

Jeannette Janssen  
Paweł Prałat (Eds.)

LNCS 4852

# Combinatorial and Algorithmic Aspects of Networking

4th Workshop, CAAN 2007  
Halifax, Canada, August 2007  
Revised Papers

Springer

Jeannette Janssen Paweł Prałat (Eds.)

# Combinatorial and Algorithmic Aspects of Networking

4th Workshop, CAAN 2007

Halifax, Canada, August 14, 2007

Revised Papers



## Volume Editors

Jeannette Janssen  
Paweł Prałat  
Dalhousie University  
Department of Mathematics and Statistics  
Halifax, NS, B3H 3J5, Canada  
E-mail: {janssen,pralat}@mathstat.dal.ca

Library of Congress Control Number: 2007941336

CR Subject Classification (1998): F.1.1, F.2.1-2, C.2, G.2.1-2, E.1

LNCS Sublibrary: SL 5 – Computer Communication Networks and Telecommunications

ISSN 0302-9743  
ISBN-10 3-540-77293-6 Springer Berlin Heidelberg New York  
ISBN-13 978-3-540-77293-4 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media  
springer.com

© Springer-Verlag Berlin Heidelberg 2007  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper SPIN: 12204793 06/3180 5 4 3 2 1 0

# Preface

The advent of the Internet has opened up a wealth of applications, but also given rise to a host of new problems. Many of those problems have led to exciting new research directions in mathematics and theoretical computer science, especially in the areas of combinatorics and algorithms. The Fourth Workshop on Combinatorial and Algorithmic Aspects of Networking (CAAN 2007) was organized to be a place where the latest research developments on all aspects of networking could be presented. The topics covered were diverse, with talks on strategies for searching in networks, for cleaning networks of unwanted intruders, on different routing strategies, and on scheduling and load balancing. The workshop started with an invited lecture by Peter Winkler of Dartmouth College, who gave a general talk on a topic related to probability, a concept central to network modeling and managing. The afternoon opened with a short invited talk by Alejandro López-Ortiz, who gave an overview of various issues in designing resilient backbone networks.

CAAN 2007 took place on August 14, 2007, at Dalhousie University in Halifax, Nova Scotia, Canada, co-located with the Workshop on Algorithms and Data Structures (WADS 2007). Three previous CAAN workshops were held in Chester, UK (CAAN 2006), Waterloo, Ontario, Canada (CAAN 2005), and in Banff, Alberta, Canada (CAAN 2004), respectively.

In response to the call for paper we received 17 submissions. Each submission was reviewed by three referees. Almost all submissions were relevant to the topic of the workshop, and most contained interesting ideas. Based on the reviews of the referees we accepted nine papers for presentation at the workshop and inclusion in this volume. The volume also includes an abstract of the invited talk by Peter Winkler and an invited paper by Alejandro López-Ortiz.

We would like to thank all those that helped to make this workshop a success, with special thanks to Anne Publicover, our administrative assistant. Thanks to the Atlantic Association for Research in Mathematics (AARMS) for financial support. Many thanks to Andrei Voronkov for providing the EasyChair conference system; with this system, managing the electronic submissions and the refereeing process has been a breeze. Thanks also to Norbert Zeh, the Local Arrangements Chair of WADS 2007. Finally, we thank all participants in the workshop, all authors of the contributed papers, and especially the invited speakers for their contribution in making CAAN 2007 into a collegial and stimulating platform for new ideas about networks.

October 2007

Jeannette Janssen  
Paweł Pralat

# Organization

## Steering Committee

Andrei Broder	Yahoo! Inc., USA
Angèle Hamel	Wilfrid Laurier University, Canada
Srinivasan Keshav	University of Waterloo, Canada
Alejandro López-Ortiz	University of Waterloo, Canada
Rajeev Motwani	Stanford University, USA
Ian Munro	University of Waterloo, Canada

