

**Congestion on the Internet: Operator responses,
economic analysis, and improving the network
architecture**

by

Steven J. Bauer

B.S., University of Tulsa (1998)

S.M., Massachusetts Institute of Technology (2001)

Submitted to the Department of Electrical Engineering and Computer
Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

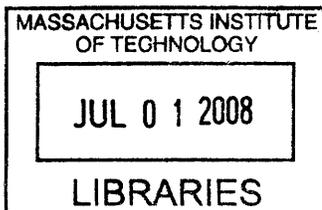
June 2008

© Massachusetts Institute of Technology 2008. All rights reserved.

Author
Department of Electrical Engineering and Computer Science
May 21, 2008

Certified by....
David Clark
Senior Research Scientist
Thesis Supervisor

Accepted by
Terry P. Orlando
Chairman, Department Committee on Graduate Students



ARCHIVES

Congestion on the Internet: Operator responses, economic analysis, and improving the network architecture

by

Steven J. Bauer

Submitted to the Department of Electrical Engineering and Computer Science
on May 21, 2008, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Our thesis presents the results of monitoring traffic management techniques of network providers. We find that in addition to complete blocking of traffic flows on some transport layer ports, providers are also actively managing traffic based upon application layer content. We document how providers are engaged in not just deep packet inspection, but deep flow inspection – tracking, shaping, and actively managing the state of individual flows.

We explain why network operators attempt to manage traffic in these manners by looking at both the technical and economic forces influencing their decisions. We argue that inefficient congestion management and contractual practices of network providers are not unexpected given the limitations of the existing Internet architecture. Indeed, managing network traffic based upon transport and application layer content are likely to become more prevalent as high bandwidth demanding applications such as video and peer-to-peer increasingly dominate the traffic mix on networks.

The existing Internet architecture leaves network operators with little architectural support for making more efficient traffic management decisions or contractual relationships. We therefore consider how making the architectural and protocol changes proposed by Briscoe et al [14, 17, 16], might help better align the economic and technical interests of traffic senders and traffic carriers. Briscoe’s proposal would enable traffic management and/or contractual relationships to be based directly on congestion information. Our contribution is to further the analysis of the economic implications of this proposal. We apply both game theory and elementary economic analysis to better understand the proposal’s advantages and limitations.

Aligning the congestion responses of end-users with the traffic management and contractual decisions of network operators is part of a larger challenge of designing the next generation Internet architecture and protocols while taking into account the economic interests of different stakeholders. Accepting a proposal such as Briscoe’s requires the research and wider networking community to be able to balance both technical and economic arguments in a design process. We examine how this integration might occur. Establishing a means of discourse is vital to the objective evaluation

of what constitutes good architecture and good research for these techno-economic problems.

Thesis Supervisor: David Clark
Title: Senior Research Scientist

Dedication

This thesis is dedicated to my family. Thank you for your love, support and patience as I have pursued this PhD.

Acknowledgments

What I write here can only hint at the debt of gratitude I owe so many people. My sincere thanks to the many friends and colleagues I have known at MIT over the years.

In particular, I would to thank my formal and informal advisers:

1. David Clark: As a researcher, I hope to pattern myself after Dave. He has an amazing mastery of both the low level technical details and the high level context of research problems. He is someone that cares passionately about not just his own work, but our field of research as a whole. He has taught me to look out at the research horizons and to see amazing possibilities for what could be there.
2. Arthur Berger: While I am indebted to Arthur for his aid on many technical and mathematical topics, the largest debt I owe Arthur is for his constant support and guidance. He has helped me navigate the sometimes rough waters of the Ph.D. program. If I someday mentor a junior member of my community, I hope I will be able to impart as much thoughtful wisdom and advice as I have learned from Arthur.
3. Peyman Faratin: Peyman has been my best and constant supporter and friend at MIT. I have spent more hours in intellectual exchange with Peyman than anyone else at MIT. These exchanges are the most valuable times in any graduate career. They sharpened my arguments and taught me how to be rigorous in thought and explanation.
4. John Wroclawski: Much of how I see the research world is through the constructively critical lens that I learned from John. The joy of doing research is there is never a shortage of new, interesting, and challenging problems. John taught me how to identify them.
5. Karen Sollins: Karen has always been the heart of our ANA research group. She pulls the group together literally, in group meetings, and figuratively, by

always being a friendly face that one can turn to for advice and support. I will always be indebted to Karen as she was the first person to shepherd my research efforts at MIT.

I would also like to thank my many colleagues, office mates, and friends at MIT. In order of when I met them at MIT, I would like to particularly acknowledge:

1. Joanna Kulik: I shared an office with Jo longer than any other student at MIT. Most of my fondest memories of graduate student life include Jo. We were always debating and pontificating, dreaming and scheming. She was the role model I looked up to as a graduate student.
2. Chuck Blake: Chuck is at the top of the list of technical influences in my life. I learned more from Chuck about operating systems, programming, programming languages, and statistics than anyone else. The number of times I looked over his shoulder at terminal windows or watched him explain a concept at the whiteboard is uncountable.
3. Tim Shepard: Tim should have been a professor. But, in lieu of a classroom, he taught me over many breakfasts and lunches a great deal about the history and traditions of networking, MIT, and research in general. I wish I had more time to soak up all Tim has to impart.
4. Mike Afergan: Mike's thesis on the economics of repeated games in networking convinced me that I should continue down the economics path. It was during his public thesis defense that I firmly concluded I would follow his lead. His work convincingly demonstrated to me the utility of looking at networking problems from an economic perspective. Indeed, much of how I view the world in general is now shaped by the economic education I have since acquired.
5. Robert Beverly: Over the last years of my graduate career, Rob has been my constant collaborator. Going forward, I hope I am lucky enough to find collaborators that are as intellectually stimulating and pleasant to work with as

Rob as been. On the few inevitable occasions during a graduate career when I grew pessimistic about my research, Rob was always a positive voice of support. “Don’t go negative” – sound advice he delivered during any times that my normally optimistic outlook seemed in danger of slipping away.

Finally, I would like to thank my wife Anne and my sons. They have been extraordinarily patient as I have enjoyed the intellectual life of a graduate student. They have never wavered in their support. Every day of my graduate career, I looked forward to returning home to see them.

Contents

1	Introduction	14
1.1	Documenting and collecting data	16
1.2	Analyzing new congestion management architectures	17
1.3	Economic arguments and engineering design	20
1.4	Summary	22
2	Motivation and the future of network congestion	24
2.1	Network capacity	25
2.1.1	Bandwidth is free	25
2.1.2	Bandwidth is scarce	27
2.1.3	Anecdotal evidence	28
2.2	Measurements of traffic growth and patterns	29
2.3	Summary	32
3	Documenting new congestion management techniques	34
3.1	Policy-based routing	34
3.2	Noticing high rates of TCP connection resets	38
3.3	Building a model of Comcast congestion management	42
3.4	Communication patterns triggering resets	43
3.4.1	Bitfield message reset	44
3.4.2	Classifier robustness	45
3.4.3	Flow state timeouts	46
3.5	Comcast's motivations	48

3.5.1	Mismatch between actual and expected traffic patterns	50
3.5.2	Myopic fairness of capacity allocation among subscribers	51
3.5.3	Long-term fairness of capacity allocation among subscribers	51
3.6	Comcast’s response	52
3.7	Summary	53
4	Limitations of the current Internet architecture	54
4.1	Costs	55
4.2	Decision mechanisms in current Internet	56
4.2.1	Classification	57
4.3	Lack of architectural support	64
4.4	Economic behaviors	67
4.4.1	Network operators	67
4.5	Summary	69
5	Overview of re-ECN	71
5.1	Re-ECN overview	73
5.2	Truthful predictions of congestion	74
5.3	Uses of re-ECN	76
5.3.1	Flow-based congestion control	77
5.3.2	Aggregate traffic shaping	78
5.3.3	Capacity provisioning	79
5.3.4	Peering	80
5.3.5	Subscriber and interprovider contracts	80
5.4	Summary	81
6	Analyzing the economics of re-ECN	83
6.1	Topologically non-uniform costs of congestion	85
6.2	Congestion marking as a signal of a negative externality	88
6.3	Congestion Games	91
6.4	Model	92

6.5	Capacity Expansion Game	94
6.5.1	Structure of the players	95
6.5.2	Intuitive example	96
6.5.3	Equilibrium player strategies	98
6.5.4	Worst case cost transfer	99
6.5.5	Average cost transfers	100
6.6	Congestion budget game	102
6.7	Summary	105
7	Accepting economic evidence in engineering debates	106
7.1	Leveraging economics within an engineering process	107
7.2	Expectations for economics within engineering	110
7.3	Applying economic theory to engineering challenges	112
7.4	Common economic misunderstandings among engineers	114
7.5	Engineering aesthetics	115
7.6	Summary	116
8	Mechanism design in networking	117
8.1	Introduction	118
8.2	Mechanism design	119
8.3	Assumptions about preferences	121
8.3.1	Assessing utility assumptions	122
8.3.2	Example: Jamming in wireless networks	124
8.4	Mechanism design for repeated games	126
8.4.1	Example: Costs in wireless-hoc networks	128
8.5	Discussion	130
8.6	Summary	132
9	Related work	133
9.1	Economics and congestion management	133
9.1.1	Models of network traffic and dimensioning network capacity .	134

9.2	General Internet economics	136
9.3	Game theory and mechanism design	137
9.4	Engineering perspectives on economic problems	138
9.4.1	Systems engineering	139
9.5	Internet measurements	140
10	Conclusions and future work	142
10.1	Future directions	143
	Bibliography	146

List of Figures

1-1	The engineering of networking protocols and architectures historically has not been strongly influenced by economic theory or analysis. . . .	21
2-1	KT Traffic vs. Revenue Growth, Per Subscriber	31
2-2	Traffic growth rates from publicly observed sites[80]	33
3-1	HFC Network Diagram	49
4-1	Comcast authorized use policy (AUP) violation letter	70
5-1	At each of the economic boundaries of the Internet, the congestion impact of traffic needs to be exposed to make traffic management decisions or contracting decisions that are economically efficient.	72
5-2	The incentives mechanisms leveraging re-ECN are designed to induce senders to truthfully report the expected congestion along a path. A shaper creates an incentive for senders to not <i>overstate</i> their view of expected congestion. A dropper creates an incentive for senders to not <i>understate</i> their view of downstream congestion.	75
5-3	“End-to-edge clearing for sender-pays charging. Two duplex flows for the same session are shown, one each in the upper and lower halves of the figure.” (Source [15])	81
6-1	Congestion that occurs in different locations in the network may have different economic costs to network providers.	87
6-2	RED marking probability	89

6-3	The congestion threshold H is increased once R_e revenue has been collected.	94
6-4	Two players, truthfully revealing their traffic demands, share of the capacity expansion cost when the $R_e = 12c$	97
6-5	By sending additional traffic, gray can reduce by 50% its overall contribution to the revenue required for capacity expansion.	98
6-6	Cost allocations from true demands in the capacity expansion game .	101
6-7	Cost allocations with strategic behavior in the capacity expansion game	102
6-8	Factor change in cost allocations with strategic behavior in the capacity expansion game	103
7-1	It is this closing of the cycle to reconnect the results to the engineering process that distinguishes economic engineering from more theoretical network economics.	108
7-2	Economic evidence or arguments influencing engineering decisions . .	109
8-1	Mechanism design viewed as a three step game of incomplete information.	120
8-2	Mechanism design where agents repeatedly play the game induced by the mechanism.	127
8-3	The challenge in wireless ad-hoc networks is that the cost of a link cannot be determined by the sender alone. Receivers are an integral part in reporting what power levels the sender must employ.	129

Chapter 1

Introduction

A fundamental characteristic of the Internet is the statistical multiplexing of traffic demands. Network links are provisioned to accommodate an expected traffic demand, not the maximum possible traffic demand. Thus, unavoidably at times, actual traffic demand can exceed link capacity. When this occurs, the network is congested – packets are dropped or delayed.

Historically, responding to congestion was primarily viewed as an end-host responsibility. End-hosts implement congestion control algorithms (such as the TCP congestion control algorithm) that adjust the volume of traffic associated with a flow in response to the perceived congestion state of the network. This cooperative form of congestion control generally allows senders to achieve a high-utilization level of network links while achieving a form of fairness that has historically been perceived as equitable.

Network operators, though, have always also strongly influenced the congestion state of a network. Their decisions regarding how traffic is routed, how links are provisioned, and how the network topology is constructed have always fundamentally influenced where congestion is likely to occur. Put another way, the decisions of network operators constructs the context and environment, the game, in which the cooperative congestion management of end-users occurs. Providers thus strongly influence what cooperative outcomes are even possible.

The network operators' ability to influence and shape the patterns and nature

of end-to-end communications has increasingly come under scrutiny. Many of the commercial network operators are perceived to have some amount of market power and thus their actions are less likely to be regulated by normal market forces. This raises the question of whether or not the congestion management and prevention actions they take are “normal,” “fair,” or “economically efficient.”

At the same time that the fairness of network providers’ congestion management actions are being questioned, the providers’ themselves are questioning whether the flow-based fairness of TCP is indeed “fair” given the changes in the traffic demands that are presented to the network. They argue that the myopic views of fairness taken by TCP does not match with many end-users’ actual views on fairness which often include some concept of fairness over longer periods of time. They also note that flow-based fairness has less relevance to actual real-world fairness if the number of flows differs significantly between end-users.

How congestion is managed and/or prevented is particularly important at this point in time as the traffic demands increasingly fill the capacity that is made available on networks. Peer-to-peer applications and video content are already responsible for a dominant percentage of traffic on the network. Both appear likely to grow rapidly in the future.

While future traffic growth will always be uncertain, all that really has to be agreed to by our readers is that traffic growth may exceed the capacity made available by the network operators at times. This already happens in many locations on the global Internet.

How this game between end-users and network operators plays out is important. It will impact end users’ perception of their network experience. It will impact their ability to adopt innovative new applications. From the network operator perspective, how congestion is managed and/or prevented will impact their ability to cost effectively operate their network. It will affect their incentives to invest in expanding capacity.

1.1 Documenting and collecting data

Our motivation in pursuing this thesis is that the previous roughly-stable status quo, or equilibrium, between cooperating end users and between end users and network operators in these congestion games is in the process of shifting. We are increasingly in uncharted territory where rationalizable norms of behavior have not yet been established by technical standards, regulatory rule, or cooperative agreement between all stakeholders.

What we have is a tussle [24] with conflicting interests of multiple stakeholders in a complex technical, economic, and regulatory space. No single dimension of this tussle space can be addressed without understanding its relationship to the other dimensions. Our thesis thus touches on more than just the technical; we employ language and tools from economics.

Some of the important questions in this tussle space include: Is the cooperative outcomes achieved by the ubiquitous TCP congestion control appropriate given the new mix of application behaviors? What are “fair” behaviors for applications over longer periods of time? What are acceptable operational practices for network operators to manage and or prevent congestion? Can traffic be blocked, degraded, or prioritized by network operators? What criteria are acceptable to differentiate traffic on? Which parties can commercial broadband providers charge (i.e. can BellSouth charge Google for access to its broadband customers)? What forms of charging are allowed and accepted? What interconnection policies are required between network providers? What role does government have in shaping and influencing market structures and operations of the commercial networking industry?

Given the complexity of this tussle, our thesis by no means “provides all the answers.” Rather we have sought to act as impartial observers, collecting and documenting facts that will inform this debate. We have sought to:

1. Collect trace evidence of and document the current congestion management and prevention behaviors of network operators.
2. Design and implement mechanisms that make it easier to monitor network op-

erators' behaviors in the future.

1.2 Analyzing new congestion management architectures

We also have sought to explore the hypothesis that a new design and re-structuring of the network architecture might help resolve this tussle. The underlying network architecture and protocols defines the very structure of these debates. These technical architectural decisions fundamentally determine the answers that are possible in the larger tussle. As architects and engineers, we can influence and change the tussle by changing the network architecture. We reiterate – the tussle arises not simply from the economic interests or policy and legal environments, but also because of the very design of the Internet architecture and protocols.

Our initial objective was to design and build a novel architecture, leveraging the tools of economics, that allowed for a better alignment of incentives between network stakeholders. However, we quickly discovered an existing approach that we believe already provides a compelling starting point – the re-ECN protocol and supporting mechanisms proposed by Briscoe [14, 17, 16].

The re-ECN protocol is a building block on which functionality and policies can be constructed. The protocol itself is technically simple – it is the policies, contracts, and traffic management behaviors that can be constructed with it that are interesting. What it creates is an architecture in which economically justifiable congestion management and prevention strategies can be devised.

The re-ECN protocol leverages and extends the Explicit Congestion Notification (ECN) [90] bits of the IP header to expose to observers along a path the sender's *expectation* of downstream congestion. The re-ECN protocol accomplishes this by adding a 're-echo' stage to the existing ECN protocol. Whenever a sender is notified of congestion on the downstream path (most likely by feedback received in a transport layer acknowledgment), the sender re-echoes that notification by 'blanking' a newly

defined RE flag (in normal operation the flag is set to 1) in the network layer header of a subsequent packet in the same flow.

In the current Internet, if regular ECN were widely enabled, one could observe whether congestion was occurring *upstream*¹ of a network interface by observing whether the congestion experienced (CE) bit had been set in the IP header of incoming packets. But it has never been possible within the network to know the *expected congestion* a packet will encounter on the rest of the path to its destination i.e. whether the CE bit will be set by some downstream router.

The RE flag exposes this expectation of congestion to the network. If expected congestion has been signaled by the sender through blanking the RE flag, any point in the network can observe whether congestion has already occurred upstream (i.e. the CE bit has already been set) or if congestion is still predicted on the remaining downstream path (i.e. the CE bit has not yet been set by any router).

Of course, without additional components, senders would have no natural incentive to accurately state their expectation of congestion on the path toward their packet's destination. This would make the information revealed by the re-ECN protocol untrustworthy in a non-cooperative environment. However, the incentives to accurately reveal this congestion information can be created by the addition of other architectural components that we will discuss later in the thesis.

If flows of traffic contained an indication of the expected downstream congestion, this would enable some interesting answers to one of the central questions that has been beguiling the network neutrality tussle – how might we distinguish between “good” or “acceptable” forms of discriminatory behaviors by network operators and “bad” or “detrimental” behaviors? The answer would be that behaviors that constrained actual congestion would be acceptable. This could also enable more efficient traffic management mechanisms and/or the establishment of more economically effi-

¹Note that the term ‘upstream’ can take different meanings in networking. When discussing flows of traffic, upstream refers to the routers that a flow has already passed through and downstream refers to the remaining routers on the path to the destination. (This is the terminology employed by Briscoe when discussing re-ECN so we mirror it here.) When discussing network topology, upstream can also describe the relationship between a tier-one and tier-two network provider. The tier-one provider would be considered the ‘upstream provider.’

cient contracts between network participants.

Let us clearly emphasize to the reader that this protocol and supporting architectural mechanisms has not been devised by us. The engineering work was led by Briscoe[17] while the theoretical foundations are primarily by Kelly [59, 47].

Our thesis instead details an analysis of the cases where the engineered system may diverge from the supporting economic theory. While we are generally supportive of the engineering design of this architecture, our contribution is important to scoping the boundaries of the economic claims.

We see our work as a constructive contribution to an architecture that has yet to be deployed. Indeed, an earlier analysis we conducted of re-ECN [7] contributed to some changes in the incentive mechanisms proposed together with re-ECN [14]. Our attack on the incentive scheme is noted in the protocol specification and mitigation techniques are discussed.

The economic claims associated with re-ECN are important to understand thoroughly. An engineering decision to implement and deploy the re-ECN protocol could be made based only on the technical arguments about the functionality it enables. However, the economic arguments are likely central to the engineering question of whether re-ECN should be deployed or not.

We therefore present an analysis of the following issues.

1. We consider the implications of topologically non-uniform economic costs of congestion. This issue is important to understanding whether or not there would ever be an economic incentive to treat aggregate flows of traffic with identical re-ECN markings differently.
2. We consider the relationship between re-ECN congestion marking and the economic notion of congestion externalities. This issue is important because it highlights cases where congestion markings placed by routers are not necessarily a signal of a negative economic externality.
3. We consider strategic behaviors by end-users in congestion games that represent scenarios that may arise on the Internet. These games are important to analyze

because they serve as predictors of potentially undesirable behaviors that may arise under re-ECN (given sufficiently sophisticated players).

1.3 Economic arguments and engineering design

Our thesis does require the reader to first comprehend the existing re-ECN protocol and supporting mechanisms and then understand our own further analysis of it. While we recognize that a common understanding and acceptance of re-ECN has not yet occurred, even readers that are skeptical of re-ECN may still find the challenge of integrating technical engineering and economic design to be intellectually worthwhile exercise.

Accepting a proposal such as Briscoe's or an analysis such as our own requires the research and wider networking community to be able to balance and integrate both technical and economic arguments and evidence in an engineering design process. We believe that the networking community is at the beginning stages of learning how to do this. Establishing a means of discourse is vital to the objective evaluation of what constitutes good architecture and good research for these techno-economic problems.

We readily acknowledge that engineering itself has always been a technical and economic activity. In networking, engineers ask: how do we trade off communication costs for storage costs in a system design? How can we improve the performance of a routing algorithm to reduce capacity costs? Indeed, one of the fundamental characteristics of modern networking – statistical multiplexing of packets on links – is born out of recognizing the cost advantages (along a variety of dimensions) of packet-switching over circuit-switched designs.

Engineers just do not tend to think of these dimensions of their work as economics; they claim it instead as a central component of what it means to do engineering. So, in a sense, engineers are already good at some economics. But this engineering tends to address a somewhat limited economic problem – optimizing a design given an engineering budget.

Economists and the theory community in computer science, on the other hand,

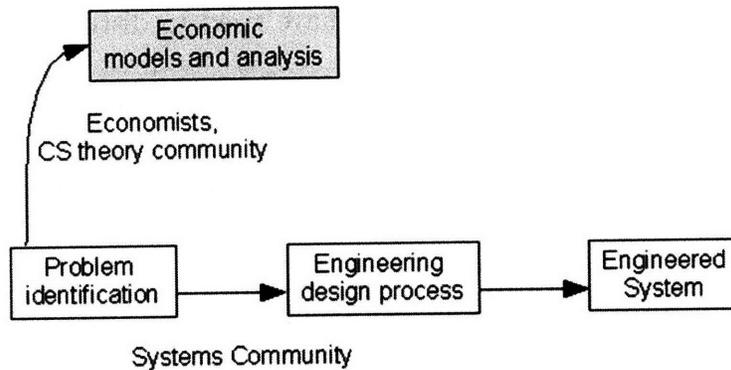


Figure 1-1: The engineering of networking protocols and architectures historically has not been strongly influenced by economic theory or analysis.

address a much wider selection of economic problems in networking. Indeed, networking has been a rich source of problems for these communities. But the rich set of network economic literature that has been produced has not influenced actual network engineering as much as it perhaps could or should have. (Figure 1-1.)

We believe this gap between economic theory and engineering practice needs addressing. Particularly in networking, if we want our designs to remain relevant in the real world, we believe we will have to increasingly *take responsibility for the economic systems that arise from engineering design*. This implies not just getting the technical details right, but also deliberately designing the engineered system so that it has desirable economic properties. In other words, we will have all the technical challenges of traditional engineering tasks, in addition to a responsibility for addressing the economics that are under the engineer's control.

This is exactly what Briscoe's technical architecture seeks to do. While it is based upon established theoretical work, it is not just a direct application of economic theory. But because it bridges the gap between theory and engineering, it is not as easy for either isolated community to understand its merits or limitations.

At the end of our thesis we have thus sought to understand some of the core misunderstandings between economic-minded and engineering-minded members of the networking community. Understanding and resolving these misunderstandings is

essential to making progress. We therefore have tried to distill the lessons we have learned from working at the intersection of economics and networking.

1.4 Summary

In short, what we want to do is take the reader from the very low level technical reality of what is occurring on today's Internet all the way through to an improved understanding of how to evaluate alternative network architectures which address both technical and economic requirements.

In writing this, we are reminded of Donald Knuth's perspective on what will hold the "tree of science" together as each individual area of study becomes more narrowly focused and specialized. He argues that "people in academic life are going to define themselves not by one specialty area, but by two sub-specialties that belong to two rather different main specialties."

Knuth argues that this means we will have "a web of interests, in which each person will serve as a bridge between different parts of the overall structure." He contends that this is important to integrating and leveraging the knowledge that is produced and discovered across the sciences.

We see this thesis as one of the bridges between the network architecture tradition in computer science and the economic community. This may not be an easy bridge to construct, but we see it as vital to the continued success of network research.

While we are optimistic that economics can provide a structured and systematic framework for reasoning about some of the hardest challenges in networking research, we also see a danger that the formal arguments and proofs of economics will result in researchers fooling themselves or inadvertently convincing others with a misapplied theory or reasoning. Just as statistics can be misleading, the rhetoric and formalism of economics has the potential to convincingly, but incorrectly, tip an engineering argument.

Any scientific and engineering discipline has this challenge. Richard Feynman commented on this in "Surely You're Joking Mr. Feynman" (pp 342-3):

...learning how to not fool ourselves of having utter scientific integrity is, I'm sorry to say, something that we haven't specifically included in any particular course that I know of. We just hope you've caught on by osmosis.

The first principle is that you must not fool yourself and you are the easiest person to fool....I'm talking about a specific, extra type of integrity that is...bending over backwards to show how you're maybe wrong, that you ought to have when acting as a scientist. And this is our responsibility as scientists, certainly to other scientists, and I think to laymen.

