Diss. ETH No. 13750

A NETWORK ARCHITECTURE BASED ON MARKET PRINCIPLES

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZÜRICH

for the degree of
Doctor of Technical Sciences

presented by
GEORGE FANKHAUSER
gfa@acm.org
Dipl. Informatik-Ing. ETH
born 18th of May, 1967
citizen of Trub/Berne

accepted on the recommendation of
Prof. Dr. B. Plattner, examiner
Prof. Dr. L. W. McKnight, co-examiner

3rd October 2000

–verlagstitel–

–CIP eintrag–

TIK-SCHRIFTENREIHE NR. 37
Diss. ETH No. 13750

George Fankhauser

A Network Architecture Based on Market Principles

TIK-Schriftenreihe

Abstract

PROVIDING basic Internet services is not difficult. All a provider needs is a link to another Internet service provider to become a member of the "network of networks". However, making profit in this business has become very hard, due to the *continually evolving array of Internet services* that are demanded by users and offered by companies. New services like Internet Telephony or Media Streaming mark a shift away from traditional best-effort services (also known as "pure data transport"). Similar to this service evolution, *structural changes* have an impact on the provider market. A trend to increased connectivity among Internet service providers has been observed for several years and persists. The distributed structure of the network is constantly changing (due to market entries, exits, mergers and splits). In this *distributed market all providers are competitors and collaborators at the same time*. Competitors because they have the same potential customers (regional), and collaborators because they rely on their connectivity to provide service to destinations they cannot reach themselves. Even customers compete when they access the network and local demand exceeds supply. Pricing the Internet and its services is so hard because accurate information about processes in this huge, distributed market is not always available or often only as an approximation.

In response to this difficult situation we present a network architecture that *improves the flow of financial incentives* between customers and Internet service providers. It is targeted at today's and future Internet protocols promising to improve service quality. Its design starts at the backbone of the network and is fitted to the highly connected, and asymmetric structure of network topology and traffic. Providers interact with each other through *Service Level Agreement (SLA) Traders* that integrate service allocation, routing and pricing functions. Towards the less connected access networks, the flow of financial incentives must not be interrupted to keep this market working. Since SLAs are built on large aggregates of traffic flows, traders can operate them efficiently. In access networks, SLAs are interfaced to profiles and descriptions of micro flows that better fit the needs of single users.

Usage-sensitive pricing functions are used to measure demand which is then signaled back to providers.

The SLA trading architecture was implemented on several simulation platforms and on real network equipment. To capture structure and traffic of the Internet, statistical information and validated models were fed into our simulations. The integration of service allocation, routing and pricing was successfully demonstrated in this environment. Using SLA traders, a *considerable gain in network efficiency* was shown. The system works best in the center of the core networks, can cope with bursty traffic aggregates and exploits highly asymmetrical traffic loads (hot spots). For access networks, a method to price micro flows has been developed. We showed its basic functionality and efficiency for current traffic loads generated at access networks. Besides our combined approach to Internet pricing, no other proposal covers the full range from the edge down to the core of backbone providers. This thesis shows not only the technical feasibility of a market managed Internet but also its economic efficiency and viability.

Kurzfassung

ES ist nicht schwierig Internet-Dienstleistungen zur Verfügung zu stellen. Alles was ein Internet Service Provider (ISP) benötigt ist eine Netzwerkverbindung (Link) zu einem anderen ISP um ein Teil des globalen Internet zu werden. Der profitable Betrieb für ISPs ist jedoch sehr schwierig geworden. Das liegt vor allem an den neuen Anwendungen und Internet-Dienstleistungen, die von den Benutzern verlangt werden. Neue Dienste wie "Streaming Media" oder Internet-Telefonie unterscheiden sich durch ihre hohen Anforderungen stark vom traditionellen "best-effort" Dienst des Internet.

Ebenso gibt es eine Tendenz zur erhöhten Vernetzung unter ISPs. In diesem verteilten Markt sind alle ISPs gleichzeitig Konkurrenten und Kooperationspartner. Konkurrenten, weil sie die gleichen möglichen Kunden (regional) haben, und Kooperationspartner, weil sie auf Verbindungen via andere ISPs bauen, um Dienste zu den Destinationen zur Verfügung zu stellen, die sie selbst nicht erreichen können.

Auch Kunden konkurrieren, wenn die Nachfrage das lokale Angebot an Bandbreite übersteigt. Eine genaue Preisbestimmung für Internetdienste ist schwierig, weil genaue Informationen über Prozesse in diesem weltweiten Markt nicht immer oder häufig nur als Näherungswert verfügbar sind.

Daher schlagen wir eine Netzwerkarchitektur vor, die finanzielle Anreize zwischen Kunden und ISPs durchgängig weitergibt. Sie zielt auf die zukünftigen Internet-Protokolle, die die Service-Qualität im Internet verbessern können. Der Entwurf dieser Netzwerkarchitektur ist im "Backbone" des Internet verankert, wo eine hohe Konnektivität zwischen den ISPs besteht. Ebenso ist die Architektur auf asymmetrische Topologien und Verkehrsmuster ausgerichtet. Dabei kommunizieren ISPs untereinander via sogenannte *Service Level Agreement Trader*. Diese Trader integrieren die Ressourcenverteilung, Wegewahl und Preiskalkulation für jeden ISP. Um die Marktfunktionen global aufrecht zu erhalten, müssen auch Teilnetzwerke am Rand des Internet miteinbezogen werden. Hierbei wird ein Ansatz verfolgt, der ein-

zelne Datenströme von Benutzern (Profile) oder sogar einzelnen Anwendungen (Microflows) beschreibt und in Service Level Agreements (SLA) integriert.

SLA Trading wurde auf mehreren Simulations-Plattformen und auf realer Hardware implementiert. Um Struktur und Verkehrsmuster des Internets zu erfassen wurden bekannte und validierte Modelle in die Simulationen miteinbezogen. Die Integration der Dienstallokation, Wegewahl und Preiskalkulation wurde erfolgreich in dieser Umgebung demonstriert. Mit SLA Tradern wurde ein beträchtlicher Effizienzgewinn auf globalem Niveau nachgewiesen. Das System funktioniert besonders gut im Zentrum des Backbones wobei variable Verkehrsmuster und asymmetrische Lastsituationen geglättet und verteilt werden. Für ISPs am Rand des Internet, d.h. solche mit geringer Konnektivität, wurde eine Methode zur Preisbestimmung von "Microflows" entwickelt die auf Auktionen basiert. Es konnte gezeigt werden, dass diese Methode mit den typischen Lastverhältnissen eines ISP mithalten kann.

Der Ansatz von SLA Trading and der Preisbestimmung von "Microflows" kombiniert die Vorteile der Skalierbarkeit im Backbone und die genaue Informationserfassung am Rand des Netzwerkes. Mit der Evaluation in dieser Arbeit wurde gezeigt, *dass Netzwerke die auf Marktprinzipien aufbauen nicht nur technisch machbar, sondern auch ökonomisch überlebensfähig sind.*

Contents

Figures & Graphs

Tables

Chapter 1

Introduction

LARGE , inter-connected networks such as the Internet are operated by a multitude of service providers. Together, they offer a tremendous amount of communication bandwidth and computing power to move network traffic to its destination. These resources must be allocated to an even larger number of users connected to any of these service providers. This allocation problem has aspects of interest to both network engineering and economics. It is especially challenging and complex if we consider the structure of the Internet:

- Providers and users act in a non-cooperative way. They compete among and against each other. Users, on the one hand, want end-to-end communication and information services that are provided by many service providers. Providers, on the other hand, provide resources (links to other providers) that are shared by many users at the same time.

- The nature of the Internet is distributed. Allocating resources across several inter-connected networks involves delay and occupies part of the resources themselves.

- Although to some extent standardized, the Internet is a highly heterogenous system, offering different network characteristics in different areas. And users, of course, have different needs and behavior which influences the traffic characteristics they generate.

Economic allocation methods like auctions or game theory have been used in various fields. For example, stock exchanges use double auctions to match bids and asks of a vast amount of sellers and buyers; limited resources like radio spectrum are sold in public auctions to provide fair and efficient allocation. The application of economic principles to Internet resource allocation has been proposed for a decade now, but has not been successfully implemented so far.

In this thesis we propose a network architecture that can be applied to the Internet in spite of the challenges mentioned above. It is built mainly on these economic principles:

- *Competition among both customers and providers*: Competition ensures the opportunity to find users with the highest willingness to pay and the providers with the lowest prices.
- Matching of demand and supply: An equilibrium between the customers' and providers' objectives has to be found to *establish a market price*.
- There is always a flow of *financial incentives associated with provided value*. These incentives maybe imprecise or delayed but they must be communicated ("no free lunch principle").
- Resources may be *traded freely*. Trading parties are restricted to the areas they reside in, defined by the physical extent of the network. There is always an overhead cost associated with trading activity.

