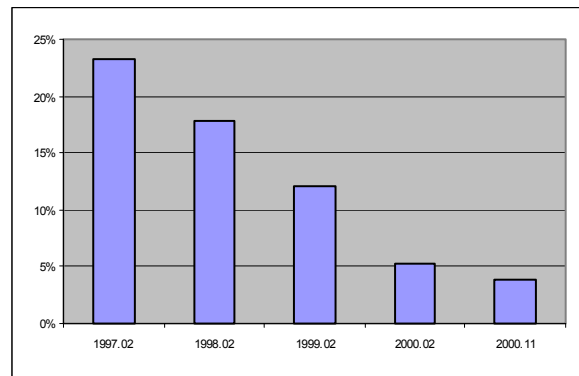


Bild 14b: WWW-Server im Internet

Eine wichtige Aussage über die Entwicklung des WWW kann aus der relativen Darstellung in Abbildung 14b abgeleitet werden. 1997 wurden über den DFN-Cache-Verbund Objekte von annähernd einem Viertel aller WWW-Server im Internet abgerufen. Heute beträgt dieser Anteil weniger als 5 %. Die Begründung für diese Entwicklung ist offensichtlich. 1997 wurde der überwiegende Anteil an WWW-Servern von Forschungseinrichtungen, Hochschulen, staatlichen Institutionen oder größeren Firmen mit internationaler Ausrichtung betrieben. Mittlerweile sind weltweit viele WWW-Server mit überwiegend regionalen Inhalten entstanden. Diese Inhalte sind für Nutzer aus dem Wissenschaftsnetz von geringem Interesse.