Thus as we begin this thesis exploring the intersection of economics and network architecture, we acknowledge to the reader that we could be wrong. There is plenty of room for debating what are valid evidence, good architecture, and sound design in this space. We welcome such debate.

Chapter 2

Motivation and the future of network congestion

The networking research community has long been grappling with the technical dimensions of congestion management and prevention. Many of the core challenges of networking research deal with these issues e.g. congestion control, routing, capacity provisioning, traffic shaping, etc.

While historically these have been technical problems, the economic dimension has become as important as the technical. One of the central challenges in preventing and/or managing congestion is the conflicting economic incentives among the network operators that carry traffic and the network participants that send and receive the traffic.

We strongly emphasize that this tussle over congestion management and prevention arises even in parts of the network that are non-commercial. We note this because it is easy for readers to assume that anyone examining such topics must be focused exclusively on a for-profit marketplace. We are not.

Any network faces the same challenges of allocating both the capacity that is available as well as the costs that arise from building, operating, and expanding networks. A well designed network architecture that accommodates conflicting interests over these decisions would be a benefit to any operational arrangement or regulatory framework. Network neutrality is not, and will not be, a tussle confined to for-profit

networks and their users.

2.1 Network capacity

Our thesis is premised on the notion that network capacity is a potentially scarce resource. We acknowledge that, to some readers, this is highly controversial. Such readers will object that network capacity, in most locations in the network, is not scarce most of the time. Such readers would argue, the cost of capacity is declining at a rate that bandwidth should be regarded as essentially "free" – any discussion of the economics of network capacity or congestion is wrongly accepting artificial constraints.

In other words, these readers reject the very premise motivating our work – that the current network architecture's different approaches to congestion management and prevention are an important topic to study. They argue that in the current Internet, any capacity constraints that do exist are artificially imposed by network operators. "If network operators just used technology X then this wouldn't be a problem." In other words, congestion is a result of what economists refer to as "artificial scarcity". Artificial scarcity is not "real" in the sense that the technology and means of production exists that could make network capacity abundant.

If this objection is true, then any *architectural* work that attempts to improve the technical and economic properties of congestion management may have little impact. Similarly, measuring and monitoring may be less a role for a scientist. Rather these roles might be better left to the regulatory regimes.

In our view, at the core of this objection is a debate about whether network capacity is indeed an abundant or scarce resource. We therefore briefly examine the history of this long running debate in the following sections.

2.1.1 Bandwidth is free

Particularly during the technological and economic boom of the 1990's, it was popular to argue that bandwidth would become virtually free [60]. In support of this

proposition, proponents pointed as evidence to the rapid technological progress made in the carrying capacity of the copper, fiber, and wireless physical mediums [48]. For the same capital investment today, over a hundred times the capacity can be bought as compared to five years ago. As the argument goes, if these trends continue, then bandwidth will be essentially “free” in the future compared to the cost of today’s technology.

However, care must be taken in the deductions that are drawn from this line of reasoning. Proponents of the “bandwidth is free” position are taking some rhetorical license to make a technological point. They are often emphasizing a shift in the relative cost of bandwidth compared to other components in a networked system [49]. In a technological and engineering debate over a networked system design such pithy statements become shorthand for a shared belief that for a given engineering budget, buying more bandwidth is cheaper, and will continue to be cheaper, than many of the technological alternatives for achieving a given performance objective.

Indeed, in many network engineering debates this position has been validated by experience. The costs of telecommunications equipment and bandwidth have historically declined rapidly. According to [87], the price of networking equipment has roughly dropped with an 80% learning curve (for every doubling in volume in equipment sold, the price of equipment declines to 80% of its previous price). This does not project the time frame over which the doubling or price declines will occur. It is simply an observation of the relationship between volume of sales and the price of equipment. The cost of bandwidth has been on a similar decline curve that is projected by some to continue [49].

Examining the “bandwidth is free” position further one finds that it is largely a hypothesis about the increasing technical capacity of links, not necessarily the cost of *operating* a network composed of such links. This is an important distinction because while buying network capacity in the form of routers or connections from an outside network provider may become cheaper with time, the aggregate cost of operating a network (is not necessarily declining as rapidly [18]. These operational costs include expenses such as buying bandwidth from upstream providers and the labor costs of

installing, maintaining and upgrading a network. One industry observer commented that there "is no Moore's law governing backhoes." Another noted that "if backhoes ignored Moore's law then service vans absolutely mocked it."

The economic reality that the costs of building and operating a network can exceed revenues is evident in the collapse of many network operators after the Internet bubble burst. The failed network operators were unable to service the debt that they had incurred in building out their networks. Their failure resulted in an oversupply of capacity that was sold for a fraction of its purchase price. (For a history of the boom and bust see Couper et al [26].)

Even today the installed supply of unused fiber, i.e. dark fiber, exceeds the demand for fiber in many places. But one has to be careful, as there is an obvious difference between installed fiber and an operational network. Just because some of the raw materials for a network are available does not imply that it is necessarily economically feasible to operate a network.

The operational practices of networks also tend to reinforce for many people the perception that network capacity is not scarce. With today's infrastructure backbone links are often lightly loaded. This "lightly loaded" perception is further reinforced by users' belief that their own usage patterns represent sporadic and low loading of the network (well below the transfer rates advertised with their broadband service.) With a mental picture of the state of the network as being often lightly utilized, many users have trouble understanding the claim that certain patterns of usage are costly. They perceive these claims instead as stemming from network operators' greed or incompetence.

2.1.2 Bandwidth is scarce

A competing conjecture surrounding network capacity for many years was that it was a very scarce capacity consumed by rapidly expanding traffic loads. Some famously claimed that traffic load was doubling every six months [26]. (For an analysis and debunking of these claims see [79].) Similar lines of reasoning were used to justify massive investment in capacity during the Internet boom years. However, the traffic

load did not reach the levels that were predicted at the time and consequentially this led to the demise of some major network operators [26].

Today predictions of sizable traffic growth are again appearing. Gilder[99] describes “exafloods” and “zettabytes” of data could come onto the networks.¹ The primary applications driving this potential demand involve video content and the peer-to-peer applications that distribute video.

Similarly, in Cisco’s System’s “Global IP Traffic Forecast and Methodology, 20062011” whitepaper[21], the projection is made that “consumer IP traffic will bolster the overall IP growth rate so that it sustains a fairly steady growth rate through 2011, growing at a compound annual growth rate (CAGR) of 46 percent and nearly quadrupling the monthly traffic run rate from 2007 to 2011.”

Odlyzko, while generally skeptical of exaggerated claims of network traffic growth, notes that the amount of data likely stored at the edges of the network dwarfs that which is actually transmitted. He notes that “most of those bits are redundant or duplicative, but ... that just a slight change in the velocity with which information circulates can have a large impact on Internet traffic.”[81]

While it is impossible to predict demand with accuracy, these scenarios are at least plausible.

2.1.3 Anecdotal evidence

The motivation for our thesis would be weaker if we needed to rely upon potentially inaccurate predictions about demand growth to create scarcity and hence economic motivation for our work. However, one does not need to speculate about future traffic loads to become convinced that network capacity is a scarce resource at some locations and is already constraining important network applications.

In a presentation to NANOG [62], the CTO of a major media broadcaster noted that they select their codec and video size based upon their belief about their typical consumer’s access pipe size. As a content producer they would like to send higher quality video. Consumers presumably would prefer the higher quality video as well.

¹A gigabyte is 10 to the 9th, an exabyte is 10 to the 18th, and a zetta byte is 10 to the 21st.

The constraining factor, then, for better Internet video today is network capacity.

Further evidence that network capacity is a primary constraint for innovative new applications was given by Norton [75]. He notes that “a few of the largest US ISPs are turning away ... n*10G [i.e. multiple ten gigabits per second] Internet video transit customers.” Norton noted one of the reasons given by a Tier 1 provider was that , “the network equipment currently deployed hasn’t been paid for and they would have to go back and argue for more (money) for a forklift upgrade.” This was striking given that the providers were turning away customers, not just potential peering partners. Not only was the Tier 1 provider evidently worried about the effect of handing off such large amounts of data to their other downstream customers, their own internal networks were not yet provisioned to handle the load.

We take these examples as evidence that potential, or desired, traffic loads are *already* ahead of available network capacity in the Internet. So while the capability of networking technology will likely continue to grow rapidly, the *deployment* of higher speed networks may not keep pace with the growth in traffic levels desired by end users and content and service providers. Congestion, or at least the potential of it, will remain a concern for the foreseeable future.

2.2 Measurements of traffic growth and patterns

Disappointingly little measurement data exists in the public sphere regarding how individual or aggregate patterns of traffic are changing over time. This is particularly true of broadband access networks – the very networks that are generally believed to be the ones mostly likely to experience network congestion.

One of our initial ambitions for this thesis had been to collect, aggregate, anonymize, and analyze data from industry partners that would provide insight into these changing patterns. The optimism with which we began that effort diminished with time. While providers initially indicated willingness to share data, at the end of day we were unable to acquire much of it.

We did not manage to enlist the support of any U.S. broadband provider. We did,

however, gain access to internal traffic data of KT, a Korean broadband provider. An analysis of the KT data is presented in [71].

The Korean market was interesting to examine as they have been one of the global leaders in broadband adoption. At the time of our study Korea topped the world in residential broadband penetration with 12 million of its 15 million households subscribing. They may serve as a bellwether of what other countries will experience as broadband adoption evolves. Some of the interesting points from our KT study included:

1. With the market approaching saturation, broadband subscriber growth rates had flattened out, dropping from 75.3% in 2000, to 11.2% in 2003, to 4.7% in 2004.
2. While revenue growth had slowed considerably, Internet usage continued to rise (Figure 2-1).
3. The growth in network usage forced KT to continually invest in expanding network capacity. According to their annual reports, KT invested over \$150M between 2000 and the time of our study in upgrading network capacity.
4. The aggregate traffic on KT's network nearly doubled every year between 2001 and the time of our study. (It has slowed since that point in time.)

We could not tell how much the rising traffic volumes translated into bottom line costs for KT or whether these were expected or unexpected costs. At the time of our study, KT had repeatedly indicated in public they would institute usage-based pricing in the coming years. According to news reports, the Korean public was (predictably) not receptive to this announcement. This would seem to indicate that, at the time anyway, KT believed that usage based pricing was necessary to cover their raising costs for handling traffic.

However, now in 2008, we cannot find any indication of usage based pricing on their web pages. According to their current promotional materials, "Megapass ADSL is the most representative broadband service in Korea, which is available across the

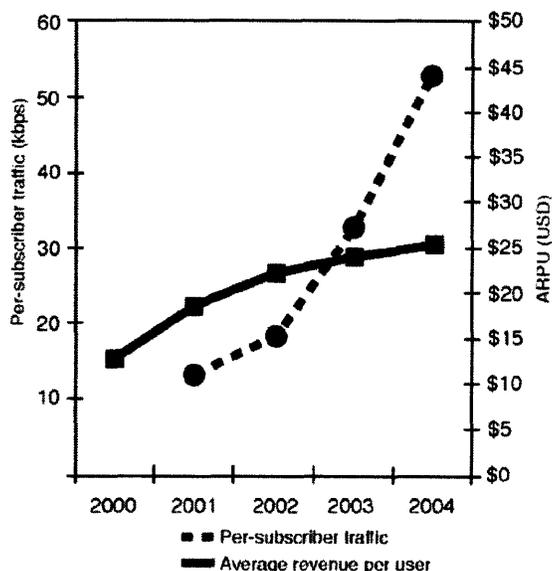


Figure 2-1: KT Traffic vs. Revenue Growth, Per Subscriber

nation. As a flat rate service, its performance vs. price is superior.”² So it seems that their most popular offering remains a flat rate service. It is unclear if they have instituted other technical means for limiting congestion from increased traffic loads.

Other researchers have collected limited data sets that provide some additional insight. But basic questions such as how much variation there is in individual user’s demands over time remain unanswered. Is it the case for instance that a “heavy-user” one month is a “average-user” the next or are “heavy-users” likely to remain heavy-users over longer periods of time? These longer term questions about the dynamic nature of usage patterns are important to questions about fairness, service plan design, and provisioning.

Andrew Odlyzko has been collecting and aggregating data that is often interesting. The following table 2.2 from Odlyzko[80] estimates the average amount of traffic sent per user in a variety of countries. His estimate of 3.0 GB of traffic per month per broadband user in the United States is in line with what we have confidentially heard as well. (Odlyzko’s estimates and data come from both official government statistics from places like Australia and Hong Kong or ISP cooperatives in the case of Japan.

²http://www.kt100.com/cy05en/Product/Service_contents.jsp?mn=1110&scode=1110&sname=Megapass%20ADSL

Australia	1.0
Western	Europe 2.3
Japan	2.6
U.S.	3.0
Hong Kong	17.0
South Korea	17.0

Table 2.1: Year-end 2007 estimates for monthly Internet traffic (GB per capita) [80]

In other cases he indicates that “figures for other countries are based on snippets of information of varying degrees of reliability, as well as confidential reports by some service providers.” [80])

These estimates are for the average user. What we have been told is that the tail of usage extends out past 150 GB of traffic per month for particularly heavy users. We present these numbers to give the reader a rough estimate of the wide variation in demand across the user population.

Figure 2-2 (generated by Odlyzko[82]) shows the traffic growth from 2002-2008 on a variety of networks. These networks range from universities to public Internet exchanges. The x-axis represents the bits/s transmitted by the network at the end of the time period. The y-axis represents the annual growth rate from the beginning of the study. The traffic loads at these sites is growing between 50 and 60 percent a year. (This growth may not be representative of the rest of the Internet. Odlyzko speculates that traffic growth at Internet exchanges, in particular, may be larger than on other networks.)

2.3 Summary

While network capacity is obviously never permanently fixed and can be expanded, *at any point in time* it should be treated as a limited resource for the purposes of an engineering design process and for the purpose of economic analysis. Capacity is a “scarce resource” in the purely economic sense of a resource where demand exceeds the supply if no price (or architectural mechanism such as TCP or the efforts of

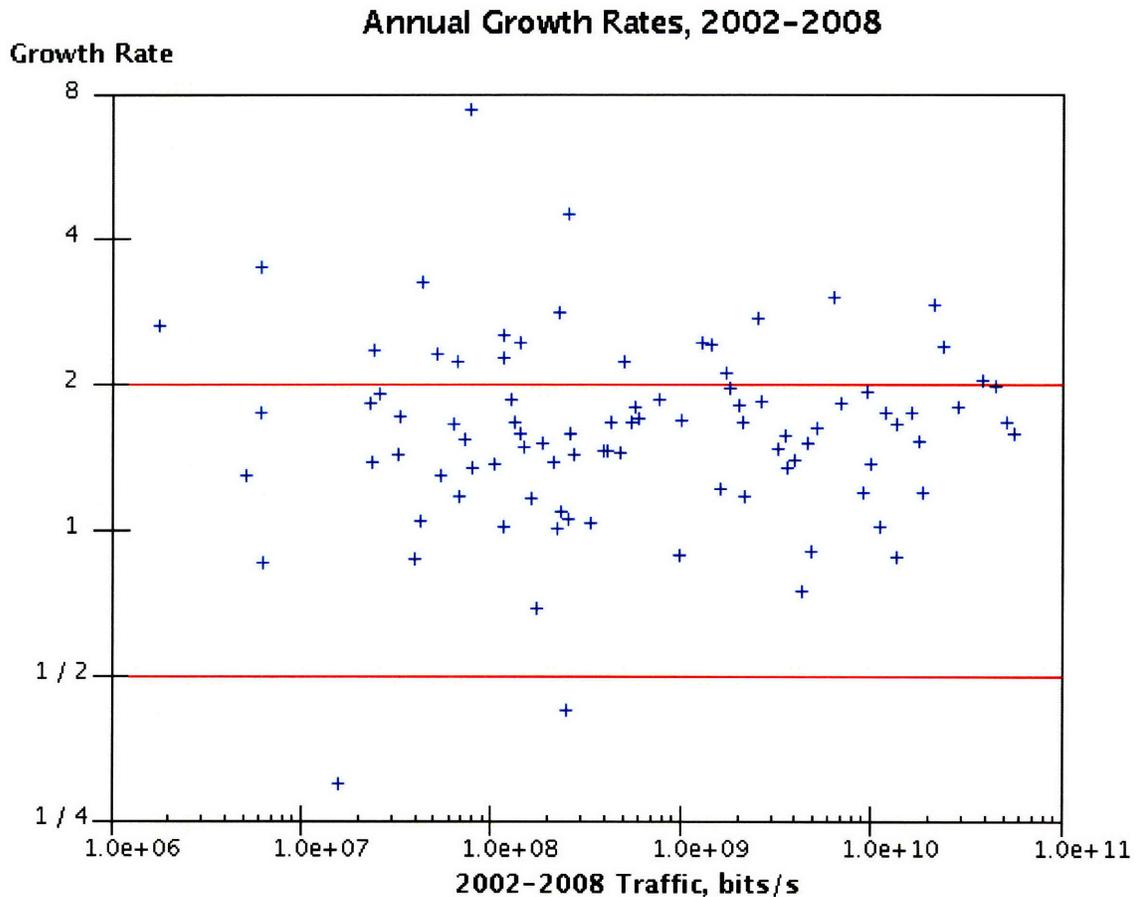


Figure 2-2: Traffic growth rates from publicly observed sites[80]

network operators) exists to regulate demand.

We therefore hope that both the data and arguments put forth within this chapter convince the reader that managing and or preventing congestion is, and will likely remain, a challenge in the future. While the KT example demonstrates that even if managing congestion is a challenge for network operators, moving away from flat-rate pricing as a solution appears to face strong end-user and customer resistance. What, then, are operators doing to manage and prevent congestion in their networks? We examine some of the responses in the next chapter.

Chapter 3

Documenting new congestion management techniques

Most network operators do not disclose the technical measures they take to manage traffic on their networks. Therefore probing a network is often the only way to discover and document an operator's policies. We specifically sought to understand ways in which congestion management was evolving. Most of our testing happened to involve the Comcast network. Fortunately, this ended up being a good broadband provider to test as they have been an early adopter of novel techniques for managing and preventing congestion. In this chapter we document some of the interesting findings from this testing. Some of our other studies, such as understanding the spoofing prevention policies or the TCP port blocking behaviors of network operators, are documented elsewhere [11, 10, 12]

3.1 Policy-based routing

Our first indication that Comcast had adopted some interesting network management techniques came from noticing that they appeared to be treating UDP and TCP traffic differently in parts of their network. We noticed this when mapping parts of the Comcast network using traceroute-based techniques.

Traceroute works by sending a series of packets constructed so that the IP proto-

col's time-to-live (TTL) field will reach zero at each of the routers along the path to the destination. Most routers will generate an ICMP TIME_EXCEEDED in response to such packets. The sender can thus reconstruct the path packets are traveling by noting each of the routers that sends back an ICMP packet.

A key thing to observe about the traceroute technique is that the TTL field that is decremented by each router is an IP layer field. We expected, and we believe most people would have, that traceroute would produce identical results regardless of the transport layer or application layer content in a packet.

However, this is not what we found. Consider the following paths that were stable over long periods of time (multiple months) between MIT and the Comcast broadband network noting in particular the third hop.

UDP path between Comcast and MIT

Tracing the path to 18.26.0.106 on UDP port 33434

- 1 c-3-0-ubr06.needham.ma.boston.comcast.net
- 2 ge-1-40-ur01.needham.ma.boston.comcast.net
- 3 **te-5-1-ar02.needham.ma.boston.comcast.net**
- 4 po-12-ar02.woburn.ma.boston.comcast.net
- 5 24.218.0.194
- 6 B24-RTR-2-BACKBONE.MIT.EDU (18.168.0.23)
- 7 MITNET.TRANTOR.CSAIL.MIT.EDU (18.4.7.65)
- 8 trantor.kalgan.csail.mit.edu (128.30.0.246)
- 9 atlas.lcs.mit.edu (18.26.0.106)

TCP path between Comcast and MIT

Tracing the path to 18.26.0.106 on TCP port 80 (http)

```
1 c-3-0-ubr06.needham.ma.boston.comcast.net
2 ge-1-40-ur01.needham.ma.boston.comcast.net
3      No response
4 po-12-ar02.woburn.ma.boston.comcast.net
5 24.218.0.194
6 B24-RTR-2-BACKBONE.MIT.EDU (18.168.0.23)
7 MITNET.TRANTOR.CSAIL.MIT.EDU (18.4.7.65)
8 trantor.kalgan.csail.mit.edu (128.30.0.246)
9 atlas.lcs.mit.edu (18.26.0.106)
```

The UDP-based traceroute tests from the client elicit a TTL expired response from all the hops. With the TCP-based traceroute tests between the same two locations, the third hop consistently does not send TTL expired messages. This is not a transient behavior. In all traces collected, the third hop does not send a TTL expired message for TCP traffic. This is true regardless of the TCP port targeted or the size of the TCP packet sent.

We found this fascinating because there is no way that we are familiar with to configure common routers such as Cisco or Juniper to respond differently to a $TTL = 0$ packet based upon the transport layer protocol type. The only possibility that we were aware of is that UDP and TCP traffic are traveling over slightly different router paths in Comcast's network.

At the time we suspected, but were unable to confirm, that this behavior resulted from Comcast's usage of bandwidth management appliances. When we were doing these traces we had already heard rumors that Comcast might be beginning to employ Sandvine's Policy Traffic Switch [94]. The evidence for this grew over time.

This choice of behavior is interesting because one might imagine that UDP-based flows would be *more* likely to be subject to traffic shaping and monitoring since

senders are less likely to adopt the cooperative congestion responses built into TCP senders. However, traffic shaping appears to be more common for TCP based applications because of the significant amounts of traffic associated with peer-to-peer applications and, increasingly, applications that transfer video over TCP.

Traffic inbound to the Comcast broadband subscriber appears to be similarly differentially treated based upon the transport layer protocol. (Notice in particular the six, seven and eight hops.)

UDP path between MIT and Comcast

Tracing the path to 65.96.238.168 on UDP port 33434

```
1 legacy26-0.default.csail.mit.edu
2 kalgan.trantor.csail.mit.edu
3 B24-RTR-2-CSAIL.MIT.EDU
4 EXTERNAL-RTR-1-BACKBONE.MIT.EDU
5 EXTERNAL-RTR-2-BACKBONE.MIT.EDU
6 ge-1-2-ar02.woburn.ma.boston.comcast.net
7 po-12-ar02.needham.ma.boston.comcast.net
8 te-9-1-ur01.needham.ma.boston.comcast.net
9 ge-0-1-ubr06.needham.ma.boston.comcast.net
10 65.96.238.168
```

TCP path between MIT and Comcast

Tracing the path to 65.96.238.168 on TCP port 80 (http)

```
1 legacy26-0.default.csail.mit.edu
2 kalgan.trantor.csail.mit.edu
3 B24-RTR-2-CSAIL.MIT.EDU
4 EXTERNAL-RTR-1-BACKBONE.MIT.EDU
5 EXTERNAL-RTR-2-BACKBONE.MIT.EDU
6      No response
7      No response
8      No response
9 ge-0-1-ubr06.needham.ma.boston.comcast.net
10 65.96.238.168
```

Again in these two samples the six, seventh, and eight hops along the path will send an ICMP `TIME_EXCEEDED` in response to UDP packets, but not in response to otherwise identical TCP packets. We reiterate that this is not an artifact of ICMP limiting options. This behavior was consistent for months at a time and independent of the probing order or other variations in the packets.

3.2 Noticing high rates of TCP connection resets

At the time we noticed the different behavior for UDP and TCP traffic, we had yet to construct any tests that definitively triggered traffic shaping or other behaviors. It only appeared to us that the packets were taking slightly different paths through the Comcast network. We assumed they were perhaps monitoring and inspecting TCP flows for measurement and analytic purposes – an interesting traffic management decision, but not something that would be controversial.

On August 17, 2007, however, *TorrentFreak*, a blog site that reports news about Bittorrent and peer-to-peer, noted that Comcast was blocking its customers from

“seeding” torrent files.¹ A client that is “seeding” a file has already downloaded or is in possession of a complete file. The seeder therefore uploads file content only.

Remaining a seeder is sometimes considered to be an altruistic action because the seeder gets no direct benefit from uploading the file content to others – they already have the entire file. Indirectly benefits do accrue to seeders. Some torrent sharing sites, for instance, require members to have a upload to download ratio greater than one. Therefore members have to remain seeders for some amount of time to remain in good standing within their communities.

Initially, it was unclear technically how Comcast was blocking their customers from seeding torrent files. They were obviously not simply adopting an outright block of the common bittorrent ports e.g. TCP ports 6881-6999. The clients could send and receive traffic on these ports up till the point they had the entire file already downloaded to their box.

Topolski was the first person that we are aware of that gave a public explanation of this behavior.² He noted that the number of bittorrent connections that were experiencing a TCP reset appeared abnormally high.

According to the official protocol specification[89], a TCP reset packet is normally generated by a receiver under three conditions.

1. If the connection does not exist then a reset is sent in response to any incoming segment except another reset. In particular, SYNs addressed to a non-existent connection are rejected by this means.
2. If the connection is in any non-synchronized state, and the incoming segment acknowledges something not yet sent (the segment carries an unacceptable ACK), or if an incoming segment has a security level or compartment which does not exactly match the level and compartment requested for the connection, a reset is sent.
3. If the connection is in a synchronized state, any unacceptable segment (out of

¹<http://torrentfreak.com/comcast-throttles-bittorrent-traffic-seeding-impossible/>

²<http://www.dslreports.com/forum/r18323368-Comcast-is-using-Sandvine-to-manage-P2P-Connections>

window sequence number or unacceptable acknowledgment number) ... and an incoming segment has a security level, or compartment, or precedence which does not exactly match the level, and compartment, and precedence requested for the connection, a reset is sent.