Ideally, these principles would work in near real-time, would be friction-free and *scale to any network size and traffic load*. However, we

Figure 1.1: *Colliding packets in the Internet of life.*

expect significant loss of the benefits of economic resource allocation induced by the limitations of the real-world system. Can we keep this loss small enough that we do not render the allocation scheme inefficient? How can we prevent that benefits are eaten up by operational overhead to drive the allocation scheme? Taking current network statistics into account, we will address this question and give directions where and how economic principles can be applied to the Internet's infrastructure.

Finally, we should keep in mind that resource allocation is a general topic that affects everybody's life. For example, we all must queue in front of counters or traffic lights, we compete for basic resources like food, energy, etc. Taking this position, this thesis concentrates on a rather narrow aspect, namely the allocation of communication and computing resources (Figure 1.1). [1]

1.1 Motivation

To motivate our work we give a few examples of allocation techniques used in today's Internet and argue why they lead to inefficiencies.

First, resource allocation for users is usually done by measuring

[1] Reprinted with permission by User Friendly Media, Inc.

peak usage. Access providers connecting individual users by dial-up lines have to over-provision their modem pools and networks. As an alternative, providers can go for less than peak allocation but they eventually end up with unsatisfied customers. Neither of these strategies seem to be very attractive. The fundamental problem here is the price plan: a constant price per time ("flat-rate") gives no feedback to users on how to use a service. In contrast, usage-based pricing provides such a feedback by limiting users to their budgets and willingness to pay.

By using different Internet services, users put a different load onto the network. Compare, for example, one user surfing the Web and another listening to music streamed in MP3-format. While the first user generates only a moderate average load, the latter is using significant bandwidth at a constant rate. Even worse, audio streaming uses the aggressive User Datagram Protocol (UDP) which preempts Web traffic that is TCP-based[2]. In this case, even usage-based pricing cannot solve the problem but a service-sensitive price could help. Today, TCP is still responsible for most of the traffic on the Internet. However, this is changing with more and better multimedia applications[3].

Besides multimedia applications that are popular with users, providers try to offer more services on the Internet by integrating traditional services like telephony. The main reasons behind this convergence trend[4] are smoother statistical multiplexing and less maintenance and personnel cost since there is only one network to care for. In Figure 1.2, a forecast is given for the convergence of voice and data network traffic [Tennenhouse, 1998]. The effect of having customers watch TV and phone via Internet is similar to the case of MP3-

[2] Transmission Control Protocol

[3] We should mention here that there is still a lot of media "streamed" and buffered via TCP. This is due to limitations of media players to cope with packet loss. Again, there is a trend to UDP streaming once these technical limitations are overcome.

[4] With digital studio formats like Digital Audio Broadcasting and High Definition Television and Internet streaming technology, we will also see more and more Radio and TV stations broadcasting via the Internet.

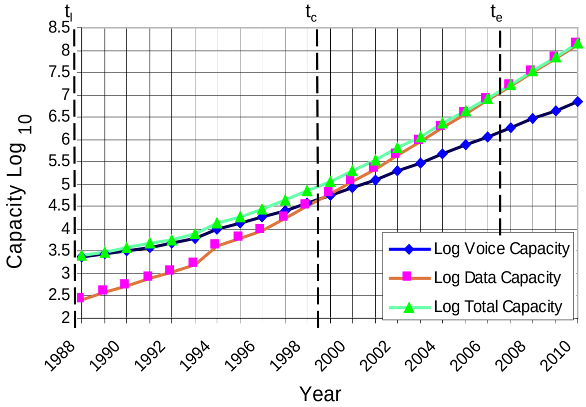

Figure 1.2: *Voice/Data convergence forecast.*

streaming as described above[5].

While these issues are largely unsolved, first attempts at providing flexible pricing and resource allocation, down to the minute timescale, are emerging. A recent example is given here:

"Carriers are now beginning to offer usage-based billing rather than flat-rate pricing. MCI WorldCom and WarpSpeed are among the ISPs offering usage-based data service. The WarpSpeed service can be turned on and off according to need, says WarpSpeed's Mil Ovan. Usage-based pricing will not only lower total telecom costs but will also significantly affect companies' business dealings, according to Ovan. [...] WarpSpeed offers T1 circuit usage for $1 per minute. The company charges an additional monthly fee of roughly $500 for a T1 circuit into WarpSpeed's local network. MCI WorldCom requires that users spend at least $500 on its private line service, which is offered at about one tenth the price of an ordinary T1 connection. Technical limitations had prevented service providers from provisioning bandwidth in a timely fashion. But now, service providers store customer information on a centralized database for auto-

[5]The main benefits of this convergence trend are low maintainance cost of fewer networks and tight integration/creation of new services.

mated provisioning and subscriber management." [McGarvey, 2000]

However, this offer is only attractive to medium to large companies with a high communication budget. Also, the service provides plain best-effort data service without any distinction whether the customer uses its connection mainly for voice, video, or just for Web browsing.

Internet services are perceived end-to-end but the provisioning process is established by inter-connecting through a number of providers. At these points of inter-connection, providers setup Service Level Agreements (SLAs) to allow traffic from peer providers to flow into their networks and vice versa. Similar to the end-users, providers often use flat-rate agreements between their networks. In more advanced configurations, a usage-based price on volume basis is used to settle accounts more accurately. Still, SLAs lack precise definition and are tedious to negotiate since everything is done manually. This inflexibility prevents providers to adapt SLAs to their needs or to seek alternative providers in times of long-term congestion.

Finally, the path that is taken by packets is only loosely bound to an optimization objective. Current practice for inter-domain path selection is to use BGP [Rekhter and Li, 1995], a path-vector routing protocol that is able to compute shortest paths. Since providers behave non-cooperatively at this level, pure shortest path selection is not an option. Instead, a policy database managed with an associated specification language is introduced [Alaettinoglu et al., 1999]. Through rules stored in that database, providers define whose traffic they like or dislike. For example, the Swiss academic provider prefers MCI to forward traffic to the US and to all commercial sites. Special rules describe additional connections to the European research networks. As we see in Figure 1.3, a trace from ETHZ to a commercial customer in Grenoble, Southern France (about 500 kilometers geographical distance!) takes a long detour across the Atlantic, wasting precious bottleneck-bandwidth and adding up much delay (this route was observed on Sunday, October 24^{th}, 1999).

Of course, this problem is easily solved by adding SLAs with French providers, presuming those providers have sufficient network capac-

```
pong:/tmp> traceroute -A -h whois.ripe.net salieri.imag.fr
traceroute to salieri.imag.fr (129.88.32.145): 1-30 hops, 38 byte packets
 1  ezwf7-tik-komsys.ethz.ch (129.132.66.1) [AS559 - ETH-NET]                           2.1 ms
 2  ezci1-fddi.ethz.ch (129.132.100.2) [AS559 - ETH-NET]                                2.3 ms
 3  swiez1-eth-switch-fast.ethz.ch (192.33.92.87) [AS559 - ETH-NETx]                    2.4 ms
 4  swiEG1-F5-1-0.switch.ch (130.59.20.211) [AS559 - SWITCH-LAN]                        2.5 ms
 5  swiNY1-A5-0-0-1.switch.ch (130.59.33.2) [AS559 - SWITCH-LAN]                      103.0 ms

 6  39.atm4-0-0.GW3.NYC4.ALTER.NET (157.130.13.61) [AS702 - UUNET-NET]                105.0 ms
 7  110.ATM3-0.XR1.NYC4.ALTER.NET (152.63.21.194) [AS702 - UUNET-NET]                 106.0 ms
 8  189.ATM11-0-0.BR1.NYC4.ALTER.NET (146.188.177.189) [AS702 - UUNET-NET]            106.0 ms
 9  137.39.23.146 (137.39.23.146) [AS702 - UUNET-NET]                                 106.0 ms

10  sl-bb10-nyc-2-3-155M.sprintlink.net (144.232.8.221) [AS311 SprintLink]            106.0 ms
11  sl-bb10-pen-6-1.sprintlink.net (144.232.9.102) [AS311 SprintLink]                 106.0 ms
12  sl-gw22-pen-0-0-0.sprintlink.net (144.232.5.22) [AS311 SprintLink]                109.0 ms
13  sl-ftpen-1-0-0-155M.sprintlink.net (144.228.179.18) [AS311 SprintLink]            109.0 ms

14  pastourelle1-backbone.opentransit.net (193.251.128.125) [AS5511 - OPENTRANSIT]  118.0 ms
15  bagnolet1-backbone.opentransit.net (193.251.128.45) [AS5511 - OPENTRANSIT]      118.0 ms
16  bagnolet2-backbone.opentransit.net (193.251.128.34) [AS5511 - OPENTRANSIT]      118.0 ms
17  bagnolet2-access.opentransit.net (193.251.128.114) [AS5511 - OPENTRANSIT]       119.0 ms

18  nio-i.cssi.renater.fr (193.51.206.33) [AS2200 - FR-RENATER2]                      119.0 ms
19  nio-n2.cssi.renater.fr (193.51.206.13) [AS2200 - FR-RENATER2]                     120.0 ms
20  grenoble.cssi.renater.fr (195.220.98.38) [AS2200 - FR-RENATER2]                   128.0 ms
21  CICG-grenoble.cssi.renater.fr (195.220.98.54) [AS2200 - FR-RENATER2]              129.0 ms

22  r-campus.grenet.fr (193.54.188.1) [AS1717 - FR-GRENET35]                          129.0 ms
23  r-imag.grenet.fr (193.54.185.123) [AS1717 - FR-GRENET32]                          130.0 ms
24  salieri.imag.fr (129.88.32.145) [AS1717 - RENATER]                                141.0 ms
```

Figure 1.3: *Example of a detour on the Internet.*

ity. However, the fundamental problem is the process: adding SLAs and changing policy database entries takes considerable time. Therefore, the current solution is only suitable for static routing. In order to be prepared to react to demand changes or to changing prices of SLAs, a more flexible approach is needed. In addition, the current solution does not distinguish different services; all traffic is routed along the same paths.