## Program Committee

Dimitris Achlioptas	UC Santa Cruz, USA
Anthony Bonato	Wilfrid Laurier University, Canada
Allan Borodin	University of Toronto, Canada
Colin Cooper	King's College, UK
Erik Demaine	MIT, USA
Thomas Erlebach	University of Leicester, UK
Angèle Hamel	Wilfrid Laurier University, Canada
Jan van den Heuvel	London School of Economics, UK
Klaus Jansen	Universität Kiel, Germany
Jeannette Janssen	Dalhousie University, Canada
Valerie King	University of Victoria, Canada
Danny Krizanc	Wesleyan University, USA
Stefano Leonardi	Università di Roma “La Sapienza”, Italy
Alejandro López-Ortiz	University of Waterloo, Canada
Lata Narayanan	Concordia University, Canada
Richard Nowakowski	Dalhousie University, Canada
Paweł Prałat	Dalhousie University, Canada
Sunil Shende	Rutgers-Camden, USA
Eva Tardos	Cornell University, USA

# Table of Contents

## Invited Lectures (Abstracts)

Luck vs. Skill (Long Invited Talk) . . . . .	1
<i>Peter Winkler</i>	
Valiant Load Balancing, Benes Networks and Resilient Backbone Design (Short Invited Talk) . . . . .	2
<i>Alejandro López-Ortiz</i>	

## Contributed Papers

Valiant Load Balancing, Capacity Provisioning and Resilient Backbone Design . . . . .	3
<i>Alejandro López-Ortiz</i>	
Cleaning Random $d$ -Regular Graphs with Brushes Using a Degree-Greedy Algorithm . . . . .	13
<i>Margaret-Ellen Messinger, Paweł Prałat, Richard J. Nowakowski, and Nicholas Wormald</i>	
Nonadaptive Selfish Routing with Online Demands . . . . .	27
<i>Tobias Harks and László A. Végh</i>	
Vertex Pursuit Games in Stochastic Network Models . . . . .	46
<i>Anthony Bonato, Paweł Prałat, and Changping Wang</i>	
Preemptive Scheduling on Selfish Machines . . . . .	57
<i>Leah Epstein and Rob van Stee</i>	
Selfish Routing and Path Coloring in All-Optical Networks . . . . .	71
<i>Ioannis Milis, Aris Pagourtzis, and Katerina Potika</i>	
A Worst-Case Time Upper Bound for Counting the Number of Independent Sets . . . . .	85
<i>Guillermo De Ita and Aurelio López-López</i>	
Improving the Efficiency of Helsgaun's Lin-Kernighan Heuristic for the Symmetric TSP . . . . .	99
<i>Dirk Richter, Boris Goldengorin, Gerold Jäger, and Paul Molitor</i>	
Combinatorial Algorithms for Listing Paths in Minimal Change Order . . . . .	112
<i>Zareen Alamgir and Sarmad Abbasi</i>	

Improving Topological Routing in N2R Networks ..... 131  
    *Jose M. Gutierrez Lopez, Ruben Cuevas Rumin,*  
    *Jens M. Pedersen, and Ole B. Madsen*

**Author Index** ..... 149

# Luck vs. Skill

## (Long Invited Talk)

Peter Winkler

Dartmouth College, Hanover, NH, USA  
`peter.winkler@dartmouth.edu`

**Abstract.** Recent legislation in the US regarding gambling over the web has led to renewed interest in the question of which games are games of skill. We take a statistical approach to the problem, defining the skill index of a game to be the average amount of playing time after which variance due to chance and variance due to skill differences are equal.

We then look at tournament results for championship-level duplicate bridge, PGA golf, and duplicate poker, as well as some simulated toy games, to see how their skill indices compare.

## Biography

Peter Winkler is Professor of Mathematics and Computer Science at Dartmouth College and Albert Bradley Third Century Professor in the Sciences.

A winner of the Mathematical Association of America's Lester R. Ford Award for mathematical exposition, Dr. Winkler is the author of about 125 mathematical research papers and holds a dozen patents in computing, cryptology, holography, optical networking and marine navigation. His research papers are primarily in combinatorics, probability and the theory of computing, with forays into statistical physics.

Dr. Winkler received his BA from Harvard summa cum laude in mathematics, then after a stint in the US Navy, his PhD from Yale as a student of Abraham Robinson and Angus Macintyre. He joined the faculties of Stanford and then Emory University, where he became Professor and Chairman of Mathematics and Computer Science. In 1989 he left academia for industry, returning in 2004.