The effect of receiving a reset packet is defined in the protocol specification as well. If a TCP receiver is in “the SYN-RECEIVED state and had previously been in the LISTEN state, then the receiver returns to the LISTEN state, otherwise the receiver aborts the connection and goes to the CLOSED state. If the receiver was in any other state, it aborts the connection and advises the user and goes to the CLOSED state.” [89]

On the Comcast network, however, the reset packets were being generated not by end hosts, but rather by Comcast itself. These TCP resets are non-trivial for Comcast to generate. For a reset packet to be accepted by a sender as valid, the source address, source port, destination address, and destination port all have to match an existing connection and the sequence number of the RST has to be within the connection’s current TCP window. If all these conditions are met, the TCP connection is ended and no more data is transmitted.

What Comcast was doing was actively monitoring TCP flows and injecting TCP reset packets for any connections of the client that involved a file they already completely downloaded. Topolski explained in a follow-up Internet posting the method he used for detecting the Comcast resets.³ His procedure relied upon the background knowledge that the number of reset connections under “normal” network operations of the bittorrent application is rather low. (A control case that establishes such background knowledge might be a bittorrent client’s behavior on a non-Comcast network.) A bittorrent application on the Comcast network would experience a large number of TCP resets.

We replicated Topolski’s procedure and confirmed that we were seeing the same behavior on a number of Comcast connection points. These included our home Comcast connection as well as the Comcast connections at two public libraries in the

³<http://www.dslreports.com/forum/r18901881->

Boston area. (We were surprised that the public libraries' connections were managed using the same techniques as our home connection given that these were presumably business-class broadband connections.)

While fairly conclusive evidence, without a packet trace from the remote clients there was the slim possibility that the remote clients were in fact generating the RST packets. Indeed, Comcast at the time seemed to be publicly denying they were responsible. While we can not say whether the following snippet from a “talking points” memo to Comcast service representatives is definitively authentic, it is consistent with what most reported hearing from Comcast.⁴

If a customer asks: I read that Comcast is limiting customer access to BitTorrent. Is this true? Respond: No. We do not block access to any applications, including BitTorrent.

To conclusively determine the origin of the RST packets, we set up a network test involving bittorrent where we controlled both the Comcast bittorrent clients and the remote clients. By capturing packets on both ends of the TCP connection, we were able to definitively conclude that the RST packets were indeed being injected by Comcast. (MIT has a direct connection to Comcast, as indicated by the traceroutes above, so we knew that Comcast must be the party injecting the packets.) We did not publish our findings at the time, but others such as the EFF publicized an identical methodology and results [32].

The definitive signature of the Comcast injected RST packet was a series of RST packets with their sequence numbers spaced apart by 12503 bytes. The motivation for the RST sequence will be explained below. Both TCP ends are sent the pair of RST packets. (Table 3.1)

⁴<http://consumerist.com/consumer/leaks/comcasts-we-dont-throttle-bittorrent-internal-talking-points.php>

Comcast	→	MIT	SYN
MIT	→	Comcast	SYN, ACK
Comcast	→	MIT	ACK
Comcast	→	MIT	Bittorrent handshake message
Comcast	→	MIT	Bittorrent bitfield message
MIT	→	Comcast	RST, seq
MIT	→	Comcast	RST, seq+12503

Table 3.1: Packet trace depicting Comcast injection of multiple TCP RST packets

3.3 Building a model of Comcast congestion management

The first question we had was why would Comcast be sending a series of RST packets. Sometimes these RST packets were not valid as they are outside the sequence window of the existing TCP connection.

The likely explanation for this behavior is that the “deep flow inspection” occurs out-of-band on a copy of the packets. When a packet arrives at a router is queued both for transmission on an output queue as well as copied and inspected on a potentially slower processing path.

This is significant because it implies that the box does not directly block or drop packets. The TCP connection state could potentially “run away” from what was being tracked by the middle box. If the ACK from the remote host for a sequence number X arrived before the RST with the same sequence number the RST packet would be disregarded as invalid. Thus by sending a series of RST packets, the middle box has a better chance of injecting a valid RST packet into the TCP communication stream. This is identical to the operational model postulated for the “Great Firewall of China.” [25]

The particular sequence number 12503 bytes into the future of the byte stream does not seem particularly significant. (We do note, though, that it happens to be a prime number.) It is a bit more than eight 1500 bytes packets into the future of the TCP stream.

3.4 Communication patterns triggering resets

Our next goal became to probe some of the rules that triggered these Comcast connection resets. There were a number of possibilities we investigated. Bittorrent clients communicate that they have the entire file already downloaded during two key protocol exchanges.⁵

The first exchange occurs when clients communicate with a “tracker.” One of the purposes of trackers is to coordinate the peer list of clients that are participating in a given torrent. When a client makes a request of the tracker it can include as one of the request parameters an event “completed” message. This indicates to the tracker that the client has successfully downloaded the entire torrent file.

The other protocol exchange that indicates a client has the entire file is the sending of the “bitfield” message. This message is sent after the handshake sequence is exchanged between two communicating bittorrent clients. The bitfield messages have the following format.

```
bitfield: <len=0001+X><id=5><bitfield>
```

The bitfield messages represent the pieces of a torrent file that have been successfully downloaded by a client. Bits that are zero indicate a missing piece, bits that are one indicate that the client has a valid and available piece. A client that has a completely downloaded file will therefore send a bitfield message to corresponding clients that consist of all ones i.e. `0xff`.

What we believe Comcast has is a rule based system that attempts to pattern matches certain protocol exchanges. The question then is what patterns would trigger the Comcast box to send connection resets.

We present below some of the patterns that we were able to detect. We find them to be interesting because they demonstrate characteristics of the Comcast box such as how long flow state is kept and how robust or fragile the patterns are to certain perturbations in the bittorrent protocol exchange.

All of our testing below was done by writing our own bittorrent pseudo-clients.

⁵<http://wiki.theory.org/BitTorrentSpecification>

This enabled us to carefully script the protocol exchanges. All of our testing was done between a node on the Comcast network and the MIT network.

3.4.1 Bitfield message reset

During the period of our testing, a Comcast RST could be generated by engaging in the following protocol exchange. (Other bittorrent exchanges have been reported to trigger the RST injection as well.) The key message in this protocol exchange is the bitfield message that indicates a complete file is hosted on the client. As soon as that message had been sent, Comcast would inject the TCP RSTs. (Table 3.2)

Comcast	→	MIT	SYN
MIT	→	Comcast	SYN, ACK
Comcast	→	MIT	ACK
Comcast	→	MIT	Bittorrent handshake message
Comcast	→	MIT	Bittorrent bitfield message

Table 3.2: Sequence of packets that would reliably trigger the Comcast injection of TCP RST packets.

At the time we did the testing in 2007, this sequence of packets would generate a Comcast RST every time we sent it. The RST injection did not appear to be time-of-day or load dependent. We were able to trigger the RST from three different Comcast locations with identical results at each location.

The following TCP RST packets would be received at our Comcast host upon completing the sequence of packets above. (Table 3.3)

MIT	→	Comcast	RST, seq
MIT	→	Comcast	RST, seq+12503

Table 3.3: Injected packets received at the Comcast host.

A similar pair of packets would also be received at the MIT host. (Table 3.4)

Comcast	→	MIT	RST, seq
Comcast	→	MIT	RST, seq+12503

Table 3.4: Injected packets received at the MIT host.

3.4.2 Classifier robustness

The next question we explored was how robust the traffic classifiers patterns were to perturbations in the data sent. Would the classifier identify flows that were not valid bittorrent protocol exchanges? Would the classifier identify protocol exchanges that were valid but somewhat atypical?

While on one hand these questions could be regarded as largely academic and not of much importance, this probing illuminates the fragility of classifying traffic flows by static patterns. It demonstrates that it is relatively easy to make minor changes in the end-host clients to evade detection.

Granted, it is likely easy to add additional patterns to the traffic classifier. But as the community’s experience with intrusion detection systems has shown, continually expanding pattern databases presents scalability and maintenance challenges. Each one of the following would likely require additional rules or a refinement of the rules at the classifier.

The Bittorrent handshake message consists of the following protocol fields.

handshake: <pstrlen><pstr><reserved><info_hash><peer_id>

1. The pstr is a fixed protocol identifier “BitTorrent protocol”. Merely changing the capitalization of this field to ”bittorrent protocol” would evade the classifier. This is an obviously small change that would be trivial for clients to make.
2. Since the Bittorrent protocol operates over TCP the bittorrent message boundaries are found within the data stream by the end hosts. Clients are indifferent to how these messages are fragmented at the network layer. The Comcast network, though, expected individual Bittorrent protocol messages to be within individual IP packets. Combining the bittorrent handshake message with the bitfield message would evade the classifier.

3. Similarly, if the bittorrent handshake message was fragmented over two IP packets, the Comcast RSTs would not be injected.
4. The RSTs could also be triggered if the `<reserved><info_hash><peer_id>` were omitted from the handshake message. It appeared that the test of the handshake message involved only identifying the `<pstrlen><pstr>` fragment. This is an example of a non-valid bittorrent exchange that caused a RST injection.

3.4.3 Flow state timeouts

The next question we explored was the flow state timeouts the Comcast classifier employed. Clearly the classifier needed to keep state about TCP flows until it could either classify a flow as being one that it was not interested in following or one that it was interested in following. The selection of these timeouts is important given the likely large number of concurrent flows through parts of the Comcast network.

The first timeout we sought to identify was how long the classifier would hold state about a connection where it had only seen the three way TCP handshake. Through a quick binary search of the timeout space we determined that the timeout value was 20 seconds. If the bittorrent handshake message was seen before 20 seconds followed by the bitfield message, the Comcast RSTs would be injected. (Table 3.5)

Comcast	→	MIT	SYN
MIT	→	Comcast	SYN, ACK
Comcast	→	MIT	ACK
Pause 19 seconds			
Comcast	→	MIT	Bittorrent handshake message
Comcast	→	MIT	Bittorrent bitfield message
MIT	→	Comcast	RST, seq
MIT	→	Comcast	RST, seq+12503

Table 3.5: TCP sequence demonstrating that flow state was kept after the TCP handshake for 19 seconds.

If the connection was paused for 21 seconds before the bittorrent handshake message was sent, the Comcast RSTs would not be sent. This provides an obvious evasion

technique for end-hosts. The traffic classifier could extend the timeouts and monitor TCP connections for longer, however these timeouts are likely selected to limit the amount of memory resources required to keep connection state on the traffic classifier. Longer timeouts would require more memory resources and impact lookup algorithms that match incoming packets to existing flows. (Table 3.6)

Comcast	→	MIT	SYN
MIT	→	Comcast	SYN, ACK
Comcast	→	MIT	ACK
Pause 21 seconds			
Comcast	→	MIT	Bittorrent handshake message
Comcast	→	MIT	Bittorrent bitfield message

Table 3.6: TCP sequence demonstrating that flow state timed-out after 21 seconds. No RSTs injections could be generated for any trace that started with the above pattern.

If the Bittorrent handshake message was sent immediately following the TCP three way handshake, the traffic classifier kept track of the flow state for significantly longer. Again using a binary search of the timeout space we discovered a five minute timeout on the flow state. Completing the triggering Bittorrent sequence in less than five minutes would cause the Comcast RSTs to be injected. (Table 3.7)

Comcast	→	MIT	SYN
MIT	→	Comcast	SYN, ACK
Comcast	→	MIT	ACK
Comcast	→	MIT	Bittorrent handshake message
Pause 299 seconds			
Comcast	→	MIT	Bittorrent bitfield message
MIT	→	Comcast	RST, seq
MIT	→	Comcast	RST, seq+12503

Table 3.7: TCP sequence demonstrating that flow state was kept after the bittorrent handshake for 5 minutes.

We did not try to overload the traffic classifier, but we did wonder how it would react to a very large number of flows being opened through the first four packets exchanged in the sequence above. Potentially, the traffic classifier would have gracefully degraded under heavy load and this would have been a non-issue.

3.5 Comcast's motivations

Comcast eventually acknowledged their employment of the RST based traffic management technique [34]. Likely because of public pressure and FCC scrutiny, they have announced they will drop this particular approach [40].

But the question is why did Comcast turn to this approach in the first place? The answer lies in the challenges of managing traffic on a cable-based broadband network. In particular, the very architecture of the network is designed with an expectation that more traffic will be flowing toward users and comparatively little traffic will be uploaded to the rest of the Internet.

The problem for Comcast was how to respond when the traffic behavior from part of their subscriber base began diverging from the expected behavior. This is a general problem in that the time scales on which traffic behavior can change are far more rapid than the time scale on which network provisioning decisions/investments can be made.

The Comcast network is, as is the case with most cable broadband systems, a hybrid fiber-coaxial network. The first link is from a subscriber's cable modem over a shared coaxial network to a cable modem termination system (CMTS). Data is transmitted over this link using one of the DOCSIS standards. Everything past the CMTS is a fiber optic link. (See Figure 3.5.)

The DOCSIS standards make use of selected frequency bands on the coaxial cable for transmitting data both upstream and downstream. The maximum usable speeds for both upstream and downstream traffic are indicated in Table 3.8. (Note the maximum usable speeds are lower than the synchronization speeds that the reader might be familiar with from other publications.) Most cable networks today make use of either the DOCSIS 1.1 standard or the DOCSIS 2.0 standard.

Comcast reported in their recent FCC filing [34] that the average number of households served by a node was 450 homes/users. We suspect the average number of actual broadband customers per node may be different than the average number of households served by a node. But this still provides a ballpark figure for understanding the

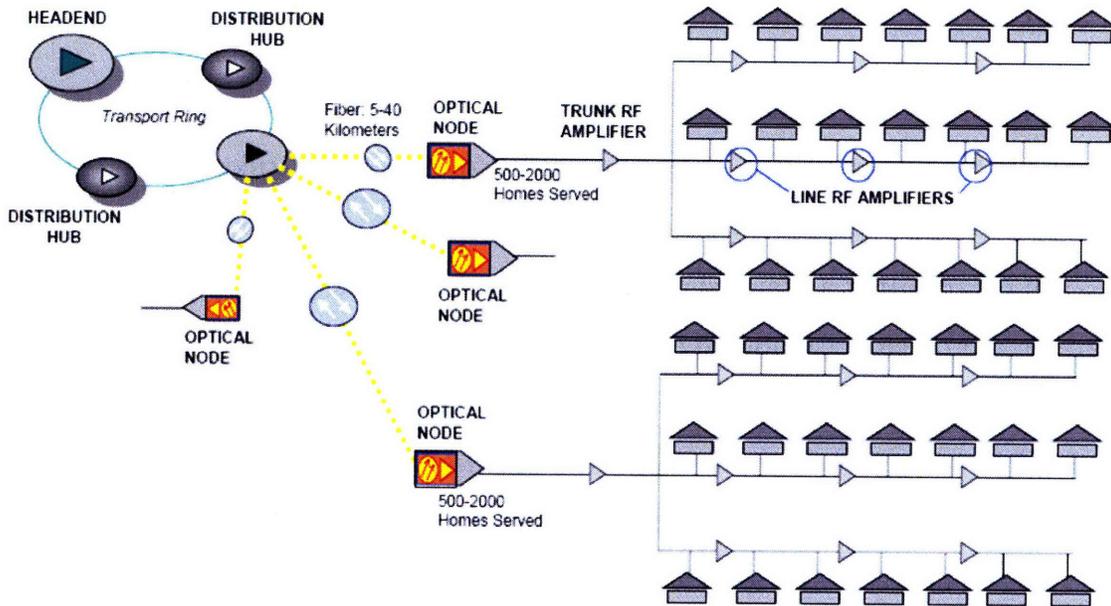


Figure 3-1: HFC Network Diagram

Version	Downstream	Upstream
1.x	38 Mbit/s	9 Mbit/s
2.0	38 Mbit/s	27 Mbit/s
3.0 4-channel	152 Mbit/s	108 Mbit/s
3.0 8-channel	304 Mbit/s	108 Mbit/s

Table 3.8: DOCSIS Speeds: Maximum usable speed

number of subscribers that share an upstream and downstream channel in at least parts of the Comcast network.

For the purposes of easy estimating, assume that 300 subscribers were sharing 9mbps of upstream capacity in at least some parts of the Comcast coaxial network. The only thing Comcast could guarantee, assuming capacity were shared equally, would be that each subscriber could get 30 kbps of upstream capacity (less than a 56.6kbps modem). Given this environment, there are three challenges that arose for Comcast:

1. A mismatch between the expected and actual aggregate traffic patterns
2. Short-term unfair allocations of capacity among subscribers

3. Long-term unfair allocations of capacity among subscribers

Each of these represents a distinct problem. The actual **aggregate** traffic pattern might match the expected traffic patterns but Comcast could still have concerns over which subscribers are contributing to the aggregate. Conversely, subscribers might be each contributing a “fair” share to the aggregate traffic, but Comcast might still perceive a problem with how to deal with a divergence between actual and expected aggregate traffic growth.

In Comcast’s case, though, they likely faced all three problems – the aggregate traffic growth was exceeding their expectations and some individual subscriber traffic over both short-term and long-term periods did not match one they believed to be fair or desirable for their customers.

3.5.1 Mismatch between actual and expected traffic patterns

The most common contract Comcast had signed with subscribers was for 6mbps down and 384kbps up maximum bitrate. This is an order of magnitude more upstream capacity than the equal 30 kbps share. Clearly, Comcast had built into their offering an expectation that subscriber traffic would be periodic. When actual traffic patterns of some subscribers began to diverge significantly from the expected traffic patterns, Comcast faced a challenge of how to respond.

Ten subscribers out of 300, or three percent, sending at 300 kbps upstream could utilize a full third of the upstream network capacity. Many of the peer-to-peer applications are capable of sustaining these upstream rates for long periods of times. While higher utilization might not directly lead to packet drops, it can cause increased delay as different subscribers contend for the available upload slots.

Even without a direct impact on subscriber’s perception of the quality of the network, many operators take the load on the network as a signal that more capacity needs to be installed. Thus unexpected growth in actual traffic can undermine the schedule of planned investments in expanding network capacity.

3.5.2 Myopic fairness of capacity allocation among subscribers

The second challenge facing Comcast was that web browsers and email clients of some customers might not be able to actually grab even their “equal” share of network capacity if they were competing with applications that open a larger number of flows. As is well known, TCP fairness is a per-flow fairness concept [16]. What Comcast may desire though is some semblance of fairness of capacity allocation across their subscriber base.

If 100 subscribers each opened one TCP connection to a remote host, each would get roughly 90 kbps of the 9mbps upstream capacity. However, if ten subscribers opened ten TCP connections each, they would collectively grab roughly 4.7mbps, or 470kbps apiece leaving the subscribers that opened only one TCP connection with 47kbps each.

3.5.3 Long-term fairness of capacity allocation among subscribers

TCP fairness has always been a rather myopic definition of fairness. It defines fairness in terms of the bitrate a flow will get at a point in time. Implicitly, a long-term notion of fairness could follow from this that a flow is “long-term TCP fair” if it is fair sharing at each moment in time.

However, there is not general agreement that this is the right notion of fairness [16]. If ten subscribers continuously had one TCP connection to a remote host, while three different groups of 90 subscribers each took turns occupying the remaining 90 sending slots, should each get 90kbps of the available 9mbps? Or should the ten subscribers that are continuously transmitting get a reduced rate?

Compounding this problem further is that peer-to-peer applications open a larger number of TCP connections to remote hosts and keep them open for relatively longer periods of time. So potentially ten subscribers have open ten connections each and are on for longer periods of time. Subscribers that are on only periodically and only open a small number of TCP flows at a time will send a comparatively small amount

of traffic.

The adoption of per-month volume limits by some network operators suggests that a longer-term of fairness is important to providers. The acceptance by broadband subscribers, such as Cox Communications subscribers in the United States⁶, suggests that at least some customers find such limitations to be acceptable.

3.6 Comcast's response

Comcast choose to respond not by per-month upload limitations though, but rather by monitoring selected applications that they had identified as generating a particularly heavy load of upstream traffic. (While it is possible they were monitoring other applications than peer-to-peer, only the peer-to-peer applications have been definitively identified by the wider community.)

They likely perceived the solution they had been using as convenient because they could deploy the box as an addition to their infrastructure without needing to make any changes to the cable modem or CMTS infrastructure. However, because of public pressure, Comcast has announced they will no longer use this technique. They will instead rely upon the ability of the CMTS to schedule the upload slots of cable modems. They will simply limit subscribers that generate heavy loads on their network.

But why didn't they take this approach in the first place? We are not privy to their decision making process, and there are likely many different factors that influenced their decision. But we can guess at some reasons.

Many of Comcast's customers are likely families where different family members make use of rather different sets of applications. Older members of a family may be more likely to upload files to their workplace while younger members may be more likely to employ peer-to-peer applications. If Comcast triggered a general upload limit on a cable modem, every member of the family's applications would be impacted. By selectively targeting peer-to-peer, Comcast may have hoped to avoid impacting the

⁶<http://www.cox.com/policy/limitations.asp>

bill-payer's applications. Comcast may have been trying to limit the load from peer-to-peer applications without limiting the ability to upload presentations or photos. While this is a subjective judgment about what application traffic to prioritize, it may indeed correspond to what a large set of customers would choose. However, no actual choice exists for customers in the residential subscription plans.

3.7 Summary

Longer-term Comcast will deploy the DOCSIS 3.0 standard starting in 2008.⁷ This will increase the amount of both upstream and downstream capacity and help alleviate some of the perceived problems with scarce upload capacity.

We expect, however, that congestion and congestion management techniques of network operators will remain a controversial issue for the foreseeable future. We have developed infrastructure and tools to make monitoring and testing easier in the future. These tools are part of our Netwatching Research Project which we describe elsewhere.⁸ This infrastructure makes it easy to share packet traces, capture packet traces at cooperating remote end-hosts, and inject packets into the network at remote locations. This is an open-source collection of tools and services.

⁷<http://arstechnica.com/news.ars/post/20080108-comcast-100mbps-connections-coming-this-year.html>

⁸<http://www.netwatching.org/>

Chapter 4

Limitations of the current Internet architecture

While the last chapter focused upon Comcast's controversial techniques for managing congestion, such challenges and fairness questions will continue to arise in other networks. Indeed, we have detected Cox Communications, another broadband provider in the Midwest of the U.S, utilizing the same techniques. This is particularly interesting given that Cox already had in place explicit volume-based upload and download limits. (This suggests that network operators may not view volume-based limits as a viable substitute for directly limiting peer-to-peer traffic.)

We believe that network operators and end-users *lack the architectural support for making fundamentally better economic decisions regarding congestion management*. There is room for an engineering effort such as re-ECN, which is informed and built upon an economic theory, to significantly improve the network architecture. The rules/policies/mechanisms that are possible *given the limitations of the current Internet architecture* lead to conflicting incentives between network operators and end users.

In this chapter, we explore the roots of this conflict. We examine the relationship between the costs that arise in the current Internet architecture and how capacity is allocated. This exploration is not a formal mathematical model of the underlying economics. While such formalism is very useful in conveying some economic argu-

ments, not all economic arguments have to adopt this form. We target this instead at engineering readers.

4.1 Costs

There are the large sunk and fixed costs associated with building and operating a network infrastructure [91]. These are the costs associated with buying routers, optics, wires, and computers, paying the operational staff, maintenance, etc. All of these tend to represent a significant cost to network operators [91].

There can be (but do not have to be) usage costs that are a function of an aggregate traffic load [91]. These costs arise, for instance, if a network operator has a usage sensitive backhaul contract for their connection a larger Internet provider. For some operators, these usage costs can be significant [18].

The next major category of costs that arise in networks are congestion costs. Congestion costs are the negative externalities associated with traffic being delayed or dropped due to traffic demands exceeding the capacity of a network [65, 29]. Congestion costs are generally considered to be the only marginal costs in networking [29]. For instance, any marginal costs associated with slightly larger power consumption while sending traffic are generally ignored as negligible

Congestion costs only directly impact those that send and receive network traffic. Network operators, however, indirectly care about congestion costs because they risk losing customers over the long term if the accumulation of congestion costs drives the value of network service below what the customer is willing to pay for a subscription. (These indirect congestion costs are the costs Comcast was arguably attempting to manage.)

Even for network operators that do not have directly paying customers (such as a university network) the network operator still has an incentive to care about the distribution of capacity and the congestion and costs among its users. This incentive arises from the network operator's obligation to "run a good network." They are, in a sense, benevolent operators attempting to maximize the social welfare.

4.2 Decision mechanisms in current Internet

Given these costs, we now consider how network capacity is allocated and how the network costs are recovered/managed in the current Internet. What rules/policies/mechanisms can determine how costs are allocated/recovered/managed? What rules/policies/mechanisms can determine how capacity is shared?

We classify the different ways in which capacity is allocated and costs are recovered in the current Internet architecture into three categories:

1. Cooperative mechanisms
2. Internal decisions made by network operators
3. Market-based mechanisms

Cooperative mechanisms are mechanisms that induce a cooperative game. Participants willingly “play by the rules” of such games – no direct penalties arise if users choose to strategically pursue their own self-interests. The outcome of cooperative games is not directly driven by economic criteria, but are rather determined by the structure of the cooperative mechanism.

Market based mechanisms are ones where the allocation of network capacity is determined through a “price” or “incentive” system. This does not necessarily imply “money” exchange as engineers might commonly envision. A price or incentive system, rather, is any system that coordinates the allocation of goods through the establishment and trading of things that are valued by the participants.