1.2 Scope

This thesis' scope are the *network and financial layers* as shown in Figure 1.4[6]. At the network layer, we look at the basic functionality, i.e. resource allocation and path selection which involves signaling, routing, queuing and scheduling. At the financial layer, we look at price determination, distributed trading and admission control algorithms that are based on pricing. Solutions developed at the financial layer are mapped to existing, extended and newly developed network protocols and mechanisms. The financial layer as an abstract overlay is recognized by the IETF. Developing and formalizing technical mechanisms to support the financial layer is one of the most significant contributions of this thesis.

This focus implies that aspects such as applications are beyond the scope of this thesis. This is also unnecessary since our network architecture supports basic and advanced Internet services and is therefore open to almost any Internet application.

It is equally important to define the scope of this work with respect to the *time scale* of events that control the network. Long time scale events include network planning activities and other structural changes. They happen with a per-month frequency. Such events are mostly based on human intervention and are processed over longer periods using statistical information from systems running at higher frequency (lower levels in Figure 1.5)[7]. Systems running at medium time scale are mainly inter-domain routing and traffic engineering which generate events on a weekly or up to an hourly rate. On shorter time scales (minutes to sub-second range) events are generated fully automatically and their dynamics become visible (e.g. closed-loop control dynamics of TCP). In this thesis we will work at the current

[6]Of course, the standard OSI protocol stack has only seven layers, but the financial and political layers can be seen at IETF meetings (usually on T-shirts)... More seriously, financial and political issues have significant impact on the networking industry and are often neglected in purely technical work.

[7]Manual operation is often supported by tools that run off-line (decision support systems, strategy planning tools, etc.).

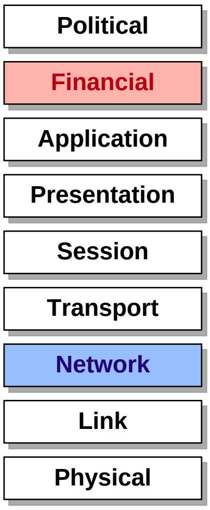

Figure 1.4: *Extended OSI protocol stack.*

boundary where automation ends and manual operation starts.

The proposed SLA trading architecture pushes the limit of automatic network control a little further (see Figure 1.5), but the ISP business will still remain turbulent and fast-paced. To quote the *ISP Survival Guide:*

> *"To put it mildly, cyberspace business is booming. There are presently more than 6'000 Internet Service Providers worldwide, and about 600 new providers are springing up each quarter. However, the ISP business is still very young and without precedent — no how-to manual or foolproof start-up recipe exists for those who want a piece of the action." [Huston, 1998]*

1.3 Objectives and Claims

From this introduction we see that there is a huge gap between the ideal, friction-free economic system and the Internet. The objective of

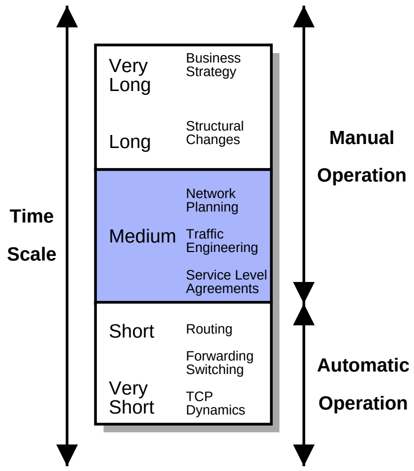

Figure 1.5: *Time scales of control events.*

this work is to design a *practical system*, built on existing and proposed standards, that extends the network's capabilities to run on economic principles. And the goal is, to design it in an efficient way to ensure its attractiveness to providers. In detail, the claims are:

- *Technical efficiency and sufficiently low complexity*: The system must be simple to implement using current software practices and run on contemporary hardware.
- *Economic efficiency and profitability*: The system should be able to move more traffic with better service quality through the networks than current systems are able to do. Or in other terms, the same load as today should be processed at a lower price. This improvement of efficiency must include all overhead cost for the market managed system.
- Large, distributed systems cannot be changed over night. We

must ensure *local deployment and testability* to make our architecture a realistic approach. As far as possible, compatibility to Internet standards should be maintained.

- A major goal of our approach is to *remove bottlenecks*. In the user's end-to-end view they can occur everywhere from sender to receiver. Therefore, the architecture must offer *high scalability* and must work in networks of different sizes and at different bandwidths.

1.4 Outline

Chapter 2 gives an *overview* of current and proposed Internet pricing methods applied to Internet transport services. The Internet's structure and some of its new services are discussed in the context of a market managed system. *Related work* in the area of Internet pricing with an emphasis on bandwidth brokers is reviewed in Chapter 3.

An overview of the *network architecture* is presented in Chapter 4. In Chapter 5 and Chapter 6 we discuss how economic principles are applied to the two most distinct areas of the Internet, the *backbone and access networks*.

The *implementation* of our architecture is described in Chapter 7. Chapter 8 reviews the performance of the systems and discusses results. Finally, our claims are reviewed and future work is considered in Chapter 9.

Chapter 2

The Economics of Advanced Internet Services

IN order to understand the issues in Internet pricing one must understand the services that are involved. This chapter gives an overview of the environment and economic principles applied in Internet pricing. New service concepts such as integrated and differentiated services are an important and integral part of this environment. Also, structural considerations affect the *decisions on where and what we should price*.

Chapter 9

Conclusion

THE network architecture for SLA trading and a prototype implementation described and analyzed in this thesis have shown that a *market-managed Internet* is not only technically feasible, but also economically viable. Most of the efficiency gained stems from the integration of pricing, routing and service allocation at the inter-domain level.

Such a tight integration is only possible among sufficiently large ISPs. Smaller providers are manly concerned with the management of end-users and access capacity. But by the selection of well specified SLAs, smaller ISPs at the edge of the Internet, connected only to a few other peers, also benefit from the SLA trading system since it enables the smaller ISPs to "see" farther into the network than just the next domain. This enables these ISPs to get crucial information about the service quality to any other destination domain.

9.1 Summary of Results

The results of this thesis form the *first integrated approach* to Internet pricing which spans a range of needs from end-users to ISPs. Previous results and proposals focused mostly on a single issue. In addition, the viability of the SLA trading approach has been demonstrated in a realistic and implementable environment that supports resource allocation techniques currently being standardized. Focusing on the routing aspect of SLA traders, we also presented and evaluated one of the very few *inter-domain QoS routing* approaches.

Global Optimization in a Non-cooperative Environment

To cope with and profit from the competitive situation between ISPs we proposed an innovative *SLA trading system* between large providers. It was shown to be a stable and self-regulating method for resource pricing and allocation in core networks. It is based on *diffserv* which is a simple yet powerful network resource provisioning framework. SLA trading provides a signaling framework for bilateral SLA negotiation between ISPs. It supports local optimization, incremental deployment, and evolving definitions of services. This is good news for providers since they can pick the mechanisms and policies they like best. And it is also good news for customers. The competition among providers will be perceptible at the edge of the network in form of lower prices and better service. The trading system encourages competition among large, well-connected ISPs. Each provider is free to choose its own strategies to pursue its local objective. For the prototype implementation, a simple, dynamic price model based on residual bandwidth was used.

While SLA trading follows the approach of a “pure market”, there may be a need for external regulation of the system (i.e. by national or international authorities). Basically, the SLA trading can be extended with additional rules. However, this renders the system more and more complex. In addition, there are regulatory constraints that are easier to implement than others: For example, regulating price levels for SLAs with any scope can by easily added to the system. But expressing any policy that is not directly related to prices is very difficult to apply.

