

Adaptives Bandbreiten-Management

DFN-Projektvorschlag (Entwurf, Version 0.1)

18. März 2009

1 Eckdaten des Projekts

Projekttitlel:

Adaptives Bandbreiten-Management (DynShaper)

Projektziel:

Ziel des Projekts ist die Entwicklung und Erprobung eines adaptiven Systems zum Bandbreiten-Management, das in Mitgliedseinrichtungen des DFN als Instrument zur Erzielung eines fairen, kostensensiblen und vertragskonformen Umgangs mit Netzressourcen einsetzbar ist.

Projektleiter:

Prof. Dr. Uwe Hübner
Technische Universität Chemnitz
Fakultät für Informatik/Universitäts-Rechenzentrum
Straße der Nationen 62
09107 Chemnitz
E-Mail: Uwe.Huebner@hrz.tu-chemnitz.de

Projektdauer:

24 Monate

2 Notwendigkeit und Ziele des Projekts

Die am Gigabit-Wissenschaftsnetz (GWiN) oder Nachfolgenetzen angeschlossenen Einrichtungen haben unterschiedliche Bedarfssituationen und finanzielle Möglichkeiten. Dem trägt der DFN-Verein durch eine Angebotsdifferenzierung Rechnung, derzeit nach Bandbreite und maximalem Transfervolumen pro Monat. Idealerweise sollten bedarfsgerechte Bandbreiten und Transfervolumina beschafft werden. Dies wird wegen beschränkter Budgets nicht immer möglich sein. Weiterhin kommt jede Definition von „Bedarf“ in Schwierigkeiten, weil es Anwendungen (und Anwender!) gibt, die potenziell nahezu beliebig große Bandbreiten füllen können.

In Reaktion auf die durch diese Situation entstehenden Konflikte werden von den Mitgliedseinrichtungen verschiedene Ad-Hoc-Maßnahmen ergriffen, z.B. (Aus-)Sperrung bestimmter

Dienste (Peer-to-Peer ...) oder Nutzergruppen (Studentenwohnheime ...). Solche recht pauschalen Verfahrensweisen laufen oft einer innovativen und ubiquitären Netznutzung zuwider.

Gesucht sind möglichst intelligente Verfahren, mit denen die Nutzer zu einer überlegten Ressourcennutzung angehalten werden, ohne aber Innovationen zu behindern. Weiterhin sollen solche Verfahren nicht erhebliche Kräfte für administrative Aktionen binden (z.B. Sperren von Nutzern, Erforschen der Merkmale der aktuellsten Peer-to-Peer-Anwendung ...).

In dem Projekt soll ein Bandbreiten-Management entwickelt und erprobt werden, das durch differenzierte Bereitstellung von Bandbreiten pro Nutzer oder Nutzergruppe die Einhaltung vorgegebener Zielgrößen (z.B. Tranfervolumina) und (technisch greifbarer) Richtlinien (*Policies*) gewährleistet. Dabei soll tendenziell ein sensibler Umgang mit den Ressourcen mit besseren Netzparametern belohnt werden.

Dieses Projekt ist im DFN-Entwicklungsprogramm dem Gebiet Netz- und Transporttechnik, Weiterentwicklung von Basistechniken zuzuordnen. Das Projekt liegt im Schnittpunkt der Aspekte Netzmanagement, Traffic Engineering und Dienstqualitäten (QoS). Der Schwerpunkt liegt auf Intra-Domain-Aspekten.

Der primäre Einsatzbereich wird aus heutiger Sicht eher innerhalb der Mitgliedseinrichtungen liegen. Das Bereitstellen praxistauglicher Rezepte für einen kostensensiblen, aber nicht innovationsfeindlichen Umgang mit Netzressourcen unter den spezifischen Bedingungen von Hochschulen kann ein wichtiges Differenzierungsmerkmal des DFN-Vereins gegenüber anderen Netzanbietern darstellen.

Es ist nicht auszuschließen, dass später ein Angebot eines solchen Bandbreiten-Managements auch als DFN-Dienstleistung nachgefragt wird. Wegen der lokal sehr unterschiedlichen Entscheidungen und Entscheidungswege zu *Policies* ist das aus heutiger Sicht aber keine primäre Zielrichtung.