The third category that we consider are the capacity and cost allocation decisions that are internal decisions made by network operators. In economic terms, these are “administrative” or “command decisions” made by a firm [13]. In many ways, these are often the most interesting decisions and the decisions that receive the most scrutiny and debate in the wider tussle discussion. Regulators, for instance, ask whether network operators that have some degree of market power are making internal decisions that are beneficial to consumer or social welfare. Are these decisions economically defensible?

	Cooperative mechanism	Network operator's internal decision	Market-driven decision process
Flow-based congestion control	✓		
Aggregate traffic shaping		✓	
Traffic engineering		✓	
Capacity provisioning		✓	
Peering	✓		✓
End-user subscription			✓
Interprovider contracts			✓
Topology layout		✓	

Table 4.1: Capacity allocation and cost/control recovery in current Internet

4.2.1 Classification

Our classification of Internet functions is presented in Table 4.1. An interesting question is how and why did we end up with these particular divisions of responsibilities? This question could be answered by noting relevant decisions in the historical record of the engineering design process that led to the current network architecture. While this would be a valid answer, we contend that economics provides a more insightful answer.

Namely, we observe that each decision (congestion control, traffic shaping, etc) is made by the parties that have/share the most information about the underlying economics associated with each decision. This does not imply that parties making the decision have “perfect” information about the economic costs or values. Indeed, in the case of the aggregate traffic shaping (an internal decision of network operators), the network operators in today’s Internet have adopted heuristics that only roughly capture part of the economic costs.

However, the degree to which economic costs are visible and known has considerably influenced whether cooperative decisions, internal decisions, or market-driven decisions are made. Flow-based congestion control is a cooperative mechanism among end-users precisely because only end-users have the information about the shared congestion state of the global network. The aggregate traffic shaping, traffic engineering, and capacity provisioning are undertaken by the network operators because

the economic information required to make each of these decisions is only available to network operators. (We examine the economically relevant information for each of these decisions in the following subsections.)

The decision processes that are actually market-driven are possible precisely because sufficient economic information required by market participants is known. However, there is often little architectural support for coordinating the exchange of economically-relevant information in the current Internet. This is the root of the current tussle in our opinion.

In the case of Comcast incident, the tussle is particular pronounced because the aggregate traffic shaping mechanism (an internal decision of Comcast) conflicts and interferes with the flow-based congestion control of end-users (a cooperative mechanism among end-users). The decisions are not only not coordinated, but are actually in conflict. The cooperative congestion control mechanism adopted by end-users suggests the users can send additional traffic while the aggregate traffic shaping mechanism of Comcast injects TCP resets to curtail the sending rate of the user.

In the following subsections we briefly examine the major decision processes that impact how traffic flows over the Internet. We identify the parties that participate in each of the decisions, discuss the economically relevant information used to make each decision, and discuss how the decisions impact the other decision processes.

Flow-based congestion control

On the shortest time scale allocating network capacity is a congestion control problem. The congestion control challenge is to efficiently allocate a fixed amount of network capacity and to allocate it in a way that is deemed “fair” among competing senders. Congestion control results in a re-allocation of network capacity among active senders on the time scale of the round trip times of network traffic (e.g. 10’s of milliseconds).

Congestion control is a cooperative mechanism for allocating network capacity implemented by end-users. Each end-user bases its decision upon its perception of the congestion state of the Internet.

Applications and end-users (often) voluntarily adopt it even though it is widely

known how to gain network share by ignoring the congestion signals from the network [95, 16]. Even though voluntary congestion control is still a prevalent behavior in the current Internet, this cooperative allocation of capacity may not be a stable behavior over time.

Some applications have already abandoned congestion control leading to concerns [42] that such behaviors could lead to the congestion collapse [41] of parts of the network. While strictly speaking still complying with the per-flow definition of congestion control behavior, an application can gain a larger fraction of a bottleneck links capacity by opening multiple flows. Each individual flow is following the cooperative protocol, but the aggregate result is that the application gains more of the bottleneck capacity than a competing application that only has one flow in the bottleneck link. This potentially calls into question whether there is an economic basis for the allocations of capacity that result from the flow-based fairness objectives that govern end-users' congestion control behaviors [16].

Aggregate traffic shaping

Indeed, network operators often take a different view of what constitutes a justifiable division of capacity between their users than what results naturally from the cooperative behaviors. In particular, network operators often adopt a view of what constitutes a “fair” allocation of network capacity that includes a longer period of time than is considered in the congestion control allocation.

Network providers base their decisions regarding how to shape the aggregate traffic from a source upon a longer-term view of the likely congestion-costs. They can know the congestion costs on their own network, but do not have visibility into the congestion costs of traffic on the global Internet. Thus they adopt estimates of the costs of traffic based upon heuristics such as the volume of traffic sent or received over time.

To implement the operator's preferred capacity allocation, shaping or policing of the aggregate traffic flow is employed (e.g. token bucket mechanisms). These mechanisms shape the allocation of network capacity by delaying or dropping some

network traffic. A typical time scale that traffic shaping operates on is on the order of network session times (e.g. tens of minutes to hours).

An example is the “Fair Access Policy” (FAP) of the broadband satellite service offered by Hughes [53]. The FAP is essentially a token bucket (with a very large bucket size) that significantly limits the download speeds of a user if they exceed the bucket size. By limiting users that send and receive the most data (Hughes claims it affects only the 5% of customers), Hughes alters the allocation of network capacity among its customers in way that they deem to be more “fair”.

A question, though, is whether the allocation that results from the network operator’s decisions are any more economically defensible than the allocations resulting from end-users behaviors. In practice, economically rationalizable allocation strategies for traffic shaping are even more important than for congestion control since all the ability to make traffic shaping decisions rest with network operators that often have some amount of market power. The decisions of network operators are binding upon users that have little recourse of action except for switching to a different network.

Traffic engineering

On a longer time scale, allocating capacity is a traffic engineering problem. In traffic engineering, a mapping of traffic demand to routes through a network is constructed. Traffic engineering establishes the maximum capacity available between points in a network. It is an allocation of existing network infrastructure resources to create a logical topology over which traffic can be routed.

Like traffic shaping, traffic engineering is again an internal decision process controlled by network operators. In today’s Internet the economic information (i.e. the costs of different routes) required to make these decisions is available only to network operators. The objectives that guide this process are determined by network operators. Though routing can change rapidly, routing paths tend to be stable over longer periods of time – hours to days.

This is important, because how traffic is routed through a network can have a

significant effect on the end-to-end performance of an application. Some networks operators, for instance, segment their VPN business traffic from their general Internet traffic. The VPN traffic flows over a network topology that provides a higher class of performance characteristics (such as lower delay or fewer drops).

Capacity provisioning

On an even longer time scale, typically on the order of months or longer, the physical infrastructure of the network can be changed. Allocating capacity at this time scale is a capacity provisioning problem. This is an internal decision of network operators based upon the cost of capacity and the expectations of network traffic behavior and how the traffic will aggregate. Capacity provisioning is, therefore, an inherently speculative activity where network resources are organized in anticipation of future traffic demand.

Key, then, to the cost effective operation of networks is that end-users “share” network resources once they are provisioned. Without this sharing of resources, and the resultant sharing of costs, it would be uneconomical for most users to communicate.

The question that arises is what to do if the expectations that formed the basis for the provisioning decisions does not align with the traffic loads that actually arrive. The satellite broadband service mentioned above, for instance, is explicitly provisioned for “normal” web and email usage [53]. Hughes claims that, if they did not adopt their “Fair Access Policy”, 5% of users could account for 50% of demand on their network. Evidence suggests that usage patterns are similar in other broadband networks [46].

Peering decisions

Another capacity provisioning process that is done in expectation of future traffic loads are the peering decisions made by a network operators regarding how they interconnect with other networks. Two networks with roughly similar amounts of traffic exchanged between them may decide to peer. This both improves the performance for end users and lowers transit costs for the network operators.

Peering is often viewed as a cooperative process. Many agreements are not formalized by written contracts but rather are arranged with a simple handshake between networks' peering coordinators [76].

However, while it can be a cooperative process, establishing a peering relationship with another provider can be competitive. Network operators adopt complex strategies to “convince” other providers to peer with them on favorable terms [76]. Illustrating the nature of this process, one network operator described the way some providers convince others to peer is by threatening to shift large amounts of traffic around between ingress points – a “peer or die” strategy.

The tussle over network capacity, therefore, is not just between end users and network operators, but between network operators themselves. Indeed, some of the regulatory bodies are beginning to question the interconnection and peering decisions made by network operators.

Interprovider contracts

Interprovider contracts, such as contracts between tier-2 network providers and their upstream tier-1 provider, are a market-based process. This is precisely because the economic information required by participants in this process is available to each party. Network capacity that is bought and sold in this market is traded with contracts that operate on the time scale of months or longer.

These contracts can be structured in a variety of ways. The simplest contracts charge a fixed rate for a given per mbps of transit capacity regardless of whether it is used to carry traffic or not. Volume-based can be more complex, including for instance charges that are determined by the level of traffic during the time period that is the most busy once 5% of the most busy time periods are discarded (95-5 contracts).

What is important about these contracts is that these contracts represent a cost to a network operator that is a function of aggregate traffic. The decision of how to distribute these costs to customers then is an internal decision of a network operator.

End-user subscriptions

Today, many network operators recover their costs by charging downstream users. (Other possibilities are to charge third parties such as advertisers or to be funded from subsidies from other sources such as tax revenues in the case of some municipal networks.) The economic information required to form these contracts is available to each party. The contracts that are signed as part of this subscription process have the effect of allocating parts of the aggregate costs of a network operator to individual downstream users.

Both end users and network operators sign these contracts with an expectation of how the other will behave. End users sign these contracts with implicit assumptions regarding all the internal decisions of network operators. End users, for instance, assume that if congestion becomes a frequent occurrence on the network, that the network provider will expand capacity, peer, or purchase additional transit capacity.

Network providers similarly sign subscription contracts with end users based upon an expectation of how the end users will behave. Network providers assume that end users will not deviate significantly or for long periods of time from the profile of the “average” customer. Users that consistently exceed the provider’s determined “fair” share of network capacity are at risk of having their contracts terminated. (See Figure 4-1.)

Similar to the broadband satellite provider mentioned previously, Comcast noted recently [107]:

”More than 99.99% of our customers use the residential high-speed Internet service as intended, which includes downloading and sharing video, photos and other rich-media. Comcast has a responsibility to provide these customers with a superior experience, and to address any excessive or abusive activities usage issues that may adversely impact that experience.”

”The customers who are notified of excessive use typically and repeatedly consume exponentially more bandwidth than an average residential user,

which would include, for example, the equivalent of sending 256,000 photos a month, or sending 13 million e-mails every month (or 18,000 emails every hour, every day, all month). In these rare instances, Comcast’s policy is to proactively contact the customer via phone to work with them and address the issue or help them select a more appropriate commercial-grade Comcast product.”

Topology layout

Finally, allocating capacity can be an infrastructure planning problem – a problem where the very physical topology of a network must be planned. This is the time-scale at which fibers are laid and operations centers and wiring cabinets are situated (i.e. multiple years). This construction of the physical topology of a network serves as the raw inputs into all the other capacity allocation problems at shorter time scales. The economic information required to determine whether or not

4.3 Lack of architectural support

We contend that the current Internet architecture does not provide much help aligning the interests of different stakeholders affected by these decisions. This is precisely because the economic information necessary to improve the decision processes is either not available or not available to the right parties in the current Internet. (As we argue in the next chapter, the most important piece of economic information that could help align the interests of different stakeholders is the congestion effects of flows of traffic.)

This is not to say that the interests of stakeholders cannot be aligned without support from the architecture. Indeed, market-based allocation mechanisms arise without architectural support in the current Internet. In our above categorization, peering, interprovider contracting, user subscriptions are all markets existing largely “outside” the technical architecture. There aren’t, for instance, standardized algorithms that make these decisions. The outcome of the market based process affects the technical architecture but most of what actually enables the market exchange

(i.e. the communication of information, the commitment mechanisms, verification, dispute resolution, etc.) is human centered.

At least in the current Internet, these non-architecturally based markets tend to operate on fairly long time periods. Contracts for network capacity, for instance, tend to run for a month or longer. This is unsurprising given that human decision makers are involved. We contend that with the appropriate architectural support, the timescales of operation of these markets could be shortened considerably. Indeed this is exactly what the re-ECN protocol could enable.

The current lack of architectural support forces most of the decisions that affect how capacity is allocated to be made internally by network operators. Other stakeholders have no ability to *directly* influence these allocation decisions. Routing, traffic shaping, and capacity provisioning are all internal decisions of the operators. Of course, just because a decision that affects how capacity is allocated is made by network operators instead of by a market based process does not imply that other stakeholders will be displeased with the outcome, but the potential exists for conflict.

This is indeed what often happens in the current tussle over managing and preventing congestion— some end users (and policy proponents and regulators) have become concerned about the decisions of network operators. End users can form overlay networks to overcome some of the limitations of this architectural constraint [23]. But the operator’s internal decisions constrain what is possible.

Our perspective is perhaps slightly different from the prevailing views on this issue. We believe that the central problem is not that network operators are making the “wrong” decisions per se. Rather, the problem is that the architecture forces network operators to make these decisions without information about the real economic costs or implications. In particular, the network operators have little visibility into which traffic sources are actually responsible for costs on their networks. At best, network operators have a rough approximation of individuals’ contributions to costs based upon measurements of the volume of traffic sent or received over some interval of time.

Further, the current Internet architecture was designed without consideration of

what constituted a “fair” or “appropriate” allocation of capacity among *senders* of traffic. Instead, the current architecture and protocols are designed to control the amount of traffic per TCP flow. It is *not* designed to control the *overall* amount of traffic from an individual traffic source.

As discussed in the last chapter, by opening multiple connections, senders can potentially increase their share of the network capacity [16]. Thus, an individual sender can be “following” the rules of the cooperative congestion process, yet receive a larger share of a bottleneck link than an otherwise identical sender.

Even if each individual sender only opens one TCP connection apiece, there is a question of is “fairness” over time. Is continuous usage “fair” compared to users that only sporadically use the network?

Given the current architecture, only network operators could coordinate any longer term notions of a “fair” allocation of network capacity. There is no architectural support for users to cooperatively implement such a longer-term behavior themselves. The challenge for providers, therefore, is to adopt a longer-term notion of “fairness” that is acceptable to users and economically justifiable to potential government regulators. However, the current Internet architecture does not provide much help in economically grounding these operational decisions.

The volume of traffic over a time interval has been adopted by some providers as a proxy metric for arbitrating longer term notions of fairness for both capacity allocation decisions and cost recovery/control strategies. But this metric can be a poor proxy for congestion costs as small volumes of traffic can contribute to disproportionately large amounts of congestion, while conversely large volumes of traffic may not contribute to much congestion on a network if it is sent at a time when there is an overall lower level of network utilization. At best, volume of traffic is an rough approximation of the amount of congestion a user is likely to contribute. Such metrics result in potentially inefficient usage of network resources. Traffic from a source may not be sent even though there would be no costs to a network operator or congestion costs for other users.

4.4 Economic behaviors

Given that the current Internet architecture does not provide much support for aligning the interest of different stakeholders, we briefly consider the potential problems that have and will continue to arise from conflicting economic interests. This summarizes the issues we raised in 2005 that anticipated the problems we are seeing today [18].

4.4.1 Network operators

One of the central problems facing network operators is that of recovering the long-term costs of building and operating the network and controlling the short-term congestion costs in ways that are acceptable to users. This becomes more challenging as capacity is expanded and the aggregation properties of network traffic are changed by evolving application and user behaviors. Faced with these problems, providers *may* be motivated to 1) suppress the emergence of non-revenue generating applications or 2) delay investments in improving network services [102].

Either effect is undesirable. Other stakeholders in the larger Internet eco-system are adversely impacted. Component and equipment makers, providers of content and application services, end-users and entire sectors of the economy such as education, government and health care would benefit from improvements in networks, yet currently these stakeholders are captive to the decisions of the network owners. The network owners must make the bulk of the investment in upgrades even though they potentially realize little return on the benefits that accrue to other stakeholders [18].

An inability to directly recover or control the costs of Internet traffic may result in the following problems [18]:

- Network providers will lack the incentive to deploy higher capacity and quality services for general Internet traffic.
- Network providers may resort to techniques that discriminate against applications and services that are viewed by the providers as unprofitable – an outcome

which potentially raises significant barriers to innovations at the edges.

- Network providers may be discouraged from providing comparable capacity and quality for general Internet traffic if they view it as a potential inroad for content and services that compete with their own revenue generating offerings.

In practice, network providers may cover some of the costs of general traffic with revenues from other services (such as content or advertising). Charging for these other services can be easier so conceivably recovering the general traffic costs there as well is a reasonable approach. However, the long term viability of covering the costs of Internet traffic with the larger profit margins from other services is questionable – Internet traffic may grow to erode the margins of the other revenue streams over the long run.

A strategy that directly recovers or controls the congestion costs in a network may be more beneficial for all stakeholders. This is exactly what the re-ECN architecture that we examine in this thesis attempts to provide.

Competition

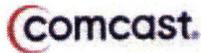
A natural question is whether competition is sufficient to adequately resolve the incentive problem in terms mutually beneficial to providers and customers (and also the long term architectural interests of the Internet.) In the broadband space, conceivably competition between cable and DSL providers (and in the future perhaps wireless, satellite or broadband over powerline) is sufficient motivation for further investment in the general Internet infrastructure.

Competition can be a powerful force in a market that fosters investment and development. But there is a real question of whether or not existing service models and cost recovery strategies will continue to scale to the higher capacity and advanced services that are increasingly technologically possible. Without an adequate cost recover strategy, no amount of competition will motivate provider investment. Competition can speed adoption, but providers will require some way of appropriately recovering or controlling their costs.

Whether competition alone is adequate is further brought into question by the wide geographic differences in the amount of provider competition. The network marketplace is a complicated market where the number of providers varies greatly across locals. One town may have a single provider, while an adjacent town will have three or more broadband access options. Even in areas that are currently competitive, the introduction of high capacity fiber to the home may eventually lead to reduced competition in the future.

4.5 Summary

In this chapter, we considered how the lack of architectural support of market based allocation mechanisms leads to potentially undesirable behaviors on the part of both end users and network operators. We conclude that the goal of redesigning the architecture to better align the incentives of both the senders and receivers of network traffic with network operators is well motivated. There is room for an engineering effort informed by economic analysis to improve the network architecture.



Comcast IP Services
7 Eisenhower Campus, 5th Floor
Cherry Hill, NJ 08002
Tel: 856-317-7272
Fax: 856-317-7219

Notice of Acceptable Use Policy Violation

Personal and Confidential

Abuse Ticket Number: [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]

Incident Type: Network, Bandwidth, Data Storage, and Other Limitations

Comcast High-Speed Internet Acceptable Use Policy (AUP) Violation –Bandwidth Usage Limitations

Dear Subscriber:

As a subscriber to Comcast High-Speed Internet service, you have agreed to use the service according to the Subscriber Agreement (<http://www.comcast.net/terms/subscriber.jsp>), Acceptable Use Policy (AUP) (<http://www.comcast.net/terms/use.jsp>), and other terms of service and policies. Our records indicate that during July 2003 your Comcast High-Speed Internet account exceeded Comcast's bandwidth usage limitations for your service. This usage activity violates Comcast's AUP. If your account continues to exceed our bandwidth usage limitations for your service, this activity could result in the suspension and ultimately termination of your Comcast High-Speed Internet account.

Excessive bandwidth usage can be the result of many different activities including, for example, commercial or business applications, peer to peer networking, newsgroup downloading, file sharing, streaming music/videos, or voice/video chat. If you are unaware of any activity like this on your Comcast account, we suggest that you speak with any other person who has had access to your Comcast High-Speed Internet service to determine if they have performed any of these activities. As the service account holder, you are responsible for any misuse of the service or violation of the AUP, even if the misuse or violation was committed by a friend, family member, or guest with access to your Comcast High-Speed Internet service account.

To avoid future violations of the AUP for exceeding bandwidth limitations, we recommend that you immediately review and modify your current Internet usage activities and practices. Additionally, you may also want to update your anti-virus program, or obtain one if you don't already have anti-virus software. We also recommend that you install a firewall if you don't already have one to help control unauthorized access to, and use of, your Comcast High-Speed Internet account. Please be aware that repeat violations of Comcast's AUP could result in the suspension and ultimately termination of your Comcast High-Speed Internet account with or without notice to you.

If you have any questions about this Notice, or would like to speak with someone about obtaining Comcast commercial Internet service to support business use of the Internet, please contact us at **877-561-9324**.

Sincerely yours,

The Comcast IP Network Abuse and Fraud Management Team

Figure 4-1: Comcast authorized use policy (AUP) violation letter

Chapter 5

Overview of re-ECN

In this chapter, we examine the re-ECN protocol. It is an example of a protocol that signals to network carriers the likely congestion impact of a flow of traffic. Integrated over time and across individual flows of traffic, the aggregate congestion impact of a source of traffic can be measured and monitored. We explore the argument that this is precisely the economic information needed to improve contractual and traffic management decisions on the Internet.

We focus on re-ECN because Briscoe has compellingly worked the challenging engineering details of making the protocol a reality. However, any protocol which makes the downstream congestion impact of traffic aggregates available at the economic boundaries of networks would have similar properties. There are no such other alternatives today. (Briscoe[16] addresses why end-host or router-based mechanisms such as TFRC[50], XCP[57], or weighted fair queuing are ill-suited for managing the aggregate congestion impact of traffic across multiple links and across time in a provider's network.)

The re-ECN protocol exposes, as an inherent part of a flow of traffic, information about the expected downstream congestion. Senders have always adjusted their sending rates of individual flows based upon their inferred congestion state of the network (c.f. TCP congestion control.) Under re-ECN, senders would explicitly reveal their flow-based inferences about the congestion state of the network. We describe the technical details of both how senders form inferences of the congestion state and how

they reveal it back to the network in the next section.

The inferred congestion impact of a source of traffic is important because of its economic significance – network congestion has real economic costs. The re-ECN protocol exposes this congestion information at the economic boundaries of the Internet where either contractual mechanisms (i.e. billing) or traffic management decisions are made. Thus, the downstream effects of congestion can be constrained (in the case of end-users) or billed for (in the case of businesses or ISPs) at precisely the locations in the network where they can be best accounted for or controlled. (See Figure 5-1.)

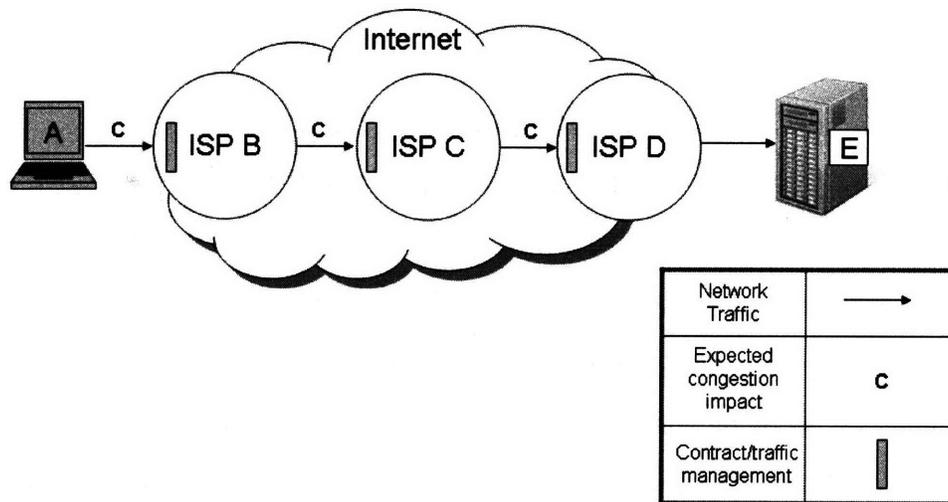


Figure 5-1: At each of the economic boundaries of the Internet, the congestion impact of traffic needs to be exposed to make traffic management decisions or contracting decisions that are economically efficient.

The decisions of the network to shape, route, or charge for traffic can therefore be based upon information that has genuine economic significance – the flow’s likely congestion cost. If the inferred congestion state of the downstream path is available at the economic boundaries of the network, individual networks can make economically rationalizable decisions about how to handle or charge for traffic.

The re-ECN protocol potentially enables an architecture in which one could distinguish, using relatively simple rules, between some forms of “good” or “acceptable” discriminatory behaviors by network operators and “bad” or “detrimental” behaviors. Namely, regulators or network operators themselves might conclude that behaviors

or contracts that are predicated upon the congestion impact of a flow of traffic could be deemed acceptable. (Traffic management procedures, such as those adopted by Comcast, that block specific applications regardless of their actual congestion impact would be suspect.) The re-ECN protocol is thus a potential architectural response to the “network neutrality” debate.

In the following sections, we provide an overview of re-ECN and supporting mechanisms and examine some of the policies that can be constructed with it. The re-ECN protocol was introduced by Briscoe [14, 17, 16]. The theoretical foundation of the work was developed by Kelly and Gibbens [59, 47]. The technical viability of re-ECN has led to ongoing work within the Internet Engineering Task Force (IETF) [14].

5.1 Re-ECN overview

The re-ECN protocol leverages and extends the Explicit Congestion Notification (ECN) bits[90] of the IP header and the congestion feedback in the transport header to expose to observers along a path the sender’s *expectation* of downstream congestion.