When not proving theorems or enjoying his family, Winkler is generally found on a squash court or playing and composing ragtime piano music. He collects puzzles both mechanical and mathematical, the latter appearing in two popular books. In some circles Winkler is notorious as the inventor of cryptologic techniques for the game of bridge, which have now been declared illegal for tournament play in most of the western world.



# Valiant Load Balancing, Benes Networks and Resilient Backbone Design

## (Short Invited Talk)

Alejandro López-Ortiz

University of Waterloo, Waterloo, ON, Canada  
alopez-o@uwaterloo.ca

**Abstract.** At any given time, the traffic on the network can be described using a traffic matrix. Entry  $a_{i,j}$  in the matrix denotes the traffic originating in  $i$  with destination  $j$  currently in the network. As traffic demands are dynamic, the matrix itself is ever changing. Traditionally network capacity has been deployed so that it can support any traffic matrix with high probability, given the known traffic distribution patterns. Recently the need for resilience and reliability of the network for mission critical data has brought the need for backbone capacity that can support all traffic matrices. In this talk we give an overview of the state of the art on networks and routing schemes with this property.

## Biography

Alejandro Lopez-Ortiz received his B.Math. degree from the National University of Mexico (UNAM) in 1989, and his M.Math. and Ph.D. from the University of Waterloo in 1990 and 1996 respectively. In his research he has combined the development of theoretical tools and efficient algorithms with real life applications. He has been a faculty member in the School of Computer Science, University of Waterloo since 2001 (promoted to Associate professor with tenure in 2004) and was Director of Research at Internap network services corporation in Seattle. His research addresses questions of both theoretical and practical relevance such as robot navigation, search engines, data streams and the internet. He is currently co-chair of the DIMACS three year special focus on Algorithmic Foundations of the Internet, jointly with Jennifer Rexford from Princeton University and Rebecca Wright of Rutgers University.

# Valiant Load Balancing, Capacity Provisioning and Resilient Backbone Design

Alejandro López-Ortiz

Cheriton School of Computer Science  
University of Waterloo  
Waterloo, ON N2L 3G1 Canada  
alopez-o@uwaterloo.ca

**Abstract.** The two main alternatives for achieving high QoS on the public internet are (i) admission control and (ii) capacity overprovisioning. In the study of these alternatives the implicit (and sometimes explicit) message is that ideally, QoS issues should be dealt with by means of sophisticated admission control (AC) algorithms, and only because of their complexity providers fall on the simpler, perhaps more cost-effective, yet “wasteful” solution of capacity overprovisioning (CO) (see e.g. Olifer and Olifer [Wiley&Sons, 2005], Parekh [IWQoS’2003], Milbrandt et al. [J.Comm. 2007]). In the present survey we observe that these two alternatives are far from being mutually exclusive. Rather, for data critical applications, a substantial amount of “overprovisioning” is in fact a fundamental step of any safe and acceptable solution to QoS and resiliency requirements. We observe from examples in real life that in many cases large amounts of overprovisioning are already silently deployed within the internet domain and that in some restricted network settings they have become accepted practice even in the academic literature. Then we survey the main techniques currently in use to compute the provisioning capacities required in a resilient high QoS network.

## 1 Introduction

In the quality-of-service literature (QoS) two main alternatives are given for achieving high QoS on the public internet. These are (i) admission control and (ii) capacity overprovisioning. In the study of these alternatives the implicit (and sometimes explicit) message is that ideally, QoS issues should be dealt with by means of sophisticated admission control (AC) algorithms, and only because of their complexity providers fall on the simpler, perhaps more cost-effective, yet “wasteful” solution of capacity overprovisioning (CO) (see e.g. [22,23,20]). AC researchers often express the hope that this situation will eventually remedy itself and that sophisticated AC algorithms will do away with the need for bandwidth overprovisioning (e.g. [8]). Only recently Menth et al. in a SIGCOMM’06 paper gave evidence that CO might not be as undesirable as previously thought [19].