In technical terms, SLA trading protocols can be implemented very efficiently. SLATP was defined as part of our prototype implementation. In addition, a new, dynamic approach to *diffserv* code point allocation was proposed. It simplifies and enhances *diffserv* PHBs by introducing the concept of "inter-AS label switching". Combined with an intra-domain label-switching method, e.g. MPLS, a hierarchical solution for label-switching is achieved.

An evaluation of SLA trading using recent, Internet-typical traffic and topology models was performed. Its main results are:

- Statistical information of inter-domain network topology suggested to apply SLA trading to the core of the network. Simulation results with up to 36 ASes support this claim.
- Non-uniform jumbo flow distributions (hot spots) are leveled out by SLA trading. For the Internet-typical range of ν-values we show a significant potential for load balancing (20% to 50% higher utilization than a statically configured network).
- Bursty traffic at different time scales hurts network utilization. Using our dynamic approach, SLAs are reduced in times of low demand and reallocated when needed. Although we showed small improvements at a short time scale, the SLA trading approach works best on the medium/long time scale.

These results have their significance on the inter-domain level and affect aggregated traffic. They provide a macroscopic view of a very complex system and, due to the nature of observation, they not only abstract but also change some of the details.

Discovering User Demand

At the edge of the Internet we introduced a pricing approach for micro flows. It serves to convert user demand directly to aggregated, financial stimuli that providers can use to ask for net-wide resources. The micro flow pricing approach is compatible with SLA traders (i.e. by conversion). In contrast to SLA traders that adjust prices themselves, users participate in auctions that determine the winners and therefore the resource allocation and pricing. Unlike classical auctions, the newly introduced Delta Auction is distributed in both time and space to accomodate for temporally scattered requests and multiple providers.

Technical overhead in terms of bandwidth, memory and processing is sufficiently low to support pricing of several ten thousand concurrent flows which is a realistic assumption when applied to access networks only.

Standards Compatibility and Best-effort Traffic Support

Although the Internet architecture for multiple services is not yet fully defined, our architecture for pricing is built on top of these early proposals. Most important, the architecture fits the existing, Internet-style concept of domains. Also, enhancements to these emerging Internet standards were presented. These add-ons do not challenge compatibility of the system but we cannot tell whether such enhancements are also endorsed by providers and equipment manufacturers.

What comes almost for free is the fact that our architecture co-exists with best-effort provisioning, routing at the inter-domain level (BGP) and pricing (including flat rate) for this "low value" service. Although we recommended to choose a usage-based pricing method over flat-rate for the access area, it is possible to integrate best-effort traffic into a service bundle. ISPs should separate their operations into a defined, SLA-based and a legacy, best-effort logical network (run by BGP). An interesting alternative we didn't explore is to fully rely on SLA traders at the inter-domain level by creating ad-hoc SLA requests based on measurements.

9.2 Review of Claims

Here, we review the four general claims we made in Section 1.3:

- Technical efficiency: This claim was shown by building the system and testing the efficiency of the technical components with respect to memory, processor and bandwidth usage. Most important, bandwidth usage, determined by message size and frequency, was sufficiently low for both SLA trading and micro flow reservation protocols with pricing capabilities.
- Economic efficiency and profitability: The fundamental question behind this thesis was whether a system can be built that has a *higher net efficiency on the global scale* than current solutions. Although the result is limited by the accuracy of the simulation model, significant gains indicate the viability of the architecture and methods.

  This positive balance can be expressed differently: If we used the network resources needed by SLA trading and access pricing for simple *overprovisioning* instead, we would gain less than by applying our market-based approach.
- Local deployment and testability: In technical terms, our network architecture can be built and operated on single routers and in single domains. However, there are *dependencies and quantitative requirements*. If a domain wants to become an SLA trader it must employ pricing methods for its access customers in any of the forms described in Section 4.3. For traders to become effective, a critical mass of market participants must be present.
- Scalability across many domains and end-to-end support: Scalability was achieved by splitting the network into core and edge domains. Expensive to solve problems like routing were only applied where appropriate and efficient. End-to-end support was achieved by tunneling existing solutions across the core networks.

9.3 Future Work

While our simulation results are very encouraging, future work should address *larger systems and more detailed network and traffic models*. Longer and larger network simulations with the goal to use about 80 class 1 and 2 ASes should be run.

Then, we should add better models of traffic distributions or actual flow measurements to the simulation process. Also, there is little known about the hot-spot distribution and its movements. More measurement and statistical analysis is needed. For example, the AS data set does not reflect multiple connections between ASes, a fact which is significant for networks spanning large geographical regions. In the core (10 to 20 best-connected ASes) many links between two ASes exist because global networks interconnect in different countries and continents. This is not directly reflected in the AS data set but could be extracted from BGP routing tables. Most of these *modeling issues* would need global, net-wide measurements.

SLA traders use *service profiles* based on bandwidth and delay metrics. An advanced service profile may also contain the burstiness of traffic. Also, SLAs are requested or offered immediately with a life time of bids. Extending SLAs with bid and ask requests that take place in the *future* could offer new possiblities for traders. However, it would also increase their complexity significantly.

There are a number of *implementation issues*. A first implementation of SLA trading on one of the upcoming *diffserv* implementations would provide further insights. In our current simulations, interfaces to local provisioning interfaces were mostly ignored. In addition, only a real-world implementation (or a node-level simulation) can show how communication and synchronization between SLA trader and boundary nodes work in detail. A real implementation in a test-bed could make use of advanced deployment mechanisms, such as Active Networks. In addition, a practical way to use SLAs in an extensible fashion were to implement and maintain SLATP in XML (Extended Markup Language). However, we should remember that a core part of an SLA must be standardized. Even if we can implement such protocol changes very quickly, an Internet Standards organization, such as ICANN, should be involved to make new features available to every ISP.

A comprehensive *security architecture* needs to be included to authenticate payments and to prevent fraud. Although discussed results already include basic authentication, further work is planned in cooperative security and low overhead authentication that is adequate for the small payments being transferred. This concerns mainly the business between customer and ISP. Also, many contracts will cross national borders and may become legal problems. However, as with current operators' contracts legal frameworks may cover many specific contracts that are changed at the rate required by SLA traders.

And finally, the customer's objective matters most in any business. Sometimes, it's not about *how fast* you get information, but about *what* you get from the global information infrastructure. Advanced and powerful communications infrastructures are just enablers of information retrieval and delivery (Figure 9.1). [1]

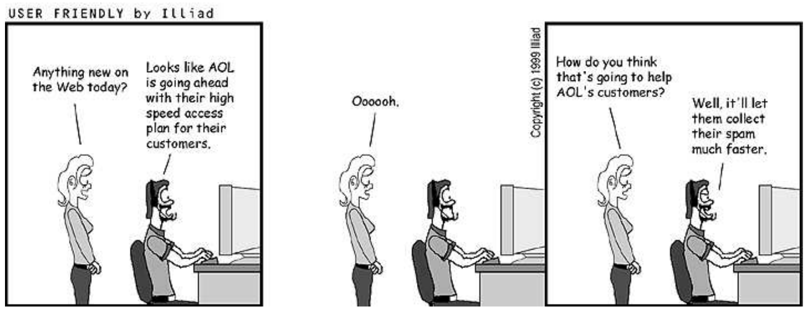

Figure 9.1: *It's not about how fast you get on the Net ...*

[1]Reprinted with permission by User Friendly Media, Inc.

Appendix A

Standardization and Deployment

A.1 SLA Trading

A flexible and dynamic market system doesn't imply a regulation-free system. It rather means "the more market the more rules."[1]

Deploying SLA trading in the real world needs a minimum set of rules to which all market participants must comply. This section is a summary of the elements used in SLA trading and gives an overview of standards needed with respect to required protocols and market rules.

In Table A.1 the rules are given that form the minimal SLA trading system. They are mandatory for an efficient and properly working sys-

[1]This quote is often read in the economic press. However, it is not as accurate as one could wish and should read "the *more complex* the market the more rules".

Table A.1: *Market rules in SLA trading.*

Non-zero, positive price	Since pricing is an integral part of the routing metric, this rule prevents service loops. Depending on the systems implementation, very small value should be avoided too.
Double-signed contracts	Cryptographically signed contracts are used for conflict resolution. Based on asymmetrical cryptography, such a system requires a PKI (public key infrastructure).
Minimum SLA ask message	Ask messages must include at least an AS destination, service quality information and the expiration of the request.
Minimum SLA bid message	Bid messages include a price, *diffserv* numbering information, duration of the offered service and expiration of the bid itself.
Use of SLATP equivalent states	Although any protocol may be selected by two ISPs the minimum SLATP states and transitions must be used (see figure 7.3 for a formal description).

tem. If some of these market rules are bent by single ISPs, they are forced out of business. For example, if an ISP generates a loop it is he who has to pay the bill for it. Or, if another ISP refuses to sign contracts he won't be able to continue a business with other ISPs that adhere to the standard practice of signing contracts. If a significantly sized group of ISPs changes these rules, however, the system will render itself less efficient.