The re-ECN protocol adds a ‘re-echo’ stage to the traditional ECN protocol. Whenever a sender is notified of congestion on the downstream path (most likely by feedback received in a transport layer acknowledgment), the sender re-echoes that notification by ‘blanking’ a newly defined RE flag (in normal operation the flag is set to 1) in the network layer header of a subsequent packet in the same flow.

In the current Internet, if ECN were widely enabled, one could observe whether congestion was occurring *upstream* of an interface by observing whether the congestion experienced (CE) bit had been set in the IP header of a packet. But it has never been possible within the network to know the *expected congestion* a packet will encounter on the rest of the path to its destination i.e. whether the CE bit will be set by some downstream router.

The RE flag of re-ECN would expose this expectation of congestion to the network. If expected congestion has been signaled by the sender through blanking the RE flag, any point in the network could observe 1) the overall flow rate and 2) whether

congestion has occurred upstream already (i.e. the CE bit has already been set) or if congestion is still predicted on the remaining downstream path (i.e. the CE bit has not yet been set by any router).

While most senders would learn of congestion along a path via feedback in a transport protocol, the re-ECN protocol is capable of supporting a series of single packet flows, such as DNS lookups. In these cases, an expectation of congestion is formed by reusing knowledge about previous congestion information along a path or from a shared proxy such as a Congestion Manager [6] that aggregates congestion information across multiple senders.

Most applications of the re-ECN protocol involve observing the average fraction of packets that blank the RE flag (thereby signaling downstream congestion is expected) compared to the average fraction of packets that are already CE congestion marked *over some period of time*. This is an important point. While a sender may not know exactly which packets will be congestion marked during an interval, senders can form good predictions, based upon congestion feedback, about the overall fraction of packets in their traffic flow that will be marked.

Only senders have to be modified to re-echo the congestion information they receive. Receivers that are modified to support the protocol can improve the protocol performance, but re-ECN works with traditional ECN receivers as well. Routers do not need to be modified; they simply set the CE bit in packets during periods of congestion. This is functionality that already exists, but is not widely turned on. (It is argued in [14] that some of the functionality enabled by Re-ECN creates incentives for network providers to begin employing ECN in their routers.)

5.2 Truthful predictions of congestion

Of course, without additional components, senders would have no natural incentive to accurately state their expectation of the downstream congestion for a given flow rate. This would make the information revealed by the re-ECN protocol untrustworthy in a non-cooperative environment. The incentives to accurately reveal this congestion

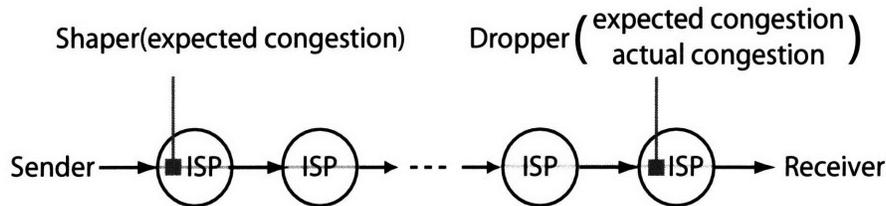


Figure 5-2: The incentives mechanisms leveraging re-ECN are designed to induce senders to truthfully report the expected congestion along a path. A shaper creates an incentive for senders to not *overstate* their view of expected congestion. A dropper creates an incentive for senders to not *understate* their view of downstream congestion.

information though can be created by the addition of other architectural components. In particular, the incentive for senders to accurately state their downstream congestion can be created through the addition of shapers and droppers in the network (see Figure 5-2). (These are the terms that Briscoe employs [17], so we adopt them here as well. Both have the effect of constraining the flow of traffic.) These shapers and droppers would be deployed by any network operator that sought to management the congestion impact of traffic. As explained in ??, these can be deployed incrementally by network operators on the Internet.

A shaper creates incentives for senders to not *overstate* their expectation of downstream congestion for a given flow rate by shaping or charging for traffic. A sender could pay more to continue to continue sending at a higher rate during periods of congestion or its sending rate could be constrained to the desired congestion response of the network operator (perhaps the TCP-friendly rate). The sending rate of TCP-friendly flow is a function of the drop rate and round trip time of a flow of traffic. Thus, given the stated downstream congestion information revealed by a sender and the flows rate, the shaper can enforce the congestion response that would correspond to the revealed level of congestion. If a sender overstates their expected downstream congestion, their sending rate could be constrained by the shaper instead of the actual bottleneck link in the network.

A dropper creates an incentive for senders to not *understate* their expectation of downstream congestion for the flow rate by dropping enough traffic to make the rate of congestion marked packets (CE) equal to the rate of packets marked with expected

congestion in a flow of traffic. Thus, if a sender marks fewer packets with an expected congestion bit than will actually be marked with congestion, some of their packets will be dropped by the network. If a sender understates their expected downstream congestion, their sending rate would be constrained by the dropper instead of the actual bottleneck link in the network.

Any deviation from truthfully stating the expected downstream congestion will lead to shaping or dropping the sender's traffic. Overall the sender's throughput to the receiver will be reduced or their billing charges will be larger. Thus senders of traffic have an incentive to accurately estimate and reveal their expectation of downstream congestion and network operators have an incentive to deploy shapers and droppers to manage the congestion impact of traffic.

5.3 Uses of re-ECN

The question, then is what can be enabled by revealing this expected congestion information? While Briscoe lays out a variety of functionality ranging from denial of service mitigation to end-to-end quality of service applications, our interest has been in the use of expected downstream congestion information as a basis for improving the economic efficiency of networking decisions.

In the last chapter, we suggested that many of the decisions impacting how traffic is handled in a network were at best only indirectly shaped by market forces or economically relevant information. Most decisions are made by cooperative mechanisms or are internal decisions of network operators. While on a much longer time scale market forces (such as customers defecting to rival providers) might slowly work to improve the economic efficiency of these decisions, directly incorporating the economically significant congestion information could improve the decision process.

The decision processes that influence how traffic is handled have the potential to be improved with the incorporation of expected congestion information. The objective is to move these decisions to being (more) market-driven (or at least based upon economically significant events.) (See Table 5.1).

	Cooperative mechanism	Network operator's internal decision	Market-driven decision process
Flow-based congestion control			✓
Aggregate traffic shaping			✓
Traffic engineering			✓
Capacity provisioning			✓
Peering			✓
End-user subscription			✓
Interprovider contracts			✓
Topology layout		✓	

Table 5.1: Capacity allocation and cost/control recovery enabled by re-ECN

In the following subsections, we briefly consider how incorporating the information revealed by re-ECN could change some of these decision processes to being driven by the re-ECN signal of congestion i.e. a signal of supply and demand for network resources.

5.3.1 Flow-based congestion control

The theoretical work by Gibbens and Kelly [47] on which re-ECN is based created an economic game in which traffic sources could be governed by new congestion control algorithms. By applying pricing to ECN markings they proved that optimal and stable allocations of capacity would naturally occur (given certain assumptions about senders' utility functions.) Senders that most valued network capacity during times of congestion would simply pay more to send their traffic. Senders would automatically adjust their sending rate to balance the utility they received from sending traffic with the current congestion price.

A re-ECN enabled network would maintain the economic incentives that led to the optimality and stability results of Gibbens and Kelly, but change the engineering mechanisms that create these incentives. Instead of facing a variable congestion priced bill (which real-world users are unlikely to accept), the incentives for users to back off during periods of congestion would be created by the shaping and dropping mechanisms introduced by network operators. Traffic that is not compliant with an

operator's desired congestion response would be shaped or routed in way that would degrade the performance of the non-compliant traffic sources.

Each network operator could establish their own policy for handling congestion marked packets. Some networks may choose to allow relatively low bitrate flows to ignore most signals of congestion – approving, for instance, VoIP applications that may not cut back their sending rates in response to congestion. Or a network might require that particularly long lived flows be more aggressive than the typical TCP response in cutting back their sending rate in response to congestion. Whatever the policy, it could be based upon economically significant events – the congestion effects of traffic flows.

Using re-ECN, network operators could establish longer-term notions of fairness by comparing the level of congestion different flows of traffic contribute to over time. By integrating the expected downstream congestion over time, the network operator can calculate the overall congestion impact of a source of traffic. This gives an economically well supported measure for comparing the marginal costs of carrying different aggregate flows of traffic from different sources.

Congestion control thus is moved from being a cooperative mechanism to one that could have economic incentives governing its adoption and use. Coupling the congestion response to the pricing or shaping policies of network providers enables new congestion responses while maintaining a coordinated, stable, and efficient allocation of network capacity [59, 16].

5.3.2 Aggregate traffic shaping

Today, traffic shaping behaviors are often a function of the overall volume of traffic from a source. They can be predicated upon the time of day or the application layer content of a flow of traffic. However, these are at best rough approximations employed by network operators to control the congestion impact of traffic coming onto a network.

Under re-ECN, however, traffic shaping could be based directly on the congestion signal embedded in the aggregate flow of traffic. The rough heuristics employed today

would not be necessary for managing congestion. The need for mechanisms such as Comcast's resetting of peer-to-peer applications TCP connections would vanish. Instead of singling out individual applications, Comcast could limit the aggregate amount of congestion that a subscriber could contribute to the network over different periods of time. This potentially enables traffic shaping policies that are less likely to be viewed as arbitrary, ad-hoc, or unjustifiably discriminatory.

5.3.3 Capacity provisioning

Today, capacity provisioning is an inherently speculative activity where network resources are organized in anticipation of future traffic demand. Those that make investments in expanding capacity must speculate whether additional capacity will actually be valued by those that will send and receive traffic. There is no architectural mechanism for a sender to indicate that an expansion in network capacity would be valued.

Economists have long argued that congestion pricing is exactly the signal that is needed to fulfill this role. By demonstrating a willingness to pay a congestion price or penalty, a sender is signaling the value of expanding the network capacity. Mackie-Mason explains that a congestion price “provides a measure of the social cost of increased usage for an given capacity, but it also determines the value of a change in capacity.” [65]

In re-ECN, the congestion impact of an aggregate flow of traffic becomes common knowledge to both the sources and the carriers of traffic. A network operator thus could establish contracts that charge more to those (perhaps other ISPs) that are contributing to congestion. This could fund the expansion of their network capacity.

Such a funding strategy does assume that traffic patterns remain similar over time, i.e. those that are contributing to congestion now would benefit from expansion in the future and those that are not contributing to congestion now would not directly benefit from expanding capacity in the future. But analogously to how some network providers today charge extra every gigabyte of traffic over a monthly quota, providers adopting re-ECN could charge extra for the congestion that a network caused.

5.3.4 Peering

While peering can be a cooperative process on today's Internet, there are strategic games that are played in inducing or denying peering as well [35]. While these type of agreements receive relatively little regulatory scrutiny today, on occasions when there is a significant de-peering event, there have been calls for regulatory scrutiny of peering agreements and requirements.

The question that is raised is what the standards are for establishing a peering relationship versus a provider/subscriber relationship. Most de-peerings occur because of a perception of asymmetries in traffic volumes. However, traffic volume is not as economically significant as the level of congestion a peering network causes.

Some peering agreements therefore might come to require congestion symmetry in addition to the other existing peering requirements (such as number of peering points, geographic diversity of peering, etc). Agreements could, for instance, establish that peer-traffic with some fraction of expected congestion would be handled according to a free peering arrangement, but if the expected downstream level crosses a certain threshold, charges would apply.

Peering coordinators may find this criteria appealing as it provides a principled requirement that limits the congestion impact of peering traffic. This controls the risks associated with accepting traffic for free. Further, by monitoring the congestion impact of peering traffic coming onto its network, a network operator has a natural metric for continually reevaluating whether the peering relationship is worthwhile. This provides an economically grounded metric for peering coordinators to justify to upper-level network management and other stakeholders their peering and de-peering decisions.

5.3.5 Subscriber and interprovider contracts

Given individual customer's obvious preference for flat-rate or simple contracts, re-ECN is unlikely to change the current billing practices for end users. Contracts may simply inform the customer that their traffic could be shaped based upon the

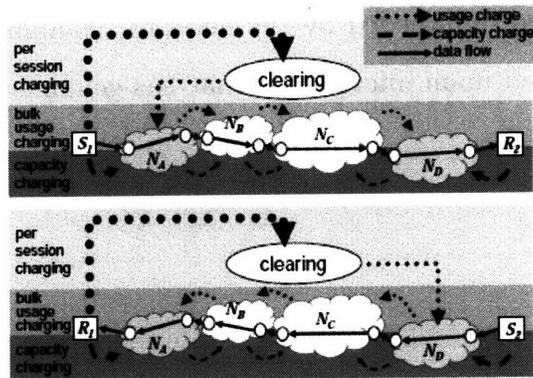


Figure 5-3: “End-to-edge clearing for sender-pays charging. Two duplex flows for the same session are shown, one each in the upper and lower halves of the figure.” (Source [15])

aggregate amount of congestion they contributed to over some periods of time.

Interprovider contracts, though, may be based more directly on the expected congestion. Briscoe examines a variety of novel commercial arrangements in [15]. He details arrangements that would enable re-ECN to be deployed incrementally on the Internet and arrangements that would enable clearing house like functionality where different parties could pay for congestion at different congestion bottlenecks in the network. (See for example Figure 5-3 for a sender pays functionality.)

The key is that re-ECN enables a richer set of interprovider relationships. By exposing the expected congestion signal in a flow of traffic, a network operator can build contractual agreements or shaping mechanisms that hedge against the risk of accepting traffic that will congest its network.

5.4 Summary

If it is indeed the case that the mapping between economic theory supporting re-ECN and the real world networks is appropriate, then the network engineering community has at hand an opportunity that could enable a significantly better economic system for the Internet. Applications and network traffic that today is rejected or shaped simply because it is caught in the crude heuristics providers elect to employ today

could instead be carried unimpeded by the network. Shaping, routing, or charging for traffic could be based upon information that has genuine economic significance – its likely congestion cost.

Chapter 6

Analyzing the economics of re-ECN

A more complete defense of re-ECN, the underlying theory, and the potential applications of the protocol can be found in the literature [17, 59, 47, 14]. What we seek to do in this chapter is to understand the limits of the economic arguments supporting a re-ECN enabled architecture. We do this in two different ways. In the first part, we scrutinize the assumptions underlying the economic arguments for re-ECN. In the second part, we develop a simplified game theoretic model of a re-ECN enabled network and analyze some of the strategic behaviors that might result in a long-term repeated game.

While we are generally supportive of the engineering design and economic potential of the re-ECN enabled architecture, the analysis we present explores the boundaries of some of the economic claims.

1. We consider the implications of topologically non-uniform economic costs of congestion. This issue is important to understanding whether or not there would every be an economic incentive to treat aggregate flows of traffic with identical re-ECN markings differently.
2. We consider the relationship between re-ECN congestion marking and the economic notion of congestion externalities. This issue is important because it

highlights cases where congestion markings placed by routers are not necessarily a signal of a negative economic externality.

3. We consider strategic behaviors by end-users in congestion games that represent scenarios that may arise on the Internet. These games are important to analyze because they serve as predictors of potentially undesirable behaviors that may arise under re-ECN (given sufficiently sophisticated players).

While this chapter could be interpreted as a partial critique of re-ECN, we position it instead as a constructive contribution to a debate about whether to deploy re-ECN and/or what scenarios are most appropriate to use it in. This is particularly important as regulatory authorities or network operators may consider re-ECN like mechanisms as solution to the network neutrality tussle.

Today, network operators like Comcast argue they are preventing congestion on their networks by adopting peer-to-peer limiting mechanisms – an obviously controversial operational strategy. By adopting re-ECN-based mechanisms instead, network operators could appeal to an argument that they are no longer singling out particular applications but are instead exclusively targeting aggregate flows that contribute to congestion. Network operators could argue that they were enforcing a form of “cost fairness” [16] over traffic sources. Application or users that contributed similar amounts of congestion (over particular periods of time) would be treated similarly if they had similar contracts.

Briscoe argues, for example, that under re-ECN operators have an economically justified basis for filtering or limiting the traffic from voice or video *if* they are unresponsive to congestion [14]:

Some network owners want to block applications like voice and video unless their network is compensated for the extra share of bottleneck bandwidth taken. These real-time applications tend to be unresponsive when congestion arises.¹ Whereas elastic TCP-based applications back

¹This quote refers to voice or video applications that are not responsive to congestion. Some voice and video applications operate over TCP and do respond to congestion.

away quickly, ending up taking a much smaller share of congested capacity for themselves. Other network owners want to invest in large amounts of capacity and make their gains from simplicity of operation and economies of scale.

Re-ECN allows the more conservative networks to police out flows that have not asked to be unresponsive to congestion—not because they are voice or video—just because they don’t respond to congestion. But it also allows other networks to choose not to police. Crucially, when flows from liberal networks cross into a conservative network, re-ECN enables the conservative network to apply penalties to its neighboring networks for the congestion they allow to be caused. [Emphasis added]

If network operators do indeed employ re-ECN as a basis for filtering or policing and it primarily has an effect on video and voice applications, this will almost inevitably lead to close scrutiny of the practice. We therefore want to carefully examine how network operators might use re-ECN in making an argument that their network operational practices are economically justifiable.

Our general conclusion is that the expected congestion information that re-ECN exposes is still an approximation of the “true” economic costs – but one that in almost all cases will be better than a volume-based approximation or the existing heuristic targeting certain classes of applications. Yet the limitations of re-ECN are important. They are both academically interesting and relevant to the real world discussions of network neutrality.

6.1 Topologically non-uniform costs of congestion

The congestion control algorithm for TCP treats signals of congestion identically. Upon detecting a packet lose or upon reception of an ECN marked packet, a sender will half its congestion window, slowing its sending rate. Senders do not, and indeed cannot, differentiate between congestion drops in different parts of the network.

This implicitly implies that congestion everywhere in the network is of equal importance. Congestion on a trans-oceanic link is treated by a sender the same as congestion on a local area network. The TCP control algorithm backs off and increases its sending rate identically in both cases.

The same implicit assumption underlies the re-ECN protocol. The re-ECN protocol exposes a sender's expectation of how many packets will be congestion marked. The economic significance of each of those congestion marks – i.e. a price (or cost depending upon how one looks at it) of congestion is known only by the network operator where the actual congestion occurs. Senders simply do not receive enough information to make a distinction between congestion marks from different parts of the network. This again, implies that all congestion in the network is of equal importance.

But the economic costs of congestion are not uniform across the network topology. Indeed, the economic costs can vary considerably. Congestion on a link that impedes the activity of a online commerce site likely has more economically significance to a network operator than congestion on a link to a local cache of software updates. Even in a non-commercial setting, such as a military network, congestion on links that a high ranking officer is using to coordinate troop movements may be more important than congestion on a link used to backup imaging data. In short, easing congestion in some parts of a network may take priority over easing congestion in other parts of a network.

But at the economic boundaries of a network, where the contract or traffic management mechanisms reside, only the existence of downstream congestion is signaled. If, as the designers of re-ECN suggest, contracts between providers become predicated on congestion information, these contracts might have different prices associated with congestion in different provider relationships.

Consider Figure 6-1 where traffic sources *A* and traffic source *B* both have traffic to send to destination *C*. Both *A* and *B* have identical amounts of traffic to send and both flows of traffic have identical congestion marks. The only difference is where in the network this congestion occurs. In the case of flow *AC* congestion occurs on

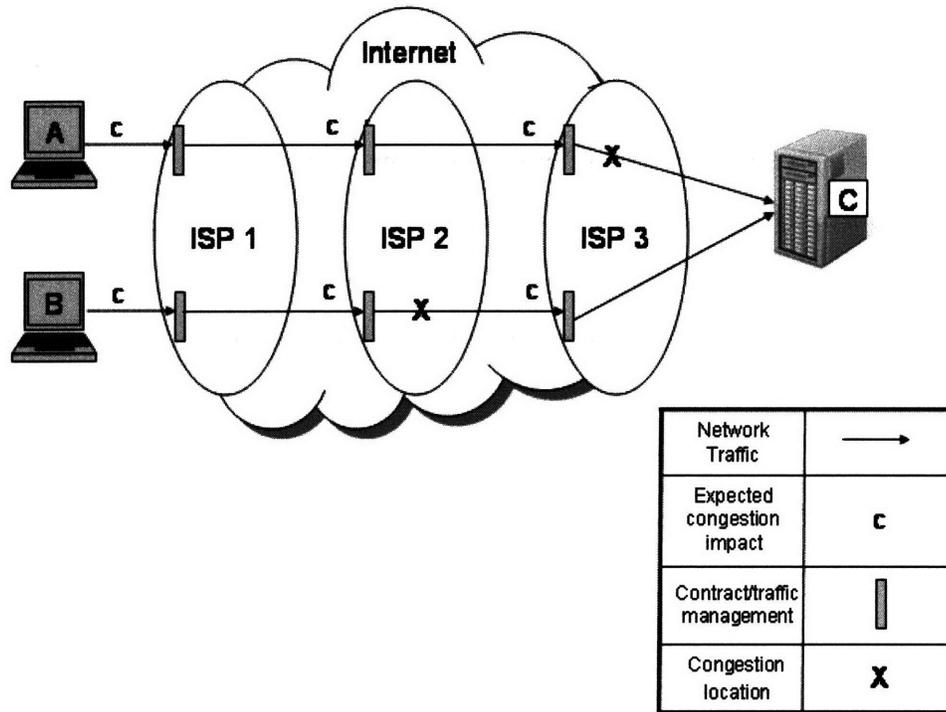


Figure 6-1: Congestion that occurs in different locations in the network may have different economic costs to network providers.

provider 3's network thereby triggering a congestion charge to provider 2. In the case of flow *BC* the congestion occurs on provider 2's internal network – congestion that is only indirectly costly to provider 2. In this case, provider 2 may have an incentive to treat the two flows differently (either in how they are traffic shaped or charged) even though both flows had the same downstream congestion signal observed at the incoming interfaces of 2's network.

An economist might comment that the problem this example illustrates is that the system is not exposing enough information about the true economic costs of congestion. One bit per packet of congestion present/not-present may not convey enough information about the economic importance of congestion on the network. The practical response, though, is that it may be a fairly good approximation.

The important engineering question is whether re-ECN congestion information will be a better approximation of real world costs than the existing approaches employed by today’s network operators? In most cases, we believe the answer is yes, but this depends upon the network topologies, traffic patterns, and assumption about the economic cost of congestion at different locations in the network.

The important point is that expected congestion should be regarded as an approximation of “real-world” costs. Operators will still have an incentive to differentiate between two aggregate traffic flows with identical flow rates and expected congestion rates. They may combine the expected congestion information with their knowledge about their different local costs for congestion to treat two otherwise identical sources of traffic differently.

So network operational behaviors that discriminate between the two should not necessarily be regarded as suspect. Regulators, for instance, should not mandate that traffic shaping mechanisms look at only re-ECN expected congestion information. Network operators should also not be ordered to interconnect or peer with simplistic criteria that treats all re-ECN information equally. (Note, nowhere do the authors of the re-ECN papers imply that any of these should occur. Indeed, we suspect they are likely fully aware of this issue. But explicitly documenting these caveats is important in our view.)

6.2 Congestion marking as a signal of a negative externality

We next consider the degree to which congestion marks are actually a signal of a negative externality that need to be “priced” in an economically efficient network. The theory behind congestion pricing or congestion based incentive mechanisms such as re-ECN rests upon the assumption that congestion marks are representative of the congestion *externality* associated with a flow of traffic. By holding traffic sources accountable for the congestion they cause, through either charging for congestion or

shaping the source's traffic, this externality is internalized by the traffic source. This is standard economic theory that leads to efficient usage of scarce resources.

However, congestion marks are not *necessarily* always a signal of an externality that needs to be “priced” for efficient usage of the network. Namely, this arises when the parts of the network are not in fact shared between multiple sources of traffic. In such cases, the economic costs of all traffic delays and drops are already being internalized by the source of the traffic – it alone is bearing the cost of the delayed and dropped traffic. Nothing needs to be priced for efficient operation of the network.

This is a result of a mismatch between the economic meaning of congestion and the technical network meaning of congestion. In networking terms a router can be considered to be “congested” whenever the average queue length over some interval of time crosses a threshold. In the active queue algorithm of RED, Random Early Detection, this threshold is used to trigger the start of ECN packet marking. (Figure 6-2)

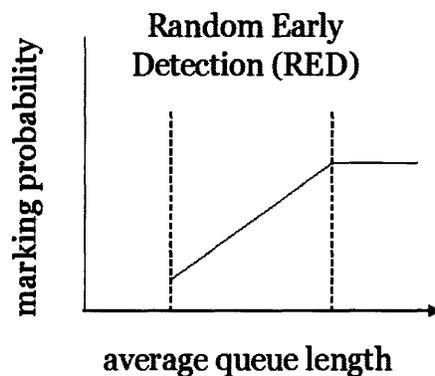


Figure 6-2: RED marking probability

Packets are marked regardless of the mix of traffic sources in the queue. Even if all the traffic is associated with one ISP or one individual user, more packets will be marked as the queue begins to build. At a technical level this is the desired and correct behavior – sources of traffic want to be notified of impending congestion.

The economic significance of these marks is less straightforward. If there were other economic entities contending for the congested resource, then the markings are representative of a flows contribution to economically relevant congestion. However,

if there are no other economic entities contending for the congested queue then there is not congestion externality in the economic sense.

This is significant because just as today where the economic costs of a volume of traffic are debated, the economic costs of congestion markings could be debated. A source of traffic could rationally object to being shaped or charged for causing congestion if that congestion does not impact traffic associated with another source. Determining the economic significance of congestion marks depends upon where congested queues arise and their mix of traffic.