Die Wahl einer vergleichsweise einfachen technischen Basis (Open-Source-Router mit QoS-Fähigkeiten) zielt auf die Vielzahl der kleinen und mittleren DFN-Einrichtungen, die das oben beschriebene Problem meist besonders hart trifft. Voruntersuchungen lassen nutzbare Bandbreiten im Bereich 100 MBit/s .. 1 GBit/s realistisch erscheinen.

Für die Dimensionierung des DFN-Backbone kann ein breiterer Einsatz solcher Techniken Auswirkungen haben, weil dann die Tarifgrenzen wohl häufiger erreicht (aber kaum überschritten) werden. Auch aus dieser Sicht liegt es im Interesse des DFN-Vereins, dass solche Arbeiten mitgestaltet werden (und nicht die Netzplaner von „cleveren“ Applikationen überrascht werden).

3 Vorarbeiten

Wir können auf Arbeiten zur Dienste-Differenzierung für einen lokalen ISP nach den Modellen der *Olympic Services* zurückgreifen. Dabei wurden Paradigmen der Energieversorger genutzt, die auch das Problem von Spitzenlast vs. abgenommene Energiemenge kennen [1].

Ein lokal geprägtes System, in dem die oben erläuterten Grundgedanken verwirklicht wurden, wird seit Ende 2001 an der Schnittstelle zwischen dem Chemnitzer Studentennetz (CSN) [2] und dem Campusnetz der TU Chemnitz betrieben [3, 4]. Diese Lösung weist einige einschränkende und vereinfachende Randbedingungen auf, wie z.B. die feste Einbindung in die konkrete

Netzsituation, eine 1:1-Abbildung zwischen IP-Adressen und Nutzern oder eine pauschale Bewertung von Caching-Gewinnen über einen Korrekturfaktor.
 Alle derzeit 1700 Nutzer im CSN werden von dem gegenwärtig realisierten System behandelt, was für die grundsätzliche Verwendbarkeit des Ansatzes spricht.

Kontakte mit anderen Hochschulen lassen erkennen, dass Problem und Lösungsansatz nicht nur für das Chemnitzer Studentennetz relevant sind, sondern an vielen Stellen. Bei der Anbindung von Studentenwohnheimen tritt das „Verbrauchsproblem“ naturgemäß besonders drastisch auf, aber auch bei Anbindungen von Fakultäten, Instituten, Rechnerpools etc. kann der hier beschriebene Ansatz hilfreich sein.

4 Technischer Inhalt

Eine universelle und mit erträglichem Aufwand an anderen Netz-Lokationen bzw. auch für ganze DFN-Einrichtungen einsetzbare Lösung soll aufbauend auf die im vorhergehenden Abschnitt beschriebene Lösung im Rahmen des hier beantragten Projekts entstehen.

Das System wird zwischen einem *externen Netz* und dem *internen Netz* mit den adaptiv zu regulierenden Nutzern angeordnet:

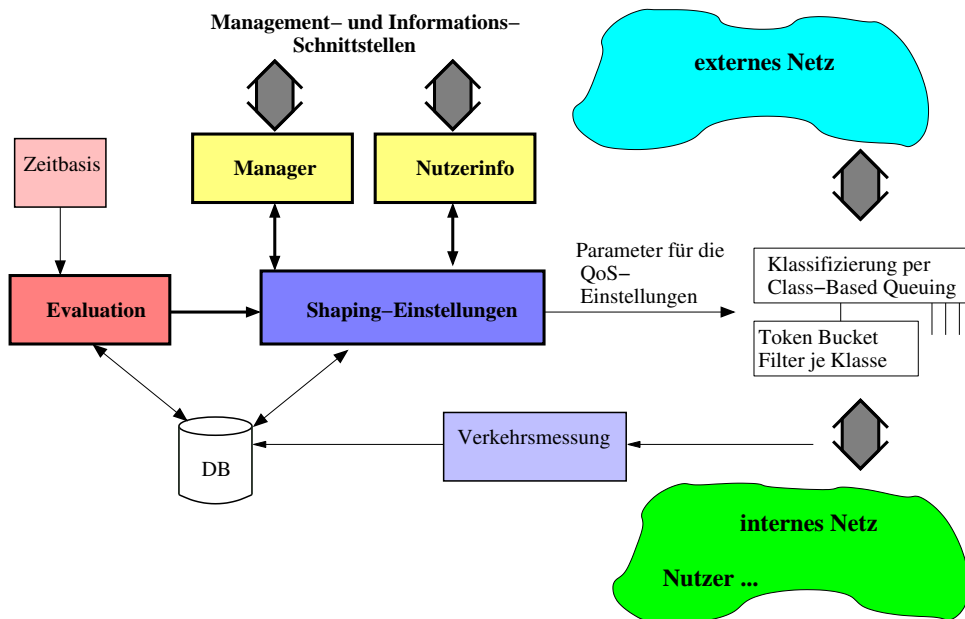


Abbildung 1: Architekturmodell des Systems

Über eine *Manager*-Komponente kann der Netzbetreiber Vorgaben machen, beispielsweise zu Nutzerklassen, Verkehrs-Obergrenzen oder Sonderbehandlungen für bestimmte Protokolle. Im operativen Betrieb sollen dann fast keine Eingriffe mehr nötig sein.

Über eine Schnittstelle *Nutzerinfo* ist es den einzelnen Nutzern möglich, z.B. das aktuelle verursachte Verkehrsvolumen abzufragen. Hier erfahren die Nutzer auch die Klasse, der sie momentan zugeordnet sind sowie die Parameter dieser Klasse (Maximalbandbreite, Verweilzeit ...).

Die Nutzer sind also nicht statisch einer Klasse zugeordnet, sondern gelangen z.B. durch hohen Ressourcenverbrauch für eine gewisse Zeit in Klassen mit stärkerer Bandbreitenbeschränkung. Der Management- wie auch der Nutzerzugang werden plattformunabhängig als HTTP/HTML-Schnittstelle gestaltet.

Die Resultate der Verkehrsmessungen gelangen in eine Datenbank, wo sie periodisch einer *Evaluation* unterworfen werden. In deren Ergebnis und abhängig von Vorgeschichte und konkreter Parametrisierung werden dann die *Shaping-Einstellungen* vorgenommen.

An dieser Stelle werden u.a. Warteschlangen-Strategien und Filterregeln festgelegt, die dann als „QoS-Einstellungen“ der QoS-Implementierung im Kern eines *Linux*-Systems übergeben werden.

Nach unseren bisherigen Erfahrungen können alle diese Komponenten in einem (neueren) Rechner auf PC-Basis betrieben werden. Eine Verteilung (per XML-RPC) ist grundsätzlich ebenfalls möglich.

Um das Potenzial und die Mechanismen zu illustrieren, sei hier als Beispiel die von uns initial verwendete Klasseneinteilung angegeben:

Klasse	Ab Menge	Bandbreite	Tage
0	< 100 MByte/Tag	nicht begrenzt	
1	100 MByte/Tag	100 MBit/s	2
2	200 MByte/Tag	50 MBit/s	4
3	300 MByte/Tag	30 MBit/s	6
4	400 MByte/Tag	5 MBit/s	8
5	500 MByte/Tag	500 KBit/s	10

Tabelle 1: Parameter der Klassen

Die konkreten Bandbreitenwerte werden dann automatisch adaptiv angepasst; wenn beispielsweise in der Semesterpause der Zielparameter *Monatslimit* nach einer Hochrechnung deutlich unterschritten würde, weiten sich die Klassen auf, d.h. die Bandbreiten werden höher eingestellt. Die angegebene Bandbreite gilt für die Summe aller Nutzer in der betreffenden Klasse. Individuelle Bandbreiten pro Nutzer sind grundsätzlich denkbar, dürften aber zu der Skalierungsschwierigkeit führen, die aus der *DiffServ/IntServ*-Diskussion bekannt ist.

Wenn ein Nutzer jetzt beispielsweise eine Datenmenge von 300 MByte/Tag überschreitet, wird er für 6 Tage in Klasse 3 eingeordnet, dann schrittweise (nicht sofort) wieder in die Klassen 2, 1, 0:

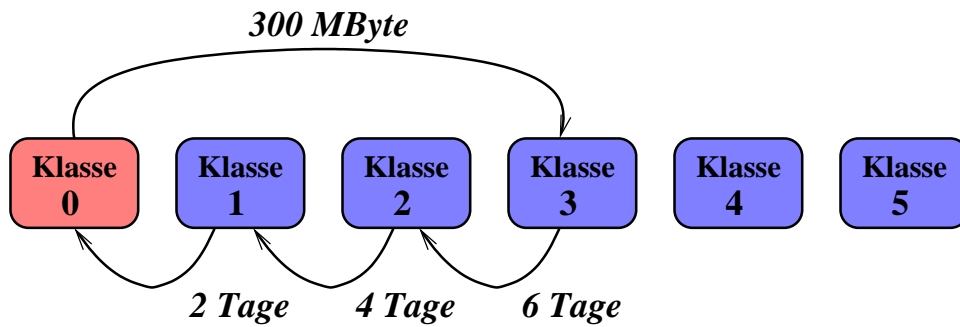


Abbildung 2: Klassenübergänge, wenn Klasse = 0

Die Bandbreitenbegrenzungen und Verweildauern in den Klassen sind so bemessen, dass die Summe der Nutzer die Zielgrößen (Monatsvolumen ...) statistisch einhalten wird.

Für „Ausreißer“ werden Sofortreaktionen implementiert, was dann auf eine so niedrige Bandbreite hinausläuft, dass keine Limits überschritten werden können.

5 Erwartete Ergebnisse des Projekts

In einem möglichst breiten Pilotbetrieb an mehreren Einrichtungen sollen Erkenntnisse gewonnen werden, die in die Entwicklung einfließen. Das betrifft nicht nur die Technik, sondern auch Regelungsrahmen (Ordnungen, *Policies*) und Argumentationen gegenüber den Nutzern. Unsere eigene Erfahrung zeigte, dass hier die Techniker nicht immer auf Anhieb die richtigen Worte finden, beispielsweise erwies sich eine an Bestrafungen und Führerschein-Punkte erinnernde Begriffswelt als unnötige Kontroversenquelle.

Es soll untersucht werden, ob und wie die vereinfachende Randbedingung einer 1:1-Abbildung von Nutzern und diesen fest zugeordneten Rechnern (wie das bei den privaten Rechnern im Wohnheim der Fall ist) aufgegeben werden kann. Als Nutzereinheit sind auf der einen Seite eventuell ganze Adressräume (Institute ...), auf der anderen Seite wechselnde (oder gar mehrere gleichzeitige) Nutzer an allgemeinen Rechnerarbeitsplätzen zu berücksichtigen.

Weiterhin wird Netzverkehr oft durch indirekte Effekte verursacht. Eine Reglementierung des direkten Weges könnte ein Ausweichen von Nutzern auf „Schleichwege“ zur Folge haben (beispielsweise offizielle und inoffizielle Proxy-Server). Hier muss vermutlich eine Kombination aus technischen und administrativen (auch sozialen?) Vorkehrungen eingesetzt werden.

Proxy-Verkehr sollte nach Möglichkeit dem echten Verursacher (und nicht dem Proxy-Betreiber) zugerechnet werden. Konsequenter wäre hier eine Weitergabe der vollen Anrechnung an den ersten Nutzer z.B. eines WWW-Dokuments. Wenn weitere Nutzer dasselbe ressourcensparend aus dem Cache bekommen, könnten die vorherigen Nutzer „Gutschriften“ erhalten. Es soll untersucht werden, welche technische Approximation solcher doch recht aufwendigen Verrechnungsverfahren praktikabel ist.

Die realisierte Lösung soll nicht nur nach Nutzern, sondern auch nach Anwendungen differenzieren können. Das primäre Kriterium wird aber der Nutzer sein müssen, da Anwendungen nur durch sekundäre Kriterien (Portnummern ...) auf einfache Weise klassifizierbar sind. Wenn

man beispielsweise in bester Absicht eine Anwendung *Videokonferenz* als Primärkriterium für eine Bevorzugung einstellt, würden nach kürzester Zeit verschiedene bandbreitenhungrige Anwendungen versuchen, dieses Selektionskriterium zu erfüllen (oft mit Erfolg).

Im Verlaufe des Projekts wird ein Betrieb bei (mindestens) 3 weiteren Nutzungseinrichtungen angestrebt. Damit wird gewährleistet, dass es Nutzer gibt, die nicht mit den Entwicklern identisch sind, ein Nachweis der Reproduzierbarkeit wird erbracht. Außerdem werden hier Anregungen für funktionelle Verbesserungen erwartet.

Die entstehenden praktischen Resultate sollen, wie auch bereits die Zwischenergebnisse, als Open-Source-Software unter GPL-Nutzungsbedingungen (GNU General Public License) [5] offengelegt und verbreitet werden. Damit sind Voraussetzungen für eine Nutzung und Weiterentwicklung auch nach Ablauf der Förderung gegeben.