In the present survey we observe that these two alternatives are far from being mutually exclusive. Rather, for data critical applications, a substantial amount

of “overprovisioning” is in fact a fundamental step of any safe and acceptable solution to QoS requirements. Indeed, a survey of common practices in the field suggests that this observation has been arrived to independently and empirically by network engineers in various settings within the Internet and otherwise, yet the QoS literature so far does not reflect this discovery nor has it attempted to explain its root causes.

We observe from examples in real life that in many cases large amounts of overprovisioning are already silently deployed within the internet domain and that in some restricted network settings they have become accepted practice even in the academic literature. In other words, distaste for overprovisioning is not a universally held belief outside the QoS domain. In fact, the telephony network which is considered a classical example of AC is in practice heavily overprovisioned and actual AC policy is rarely relied upon even though it is deployed on the network [21]. Then we survey the main techniques currently in use to compute the provisioning capacities required in a resilient high QoS network. We term this amount rightprovisioning. Lastly, we give reasons why QoS over a rightprovisioned network has different needs and goals than those currently addressed by admission control and other such mechanisms.

## 2 Internet QoS

The two main mechanisms for achieving a desired level of service on the internet are admission control and capacity overprovisioning. QoS on the network allows the user to make choices as to the level of service it requires. Typical parameters are: data rate (bandwidth), availability, end-to-end delay (latency), variation of end-to-end delay (jitter), and packet loss rate [8].

### 2.1 Capacity Overprovisioning

Capacity overprovisioning consists in increasing available bandwidth until it is large enough to sustain the vast majority of peaks in demand. Depending on the level of reliability desired this can be as low as 25% above average data rate to handle 95% of all traffic demands without loss, 50% extra bandwidth to carry 99% of traffic, and double the average bandwidth to meet 99.99% or higher of all traffic demands without loss (see e.g [32]). This last choice, meaning 50% utilization of the pipe, is often anecdotically referred to as the upper limit of utilization currently acceptable by large ISPs, with the load on an average link often being well below that [7,8].

In contrast, in the QoS literature overprovisioning is considered a simple but wasteful solution to QoS demands. For example, to quote from a computer networks textbook [22]:

Overprovisioned services keep the network infrastructure simple (no additional tools and configurations) but are wasteful as 60-70% of potential network resources are not in use. Under such conditions the best-effort

service on a standard IP network turns out to be good enough for all network applications including time-sensitive ones.

Indeed the term “overprovisioning” itself has the implication that more capacity than what was required was provisioned and hence it ends up being wasted. Yet, subutilization of a resource alone does not imply it was *over*provisioned. In fact, most mission critical applications such as avionics routinely rely on highly redundant configurations, which under normal operational procedures are not used. For example an ocean liner arriving safely to port did not utilize its life boats, yet no one would argue that they were thus “overprovisioned”.

## 2.2 Admission Control

Admission control is mostly about using resource reservation and limits on traffic volume to preven overload on the network. It is predicated on the basis that not all network traffic is time-sensitive and mission critical. The AC alternative to overprovisioning is denying resources to non crucial flows. Typical examples of time-sensitive traffic are real time flows (e.g. video/audio streaming, IP telephony) and high value transactions (stock trades, last bid at an online auction). Packets are assigned a priority value with higher priority packets being given preferential service. Yet a look at the historical development of the internet suggests that, over the years, the majority of the traffic overtime has become more time sensitive and mission critical. Recall that in the original internet the majority of traffic was smtp (email) and nntp (usenet) based. These protocols have acceptable delay tolerances from several minutes to as long as days. Web traffic which is served interactively has acceptable delays in the 10 second or less range. VoIP and other streaming traffic have subsecond delay tolerances.

As more of the nation infrastructure migrates to the public internet, a disruption in the network has larger consequences. The financial, defense, telephone, commerce, government, and business infrastructure now rely on the availability of the Internet to operate properly. Even a seemingly non-mission critical application such as a standard home network connection which might have been initially deployed for one parent’s non-time sensitive email (smtp) traffic later on became used by the kids for highly time-sensitive gaming and audio streaming as well as by a parent bidding in online auctions for objects worth thousands of dollars, and as of recently is being used as a carrier for VoIP services which means that emergency calls (911 or to the family doctor) are routed over it. These last type of calls are both time-sensitive *and* mission critical. Thus, it is not far-fetched to envision a world in which the majority of the traffic will be labeled as time sensitive and hence the savings from AC would be minimal, since not many flows can be dropped. This would make packet classification schemes at admission control points progressively more difficult and less useful, the majority of the traffic is critical to start with.