This is particularly significant in the edge networks. The number of flows of traffic sharing bottlenecks links can be relatively small. Indeed the primary upload bottleneck many broadband subscribers face is their own cable modem. If the queues in those devices were to ECN mark flows, the ECN marks would not necessarily be representative of any economic externality. The subscriber is bearing the full cost of any congestion delays or drops that occur in that queue.

For instance, if Comcast were to ECN mark packets that build up in the queues of individual subscriber's cable modems waiting for upload to the CMTS, these ECN marks might not be a signal that other users' packets are being delayed or dropped. The administrative caps placed on the upload capacity of a cable modem can cause the queue of packets waiting to be transmitted to build independent of the overall load on the access network.

Of course, the queue length in a subscriber's modem can be coupled to the larger load on the access network since the CMTS schedules upload slots for each modem. During times of congestion, queues will build in each modem as each waits for an upload opportunity. But since the queue can also build even when the larger access network is not congested, it could be difficult to distinguish whether a congestion mark is a signal of a negative economic externality or just a signal that the local non-shared queue is congested.

6.3 Congestion Games

We now turn to examining how users sending traffic in such an congestion-based incentive environment might react. We believe this analysis is novel – the primary incentives that are commonly considered are the incentives of network operators to over-state the amount of congestion in their networks, thereby artificially raising their own revenues. We, however, were interested in what would occur if providers revealed their true costs but users could be strategic regarding whether or not they revealed their true demands.

In particular, we examine whether a congestion-based mechanism is dominant strategy incentive-compatible [45] over long-term repeated interactions where by long-term we mean days, weeks, or months. In our model, incentive compatibility implies that a player cannot improve their utility through strategic gaming, making traffic demands that do not reflect their true needs.

We focus on incentive compatibility as a desirable property because it simplifies the decision making process for players. If a mechanism is incentive compatible, players only have to consider their own needs and actions without concern that the actions of other players may be compromising their own welfare.

We demonstrate that congestion-based incentives are not necessarily long-term incentive compatible. Long-term strategic players can have a potentially perverse incentive to congest a network in the short-run. The intuition behind this perverse incentive is that strategic players are willing to pay a small short-term penalty for a larger long-term benefit. The long-term benefits we consider arise in scenarios that may be common on the Internet.

Note that a player is an abstraction of any entity that controls or generates traffic load. In particular, a player is not just a human user. Humans, we readily concede, are unlikely to directly engage in strategic gaming of a congestion mechanism. Congestion-strategic behaviors, though, could arise from applications, operating systems, large enterprises, overlay networks, or Internet providers.

Our model-based conclusions leave open the question of whether strategic gaming

of congestion-based incentive mechanisms is a theoretical concern only or is an actual, practical problem. Our results might be viewed as unsurprising to some theoreticians. It is well known that achieving such a long term incentive compatibility is a difficult theoretical problem. However, even with this caveat, we still find the existence and outcomes of such games to be interesting.

We consider scenarios in which a user would interact with a congestion-based incentive mechanism repeatedly.

1. A capacity expansion game
2. A congestion-budget game

6.4 Model

To conduct this analysis we model a congestion-based incentive mechanism that, like re-ECN, relies upon marking packets. This model is intended to represent the congestion-based incentive mechanisms that leverage the capability of routers to set the explicit-congestion-notification (ECN) bit in IP packets.

We define a congestion-based incentive mechanism C operating on a link in a provider's network. We adopt a discrete model of time, and consider the mechanism computes a decision in each period.

The operation of the mechanism is described as follows. The operator picks some congestion threshold H . If more than a congestion threshold H of demand arrives, some of the arriving packets are marked as a packet contributing to congestion. If less than the congestion threshold of demand arrives then no packet during that period receives a congestion mark. The mechanism is entirely myopic; the number of packets in preceding periods has no effect on the congestion marks in any other time periods. This is patterned after the model presented in section 9.4.1 of Courcoubetis [29] for calculating the sample path shadow price.

Formally, let:

$N : 1, \dots, n$: a set N of n users

$i \in N$: user i

θ : a vector of true traffic demands for users

$\hat{\theta}$: a vector of revealed traffic demands for users (# of packet actually sent)

$\theta_i \in \theta$: user i 's true traffic demands

$\hat{\theta}_i \in \hat{\theta}$: user i 's revealed traffic demands

$\hat{\theta}_{-i}$: vector of revealed traffic demands except i

H : congestion threshold

c : vector of marked packets for users

$c_i \in c$: number of congestion marked packets for user i

Then, we define our congestion based incentive mechanism as:

A congestion based incentive mechanism C: A congestion based incentive mechanism is a mapping from the vector of *revealed* types (traffic demands) and congestion threshold H to a vector consisting of the number of packets marked for each user:

$$C : \hat{\theta}, H \rightarrow c$$

$$c_i = C_i(\hat{\theta}_i, \hat{\theta}_{-i}, H) = \begin{cases} \omega(\hat{\theta}_i) & \text{if } \sum_{j \in N} \hat{\theta}_j > H \\ 0 & \text{if } \sum_{j \in N} \hat{\theta}_j \leq H \end{cases}$$

where $\omega(\hat{\theta}_i)$ is an increasing function in the number of packets that are sent. The result of the function represents the number of packets that are marked. For simplicity, we let $\omega(\hat{\theta}_i) = |\hat{\theta}_i|$ for the remainder of the paper (i.e. all packets during a congested period receive a mark.) Our results hold for any marking function where the probability of marking a packet is identical between all packets sharing the same congestion period.

In general, a player's utility is lowered in the single shot game induced by the

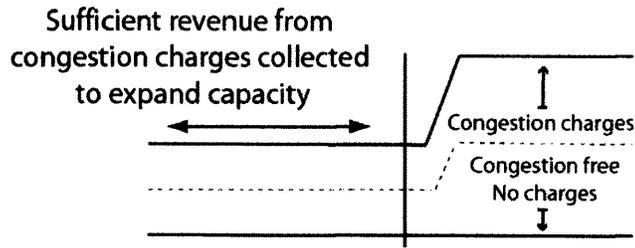


Figure 6-3: The congestion threshold H is increased once R_e revenue has been collected.

mechanism C if their packets receive congestion marks. For instance, a player may have to pay a congestion charge for any congestion-marked packets or congestion-marked packets may be subject to traffic shapers or queue management policies that effectively lower the user utility. Thus, users in a single shot game induced by the mechanism C do not have an incentive to overstate their demands.

6.5 Capacity Expansion Game

We first consider a capacity expansion game. This game is meant to be representative of a provider that funds the expansion of network capacity through charging users that contribute to network congestion.

Consider a network provider that adopts the congestion-based incentive mechanism C . The provider adopts an initial congestion threshold H and charges a player p monetary units for every congestion-marked packet.

Consider also that the provider doubles the congestion threshold H when some variable amount of revenue has been collected. (See Figure 6-3) The desired level of revenues collected before capacity is expanded is given by the exogenous variable R_e that is determined by the provider (for instance it may be the cost of upgrading the congestion threshold H .)

We model the resulting scenario as a stochastic game² where the state variables are the congestion threshold H^t and R^t (we denote the time period with a superscript

²Stochastic game are essentially just repeated games with state variables that can change over time

on variables here and in the rest of the paper). The state variables are transformed according to the transition function:

```

if  $R^t \geq R_e$  then
   $H^{t+1} \leftarrow 2H^t$ 
   $R^{t+1} = 0$ 
else
   $R^{t+1} \leftarrow R^t + p \sum_i c_i$ 

```

6.5.1 Structure of the players

The true type θ_i of a player in our game is the player's true traffic demands at every stage of the game. Since we adopt a discrete model of time, traffic demands are in terms of number of packets. The player's true type is the traffic demand of the player independent of others' demands i.e. the player is not strategizing over anticipated or real effects of their traffic demands on the structure of the game or on other players. The player is willing to pay the congestion penalty for their true demand if congestion happens to occur.

We denote the true type of player i , as a vector θ_i where each element of the vector represents individual demands in a time period:

True type: $\theta_i = (\theta_i^1, \theta_i^2, \theta_i^3, , \dots)$

The revealed type $\hat{\theta}_i$ is the observed traffic load placed on the network by player i .

Revealed type: $\hat{\theta}_i = (\hat{\theta}_i^1, \hat{\theta}_i^2, \hat{\theta}_i^3, , \dots)$

Revealed demand of a user will always be greater than or equal to the true demand of the user. The revealed type thus consists of both the true type traffic demand at every time period and any traffic the user might send to improve its overall utility. This additional traffic could be traffic that has no inherent utility to the user or it could be traffic that has some value to the user but not enough value to justify its potential congestion penalty in any single shot game.

User types then induce preferences over the mechanism's decisions. Preferences in turn are represented by a quasi-linear utility function that sums the cost-adjusted

valuations of a player's demand given their true type and their revealed type at each period of play. So we say:

$$\begin{aligned}
 v_i(\theta_i^t) & : \text{ value to user } i \text{ of their true demand at time } t \\
 u_i^t = v_i(\hat{\theta}_i^t) - pC_i(\hat{\theta}_i^t, \hat{\theta}_{-i}^t, H^t) & : \text{ utility to user } i \text{ of their revealed demand at time } t
 \end{aligned}$$

We assume that users are rational and maximize utility (i.e. they minimize their costs.) Note that the dependency between players arises because of the congestion cost is a function of the actions of all players in the game.

We assume that users are capable of calculating the long-term influence of their demands on either other players or on the state variables of the stochastic game. Let Δu_i^t be the difference in payoff that results in a period t from revealing $\hat{\theta}_i^t$ instead of θ_i^t .

$$\Delta u_i^t(\hat{\theta}_i^t) = (v_i(\hat{\theta}_i^t) - pC_i(\hat{\theta}_i^t, \hat{\theta}_{-i}^t, H^t)) - (v_i(\theta_i^t) - pC_i(\theta_i^t, \hat{\theta}_{-i}^t, H^t))$$

In any single shot game, Δu_i^t would be strictly negative. However, the cost that a user faces in any time period is a function of variables that can be affected by the revealed types $\hat{\theta}^s$ in a previous time period s . Because of the coupling of a user's utility to these previous time periods, the longer term effects of a player's revealed demand differing from their true demands can be positive. We therefore define a *non-trivial long-term strategic player* as a player at a point in time s whose present value of future increased payoffs from not revealing its true demand is strictly positive.

$$\sum_{t=s}^{\infty} \delta^{t-s} \Delta u_i^t > 0$$

6.5.2 Intuitive example

At first glance sending additional traffic under a congestion pricing regime seems like it could only increase a player's congestion charges. However, the following example demonstrates the potentially perverse incentive for players to cause congestion by

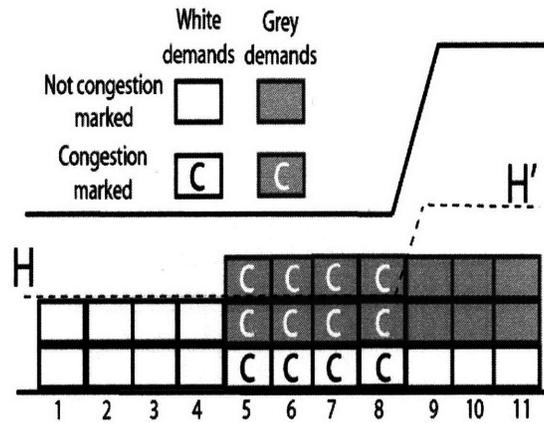


Figure 6-4: Two players, truthfully revealing their traffic demands, share of the capacity expansion cost when the $R_e = 12c$.

sending additional traffic.

Consider two players (perhaps competing ISPs) with the underlying “true” traffic patterns represented in Figure 6-4. Time periods, for instance, may be months. The network sets the initial congestion threshold to be two ($H = 2$) and sets the capacity expansion threshold R_e to twelve units.

If both players reveal their true type and send only the traffic they actually value, the network provider will collect the revenue for capacity expansion within eight time periods. In this case the gray player will end up paying two-thirds of the capacity expansion cost and the white user will end up paying one-third the capacity expansion cost. The gray player ends up paying twice as much as the white player, even though, up to the point of the expansion, the white user sent 50% more traffic than the gray user.

A long-term strategic gray player then would benefit from sending additional traffic to cause congestion when white would normally be able to send freely. By sending additional traffic during the first four uncongested periods, the gray player is able to cause the capacity expansion to occur within half the amount of time and only pay one third the capacity expansion cost as opposed to two-thirds. Thus the gray player would benefit by not revealing its true type assuming that the discounted cost of paying earlier is less than the longer term benefit. (See Figure 6-5).

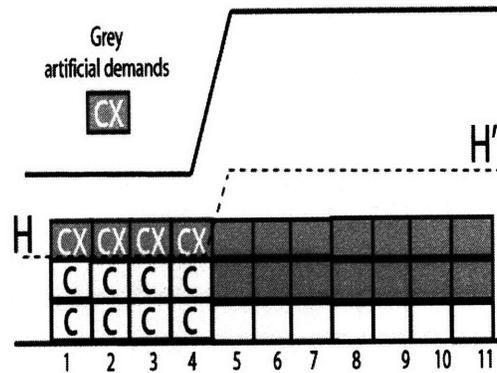


Figure 6-5: By sending additional traffic, gray can reduce by 50% its overall contribution to the revenue required for capacity expansion.

The intuition behind this perverse incentive is that long-term strategic players can extract a matching donation from other players that would not normally be paying a congestion charge. From the perspective of players paying congestion charges, players that send during uncongested periods are free-riding on the network.

6.5.3 Equilibrium player strategies

The equilibrium strategy of a long-term player with no positive expected present value of revealing a demand type is to reveal their true demand in every stage game. A truthful strategy must necessarily also be played by users fundamentally incapable of forming expectations over future outcomes of the game. Both types of players appear to be myopic – each interaction with the mechanism in a repeated game is equivalent to playing the single shot game.

A non-trivial long-term player, however, will reveal any demand that maximizes its long-term utility (i.e. minimizes the players contributions to the capacity expansion cost). The optimal demand level in stage t is computed considering the effect of the demand over the horizon game determined by the players discounting function. Player strategies in a repeated game with perfect information are given by the subgame perfect equilibrium (SPE) solution concept which can be computed using the dynamic programming technique of backward induction [9].

Calculating the SPE strategy occurs as follows. Consider the last congestion period t_c (subscripts c and u here represent congested and uncongested periods) before capacity expansion. Identify the set of users, J in that period that would benefit from making the last uncongested period t_u before capacity expansion a congested period. Calculate the benefit of shifting the congested period taking into account the time value of money by including a discount factor δ , that decreases the amount of traffic a user is willing to send the further in time traffic is shifted. To make the period t_u congested, an amount of additional traffic D_a equal to the congestion amount in period t_c minus the actual aggregate volume of traffic in period t_u must be transmitted by the set of users J . The set of users J split D_a in proportion to their contribution to congestion in period t_c . This process is repeated working backward to calculate the optimal level of demand revealed by all players in the current period.

It is unlikely, of course, that perfect information would be available to players. Hence users can never directly execute the above optimal strategy. However, many players on the Internet today can form reasonable expectations over their own future demand levels and the course-grained aggregate demand.

For instance, diurnal traffic patterns are well known. Given this predictability, consider a corporation with large daily off-site data transfers playing the capacity expansion game. If these data transfers would predictably exceed the congestion threshold H , the corporation would have an incentive to send their traffic during otherwise uncongested periods with the highest traffic demands from other users. This, of course, comes with the non-monetary expense of longer transfer times. The benefit to the corporation, however, is that the capacity is expanded sooner and the maximum amount of money is extracted from others users for the expansion cost.

6.5.4 Worst case cost transfer

While shifts in the capacity expansion costs among players are necessarily dependent on the patterns of true demand, user expectations, etc, we define the *worst case cost transfer* as the maximize amount of a capacity expansion cost R_e that can be shifted between strategic players for any pattern of demands.

We show that all but a fraction, ϵ , of R_e costs can be transferred to a player that would incur no costs under a non-strategic game. Consider a two player capacity expansion game between two players designated *no-cost*, and *all-cost*. Assume that *no-cost* is the only player with true demands for the first k periods of the game. The no-cost player's demands are ϵ below the congestion threshold H in every period.

Now assume that in the next k periods, until a capacity expansion occurs, the only true demands are of the *all-cost* player which exceeds the threshold H in every period. Assume also that, had the first k periods of the game been congested, R_e would have been collected from the *no-cost* player thus causing a capacity expansion. Assume that the all-cost player would never exceed the new congestion threshold H and bear none of the capacity expansion costs.

The strategy of the utility maximizing all-cost player then would be to send ϵ traffic during each of the first k periods causing those periods to be congested. No-cost then bears all but $k\epsilon$ of the R_e costs. Since k can be any integer, a $k = 1$ provides an example of the worst case cost transfer of $R_e - \epsilon$.

6.5.5 Average cost transfers

The worst case cost transfers are useful for determining boundary conditions but do not reveal the distributional information of how costs might be transferred among users with traffic demands representative of players on today's Internet. To gain insight, we therefore conducted a simulation study.

In particular, we were interested in understanding how long-term strategic play shifts costs between light users and heavy users. We define light users as users that bear a smaller fraction of the capacity expansion cost if no users strategically game the mechanism C . Conversely, we define heavy users as users that bear a larger fraction of the capacity expansion cost if no users are strategic. The input to the simulation is a traffic matrix of users' true traffic demands where each element of the matrix is the demand of a user in a time period. We generate this traffic matrix using a probabilistic model of when each user is online coupled with a heavy tailed distribution of traffic demands when the user is online. The other inputs to the

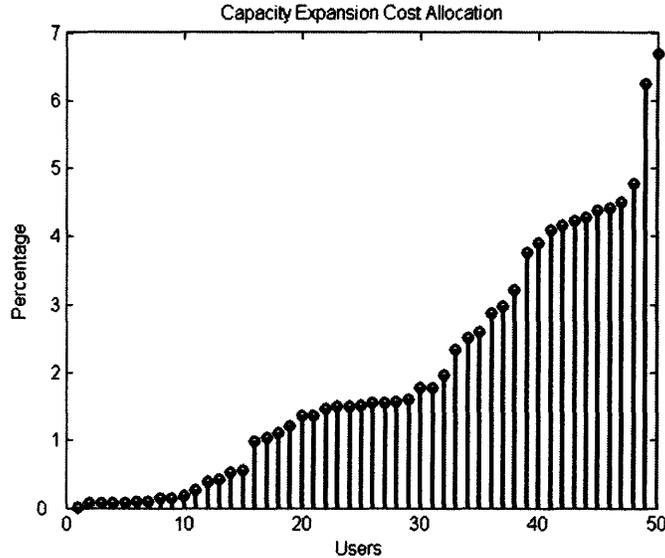


Figure 6-6: Cost allocations from true demands in the capacity expansion game

simulation are the R_e , congestion threshold H , and cost p .

We present a representative example of changes in congestion costs that occur in a population of 50 users. Figure 6-6 presents the cost allocation across the 50 users calculated employing users' true traffic demands. (We sort the users in all figures by their contribution to the capacity expansion cost determined from true traffic demands. Thus, the user index remains the same in all figures.) The user contribution ranges from small fractions of one percent to the heaviest users which contribute over six percent of the capacity expansion cost.

We then calculate the cost allocation changes that occur under strategic behavior using the optimal strategy algorithm outlined in 6.5.3. We present the results in Figure 6-7. As can be seen, there is a visible leveling of the cost allocation across the user population. The light users pay more while the heavy users pay less.

In Figure 6-8 we examine the factor change in each user's contribution to the capacity expansion cost. Each user is color coded according to whether they are better off (black) or worse off (white) under strategic behavior. What we observe is that the light users end up paying a congestion bill that is many times (in a number of cases an order of magnitude larger) than what they would have paid if users did

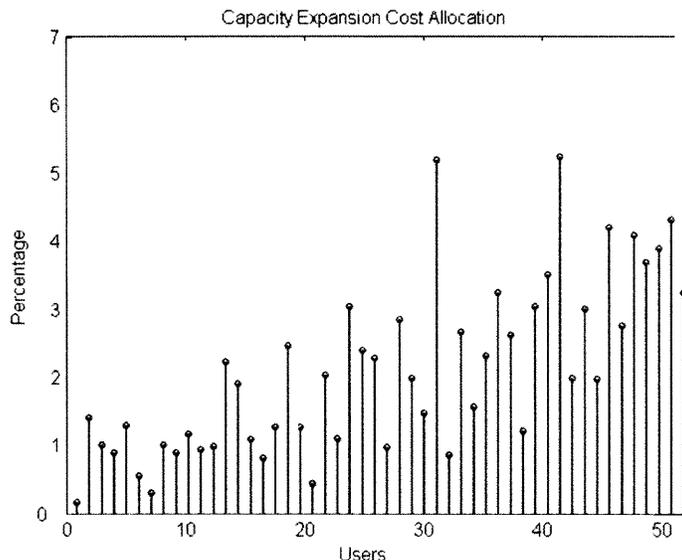


Figure 6-7: Cost allocations with strategic behavior in the capacity expansion game

not engage in strategic behavior. Note that were the users in white unilaterally not to engage in strategic behavior they would be even worse off as the strategic behavior strictly minimizes their contribution to the capacity expansion cost.

This suggests that the users bearing the largest costs of a capacity expansion not only have the most to gain, but that they in fact can successfully gain the most as well. Strategic behavior on the part of the heaviest users then can also be seen as a relative leveling of the costs across the user population (the heavy users costs are decreased and the light users costs are increased.)

6.6 Congestion budget game

The next game we consider is a congestion budget game. The major differences in this game are that providers do not expand network capacity; users, however, have a congestion budget that limits the amount of congestion they can contribute to over defined intervals of time. After the congestion budget has been exhausted, the player no longer transmits in an interval. For instance, a user might have 30 monetary units to spend every month.

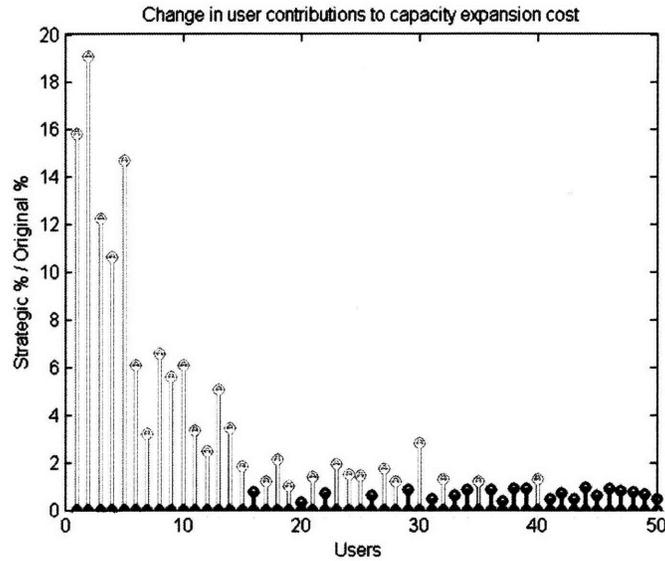


Figure 6-8: Factor change in cost allocations with strategic behavior in the capacity expansion game

In the congestion budget game a provider adopts the congestion mechanism C and charges a price p for each marked packet. Also consider a set of users with true traffic demands and utility functions as specified in the capacity expansion game. Each user also possesses a known congestion budget B that refills on every k th period where k is an integer. If in any time period t , the user exhausts their congestion budget, they reveal a demand of zero in every subsequent stage of the game until the next budget refill.

The long-term strategic interests of players in this game are very similar to players in the capacity expansion game. By strategically raising the congestion costs of others in the game, a long-term strategic player can drive competitors out of the game until the next budget refill if such an action would benefit their own ability to send traffic. The calculated benefit of temporarily driving a competitor from the game is that a player can then transmit their own demands with lower congestion costs until the next budget refill.

Similar to the capacity expansion game, the subgame perfect equilibrium of this game is calculated using backward induction. We emphasize again that this assumes

perfect information. Consider the last congestion period t_c and the last preceding uncongested period t_u (subscripts c and u again representing congested and uncongested periods) before the next budget refill. Only players that have a budget remaining are capable of sending during period t_c . Identify the set of players, J that can transmit during period t_c only if t_u becomes congested. This set of players is identified as the set of players that can transmit during period t_c given congestion during period t_u minus the set of players that can transmit during period t_c even if t_u is uncongested. To make the period t_u congested, an amount of additional traffic D_a equal to the congestion amount in period t_c minus the actual aggregate volume of traffic in period t_u must be transmitted by the set of users J . The set of users J split D_a in proportion to their desired contribution to congestion in period t_c . This process is repeated working backward to calculate the optimal level of demand revealed by all players in the current period.

It is again unlikely, of course, that perfect information would be available to players. Hence users can never directly execute the above strategy. However, some players on the Internet today may be able to form reasonable expectations over their own future demand levels and the course-grained aggregate demand and budgets of others.

Consider, for instance, a congestion budget that a content hosting company establishes for its customers. These budget levels would be known as they would be advertised as part of the terms of service. Customers of content hosting companies might come to learn the traffic patterns associated with their competitors for network capacity and be able to strategically game them. Imagine perhaps online betting sites co-hosted at facilities in the Caribbean.

One online betting company might have the incentive to cause congestion during its competitor's peak hours to try and drive the competitor off the network or at least significantly raise its traffic costs. Of course, the threat of retaliation might keep such behavior in check. But during periods of time that the network is nearly congested, predominately with the traffic of one player, the incentive of the other player to add a bit of traffic to tip the network into a congested state will exist.