6 Relation zu anderen Projekten/Standardisierungen

Relevante IETF-Standardisierungen und Standardisierungsansätze lassen sich bei den Mechanismen erkennen:

- *DiffServ* ist als Basismechanismus anerkannt, allerdings ist relevanter praktischer Einsatz noch kaum zu erkennen [6, 7]
- Die Standardisierung zu den AAA-Schnittstellen (*Authentication, Authorization, Accounting*) ist noch im Fluss und wird, soweit sinnfällig, Berücksichtigung finden [8]
- Eine Verwandtschaft mit dem hier vorliegenden Vorschlag ist bei dem Ansatz *Scavenger Service* zu erkennen, der offenbar in Backbones die Nutzbarmachung von anderweitig ungenutzter Bandbreite „unterhalb“ von *Best Effort* zum Ziel hat [9, 10]
- Relationen gibt es ebenfalls zu der Entwicklung *Resource Allocation Protocol (RAP)*, wo in größerem Maßstab *Usage Policies* Berücksichtigung finden sollen [11]

Der Fokus vieler Projekte im QoS-Umfeld liegt auf der Unterstützung missionskritischer Anwendungen bzw. „höherer“ Dienste, oft im Audio- bzw. Videobereich:

- Quasar (*Uni Stuttgart u.a.*) [12]
- Q-WiN (*Uni-Leipzig u.a.*) [13]
- TEQUILA, TF-NGN, SEQUIN ... [14, 15, 16]

Als primäres Ziel gelten typischerweise Servicegarantien. Während die Zielstellungen der genannten Projekte damit deutlich von dem hier vorliegenden Ansatz abweichen, ist eine Übernahme von Erkenntnissen und Erfahrungen im Bereich der QoS-Basismechanismen sinnvoll und vorgesehen.

7 Arbeitsverlauf/Meilensteine

Arbeitspaket 1

Analyse der Ausgangssituation sowie der technischen und administrativen Schnittstellen der Bandbreiten-Management-Lösung

Aufwand: 2 PM

M1: Analysebericht/Schnittstellenbeschreibung im (öffentlichen) Projekt-Web

Arbeitspaket 2

Entwicklung und Erprobung einer weitgehend parametrisierbaren und von lokalen Abhängigkeiten befreiten Version der Bandbreitenmanagement-Implementierung; Herstellung einer mit wenig Installationsaufwand nutzbaren Binärdistribution

Aufwand: 4 PM

M2: Nutzbare Quelltext- und Binär-Distribution im (öffentlichen) Projekt-Web

Arbeitspaket 3

Dokumentierung von Installation und Betrieb, Empfehlungen für die Einführungsprozeduren, Herstellung von erläuternden Informationen für die Netznutzer (*Was wird gemessen und bewertet? Welche Empfehlungen gehen an die Netz-Nutzer?*)

Aufwand: 3 PM

M3: Installationsanleitung und weitere Erläuterungsdokumente im (öffentlichen) Projekt-Web

Arbeitspaket 4

Analyse und Vergleich von Verfahren zur Zuordnung von Nutzer-Identitäten und verursachtem Netzverkehr; Integration der Erkenntnisse in die Implementierung

Aufwand: 3 PM

M4: Analysebericht und neue Quelltext-Distribution im (öffentlichen) Projekt-Web

Arbeitspaket 5

Analyse und Vergleich von Verfahren zur fairen Zuordnung des Verkehrs bei Nutzung von Proxys bzw. Caches; verfügbare oder zu schaffende Schnittstellen der Proxys/Caches; Praktikabilität von „Gutschriften“; Integration der Erkenntnisse in die Implementierung

Aufwand: 3 PM

M5: Analysebericht und neue Quelltext-Distribution im (öffentlichen) Projekt-Web

Arbeitspaket 6

Integration der Differenzierung nach Anwendungen (als sekundäres Kriterium), Gewinnung von Erfahrungen in der praktischen Anwendung dieses Kriteriums für verbreitete Audio- und Video-Streaming- bzw. Konferenzanwendungen

Aufwand: 3 PM

M6: Neue Quelltext-Distribution im (öffentlichen) Projekt-Web

Arbeitspaket 7

Messungen zum Durchsatzverhalten, Hard- und Software-Optimierungen, z.B. Parallelierbarkeit (*die grundsätzliche Nutzbarkeit im Bandbreitenbereich bis 100 MBit/s ist von Anfang an gegeben, hier geht es um die Auslotung der Obergrenzen*)