Table A.2 lists the related technical standards and requirements. They support the respective market rules. In certain cases, e.g. global time synchronization, Internet standards exist that are ready to use in this context. Other properties, such as security, will define how contracts are double signed. In this case, an existing Certification Authority (CA) is used or an accepted standards body (e.g. ICANN) becomes a CA for all ISPs.

Table A.2: *Technical standards in SLA trading.*

Property	Scope	Reason
Time	Global	Since end-to-end paths can take any length and span any geographical extent, SLAs with start time and duration, as well as bids with a validity time, must have the same time basis among all ISPs participating in an and-to-end service.[a]
Currency	Global; exchanges allowed	Although currency maybe exchanged between neighbouring ISPs, their rates must be well known.
AS numbers	Global	Paths, destinations must be unique; ASes announce network prefixes to maintain reachability of single IP hosts.
Bandwidth, delay and other service quality metrics	Global; exchanges allowed	End-to-end services are based on a single metric that is given by the customer. As long as an exchange exists, ISPs may switch to equivalent metrics.
SLA trading protocol	Local between two ISPs	A common container, or transport mechanism is needed to exchange SLA messages. Although the message's content has global significance, the protocol used may be selected independently by every pair of ISPs. However, a single standard protocol could help save some development overhead[b]
Security	Local	Security standards must be used to ensure proper authentication and key exchange procedures. Minimum strength of the algorithms, according to current crypto analysis, must be employed. A widely used key and certificate format should be used (e.g. X.509v3, pgp, ...).

[a]The standard Internet time protocol NTP is a good solution (RFC 1305).
[b]Remote object technology or active networks may help here too.

A.2 Early Adopters of SLA Trading

One of the interesting practical questions of this thesis is: "Who should adopt the SLA trading approach?". First we give a listing of the properties that such early adopters should have, then, potential candidates are described.

1. ISPs with better than average connectivity ($d_v > 10$) and willing to compete with others.
2. ISPs offering service quality at the network layer (other than traditional best-effort IP service).
3. ASPs (Application Service Providers) that require service quality to run advanced applications on their networks (mostly real-time services). This includes smaller ISPs depending on other, larger ISPs to provide such advanced services.

A first category of ISPs is characterized by its exceptional connectivity. This includes larger as well as smaller ISPs, with a lower bound on traffic that justifies investing in the required lines needed. Such ISPs may have significant transit traffic that has to be exported to other providers (*export-oriented* SLA trader). Being connected to many other ISPs offers a greater choice in the local "mini-market". Other ISPs in this category may specialize in offering certain paths (e.g. inter-continental) at competitive prices (*import-oriented* SLA trader). In general, all backbone ISPs are candidates in this category since none of them has really global reach.

The second category concerns large providers that want to make a business of providing guaranteed service to customers with needs for high availability and high quality. As soon as customers need these services where the provider is not physically present, SLA trading can help to extend the ISPs range (geographically) by interconnecting with other ISPs in these regions with the same business background. Using SLA trading's automatic updates for demand and price, participating providers do not have to pay for peak allocation when demand is low. Examples of this type of ISPs are VPN providers with global customers.

Finally, ASPs may want to focus on their core business and need providers that offer the underlying network service. SLA trading is a solution for such businesses that allows for extremely short-termed contracts and therefore a lower risk and lower entrance barrier. Prominent examples are IP telephony operators and other interactive streaming media providers (basically all digital broadcasters).

A.2.1 Examples

Internet2

Internet2 (I2) is an organization formed by over 170 US-based universities and some industrial partners[2]. Its goal is to provide a next generation IP infrastructure to its members. Besides advanced applications and services, the focus is set on the advanced network infrastructure and new networking capabilities.

Part of this strategy is the *Internet2 QBone* which supports mainly Quality of Service and multicast in IPv4 as well as IPv6 mode[3]. Currently, *diffserv* based services and the notion of managed QBone domains are introduced. The architecture expects such domains to operate bandwidth brokers that negotiate SLAs on the basis of service level specifications. SLA traders follow exactly the same approach but add the aspect of pricing to the negotiation process. Besides increasing the efficiency, pricing based on SLA traders adds the possibility to charge each partner of I2 according to usage of high-end services on the QBone. The importance of such an approach has been recognized by the consortium and work has been initiated to close the gap between economic research and technical implementation[4].

[2] `http://www.internet2.edu`
[3] `http://qbone.internet2.edu/arch`
[4] The *Internet2 Economics Working Group* has been recently formed.

SLA Trading on CIX

The current model of the Commercial Internet Exchange (CIX) is simple: Providers connect at a single point (in Palo Alto, California) and exchange each other's traffic without condition and at no cost. This model represents the approach of complete peering. A small annual fee (about $10k) has to be paid for operating the CIX router (CIX itself is a non-profit organization)[5].

The basic idea of CIX is to increase connectivity and shorten AS paths by connecting most ISPs directly to each other. The CIX router, with its high connectivity (and all connected ISPs), resembles our mini-markets. However, CIX is still a single, centralized exchange point. Also, the lack of settlements prevents a fair and usage-based charging of services.

While the basic idea of exchange points is valid it has been only a moderate success so far. Combining CIX points with SLA traders, however, could change many aspects[6]. In the following, we list the changes and their effects of SLA enhanced CIX points:

- Having a rule that forbids payments between ISPs has the advantage of yielding a simple system but it limits membership to ISPs that "feel" to be about equal or think they profit from the system (i.e., ISPs exporting more transit traffic than importing traffic). Adding SLAs and a price for settlement opens up a CIX point to any ISP that is interested in buying or selling any service.

- Each exchange point adds connectivity at a lower price compared to direct connection between each pair of ISPs (this pairing was used as a basic model in this work). Organizational overhead for running the exchange point infrastructure is not large

[5] `http://www.cix.net/cix/about-cix.html`

[6] The "no settlement"rule in CIX would have to be changed first, of course. Whether such a policy change would be embraced by CIX directors and members, cannot be discussed here.

since simple, full connectivity is configured. Each connected ISP builds for every connected peer a virtual network interface (on the same physical trunk) to make this topology compatible with SLA trading.

- Today, a single CIX point exists. SLA trading, however, works best in a decentralized fashion. Therefore, adding more exchange points is a must to introduce SLA trading on CIX-style networks[7].

- Finally, today's CIX model forces all members to peer with each other (all or nothing situation). Using SLA trading, a choice to select partners would be added.

[7] Note that a trial would be still possible on the existing exchange point.

Appendix B

AS by Connectivity

In the following table we give the best-connected AS numbers, out-degree and network names that were used in the evaluation of this work. These ASes correspond roughly to the class 1 and 2 networks [Fang and Peterson, 1999]. Data is from [moat, 1999] and names were translated by `whois.radb.net`.

701	1186	Alternet
3561	692	Cable & Wireless (CW)
1239	573	SprintLink Backbone
1	301	GTE Internetworking
7018	271	AT&T WorldNet Service Backbone
2548	248	DIGEX-AS
2914	244	---
6453	149	Teleglobe Canada Inc.
293	138	ESnet
6347	133	Savvis - St.Louis Region
2828	131	Concentric Network Corporation
1740	127	CERFnet
702	115	UUNET
2497	109	IIJNET
1755	101	EBONE AS
721	97	---
5696	92	GoodNet AS
209	88	Qwest Communications
3549	87	Frontier GlobalCenter AS3549
145	83	---