6.7 Summary

In this chapter, we have explored some of the economics of a re-ECN enabled Internet. We have discussed how the congestion markings placed by routers do not directly correspond to economic costs of network operators. We have also demonstrated how some users might have a perverse incentive to cause congestion in a network architecture that enforces congestion behaviors with a congestion-based incentive mechanism. By strategically gaming the mechanism, players can create or transfer costs to other players. As we demonstrate in the capacity expansion game, the worst case cost transfer of a capacity expansion cost can be arbitrarily large.

While these are likely explorations of largely hypothetical corner cases, these may call into question whether the “right” set of players are actually bearing the costs. Some might argue that the principle that congestion costs should be only attributed to those that directly cause congestion is weakened when players can indirectly manipulate the mechanism. A player may be free-riding in an uncongested stage only because they caused congestion in an earlier stage.

Like in the security community, expecting a “perfect” engineering designs can unnecessarily hinder the deployment of a predominantly very “good” design. In our view, while not “perfect”, re-ECN represents a good opportunity to improve the economics of the Internet.

Chapter 7

Accepting economic evidence in engineering debates

In this chapter, we turn to a more general problem facing networking researchers as they consider proposals such as re-ECN. Accepting a proposal such as Briscoe's requires the research and wider networking community to be able to balance and integrate both technical and economic arguments and evidence in an engineering design process. However, we have yet to establish objective evaluation metrics for what constitutes good architecture and good research for these techno-economic problems.

In general, when the economics of networks are discussed and debated within the system community, the most common approach currently is an analysis or argument by analogy. Commonly employed analogies in discussions over the economics of networks include references to:

1. Phone networks
2. Road networks
3. Postal networks
4. Electrical grid networks
5. Water distribution systems

6. Airline systems

From our perspective, arguing network economics by analogy is a sign of the relative immaturity of the community's ability to reason about the economics of packet networks themselves. While there are similarities between packet networks and these types of networks (and some lessons to be learned), there are also important differences in the underlying economics.

Based upon our experience, a typical engineering discussion in which reasoning about the economics by analogy is employed tends to devolve into a debate about how the analogy being applied is imperfect. By definition, an analogized network is never going to have identical technical and economic characteristics. This almost unavoidably leads to debate regarding the significance of the differences that exist.

Moreover, the economics of the phone, road, postal networks, etc. are themselves far from simple. Entire books and a vast body of economic literature have been written about them [63, 33]. Reasoning about the economics of packet networks through such analogies then is not likely to, and, in our experience does not seem to, lead to a simpler analysis or clearer conclusions.

It is for these reasons that we do not appeal to these often cited analogies. These reasons also leads us to conclude that the most productive way forward for the network engineering community is *not* to rely primarily on analogies in the economic engineering process.

7.1 Leveraging economics within an engineering process

How then, should the engineering community systematically reason about the economic dimension of the systems we design? While there exists a great deal of diversity in how engineering challenges are approached [103], there are a number of commonalities underlying any integration of economic design and evaluation into an engineering process.

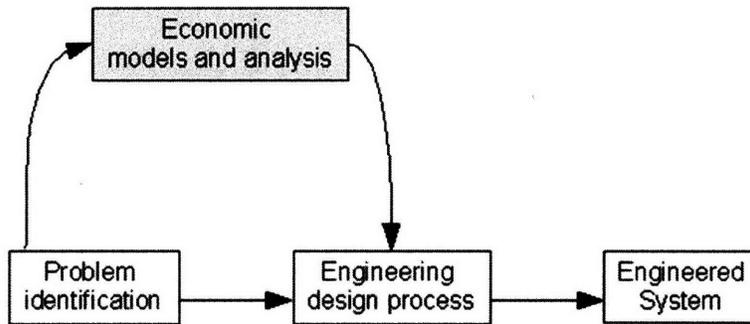


Figure 7-1: It is this closing of the cycle to reconnect the results to the engineering process that distinguishes economic engineering from more theoretical network economics.

This thesis is not an attempt to argue that there is one optimal approach. However, we can distill and address some of the common themes that arise in any engineering process that includes an economic dimension.

The application of economic tools usually involves a process of abstracting ‘up’ to a model that includes simplifying assumptions about the real-world. The economic analysis then produces a set of results that must be integrated back ‘down’ into the real-world engineering process. (Figure 7-1) It is this closing of the cycle to reconnect the results to the engineering process that distinguishes an engineering process concerned with economics from more theoretical network economics. Figure 1-1 presents an abstracted depiction of the process of moving from problem identification¹ through engineering process to a final engineered system.

We see a potential for these economic analyses to convince engineers to make different design choices based upon the presentation of evidence from an economic analysis of the system. In other words, if only technical arguments were brought to bear, then a design choice or architecture ‘A’ would be chosen. But if economic *and* technical arguments are considered then design choice or architecture ‘B’ will be chosen. Critical design choices, then, may be changed by the integration of economic analysis in the engineering process. (Figure 7-2.)

¹Problem identification in networked system can be aided by economic analysis as well.

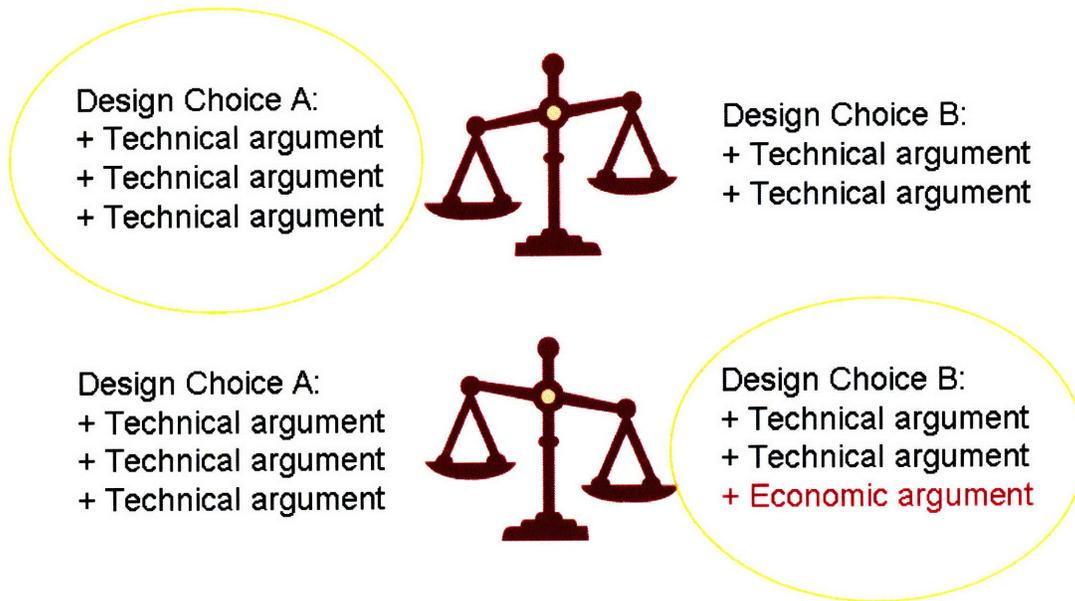


Figure 7-2: Economic evidence or arguments influencing engineering decisions

If we are wrong, and engineers inevitably make a design or evaluation decision based upon technical criteria alone, then the potential worth of economics in an engineering process is greatly diminished. We see economic analyses as only worthwhile if the engineering community weights the evidence produced concurrently with other technical factors and it becomes an influential part of the engineering decision process.

The question then is, what makes an economic argument sufficiently convincing to engineers that it has the potential to alter their design decisions? In the following sections, we address this issue and some of the common impediments to the acceptance of economic arguments in current engineering methodologies and culture. We consider:

1. The basic expectations regarding what economics can contribute to an engineering process.
2. Differences in how the applicability of a theory or claim is evaluated in engineering and economic community.
3. Some common misunderstanding on the part of engineers that hinder the acceptance of economic evidence in a design process.

4. A cultural bias against economics within the larger network engineering community.

7.2 Expectations for economics within engineering

Engineers will be misled and disappointed if they expect or are promised more than the application of economics to engineering problems can deliver. Understanding the potential and limitations of how economics can influence an engineering process then is integral to the credibility of economic analysis within the engineering community.

We deliberately say that economics has the potential to "influence" engineering decisions – at the end of the day, an engineering decision still has to be made. Applying economic tools can make this decision process more systematic and rigorous for certain problems, but economic theory cannot directly dictate an engineering design that is unequivocally "better" than an alternative design.

The reason for this is simple. Economic theory, presented in its conventional mathematical form, is an applied math. Everything in the conclusions follows logically from the initial setup and construction of the scenario being analyzed. Of course, not all economic arguments are presented using their mathematical form. But the majority of economic arguments can, and have been, extensively modeled mathematically. This is exactly what makes economic tools powerful and rigorous.

But similar to statistics, another powerful and rigorous applied math, economics can also be employed in a fashion that is unintentionally misleading. If the initial setup of the problem or assumptions are not 'good' abstractions of the real-world, then an economic analysis will be rigorously analyzing a non-applicable problem.

Economic arguments, thus are not trump cards that automatically win an engineering debate. Rather, the validity and weight of an economic argument must be scrutinized as rigorously as technical arguments are in an engineering design process.

Engineers should expect design principles and not finished designs from economics. The design principles that will guide economic engineering will be very similar to the numerous other design principles that engineers rely upon every day. In networking,

the end-to-end arguments [93] are one of the most commonly invoked design principles. There are many others as well – arguments about modularity, layering, and the use of indirection all guide the construction of overall system architecture. Design principles also guide how the engineer thinks about security, for instance, the separation of a system into the minimal parts that must be trusted and the parts that can run untrusted.

What is valuable about these design principles, including economic design principles, is that they serve as guides, not that they dictate concrete designs. Engineers are very accustomed to balancing and trading-off the concerns highlighted by different design principles. This suggests that the fact that economics theory may not provide a final concrete engineering design is not a problem, but in fact is desirable and to be expected.

In our work, we have not attempted to distill large numbers of economic insights into how network systems should be structured into pithily named design principles. Rather, we have sought to demonstrate through concrete examples how engineering design can and should be informed and guided by economic reasoning. We expect that the rhetorical shorthand for such guidance will naturally acquire common names within the engineering community over time.

Concrete designs motivated primarily by economic theory are rarely impressive when judged from an engineering perspective. We see this most often when a concrete system is blindly engineered so that an author can satisfy the requirements of an economic proof (see chapter 8 and our previous work for some examples [7].) ² An author is then able to provide a proof that the protocol is, for instance, incentive-compatible under some (often fairly fragile) assumptions about how users will behave. The perceived gain, however, in the economic dimension of the system is offset by the loss of important other technical considerations.

²The engineering community, ourselves included, should recognize that not every concrete system architected by an economic-engineered individual is intended to be judged by the full set of engineering criteria. The engineering community should perhaps be more tolerant of those that are attempting to advance economic knowledge even when they may appear to be making rather strange engineering decisions. The theorists are perhaps learning engineering-thinking just as the engineers are learning economic-thinking.

In short, a well-engineered system that satisfies economic desiderata will have been shaped by economic theory but not prescribed by theory. Engineers will be misled and disappointed if they expect or are promised more.

7.3 Applying economic theory to engineering challenges

Economics is as complex and rich a field as engineering and understanding what economic tools to apply and what a theory is really saying is a non-trivial task for those that are not engaged in economic analysis full-time. Economic proofs are often the output of an economist's efforts, however they are the inputs to an engineering effort. Engineers, therefore, need to understand which theories apply in which contexts and what economic tools are most suitable for answering particular challenges.

Without this guidance, economic theorizing can be distracting to engineers. The nature of engineering demands that engineers make decisions even if theory is not available to guide them [103]. In practice, engineers do not fret over this and are accustomed to still making progress.

In evaluating economic evidence as part of an engineering design process the natural inclination of engineers is to consider the assumptions that are made by a theory (as we ourselves did [7]). Why do engineers do this? Simply because assumptions are central to engineering activities – most design requirements and engineering activities follow from a set of assumptions.

For instance, in network planning an engineer makes an assumption about the maximum traffic load a network will be required to carry. The engineer then (perhaps) doubles this load for a margin of safety in specifying the design requirements for a system. Similarly, a protocol engineer might make an assumption about how many routing updates per unit of time that a peer can handle and then designs a routing protocol accordingly. Engineers become methodologically wired over time with a sense of what assumptions are 'good,' 'acceptable,' and 'realistic' for a given situation.

Evaluating assumptions becomes central to what engineers do when evaluating a design.

But this highlights a basic methodological difference between the economists and engineers. For many years an essay by Friedman [44] was handed out to first year graduate students in economics. This essay argued why economic models and arguments should be evaluated only on the basis of the success of their predictions.

“Truly important and significant hypotheses will be found to have ‘assumptions’ that are wildly inaccurate descriptive representations of reality, and, in general, the more significant the theory, the more unrealistic the assumptions.”

and

“The relevant question to ask about the ‘assumptions’ of a theory is not whether they are descriptively ‘realistic,’ for they never are, but whether they are sufficiently good approximations for the purpose in hand. And this question can be answered only by seeing whether the theory works, which means whether it yields sufficiently accurate predictions.”

This leaves those interested in engineering guided by economics with a challenge of how to evaluate the strength of a economic claim regarding a proposed system design. The engineer does not want to build multiple systems to test which economic theory is correct. The goal of the engineer is build something useful, not test theory.

There are no easy or immediate answers to this problem. Rather experienced-backed practice will guide engineers in understanding what economic tools and theories have been successful in which contexts and for what problems. Again, though, this is consistent with the rest of engineering practice. There are no magic answers for understanding when to employ modularity, layering, indirection, or any of the other tools of engineering either. The community is simply at the beginning of understanding how to do engineering guided by economics.

For this reason, chapter 8 is devoted to exploring the economic tool of mechanism design for the Internet. This work contributes to building up a bit of case history within engineering that highlights the applicability of one class of economic tools.

7.4 Common economic misunderstandings among engineers

Another common element that we have observed when attempts are made to integrate an economic analysis and engineering process is a misunderstanding of the basic terminology between the economic and engineering fields. In practice this appears to be a significant hindrance to the acceptance of economic evidence. The natural inclination of engineers is to make an engineering decision based upon technical criteria alone, so even a small misunderstanding can often lead an engineer to dismiss the economic argument.

In this section we do not attempt to reconcile all the language differences between the economic theorists and engineers. However, we will illustrate the point with a number of misunderstandings that we have encountered.

The first misunderstanding centers on a central concept to much of economics – that of an equilibrium. When presented with an economic argument for an equilibrium, engineers sometimes (mis)interprets it as a claim that all parties in the real-world system will act according to the behaviors prescribed by the equilibrium. This sense of equilibrium tends to be seen by engineers as an unwarranted certainty about what behaviors will arise.

However, another real-world interpretation of an economic model's equilibrium is a description of the forces that will tend to influence participants behaviors. When explained this way, a discussion of equilibriums tends to be more palatable to engineers in our experience. This interpretation of equilibrium admits the uncertainty of the real world that engineers are all too familiar with, while still capturing the economic insight that predictable economic forces influence behaviors in the real world.

Another example that illustrates the language gap between the economic theorists and engineers arises over the concept of rationality. Theorists have a tendency to label participants that do not behave in the manner consistent with a model of rational behavior as 'irrational.' 'Irrational' tends to be (mis)interpreted as the theorist saying that if such behaviors arose, they could only be the result of an unpredictable process and are therefore not something that can be designed for or accommodated anyway.

In our experience, this is grating to engineers that are again aware that the behaviors that might be termed 'irrational' are likely to arise all too often because of malice, misconfiguration, or simple differences between the real world and the particular model of rationality at hand. For an engineer, though, these 'irrational' behaviors must be accounted for in an engineering design. Incidentally, behavioral economists have similarly noted that "it is questionable whether functional behavior that violates consistency principles should be called 'irrational.'" [92]. We therefore avoid employing the term irrational when applying economics to real-world engineered systems and readily admit that economic models do not capture all real-world behaviors. This does not negate the value of the economic theory however.

Recognizing that some disagreements and misunderstandings of economic theory lies at the semantic level is useful in making progress toward their resolution. We emphasize that there really is common ground. Finding it, though, requires working through the differences and will simply require time and engagement of both the economic theorists and engineers in networking.

7.5 Engineering aesthetics

Finally, there is often an underlying sentiment among engineers we have observed that talking about economics in sense 'corrupts' an engineered system. Economics destroys or damages the 'simple', 'elegant', or 'beautiful' engineering design. This is a cultural bias against economics within the larger network engineering community.

Indeed, we will readily admit our engineering sensibilities are offended by some designs inspired by economic theory – deliberately causing interference in an ad-

hoc wireless system to induce other nodes to forward ones traffic (i.e. incentive compatibility) is one example [64] we considered in our previous work [7]. We have thought on such occasions if economics *really* dictated that a system be engineered in such a way, we would rather not see the system come into existence at all. We would choose to work on something else instead.

Accommodating economics complicates an engineering *design*. Adding additional requirements almost always makes an engineering design task more difficult. It is not surprising that economic requirements do as well. However, while accommodating economics at the design stage may complicate that particular step, it simplifies the deployment and operational stages. The ‘beauty’ of economic engineering is not found in a system’s design, but rather in the system’s operation.

7.6 Summary

What we have discussed in this chapter is not a step-by-step recipe for how to approach the engineering of economics in designed systems. Indeed, we do not think that any single narrow methodology would be appropriate or convincing. Rather, we see the methodological challenge as being centered around engineer’s understanding and weighting of economic evidence and arguments within existing engineering processes. We have therefore addressed some of the common themes that have arisen for us and that we anticipate will often arise as economic challenges in networking are addressed.

Chapter 8

Mechanism design in networking

This chapter explores the economic tool of mechanism design[55] and its application to network protocol and architecture design. This work contributes to building up a case history within engineering that highlights the applicability of different economic tools. We consider mechanism design because the networking research community increasingly seeks to leverage it to create incentive mechanisms that align the interests of selfish agents with the interests of a principal designer.

To apply mechanism design, a principal designer must adopt a variety of assumptions about the structure of the induced game and the agents that will be participating. We focus in this chapter on assumptions regarding agent preferences and non-repeated vs. repeated games. As we demonstrate, such assumptions are central to understanding the degree to which theoretical claims based upon mechanism design support architectural design decisions or are useful predictors of real-world system dynamics. This understanding is central to integrating the theoretical results from mechanism design into a larger architectural discussion and engineering analysis required in networking research. We present two examples that examine how the valid theoretical claims of [64, 106] relate to a larger, architectural discussion. We conclude with a discussion of some general criteria for designing and evaluating incentive mechanisms for complex real-world networks like the Internet.

8.1 Introduction

The networking community has often designed architectures and protocols that rely on the cooperative behaviors of participants (e.g. TCP). The field of mechanism design, though, suggests that network architectures and protocols can be designed that align the interests of non-cooperative selfish agents with the interests of a mechanism designer. This then would seem to be a powerful theory in which strong claims could be made of agent and system behaviors. However, strong claims are contingent upon many assumptions about selfish agents that may not hold in practice. Claims are weaker if a mechanism only aligns the incentives of the subset of selfish agents that happen to match a principal's underlying assumptions.

While adopting simplifying assumptions can enable a network architect or protocol designer to prove theoretical properties such as the incentive compatibility or efficiency of a mechanism, these results may not always be useful predictors of actual agent behaviors or system dynamics when a mechanism is deployed in practice. While the simplifying assumptions of any theory are often easy to criticize from a practical perspective, the point of this chapter is to not to criticize. Rather we hope to further the use of mechanism design by the networking community by promoting a better understanding of the contexts in which mechanism design succeeds (or fails) in practice to improve network and protocol designs.

While mechanism design requires simplifying assumptions – rationality, common knowledge – we focus on what mechanism designers can know about agent preferences and what they assume about whether the induced game will be a single-shot or repeated game. We focus on these assumptions because they tend to influence how applicable the theoretical results of mechanism design are in practice for the networking community.

The first class of assumptions we examine is the structure and type of agents' preferences. We consider how applicable various assumptions about agents' utilities are in practice. While the majority of agents participating in a game induced by a mechanism may match a designer's assumptions, it is likely that at least some agents

will fail to conform to a mechanism designer's expectations in networks as large, complex and diverse as the Internet.

The second class of assumptions we examine deals with whether the game induced by a mechanism is part of a larger repeated game. In mechanism design the induced game is typically analyzed as a single-shot game. However, in networking, agents will interact repeatedly with mechanisms experiencing, over time, multiple mechanism outcomes. We therefore consider mechanisms which induce an outcome in a stage-game of a larger repeated game. It is well known that the equilibria of repeated games can be different than the equilibria of single-shot games [45]. Indeed, previous research in the networking community has noted the effect of repeated play on various routing mechanisms [2]. This chapter considers the implications of the folk theorem [45] for mechanism design in any repeated context – a context that is very common in the real-world networking environments.

The rest of this chapter is organized as follows. In §2 we discuss mechanism design for the Internet. In §3 we examine the assumptions about agent's preferences and provide an example of a mechanism, Re-Feedback [17], in which assumptions about the agents play a critical role in understanding the incentive compatibility claims. In §4 we examine assumptions about the number of times an agent plays the game induced by a mechanism and examine the implications for an ad-hoc routing and forwarding protocol [106] that is designed to be incentive compatible. In §5 we discuss the impact of our arguments on mechanism design for the Internet. In §6 we conclude with a summary.

8.2 Mechanism design

Since this chapter is targeted primarily at the engineering community we begin with a brief review of mechanism design. As described in Fudenberg [45], mechanism design can be viewed as a multi-step game of incomplete information where agent “types” are private information. (See Figure 8-1.) In the first step of the game the mechanisms designer, or principal, designs a “mechanism”, “contract” or “incentive scheme.” The

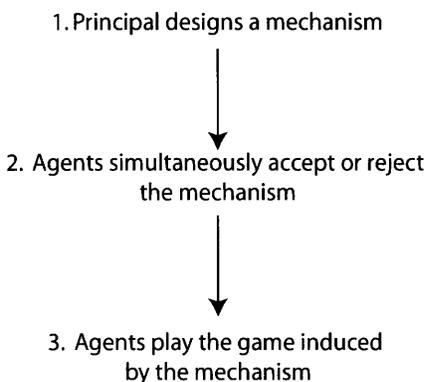


Figure 8-1: Mechanism design viewed as a three step game of incomplete information.

objective of this mechanism is to elicit “messages” or “behaviors” from agents such that the mechanism’s designer, or principal’s expected utility is maximized. In the case of a benevolent principal, the expected utility that is maximized is some notion of social welfare. As network architects we often optimistically view ourselves as such benevolent principals, designing mechanisms to improve some notion of overall social welfare for the network.

In next step of the game, each agent either accepts or rejects the mechanism designed by the principal. Agents that accept enter the third step and play the game induced by the mechanism. Playing the game entails sending messages that are selected based upon an agent’s private “type.” In a networking context, one can interpret sending a message to a mechanism as engaging in a behavior that is observable to the network providers or other network participants.

The outcome of the game induced by the mechanism is called an “allocation” or “decision” which is computed by the mechanism from the agent messages. The allocation consists of an assignment of goods and transfers of numeraire [55]. The allocation, for instance, in a Vickrey-Clarke-Groves (VCG) based lowest-cost routing mechanism [36] consists of a selection of routing path and numeraire transfers of monetary payments to each of the nodes on the lowest-cost path. More generally, numeraire can be in terms of anything the agent values; often these are monetary transfers, but they can also be tokens that are valuable within the context of the mechanism. For instance, in mechanisms designed for the Internet these tokens might

represent the right to transmit in a wireless network or they might be tokens employed in a traffic-shaping token-bucket.

The problem facing the mechanism designer is how to construct the message space and allocation rules such that it is in the interest of agents to truthfully reveal the private information that the principal conditions its allocation decisions upon. Said another way, the mechanism must be designed to align the interests, behaviors, and actions of the agents with the interests of the mechanism designer.

In designing a mechanism, the principal is assumed to have some leverage over the agents, which influences the agents' choice of messages or behaviors. This leverage is rooted in the principal's control over how goods and transfers are allocated to the agents. When designing a mechanism for the Internet then, an important question is what can a principal assume with confidence about agent preferences? The answer to this question is critically important to both the design of the mechanism as well as the equilibria that will result in the induced game. These are the topics that we consider in the following section.

8.3 Assumptions about preferences

In the game of mechanism design, each agent has a private type θ_i that determines the agent's preferences over different allocations.¹ (Though we do not rely upon it heavily, we include some of the formal notation here so that any economic readers can readily identify with our discussion.) Agent types are assumed to be drawn from a known set of types $\theta_i \in \Theta_i$. The vector of all agents types is denoted as $\theta = (\theta_1, \dots, \theta_I)$ drawn from a vector of possible types $\theta \in \Theta_1 \times \dots \times \Theta_I$ with a probability density function $\phi(\cdot)$. Each agent is also assumed to be an expected utility maximizer where the agent's utility function for an allocation k from the mechanism is denoted $u_i(k, \theta_i)$.

Many mechanisms designed by the networking community have further assumed

¹See (23.B) "The Mechanism Design Problem" [69], for a more complete introduction to the assumptions summarized in this section.

that the structure of agents' utility functions have a quasilinear form:

$$u_i(k, \theta_i) = v(k, \theta_i) + (m_i + t_i) \tag{8.1}$$

where m_i is the agents i 's initial endowment of the numeraire, k is the allocation of the good, $v(\cdot)$ is the agent's valuation function for the allocation, and t_i is transfer of numeraire to/from agents. Quasilinear utility functions are popularly adopted because utility can be transferred across agents through transfers of the numeraire.