Aufwand: 3 PM

M7: Messergebnisse und ggf. neue Quelltext-Distribution im (öffentlichen) Projekt-Web

Arbeitspaket 8

Unterstützung der Einführung des Systems an mindestens drei weiteren Einrichtungen; Einbeziehung der Erfahrungen in Weiterentwicklungen des Systems (*praktisch wird dieses Arbeitspaket parallel und fortlaufend zu den übrigen Paketen realisiert werden*)

Aufwand: 3 PM

M8: Nutzung an mindestens drei weiteren Einrichtungen, Dokumentierung der Nutzungsfälle im (öffentlichen) Projekt-Web (bei Einverständnis der Nutzer)

Arbeitspaket 9

Erarbeitung einer Projektabschluss-Version und eines Abschlussberichts

Aufwand: 2 PM

M9: Abschlussbericht und neue Quelltext- und Binär-Distribution im (öffentlichen) Projekt-Web

8 Mittelplanung

Personal

1 wiss. Projektmitarbeiter, 24 PM:	90.000 EUR
Stud. Hilfskraft (10h/Woche, Eigenanteil):	8.000 EUR
Management, 2 PM (Eigenanteil):	12.000 EUR

Hardware

Rechnerarbeitsplatz Projektmitarbeiter (Eigenanteil):	3.000 EUR
Experimentier- und Mess-Hardware (3 PCs, Linux):	5.000 EUR

Reisekosten

zu Testpartnern, DFN-Betriebstagung, Geschäftsstelle (8 x pro Jahr, je 1 Mitarbeiter): je 200 EUR zu 1 x Tagungen/IETF pro Jahr: je 2000 EUR	
Reisekosten gesamt:	7.200 EUR

Infrastruktur

8% lt. Verwaltungsvorschrift:	10.900 EUR
-------------------------------	------------

Gesamtkosten des Projekts:	136.100 EUR
Eigenanteil:	25.000 EUR
DFN/BMBF-Anteil:	111.100 EUR

Literatur

- [1] Dietrich Thie, Torsten Naumann: „Klassenbasiertes Ressource-Management für Internet-Einwahlpunkte mit zeitabhängiger Tarifierung“, PIK 22 (1999) 1, 56-60
- [2] CSN - Chemnitzer Studentennetz
<https://www.csn.tu-chemnitz.de/>
- [3] Jan Horbach: „Diplomarbeit - Dynamische Bandbreitenbeschränkung mit QoS“, TU Chemnitz, Professur Rechnernetze und verteilte Systeme, 2001
<http://archiv.tu-chemnitz.de/pub/2001/0100/>
- [4] „[CSN] Beschränkungen des übertragenen Datenvolumens“, Chemnitzer Studentennetz, 2001
<https://www.csn.tu-chemnitz.de/info/limits.html>
- [5] GNU Licenses
<http://www.gnu.org/licenses/licenses.html>
- [6] Geoff Huston: „Internet Performance Survival Guide“, Wiley Computer Publishing, 2000, ISBN 0-471-37808-9
- [7] Geoff Huston: RFC 2990 „Next Steps for the IP QoS Architecture“, November 2000
<http://www.ietf.org/rfc/rfc2990.txt>
- [8] Authentication, Authorization and Accounting (aaa) IETF-WG
<http://www.ietf.org/html.charters/aaa-charter.html>
- [9] QBone Scavenger Service Design Team (QBSS DT)
<http://qbone.internet2.edu/qbss/>
- [10] Internet2 QoS Working Group
<http://www.internet2.edu/qos/wg/>
- [11] Resource Allocation Protocol (rap), 4.12.2001
<http://www.ietf.org/html.charters/rap-charter.html>
- [12] Quasar - Quality of Service Architecture
<http://www.ind.uni-stuttgart.de/Content/Quasar/>
- [13] Q-WiN - Class of Service Konzepte im Gigabit Wissenschaftsnetz
<http://rnvs.informatik.uni-leipzig.de/~q-win/>
- [14] Traffic Engineering for Quality of Service in the Internet, at Large Scale
<http://www.ist-tequila.org/>
- [15] TF-NGN
<http://www.dante.net/tf-ngn/>
- [16] SEQUIN
<http://www.dante.net/sequin/>