This suggests that as more data exchanges migrate to the Internet infrastructure, the need for higher reliability will further increase while the ability to differentiate between types of traffic will continue to decrease.

### 3 Rightprovisioning

Capacity overprovisioning is common place in the current internet [1,8,18,12]. AC based solutions remain unused while anecdotal evidence suggests that CO is the preferred method for QoS delivery in the commercial internet. Currently QoS due to CO is such that no packets are dropped in the backbones [8,15,4]. Packet loss occurs mostly in the interface between the end points of the network and the large ISP providers. As providers have focused on ensuring that there is sufficient deployed capacity rather than on implementing admission control solutions. ISPs will go to the extent of delaying by several months the start of connectivity for a new customer to ensure that there is enough capacity on the network to support the bandwidth demands of the new customer (this can be argued is a crude form of admission control). In other words, currently ISPs find that CO is a cost effective way to achieve QoS.

While most of the literature is critical of CO as a solution of QoS, recent developments suggest that even in theory its performance is better than originally thought. Bhagat observes that in certain settings overprovisioning seems to be a better answer to the performance needs from users, and indeed he goes as far as questioning the need for admission control based QoS solutions [6]. In a recent breakthrough paper in SIGCOMM'06 Menth et al. [19] show that if overprovisioned capacity is also used to achieve resilience against network failures, then the demands in terms of bandwidth of failure-resilient AC and CO schemes are comparable, as the overprovisioned capacity can be deployed for various uses depending on the type of congestion and/or failure detected. In sum, so far we have argued that

1. selective admission as required by AC is becoming increasingly less of an option at the backbone level since traffic is increasingly time and mission critical,
2. that CO in large trunks is already in place and provides excellent QoS within the core of the network,
3. that as such its effectiveness is well supported by established practice, and that
4. the academic literature has started to explain why CO is such an effective solution.

The question then remains what is the proper level of overprovisioning, i.e. rightprovisioning. Currently the model most commonly in use is a statistical guarantee of the probability of connection denial. We argue that the right metric is to provide enough capacity so that any valid traffic matrix can be realized.

**Definition 1.** *Formally, let  $e_1, \dots, e_n$  be  $n$  end points in the network each with a send and receive capacity  $s_i$  and  $r_i$  respectively. A traffic matrix  $A = [a_{ij}]$  contains in entry  $a_{ij}$  the instantaneous amount of traffic from node  $e_i$  destined to  $e_j$ .*

**Definition 2.** *A given traffic matrix is said to be valid if  $\sum_{j=1}^n a_{ij} \leq s_i$  and  $\sum_{i=1}^n a_{ij} \leq r_j$ . That is no node is attempting to send more data than it has uplink provisioned capacity for and no node is being sent more data than it has contracted capacity to receive.*

In the past providers have deployed enough capacity to handle the average traffic matrix or a percentage of traffic matrix configurations (say 95% of the time the traffic matrix should cause no loss in traffic). Since the aim is to provide connectivity for the worst case traffic matrix we need to determine what is the minimum or most cost efficient capacity that satisfies this requirement. We could simply consider the sum of all contracted capacity by users, however this does not take into account that currently connectivity is provided in an average fashion, typically at a certain average rate per month with a maximum burstable rate.

In the new regime, two types of traffic would be provisioned. Traffic of type *A*, which is mission critical and always available at the contracted capacity and traffic of type *B*, at an average contracted capacity but rate-controlled depending on connectivity characteristics. In essence this could be thought as rate modulation over a pipe carrying type *B* traffic, not unlike in nature and effects to that performed by a modem in the presence of high levels of line noise. Observe that this establishes a very simple form of admission control. Traffic of type *A* would be unavailable at most on the order of subsecond to few seconds per year range (seven to eight nines of reliability). At the same time the entire contracted capacity should be generally available, with traffic of type *B* being flow rate controlled in the order of a half a minute to a few minutes a year (five to six nines range). This last is the current level of service reliability that the telephone network claims to have, even though arguably telephone traffic is less time critical than many of the current uses of the network. It is worthy of note that the telephone network operates at 33% capacity [21] and that the amount of admission control is minimal. For example “on Monday, Dec. 2, 1991, which was the busiest day for the AT&T network until then, of 157.5 million calls, only 228 were blocked on intercity connections” (from [3] as quoted by [21]). Our proposal parallels this design choice.