286	83	EUnet Backbone AS
6461	76	Primary AS for Abovenet
1673	71	---
3300	69	AUCS Communications Services v.o.f.
4200	68	AGIS (Apex Global Information Services, Inc.)
1221	67	TELSTRA-AS
5459	66	LINX-AS
5646	66	NAP.NET, LLC (ASN-NN-CNSM)
3257	61	Nacamar Global ASN
4000	60	Global IP
174	60	Performance Systems International, Inc
7474	56	Optus Communications
5650	54	Electric Lightwave Inc, 800-622-4354
1849	51	UUNET UK (formerly PIPEX, Public IP EXchange)
2516	50	Kokusai Denshin Denwa Co., Ltd.(KDD)
11042	48	---
3333	46	RIPE NCC
4637	45	Hong Kong Telecom
3356	45	GEONET
2041	44	CRL-GATE
3967	44	Exodus Communications
4926	43	---
8342	43	Rostelecom Internet Center’s Autonomous System
4230	42	Embratel Brazil
3951	42	ICONNET
1225	42	VERIO OHIO AS
7170	41	DISC/DREN
2764	40	connect.com.au pty ltd
6172	39	HOME-NET-1
8918	38	Carrier1 Autonomous System
3303	37	Swisscom
5006	36	Minnesota Internet eXchange / Regional Network
6467	35	American Communications Services Inc
816	34	UUNET Canada (ASN-UUNETCA-AS4)
1273	34	Cable & Wireless ECRC GmbH
5683	34	AT&T Global Network Services - Americas
2685	34	Epoch Internet - Formerly HLC Networks
4565	33	Paradox Digital Pty. Ltd
7586	33	Netaxs AS
4969	32	ICM-Atlantic
1800	32	CAIS Internet
3491	32	DACOM Corporation in Korea
3786	31	Sprint Canada Inc.
3602	31	UUNET Germany
1270	31	INSnet-EUROPE-AS
5378	30	TELECOM ITALIA
6664	30	BTnet UK Regional network
2856	30	Teleglobe Autonomous System in Europe
8297	30	Concert Internet Plus European core Network
5400	29	TeliaNet Europe
1299	29	Primary AS for LinkAGE Online Ltd.
4058	29	ICG PST, Inc.
2551	29	Verio / New-York.Net
3384	29	Bell Canada Backbone
577	28	Bell Nexxia Backbone
3320	28	Deutsche Telekom AG
297	28	NASA Internet
3215	27	RAIN
7037	27	Networking AS for GST Internet, Inc.
4694	27	Cable & Wireless IDC Inc.
1717	26	RENATER
8933	26	The TEN-155 IP Service
6503	26	AVANTEL
6765	26	Internet Network Services
2905	25	UUNET Internet Africa
6755	25	ASN - TURNET

Acknowledgements

I WOULD like to thank Bernhard Plattner for his continuous support during my time as both a research assistant and PhD student. I really enjoyed the freedom to pursue the idea of this thesis and many other research topics. I'm greatly indebted to Lee McKnight for taking the job as a co-examiner, for his helpful reviews and for pointing me to important non-technical issues.

Thanks to those people who reviewed all or part of this work: Marcus Brunner, Christina Class, Marcel Dasen, Ulrich Fiedler, Polly Huang, Peter Marbach, Lukas Ruf, Rolf Stadler, Thomas Walter and Nathalie Weiler. Special thanks to David Schweikert who contributed significantly to the SLA trading concept and was just a great office mate[1]. Thanks to Brad Cox for inspiration on trading of non-tangible goods and to Tony Przygienda for pointing me to inter-domain QoS routing issues.

I would like to thank all of my colleagues at the ETH Zürich Networks Lab, including the many that already left, for the inspiring, incoffeinated discussions and "joint ventures" that resulted in lots of good ideas and in quite a few pieces of useful software.

Kudos to the many brave students who wrote a diploma thesis and/or a semester project under my supervision. The good work aside, I believe they all had a lot of fun and learned the one thing or the other. And so did I!

Thanks to Illiad and the fine folks at *Columbia Internet*[2] for the permission to use some of his great comic strips (`http://www.userfriendly.org`).

[1] I should also mention that it was he who convinced me to install Debian GNU/Linux instead of Redhat `;-)`

[2] "the friendliest, hardest-working, and most neurotic little ISP in the world"

Biography

I WAS born in Lucerne, Switzerland on May 18th 1967 where I also attended primary school and college (Matura Typ C). After some working experience in electrical and computer engineering I started my studies in computer science at ETHZ where I received the diploma in 1995.

After working for several companies and having founded a start-up in software engineering I joined the Networks Laboratory at ETHZ to work on research projects in the fields of joint source/channel coding of video for wireless transmission, next generation protocols for mobile Internet services, operating systems and pricing of Internet service providers. In 1998, I started my work on the SLA Trading method.

I joined another start-up company after my PhD studies to work on wireless/mobile networking and advanced embedded operating system software.

Bibliography

[Alaettinoglu et al., 1999] ALAETTINOGLU, C. ET AL. Routing Policy Specification Language (RPSL). RFC 2622. June 1999.

[Anania and Solomon, 1997] ANANIA, L. AND SOLOMON, R. J. Flat–The Minimalist Price. Internet Economics, MIT Press, Cambridge, Massachusetts. 1997.

[Apostolopoulos et al., 1999] APOSTOLOPOULOS, G., GUERIN, R., KAMAT, S., A.ORDA, PRZYGIENDA, T., AND WILLIAMS, D. QoS Routing Mechanisms and OSPF Extensions. RFC 2676 (Experimental(, Internet Engineering Task Force. August 1999.

[Apostolopoulos et al., 1999] APOSTOLOPOULOS, G., GUERIN, R., KAMAT, S., ORDA, A., , AND TRIPATHI, S. Intra-Domain QoS Routing in IP Networks: A Feasibility and Cost/Benefit Analysis. *IEEE Network Magazine*. September 1999.

[Bernet, 1999] BERNET, Y. Usage and Format of the DCLASS Object With RSVP Signaling, draft-bernet-dclass-01.txt. June 1999.

[Bernet et al., 1998] BERNET, Y., DURHAM, D., AND REICHMEYER, F. Requirements of Diffserv Boundary Routers, draft-bernet-diffedge-01.txt. November 1998.

[Blake et al., 1998] BLAKE, S., BLACK, D., CARLSON, M., DAVIES, E., WANG, Z., AND WEISS, W. An Architecture for Differentiated Services. RFC 2475. Dec. 1998.

[Braden et al., 1993] BRADEN, R., CLARK, D., AND SHENKER, S. RSVP: A new resource reservation protocol. *IEEE Network*. September 1993.

[Braden et al., 1994] BRADEN, R., CLARK, D. D., AND SHENKER, S. Integrated Services in the Internet Architecture: An Overview. RFC 1633. June 1994.

[Braden et al., 1997] BRADEN, R., ZHANG, L., BERSON, S., HERZOG, S., AND JAMIN, S. Resource Reservation Protocol (RSVP) - Version 1 Functional Specification. RFC 2205. September 1997.

[Breslau and Shenker, 1998] BRESLAU, L. AND SHENKER, S. Best-Effort versus Reservations: A Simple Comparative Analysis. In *ACM SIGCOMM 98*. September 1998.

[Briscoe, 1999] BRISCOE, B. The Direction of Value Flow in Multi-Service Connectionless Networks. BT Labs Technical Report. September 1999.

[Brownlee, 1997] BROWNLEE, N. Traffic Flow Measurement: Meter MIB. RFC 2064, Internet Engineering Task Force. January 1997.

[Castineyra et al., 1996] CASTINEYRA, I., CHIAPPA, N., AND STEENSTRUP, M. The Nimrod Routing Architecture. RFC 1992. August 1996.

[Cheliotis, 1999] CHELIOTIS, G. A Market-based Model of Bandwidth and a New Approach to End-to-End Path Computation with Quality Guarantees. In *MIT Workshop on Internet Service Quality Economics*. Dec. 1999.

[Chen and Nahrstedt, 1998] CHEN, S. AND NAHRSTEDT, K. An Overview of Quality of Service Routing for Next-Generation High-Speed Networks: Problems and Solutions. *IEEE Network*. November/December 1998.

[Chu, 1999] CHU, K. User Reactions to Flat Rate Options under Time Charges with Differentiated Quality of Access: Preliminary Results from INDEX. In *MIT Workshop on Internet Service Quality Economics*. Dec. 1999.

[Clark, 1996] CLARK, D. Combining Sender and Receiver Payments in the Internet. Telecommunications Research Policy Conference. October 1996.

[Clark and Fang, 1998] CLARK, D. D. AND FANG, W. Explicit Allocation of Best-Effort Packet Delivery Service. *IEEE/ACM Transactions on Networking*, Vol. 6, No. 4, pp. 362–373. August 1998.

[Cocchi et al., 1993] COCCHI, R., SHENKER, S., ESTRIN, D., AND ZHANG, L. Pricing in Computer Networks: Motivation, Formulation, and Example. *IEEE/ACM Transactions on Networking*, Vol. 1, No. 6, pp. 614 – 627. December 1993.

[Courcoubetis and Siris, 1999] COURCOUBETIS, C. AND SIRIS, V. Managing and Pricing Service Level Agreements for Differentiated Services. In *IEEE/IFIP IWQoS 99*. May 1999.

[Crawley et al., 1998] CRAWLEY, E., NAIR, R., RAJAGOPALAN, B., AND SANDICK, H. A Framework for QoS-based Routing in the Internet. RFC 2386. August 1998.

[Danielson and Molka-Danielson, 1999] DANIELSON, K. AND MOLKA-DANIELSON, J. Admission and Preemption in Auction-based Reservation Networks. MIT Workshop on Internet Service Quality Economics. December 1999.

[Decasper et al., 1998] DECASPER, D., DITTIA, Z., PARULKAR, G., AND PLATTNER, B. Router plugins: A modular and extensible software framework for modern high performance integrated services routers. In *Proceedings of ACM SIGCOMM '98*. September 1998.

[Durham et al., 2000] DURHAM, D. ET AL. The COPS (Common Open Policy Service) Protocol. RFC 2748. January 2000.