Interestingly enough, worst-case traffic matrix  $n \times n$  capacity already exists in certain network settings. In the LAN the proper amount of overprovisioning has evolved to be such that, given  $n$  nodes on an Ethernet, a complete set of  $n/2$  disjoint pairs can communicate at full speed. Recall that this was not always the case, as the original co-axial ethernet only had sufficient capacity for a single pair to communicate freely at full capacity without collision; eventually star switches with higher capacity buses became commonplace, and currently common  $n \times n$  crossbar or Beneš network switches have the ability to sustain  $n/2$  disjoint pairs of communication [2]. Similarly Network Access Points (NAPs) as well as cores of large corporate networks often consist of an optical ring providing enough capacity for all possible crossconnects. This is not unique to the internet. In the 1970s telephone networks deployed switches with  $n \times n$  capacity at certain critical points of the infrastructure [26].

For statistical guarantees the law of large numbers can be used to determine the maximum simultaneous demand that may originate, on the aggregate, from a neighborhood of nodes sharing an entry point to the ISP backbone. This is repeated for all entry points into the backbone and then a full  $n \times n$  bandwidth capability over those averages can be deployed. The size of such an  $n \times n$  network is well understood. We discuss in detail the various known alternatives in Section 5.

Lastly, as Menth et al. observed, redundant equipment can be deployed for multiple purposes, so long as the probability of failure of such equipment is independent [19]. This amortizes the additional cost of redundant equipment. In particular redundant capacity can be used to circumvent router and link failures (digging). This has been observed to reduce the amount of apparent “subutilization”. As well, secondary sources of traffic which can be quenched at the source point can be sent over the spare capacity. Examples of this are CDN content and remote backup data which are resilient under short time delays. Anecdotal evidence suggests that spam traffic is delivered at off-peak times by certain ISPs using deployed overcapacity.

## 4 QoS and AC in a Rightprovisioned World

Observe that we do not claim that overprovisioning at the backbone is sufficient to achieve all QoS requirements, nor would it make AC trivial. This is in contraposition to claims to that end in the literature, e.g. “only when the ratio of resources at the edges of a network to those available in the core of a network becomes high is the problem of service differentiation interesting, [...] when this ratio is low, any QoS mechanism appears redundant as most users receive the service they require anyway, and so the cost introduced by a QoS scheme appears unjustified, and research into QoS mechanisms appears unnecessary” [8].

For one, as the network is used for more life-critical operations such as VoIP phone calls (911), financial transactions (stock exchange), remote surgery, and air traffic system, perhaps even carrier grade reliability is not good enough. It is not hard to envision demands for reliability reaching into the 99.999999% range (in fact today it is possible to provision bandwidth with a stated 100% reliability guarantee in the sense that *any* amount of downtime is contractually heavily penalized). Such high levels of reliability will require overprovisioning, multihoming, redundancy, admission control and intelligent routing, though the types of solutions required, their price/performance ratio and their goals change. As well, end users will still, on occasion, attempt to send or receive more time critical data that is feasible given their available network connectivity. Admission control in such situations will be needed to prioritize say, a 911 VoIP call (type *A* traffic) over downloading email (type *B* traffic).

Admission control starts from the assumption that congestion will always take place at the edge given the reduced capacities of the endpoint as compared to the capacity of the entire network (i.e. the need to send or receive more data than

what we have capacity for). What this work argues is that congestion should only take place at the edge and that CO is the way to ensure this.

The model we propose assumes that all packets reaching the network core are assumed to be critical and hence failure of delivery is not an option. Within the core there would be no differentiated services with AC taking place as a weak and simplified form of resource reservation: if the packet is admitted, it can be delivered. The end node would send data in one of two modes: normal mode in which all traffic is accepted without need for any AC intervention and exceptional mode in which the application/user is alerted of a temporary service disruption and given the choice to proceed with the communication at full speed or throttle down for a few seconds (type *A* or *B* classification). Incentives such as price differentials can be built in to ensure that the user delays non-essential traffic.