[Economist, 2000] How to Be Perfect. *The Economist*, Vol. 354, No. 8157, pp. 82. February 2000.

[Edell et al., 1995] EDELL, R. J., MCKEOWN, N., AND VARIYA, P. P. Billing Users and Pricing for TCP. *IEEE Journal on Selected Areas in Communications*, Vol. 13, No. 7, pp. 1162 – 1175. July 1995.

[Faloutsos et al., 1999] FALOUTSOS, M., FALOUTSOS, P., AND FALOUTSOS, C. On Power-Law Relationships of the Internet Topology. In *ACM SIGCOMM 99*. Sept. 1999.

[Fang and Peterson, 1999] FANG, W. AND PETERSON, L. Inter-AS Traffic Patterns and Their Implications. In *Global Internet 99, Rio de Janeiro, Brazil*. December 1999.

[Fankhauser et al., 1999] FANKHAUSER, G., DASEN, M., WEILER, N., PLATTNER, B., AND STILLER, B. WaveVideo – An Integrated Approach to Adaptive Wireless Video. *ACM Monet, Special Issue on Adaptive Mobile Networking and Computing*, Vol. 4, No. 4, pp. 255–271. November 1999.

[Fankhauser and Plattner, 1999] FANKHAUSER, G. AND PLATTNER, B. Bandwidth Brokers as Mini-Markets. In *MIT Workshop on Internet Service Quality Economics*. Dec. 1999.

[Fankhauser et al., 1998] FANKHAUSER, G., STILLER, B., VÖGTLI, C., AND PLATTNER, B. Reservation-based Charging in an Integrated Services Network. 4th INFORMS TCOM Conference, Boca Raton. March 1998.

[Feldmann et al., 2000] FELDMANN, A., GREENBERG, A., LUND, C., REINGOLD, N., AND REXFORD, J. NetScope: Traffic Engineering for IP Networks. *IEEE Network Magazine*. March/April 2000.

[Foser, 1999] FOSER, M. Pricing und Routenselektion für RSVP Flows in einer Simulationsumgebung, Semester Thesis, Computer Engineering and Networks Lab, ETH Zürich. February 1999.

[Gupta et al., 1994] GUPTA, A., STAHL, D. O., AND WHINSTON, A. B. Managing the Internet as an Economic System. Technical Report CISM, University of Texas at Austin. July 1994.

[Heinanen et al., 1999] HEINANEN, J., BAKER, F., WEISS, W., AND WROCLAWSKI, J. Assured Forwarding PHB Group. RFC 2597. June 1999.

[Herzog et al., 1997] HERZOG, S., SHENKER, S., AND ESTRIN, D. Sharing the "Cost"of Multicast Trees: An Axiomatic Analysis. *IEEE/ACM Transactions on Networking*, Vol. 5, No. 6, pp. 847–860. December 1997.

[Huston, 1998] HUSTON, G. *ISP Survival Guide*. John Wiley & Sons. 1998.

[IETF tewg, 1999] IETF Traffic Engineering Working Group Charter. `http://www.ietf.org/html.charters/tewg-charter.html`. November 1999.

[Jacobson et al., 1999] JACOBSON, V., NICHOLS, K., AND PODURI, K. An Expedited Forwarding PHB. RFC 2598. June 1999.

[Karsten et al., 1998] KARSTEN, M., SCHMITT, J., WOLF, L., AND STEINMETZ, R. An Embedded Charging Approach for RSVP. In *Sixth International Workshop on Qality of Service (IWQoS'98), Napa, CA, USA*. May 1998.

[Korilis and Orda, 1999] KORILIS, Y. AND ORDA, A. Incentive Compatible Pricing Strategies for QoS Routing. In *INFOCOM 99*. March 1999.

[Labovitz et al., 1997] LABOVITZ, C., MALAN, G. R., AND JAHANIAN, F. Internet routing instability. In *Proceedings of ACM SIGCOMM '97*. 1997.

[Lazar and Semret, 1997] LAZAR, A. A. AND SEMRET, N. Auctions for Network Resource Sharing. Technical Report 468-97-02, CTR, Columbia University. February 1997.

[Li and Rekhter, 1998] LI, T. AND REKHTER, Y. A Provider Architecture for Differentiated Services and Traffic Engineering (PASTE). RFC 2430. October 1998.

[Löffler et al., 1997] LÖFFLER, S., CHRIST, P., AND LORANG, M. Web-based Internet Traffic Analysis Using Flows. 6th Workshop on High Speed Networks (HSN '97). October 1997.

[MacKie-Mason and Varian, 1994a] MACKIE-MASON, J. AND VARIAN, H. Generalized Vickrey Auctions. July 1994.

[MacKie-Mason and Varian, 1994b] MACKIE-MASON, J. AND VARIAN, H. Some FAQs about Usage-Based Pricing; University of Michigan. August 1994.

[MacKie-Mason and Varian, 1995] MACKIE-MASON, J. AND VARIAN, H. Pricing Congestible Network Resources. *IEEE Journal on Selected Areas in Communications*, Vol. 13, No. 7, pp. 1141 – 1149. July 1995.

[MacKie-Mason, 1997] MACKIE-MASON, J. K. A Smart Market for Resource Reservation in a Multiple Quality of Service Information Network. Technical Report, University of Michigan. September 1997.

[Maclachlan, 1998] MACLACHLAN, M. AOL: Users Hate It, Wall Street Loves It, CMP Techweb News. February 1998.

[Makris, 1999] MAKRIS, J. Bandwidth Brokers, Not Exactly Nasdaq. Data Communications Magazine, `http://www.data.com/issue/990507/brokers.html`. May 1999.

[McGarvey, 2000] MCGARVEY, J. The Debut of Pay-Per-Use Data Services. *Interactive Week*, Vol. 7, No. 4, pp. 16. January 2000.

[McKnight and Bailey, 1997] MCKNIGHT, L. AND BAILEY, J. Internet Economics. MIT Press, Cambridge, Massachusetts. 1997.

[McKnight and Leida, 1998] MCKNIGHT, L. W. AND LEIDA, B. A. Internet Telephony: Costs, Pricing, and Policy. In *Telecommunications Policy*, Volume 22, pp. 555–569. August 1998.

[moat, 1999] Global ISP Interconnectivity by AS number. `http://moat.nlanr.net/AS/background.html`. June 1999.

[Neilson et al., 1999] NEILSON, R., WHEELER, J., REICHMEYER, F., AND HARES, S. Eds. *A Discussion of Bandwidth Broker Requirements for Internet2 Qbone Deployment*. Internet2 Qbone Bandwidth Broker Advisory Council: `http://www.merit.edu/working.groups/i2-qbone-bb`. Work in progress, August 1999.

[Nichols et al., 1998] NICHOLS, K., BLAKE, S., BAKER, F., AND BLACK, D. Defintion of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers. RFC 2474. December 1998.

[Nichols et al., 1997] NICHOLS, K., JACOBSON, V., AND ZHANG, L. A two-bit differentiated services architecture for the internet. `http://www-nrg.ee.lbl.gov/papers/bitarch.pdf`. November 1997.

[ns-2, 1998] Network Simulator 2, Version 2.1 beta 4. `http://www-mash.cs.berkeley.edu/ns`. 1998.

[nw, 1999] DNS-based Host Count Survey. `http://www.nw.com`. July 1999.

[Pan et al., 1999] PAN, P., HAHNE, E., AND SCHULZRINNE, H. BGRP: A Tree-based Aggregation Protocol for Inter-Domain Reservations. Columbia University CS TR No. CUCS-029-99. December 1999.

[Paxson, 1997a] PAXSON, V. Fast, Approximate Synthesis of Fractional Gaussian Noise for Generating Self-Similar Network Traffic. *ACM Computer Communication Review*, Vol. 27, No. 5, pp. 5–18. October 1997.

[Paxson, 1997b] PAXSON, V. Measurements and Analysis of End-to-End Internet Dynamics. PhD Thesis, University of California, Berkeley. April 1997.

[Postel, 1981] POSTEL, J. Internet Protocol. RFC 791, Internet Engineering Task Force. September 1981.

[Reichl, 1999] REICHL, P. Dynamische Verkehrs- und Preismodellierung für den Einsatz in Kommunikationssystemen. Dissertation, Technische Hochschule Aachen. December 1999.

[Rekhter and Li, 1995] REKHTER, Y. AND LI, T. A Border Gateway Protocol 4 (BGP-4). RFC 1771. March 1995.

[Rigney, 1997] RIGNEY, C. RADIUS Accounting. RFC 2139, Internet Engineering Task Force. April 1997.

[Rigney et al., 1997] RIGNEY, C., RUBENS, A., SIMPSON, W., AND WILLENS, S. Remote Authentication Dial In User Service (RADIUS). RFC 2138, Internet Engineering Task Force. April 1997.