Given the reliability needs detailed above this would occur with a very low probability, in the range of thirty seconds to a few minutes of service disruption per year. Such a rare occurrence means that only the simplest of differentiated services and admission control policies can be justified from the perspective of economic viability. As it has been observed [8] a weak form of AC already takes place in the edges in that providers delay customer activation to ensure that enough capacity is present to satisfy demand. This is a crude yet effective form of denying a transmission request.

As well routing in an overprovisioned network is more complicated as the multiplicity of paths allows for an intelligent choice. This determination does not involve the end point as the network makes best effort for all packets.

## 5 Valiant Load Balancing and Beneš Networks

Claude Shannon pioneered the study of networks that support  $n \times n$  communication pairs. He proved that if the proving that a fabric of  $n \log n$  switches is necessary so long as total deployed capacity is linear. Beneš introduced the later termed Beneš networks which match Shannon's lower bound switch [2,10,11,26]. Arora et al. combined Beneš networks with the butterfly network to obtain a similar topology that supports all cross connects in an online fashion, while preserving the efficiency in terms of deployed capacity.

If there are no restrictions in the total deployed capacity then other alternative realizations are possible. Two of the most common being a central high capacity ring and the  $n \times n$  crossbar. Pippenger extensively studied the topology of telephone switches that support  $n \times n$  connection patterns [24,25,26,27,28].

Valiant proposed an elegant network topology in the context of interprocessor communication networks for parallel computers [33,34]. The network starts with the complete graph on  $n$  nodes which trivially can support all independent connection pairs. However it is an inefficient solution as it requires  $n^2$  contracted capacity. Valiant's key observation is that a two phase communication protocol on a complete graph in which every link has capacity  $2/n$  suffices. This reduces the total deployed capacity to  $2n$  which is a constant times the deployed capacity.



Extensions and generalizations of both Pippenger's and Valiant's work have been the subject of intense study within theoretical computer science [9,14,2] as well as the networks community [17,30,31,29]. The field is now referred to as a Valiant Load Balancing Network and/or as a Virtual Private Network load balancing. The term VPN comes from the fact that VPNs were one of the earliest users of the internet requiring high degree of reliability. Shepherd et al. and Prasad et al. have run simulations to determine the effect of VPN load balancing in existing networks, and have observed that peak traffic loads are lowered down while resilience is improved. Many open questions remain, among them

- how to efficiently design an overprovisioned network under realistic cost measures?
- design an overprovisioned network which readily scales under incremental growth?
- given a pre-existing network infrastructure compute the lowest cost links that must be added for the network to support the worst case traffic matrix
- how to add links to an existing network infrastructure in a way that they can serve the dual purpose of worst case traffic matrix provisioning and resiliency under link cuts?
- how to implement the desired routing patterns using the current routing protocols (BGP/IGP/OSPF)?

## 6 Conclusions

We have argued that given the evolution path of internet traffic, higher levels of reliability will be required. As such admission control schemes which refuse connections are no longer feasible. At the same time we give evidence that capacity overprovisioning with a high probabilistic guarantee of delivery for  $n \times n$  traffic is already in place in the internet, though not generally recognized. We also observed that in other network settings such large capacity has been openly, purposely deployed with the full acceptance of theory and practice. We noted that “overprovisioned” capacity can be put to other uses as others have shown [19], and CO is more efficient than generally believed. We give general bounds on the amount of traffic that is required for service guarantees and we term this *rightprovisioning* the network. Lastly we argued that there is still need for QoS and AC policies at the network edge.

## References

1. Armitage, G.J.: Revisiting IP QoS: Why do we care, what have we learned. ACM 2003 RIPOS Workshop Report, ACM SIGCOMM Comp. Comm. Rev. vol. 33(5) (October 2003)
2. Arora, S., Leighton, F.T., Maggs, B.M.: On-line Algorithms for Path Selection in a Nonblocking Network. In: Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, pp. 149–158 (May 1990)