[Ruutu and Kilkki, 1998] RUUTU, J. AND KILKKI, K. Simple Integrated Media Access (SIMA) with TCP. 4th INFORMS TCOM Conference, Boca Raton. March 1998.

[Savage et al., 1998] SAVAGE, S. ET AL. Detour: a Case for Informed Internet Routing and Transport. University of Washington, Seattle, Technical Report UW-CSE-98-10-05. October 1998.

[Schelen et al., 1999] SCHELEN, O., NILSSON, A., NORRGARD, J., AND PINK, S. Performance of QoS Agents for Provisioning Network Resources. In *IEEE/IFIP IWQoS 99*. May 1999.

[Schweikert and Fankhauser, 1999] SCHWEIKERT, D. AND FANKHAUSER, G. *flowsim*, A Flow-based Network Simulator. `http://www.tik.ee.ethz.ch/~fs2`. 1999.

[Schweikert et al., 1999] SCHWEIKERT, D., FANKHAUSER, G., AND WEILER, N. Elegant Reservation Protocol Suite, `http://www.tik.ee.ethz.ch/~erps`. 1999.

[Semret et al., 2000] SEMRET, N., LIAO, R., CAMPBELL, A., AND LAZAR, A. Peering and Provisioning of Differentiated Internet Services. In *Proceedings of INFOCOM 2000*. March 2000.

[Shaikh et al., 1999] SHAIKH, A., REXFORD, J., AND SHIN, K. G. Load-Sensitive Routing of Long-Lived IP Flows. In *ACM SIGCOMM 99*. September 1999.

[Shenker, 1995] SHENKER, S. Some Fundamental Design Decisions for the Future Internet. *IEEE Journal on Selected Areas in Communications*, Vol. 13, No. 7, pp. 1176–1188. July 1995.

[Shenker et al., 1996] SHENKER, S., CLARK, D., ESTRIN, D., AND HERZOG, S. Pricing in Computer Networks: Reshaping the Research Agenda. *ACM Computer Communication Review*, Vol. 26, No. 2, pp. 19–43. April 1996.

[Shreedar and Varghese, 1995] SHREEDAR, M. AND VARGHESE, G. Efficient Fair Queueing Using Deficit Round Robin. In *ACM SIGCOMM 95*. August 1995.

[Stahl et al., 1997] STAHL, D. O. ET AL. Congestion Pricing More Profitable than AOL Pricing. Center for Research in Electronic Commerce, `http://cism.bus.utexas.edu`. 1997.

[Sun and Langendoerfer, 1998] SUN, Q. AND LANGENDOERFER, H. A new distributed routing algorithm for supporting delay-sensitive applications. *Computer Communications*, Vol. 21, No. 6, pp. 572–578. May 1998.

[Tennenhouse, 1998] TENNENHOUSE, D. From Internet to Active Net. INFOCOM '98 Panel Session, San Francisco. April 1998.

[Traina, 1995] TRAINA, P. BGP-4 Protocol Analysis. RFC 1774. March 1995.

[uunet, 1997] UUNET Press Release, Uunet Details Peering Strategy, Changing Internet Economics Prompt New Policy. `http://www.uu.net/press`. March 1997.

[Vickrey, 1961] VICKREY, W. Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance*. 1961.

[Wang and Crowcroft, 1996] WANG, Z. AND CROWCROFT, J. Quality-of-Service Routing for Supporting Multimedia Applications. *IEEE JSAC*, Vol. 14, No. 7, pp. 1288–1234. July 1996.

[Wirth, 1980] WIRTH, N. *Programme und Datenstrukturen*. Springer. 1980.

Colophon

This book was typeset using faces from the Palatino and Helvetica font families. Production was performed by LaTeX using styles from various sources with my own additions and flavors.

TIK-Schriftenreihe

Nr. 1
Kausale Beschreibung von Systemen
Hans Peter Gisiger
1993, ISBN 3-7281-2050-2

Nr. 2
Compiling Sisal for the ADAM Architecture
Srdjan Mitrovic
1993, ISBN 3-7281-2051-0

Nr. 3
Compiling the Behaviour Part of LOTOS
Eric Emile Dubois
1994, ISBN 3-7281-2054-5

Nr. 4
The ADAM Architecture and its Simulation
Olivier Maquelin
1994, ISBN 3-7281-2113-4

Nr. 5
GIPSY: Ein Ansatz zum Entwurf integrierter Softwareentwicklungs-systeme
Reto Marti
1994, ISBN 3-7281-2114-2

Nr. 6
Teilautomatische Herstellung von Codegeneratoren fuer Mikrokontroller
Hans Baerfuss
1994, ISBN 3-7281-2121-5

Nr. 7
SVOX: The Implementation of a Text-to-Speech System for German
Christof Traber
1995, ISBN 3-7281-2239-4

Nr. 8
Transparent Fault-Tolerance for Process Control Systems
Thomas Siegrist
1995, ISBN 3-7281-2242-4

Nr. 9
Routing in Weitraum-Rechnernetzen
Horst Hodel
1996, ISBN 3-7281-2216-5

Nr. 10
Execution Analysis of Real-Time Software Running on Heterogeneous Distributed Computer Systems
Andreas Danuser
1995, ISBN 3-7281-2284-X

Nr. 11
Prosodische Modifikation von Sprachsegmenten fuer die konkatenative Sprachsynthese
Beat Pfister
1996, ISBN 3-7281-2316-1

Nr. 12
A Framework for Dynamic Protocol Configuration
Thomas Peter Plagemann
1996, ISBN 3-7281-2334-X

Nr 13
Execution Architecture of the Multithreaded ADAM Prototype
Philipp Faerber
1996, ISBN 3-7281-2384-6

Nr. 14
Towards Computer Supported Audio Conferencing
Christoph A. Burkhardt
1996, ISBN 3-7281-2386-2

Nr. 15
Comparison of Classic and Hybrid HMM Approaches to Speech Recognition over Telephone Lines
Hans-Peter Hutter
1997, ISBN 3-7281-2424-9

Nr 16
An Object Oriented Petri Net Approach to Embedded System Design
Robert Esser
1997, ISBN 3-7281-2416-8

Nr. 17
Theory of Evolutionary Algorithms and Application to System Synthesis
Tobias Blickle
1997, ISBN 3-7281-2433-8

Nr. 18
Eine objektorientierte Softwarearchitektur fuer Multimedia-Kommunikationssysteme
Daniel P. Ingold
1997, ISBN 3-906469-00-X

Nr. 19
Verifikation von Implementationen abstrakter Datentypen fuer sicherheitskritische Systeme am Beispiel von Oberon
Daniel Schweizer
1997, ISBN 3-906469-01-8

Nr. 20
Group and Session Management for Collaborative Applications
Erik Wilde
1997, ISBN 3-8265-2411-X

Nr. 21
Project GIPSY: Facing the Challenge of Future Integrated Software Engineering Environments
Tobias Murer
1997, ISBN 3-906469-02-6

Nr. 22
Rapid-Prototyping eingebetteter Systeme
Ruedi Mattmann
1997

Nr. 23
A Multipoint Communication Architecture for End-to-End Quality of Service Guarantees
Daniel Bauer
1997, ISBN 3-8265-2638-4

Nr. 24
On the Emulation of Impairments in ATM-Networks
Hans Kurmann
1997, ISBN 3-906469-03-4

Nr.25
Emulation of LAN MAC Service for ATM Networks
Andreas X. Meier
1997, ISBN 3-906469-04-2

Nr.26
Controlling Segmental Duration in Speech Synthesis Systems
Marcel P. Riedi
1997, ISBN 3-906469-05-0

Nr. 27
Ein Komponentenmodell fuer die Verbindungssteuerung in Multimedia-Netzen
Stefan Frey
1997

Nr. 28
The Architecture of an Interactive Multimedia Communication System
Urs Roethlisberger
1998

Nr. 29
Internet-wide Software Component Development Process and Deployment Integration
Daniel Scherer
1998, ISBN 3-906469-5

Nr. 30
Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications
Eckart Zitzler
1999, ISBN 3-8265-6831-1

Nr. 31
Fast Longest Prefix Matching: Algorithms, Analysis, and Applications
Marcel Waldvogel
1999

Nr. 32
Synchronisation in Distributed Multimedia Applications
Christina Class
1999, ISBN 3-8265-6906-7

Nr. 33
Parallelising Molecular Dynamics for Message Passing Systems
Martin Gerber
1999, ISBN 3-8265-7225-4

Nr. 34
Service Management in a Telecom Environment based on Active Network Technology
Marcus Brunner
1999

Nr. 35
Dynamic Security in Communication Systems
Germano Caronni
1999

Nr. 36
Symbolic Methods Applied to Formal Verification and Synthesis in Embedded Systems Design
Karsten Strehl
2000, ISBN 3-8265-7193-2