A key assumption is that the probability density function $\phi(\cdot)$, the complete sets of possible types for each agent $\Theta_1, \dots, \Theta_I$, and the structure of the utility functions $u_i(\cdot, \theta_i)$ are all common knowledge. In other words, all of this information is known by all agents and the principal. The *only* private information is the actual type of each agent θ_i .

By assuming that the complete set of possible types for each agent and structure of all utility functions are known, the principal can theoretically anticipate the effect of incentive mechanisms upon agent behaviors. If all these assumptions are sound, i.e. the assumptions accurately represent agents' utilities and types in the real world, then the mechanism designer can with confidence predict and describe the behaviors and equilibria in the game induced by a mechanism.

8.3.1 Assessing utility assumptions

The question, from a network architect's perspective, is what should be assumed about agents' utilities in network environments? In this section we consider mechanisms that assume agent utility functions are composed of terms representing the value of monetary and/or network goods to an agent.

Monetary utility terms

A monetary term that captures an agent's increase in utility with increased monetary assets seems fairly safe to assume in a utility function (Equation 8.1). Most selfish agents would seemingly prefer a larger amount of monetary goods to a smaller amount.

However, even this relatively safe assumption may not hold when considering the time value of money. An real-world agent may be willing to incur a smaller short-term loss for a larger long-term gain. If real-world agents are willing to incur losses within the induced game for longer-term gains realized in a different larger or repeated game that includes the induced game, a mechanism designer has more limited leverage to shape the agents behaviors through monetary incentives. In effect, the real-world agents will not be playing the game the mechanism designer intended.

This is interesting because it suggests that even monetary rewards or penalties may not create the incentives expected by a principal. While perhaps most likely to occur over monetary goods, a tolerance for short-term losses for longer-term benefits also potentially effects how real-world agents value network characteristics such as the ones discussed in the next section.

Network-based utility terms

The next class of terms in an agent's utility function (Equation 8.1) we consider are ones that represent the value of network goods. These are terms that represent the value of network performance characteristics such as throughput, latency, and loss as well as more general goods such as transmission or access privileges on a network. It is often assumed that agents' utilities are an increasing function of improvements in the network good. Assuming that agents have traffic they want to send or receive on the network, such assumptions seem at first plausible.

However, a lesson from years of quality of service research in the networking community is that simplistic models of agent utilities are inadequate in the real-world [5]. Assuming that agents always value improvements in any one metric of network performance, such as throughput, latency, or network access fails to describe any one individual agent let alone being a good model for all agents on the network [5].

Moreover, in most networks today, agents are actively seeking to send or receive traffic only a small fraction of the time. During these periods, incentives leveraging the fact that agents value improvements in network performance will be effective. But,

this raises the question of what governs and motivates an agent's behavior during other times? One is tempted to answer that real-world agents will just cooperate during periods when they themselves are not selfishly invested in how the network is performing. But such a response weakens the strength of claims that can be made regarding a mechanism, particularly in networked environments, such as we have today, where actively malicious behaviors are common.

Finally, we note that a mechanism deployed in the real world cannot selectively admit only those agents that are well modeled by the utility functions assumed by the principal. Agent utilities and types are inherently private information. Agents that do not conform to a principal's assumptions may participate in the induced game. Therefore, the claims that can be made for a practical mechanism must be seen as limited to the subset of selfish agents that match a principal's underlying assumptions.

8.3.2 Example: Jamming in wireless networks

In this section, we provide a concrete example where the assumptions about agents play a critical role in understanding the incentive compatibility claims. We selected this paper [64] as an example because it leverages economic theory in support of a what is claimed as a real-world engineering design effort and was presented at a premier venue that brings together the economic and engineering communities. We first provide a brief overview of objectives.

Objectives and incentives

The authors seek to design a system such that all nodes in an ad-hoc wireless network can create an incentive for their neighbors to forward their packets. This is a common objective in the wireless networking research community. Other attempts to create such incentives have relied upon detection and isolation of nodes that fail to cooperate (e.g. Mahajan[67]) or some form of a payment scheme between nodes (e.g. Andereg[4]).

The authors of the paper identify situations where asymmetric demands between nodes in a wireless network leads to certain nodes not having any natural incentives to forward other nodes traffic. An easy to imagine scenario is along the edges of an ad-hoc wireless network where the node that is the second-hop toward the interior of the network does not have any natural incentive to forward the traffic of one of its leaf neighbors.

The authors suggest that an incentive can be created for neighboring nodes to forward traffic if a node engages in a strategy of “jamming” the wireless channel if its neighbors fail to forward its traffic. They suggest this jamming could be accomplished by sending meaningless broadcast packets. Given the radio technology that is the focus of this effort (802.11), such interference could be technically effective in disrupting communications.

Claims and analysis

They claim and prove that a subgame perfect Nash equilibria exists where incentives to forward traffic are created if nodes that do not forward are jammed. In other words, the best strategy for all nodes as predicted by the model is to forward others traffic at all times. If this strategy is not followed, the model predicts that the non-cooperating nodes will be worse off because sending of their own traffic will be jammed.

There are technical, regulatory, legal problems with this proposed jamming, but even from the narrow issue of creating the authors’ desired behavior, the proposal is strongly dependent upon an assumption about what nodes value. The explicit assumption underlying the model is that nodes value continuous connectivity and network throughput. Jamming then can deprive network nodes of something that they value thereby creating an incentive to forward other’s traffic.

Jamming, however, is less effective if nodes in the real world do not conform to the modeling assumptions that all nodes value connectivity and throughput at all times. Indeed, as appears to be the most common case today, nodes in an ad-hoc wireless network have very intermittent communication i.e. they value connectivity and throughput sporadically.

Consider a neighbor node that fails to forward an edge node's traffic in an ad-hoc network. In the model, the edge node should "punish" the neighbor by jamming the wireless channel. If this jamming is done immediately upon detection of the failure to forward traffic, it will only be effective if the other node is actively attempting to communicate. Otherwise, it will simply waste any resources such as the battery power of the node attempting jamming.

If the jamming node attempts to be strategic in when it applies the punishment for a previous failure to forward, say by waiting for the neighbor node to initiate a communication, the neighbor node can either 1) wait out the punishment phase or 2) initiate a fake connection soon after it fails to forward the other nodes traffic so as to induce the punishment at time that will have no real effect on its own communication needs.

This example illustrates that theoretically provable system-wide properties should be examined to understand their real-world impact. In this case, our own engineering judgment is that designing such a behavior in a real world networked system would be highly undesirable and not nearly as effective as the economic argument presented suggests.

8.4 Mechanism design for repeated games

Classically, mechanism design is viewed as inducing a single-shot game. However, when mechanism design is applied to the construction of protocols and architectures for networking problems, it is actually more likely the same agents will be repeatedly playing the game induced by the mechanism.

From this perspective, the outcome of a game induced by a mechanism must be seen as the outcome of a stage-game i.e. one iteration of a single-shot game, in a larger repeated game. However, the effect of incentives in a single-shot game can be different than in a repeated game. The classic example of this is the prisoners' dilemma game. The only equilibrium in the single-shot game is for each prisoner to defect and take a plea bargain. However, in the repeated game, staying silent can

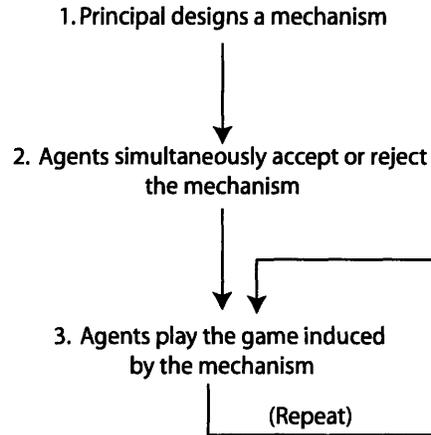


Figure 8-2: Mechanism design where agents repeatedly play the game induced by the mechanism.

also be an equilibrium strategy [45].

In general, any mechanisms designed by the networking community that are repeatedly played must be analyzed as repeated games. This entails that important theoretical results in repeated games must be considered – namely the “folk theorems” for repeated games (see Fudenberg [45] for formal statements of the folk theorems in repeated games).

The folk theorems assert that, if players are sufficiently patient, any individually rational, feasible outcome can be enforced by an equilibrium. To be individually rational, players select actions in each stage game that minimizes the maximum possible loss that they will face in the overall repeated game. A feasible outcome is one in which the rationality condition is satisfied for all agents. Thus, in a repeated game, almost any outcome can be an equilibrium outcome [45].

But since any feasible outcome can be supported for the repeated game, this raises the question of how much influence the incentives of the mechanism designer inducing each stage game has over the overall equilibrium of the repeated game. Consider again the classic prisoners’ dilemma cast as a mechanism design problem. The principal representing the justice system wants to allocate prison sentences in such a way that induces guilty suspects to defect from their partners and tell the truth about their crimes i.e. the principal wants to design an incentive compatible mechanism.

However, in the context of a repeated game, prisoners will always maximize their utility by continuously remaining silent in each stage-game. Such an equilibrium can be enforced, for instance, if agents adopt tit-for-tat or grim trigger strategies that punish any agent that ever defects [45]. If the penalty allocated to two prisoners that both remain silent is always lower than the penalties if they both defect, then the principal representing the justice system cannot design an incentive compatible mechanism that will induce the prisoners to talk.

In different contexts, though, a principal can, to a degree, influence the equilibrium of the repeated game through the design of the messages that each agent can send to the mechanism. The work of Afegan [2], for instance, considers the effect of protocol periods and field granularity on the equilibrium price computed by a routing protocol. We are unaware of general results in this area; it appears that each mechanism must be analyzed individually in the context of a repeated game to understand what effect the control of incentives in each stage game will have over the equilibria in the repeated game.

8.4.1 Example: Costs in wireless-hoc networks

To illustrate the importance of considering repeated games in the engineering of network protocols we consider the incentives created in ad-hoc wireless networks by the protocols described by a paper in one of the premier conferences on wireless and mobile networking [106]. We focus in particular on the protocols for the routing stage. For specific details of the protocols, consult the paper [106].

Ad-hoc routing and forwarding protocols

The goal of the routing stage is to compute the true costs, i.e. the power levels required for transmission, for each link along a path in the ad-hoc network. Based upon these costs, the price paid to each node on the lowest cost path is computed similar to the VCG-like mechanism presented in [4]. The basic intuition behind this calculation is to pay each node the cost of the second best path through the network

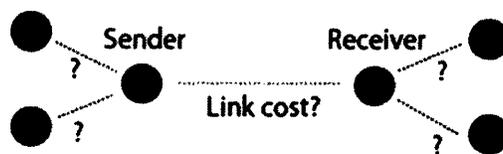


Figure 8-3: The challenge in wireless ad-hoc networks is that the cost of a link cannot be determined by the sender alone. Receivers are an integral part in reporting what power levels the sender must employ.

that does not include them.

The challenge in wireless ad-hoc networks, the authors note, is that the cost of a link cannot be determined by the sender alone. Receivers are an integral part of reporting what power levels the sender must employ for a successful transmission. In certain circumstances, receivers have an incentive to misreport the power levels that a sender requires for transmission. (Figure 8-3.) (See section 4.1 of [106] for an elaboration of these cheating behaviors.)

To address this challenge the protocol designers employ a cryptographic solution that involves sending multiple messages, encrypted under a key shared with a third party, at increasing power levels. The receiver transmits all received messages to the third party that decrypts the messages and computes the true cost of each transmitting node.

Claims and analysis

A series of claims are made regarding the behaviors that these routing and forwarding protocols will induce [106].

- “We ... design the first incentive-compatible, integrated routing and forwarding protocol in wireless ad-hoc networks.” (p.13)
- “We show that following the protocols is a dominant action for [the routing stage.]” (p.13)

In other words, the authors are making a strong economically-based prediction of the behaviors that will result if the real world system was engineered with their

mechanism. They are claiming that nodes will have no incentive to deviate from a truthful strategy.

These claims, however, are based upon an analysis of a single-shot game induced by the protocols. However, routing and forwarding in an ad-hoc network will unquestionably be a repeated game. In the real-world, routing algorithms continuously recalculate paths as nodes come and go from the ad hoc network and the connectivity of the network changes.

As Afergan [2] notes, agents can advantageously deviate from the behaviors they would exhibit in a single-shot version of a VCG-based routing game. Namely, agents that collude (either implicitly or explicitly) have an incentive to reveal costs that are higher than their true costs so that they can enjoy larger payments from the mechanism. The designer's goal of truthful revelation of costs is potentially thwarted in any repeated VCG game.

This is interesting because the cryptographic approach taken in this protocol represents considerable engineering effort to align the incentives of a single-shot game. Each node in the network has to be capable of supporting cryptographic operations – often a task that is expensive in terms of computational resources. Each nodes also sends a series of additional messages each time the routing cost calculation is performed – a pure overhead cost that drains power resources (which are often in limited supply in ad hoc networks.)

If, in fact, the game is repeated, admitting other agent behaviors, the effort at aligning the incentives in the single-shot game appears significantly less worthwhile in the context of an engineering cost analysis. The addition of the crypto mechanisms would have little real world impact on the strategic behaviors of nodes.

8.5 Discussion

This chapter can be seen as an examination of applying economic theory to engineering practice. While simplifying assumptions are crucial to employing theory and models, this necessarily entails that any model will not capture all details of the real-

world. What is crucial is to understand when theory and models provide support and understanding of a system design versus when they are no longer applicable.

The theory of mechanism design can “raise the bar” of networking and protocol designs even if it does not accommodate all the types of selfish agents on the Internet or perform exactly as expected in a repeated game. We emphasize that we do not consider a mechanism designed for the Internet to be a failure simply because one can construct agents that do not meet the principal’s assumptions. Creating an incentive mechanism that aligns the interests of a subset of agents is an improvement over a design that assumes full cooperation.

But understanding the real-world limitations of theoretical claims is important from an architectural and system engineering perspective. It is this understanding of the real-world limitations that enables the theoretical claims to be integrated into a larger architectural discussion and engineering cost analysis. While there is not a rigorous framework in which to conduct this discussion and analysis, we offer the following criteria for designing and evaluating incentive mechanisms for complex real-world networks like the Internet.

1. *Explicitly state assumptions:* Understanding the implications and applicability of mechanism design requires the underlying assumptions to be explicitly stated so that they can be analyzed and judged for soundness.
2. *Design defensively:* Network architectures and protocols should not rely upon incentives derived from mechanism design alone to ensure that desirable system dynamics are achieved. At least some agents in any network environment will not conform to the assumptions made in a theoretical model.
3. *Understand the limitations of simple models of utility:* Assuming that agents always value improvements in any one metric of network performance, such as throughput, latency, or network access is not a applicable model of real-world agents.
4. *Analyze the repeated game:* Many mechanisms designed for networks will, in

fact, be repeatedly played. The incentives created by this repeated play must be analyzed as a repeated game.

8.6 Summary

This chapter focuses attention on the underlying assumptions about agents and how they will interact with mechanisms in complex networks like the Internet. We have emphasized that strong economic claims are contingent upon assumptions about the selfish agents and how they will interact with a mechanism. We have suggested that claims should perhaps be interpreted more narrowly if mechanisms only aligns the incentives of a smaller subset of selfish agents that match a principal's underlying assumptions. But we emphasize that creating an incentive mechanism that aligns the interests of a subset of agents can still be seen as an improvement over a design that assumes full cooperation from all agents. Finally, we emphasize that mechanism design cannot be a substitute for a systems engineering perspective.

Chapter 9

Related work

Our thesis sits at the intersection of two areas – literature that applies economic theory to networking problems and the traditional computer science systems community literature on networks. This chapter looks at the literature that has been most relevant to our work.

9.1 Economics and congestion management

A paper by Briscoe [16] (deliberately provocatively written) takes the network engineering community to task for adopting flow-rate fairness as the basis for determining how network capacity should be allocated. He argues that there is no basis in economics or theories of justice for enshrining flow-rate fairness as a desirable technical objective. While we disagree with some aspects of this paper, it presents a sophisticated argument for why the core network architecture and protocols should support economic-based allocations methods for network resources.

Feigenbaum et al. [36] describe a mechanism that creates incentives for network providers to reveal their true costs (congestion or otherwise) using a standard routing protocol such as BGP. Some of our work looks at the dual of this problem. We assume that providers reveal their true costs and we consider the incentives of users to reveal their true demands. We believe we are the first to address demand-side incentive issues for Internet congestion mechanisms.

A part of our work focuses on incentive-compatibility because of a normative judgment of what properties a congestion-based incentives mechanism ought to possess. Other desirable properties of cost sharing are considered in the work of Moulin et al [72] on axiomatic cost sharing. Directly related is Moulin's notion of demand monotonicity which states a user cannot decrease their share of the costs by increasing their demand.

Some of our work can also be seen in a similar vein to past literature that has criticized Internet congestion-pricing research [96]. Congestion pricing is often criticized for creating artificial incentives for network providers to under-provision their networks thereby raising their revenues. Part of our work suggests incentive problems are created on the demand-side even if supply-side incentives can be addressed.

Many others have noted that it is technically difficult to implement congestion pricing in a cost effective manner [78]. This criticism though is not fundamental and can be potentially negated by innovative engineering solutions. Work on creating congestion mechanisms has ranged from the theoretical [65] to technically viable mechanisms [17].

Congestion-based incentives are criticized for their perceived failure to match the structure of incentives that users prefer. Historical evidence of user preferences and the evolution of pricing is presented by Odlyzko who notes that service plans, pricing, and user incentives evolve toward simpler structures over time [78]. However, congestion-based incentives can be packaged in ways that are more acceptable to users. Briscoe et al, for instance, demonstrates how users can be presented with a flat-rate congestion-based service plan [15]

9.1.1 Models of network traffic and dimensioning network capacity

Central to the economic and technical question of how network resources are allocated over different time scales is the nature and aggregation behavior of network traffic. While mathematical characterizations of Internet traffic are often contentious topics

(see [104]) the basic consensus is that Internet traffic is self-similar at different time scales [43]. Analysis of traffic measurements has shown that a large portion of traffic exhibits self-similar behaviors. [101, 43, 85] provide overviews of the large body of literature and results in this space. One has to be careful how these models are leveraged. As noted in [27] commonly made assumptions, such as independence of traffic flows, are often incorrect and can have dramatic impacts on the conclusions drawn.

The concept of effective bandwidth attempts to quantify the resource usage of a variable bit traffic source by summarizing the statistical characteristics of a traffic source over different time and space scales [58]. Knowing the effective bandwidth of a traffic source could be useful in provisioning network links. For example, a stream with an effective bandwidth of two can be replaced by two streams with an effective bandwidth of one. Experimentation with real traffic ([61, 27] and theoretical results ([58, 28]) have shown that the effective bandwidth of a stream depends on the network resources (link and capacity resources) and surrounding traffic mix. Quantifying the resource usage independent of these factors leads to under or over estimating the required resources. The most widely used formulation of effective bandwidth in the academic literature is given by Kelly [58].

Effective bandwidth though is not as meaningful for elastic data applications that do not have an inherent transmission rate but rather adapt to the available bandwidth in a system. [31] takes a simulation based approach in analyzing the aggregation properties of bursty Internet traffic while [105] provides an analytic look at the aggregation properties of bursty Internet traffic. As the volume/month approach for relating individual costs to aggregate costs becomes less applicable, approaches such as those mentioned here that explicitly take into account the bursty behavior of traffic will be required.

While these models of network traffic capture the behavior of traffic over a range of time scales, these time scales are on the order of multiple days or weeks. There is less consensus on the long-term (i.e. months or years) *trending* patterns of network traffic. Odlyzko [79] examines some of the myths of astronomical Internet traffic growth

rates during the Internet boom years and presents evidence regarding past and future growth rates. In [46, 20], the authors present evidence of the traffic growth in the residential broadband market of Japan. This is a particularly interesting study since many Japanese broadband users have higher access capacity links into the Internet and may be bellwethers for how demand will evolve in other locations around the globe.

9.2 General Internet economics

In this section we focus on some of the economics literature that has been read most widely by the engineering side of the networking community. This provides some context for understanding the economic arguments that many networkers are familiar with and some context for understanding the resistance of traditional engineers to consider economic engineering.

Based upon our experience, pricing is the first topic that most engineers think of when the economics of networks is discussed (for a good overview see the included papers in [70].) A particularly strong focus of the economic literature has been various forms of congestion-based pricing. There are two main forms of this congestion-based pricing. The first is motivated by the “Smart-Market” style pricing where network end-users “bid” for how much they value each packets or streams of packets [66]. In the second approach, users do not reveal their valuations to the network but rather respond to signals from the network regarding the current “price” for sending traffic. Each ECN congestion-marked packet, for instance, may be charged a fixed price [15].

The practical influence of this literature to this point in time has been limited. The engineering perception is that economists have been insisting on the importance of adding more pricing information to the Internet even though tremendous growth and success have followed from not adhering to this particular piece of economic engineering advice.

Engineers’ most common reference to rebut those that advocate more complex pricing mechanisms is the work of Odlyzko and his co-authors who are famous for

emphasizing the importance and trend toward simple pricing for high volume transactions like would exist in a network[77, 78]. Odlyzko’s work provides an economic and technological history that highlights the strength of engineering arguments for keeping such systems simple. These are, in essence, historical lessons that engineering concerns for simplicity can trump economic arguments for more economically efficient, but complex, systems. ¹

9.3 Game theory and mechanism design

Two other fields of economics, game theory [45] and mechanism design [55], have been extensively applied by the theory community to networking problems. While the engineering community has often designed architectures and protocols that rely on the cooperative behaviors of participants (e.g. TCP). Mechanism design and game theory, though, suggests that network architectures and protocols can be designed and evaluated in ways that consider the interests of non-cooperative selfish agents. These then appeared to be powerful theories in which strong claims could be made of agent and system behaviors.

The work on Distributed Algorithmic Mechanism Design (DAMD) [38, 37] emphasizes the importance of the algorithmic properties of mechanisms designed for the Internet. They introduce the notion of protocol-compatibility which focuses on two aspects of the practical feasibility of a mechanism: the computational tractability and deployability of a mechanism.

The work of Afergan et al. [1, 2, 3] emphasizes the importance analyzing networking problems as repeated games. One of the focuses of this work is that the mechanism designer can influence the equilibria that occur in an incentive-based routing mechanism by controlling some of the protocol parameters such as the period lengths and granularity of protocol fields [2].

¹Illustrating a bit of the gulf between economists and engineers, one economist’s response in a class conversation to an engineer that was citing Odlyzko’s work was (we are paraphrasing, but to the best of our recollection): “Economists rely upon analysis, not anecdotes. Anyone that makes billion dollar investment decisions today based upon what happened to England’s postal system [78] in the 1600’s is an idiot.”

Practical experiences applying mechanism design and game theory to networking problems are reported in Mahajan et al. and Huang et al. [52, 68]. Both note that theory does not necessarily apply as completely or easily as one might initially have hoped.

Establishing the fidelity between model design and implementation is a recognized problem in the economic literature. The work of [97] considers how to prove, under certain assumptions, that an implementation of a mechanism in real-world system will match a designer’s specification. Our work similarly considers the fidelity between model and real world system from an engineering perspective.

If a mechanism designer can adaptively re-implement a mechanism repeatedly, learning the participating agents’ stationary preferences over time, then any outcome the designer cares about can be implemented in dominant-strategies (a strong solution concept in economic theory) [56]. In the case of most networking problems, though, changing a mechanism that is deployed in the real world is often very difficult so, unsurprisingly, perfection is only possible in theory.

9.4 Engineering perspectives on economic problems

The computer science systems community has recognized the open challenges of networking are not purely technical in nature. Our work complements and re-enforces the messages found in the literature cited here that the inclusion of design and evaluation criteria that are not purely technical is vitally important to successful engineering efforts.

Accommodating different stakeholders with potentially adverse interests is identified as a “tussle” by Clark et al [24]. This paper deliberately does not cast these “tussles” in economic terms. They do note economic considerations are a strong force shaping these tussles, but they catalog other tussle forces as well. (We do recognize that some economists would argue that all social issues can best be analyzed as eco-

conomic problems.) While we recognize the power of economics, other frameworks for reasoning such as legal, moral, or philosophical may be more useful in approaching some parts of networking tussles. Our own focus, however, is on networking tussles that can be best approached using the tools of economics.

Beyond problem identification, though, the tussle paper argues for technical design principles that can guide engineering and architectural work. They argue that the design process should ‘design for variations in outcomes’ and “modularize along tussle boundaries”. (See the paper for concrete examples of these design principles in practice.) These are designed to shape and influence the engineering methodology of networking work.

Learning from past failures is central to how engineering knowledge is advanced [88]. Networking research has had its share of (perceived at least) failures. While perhaps the formal literature in this area is not as widely read by engineers, the engineering community has a general understanding that the failures of many network engineering efforts have both a technological and economic origins. We take this as a good sign that engineering-minded networkers respect the importance and need to understand the larger economics issues.

Workshops have been convened to summarize the engineering lessons learned from the perceived failures of protocol efforts (e.g. [5]). Bell [8], for instance, emphasizes the balance required between engineering and economic concerns in the design of quality of service protocols (QoS).

In the space of multicast protocols, that have been similar efforts to understand and correct the perceived mismatch between the economic incentives and technical effort required by network operators [51, 39]. This body of work illustrates how understanding the economic incentives of players can dramatically influence how a networked system is designed.

9.4.1 Systems engineering

Closely related to our efforts is the field of engineering system design (also called systems engineering)[30, 100, 19]. This field looks at how systems can and should

be designed when multiple engineering fields (and other) concerns are simultaneously present. Our ambition is not as broad, but this community has valuable experience and models for how to transfer and integrate the tools and insights of multiple disciplines in building better systems.

We emphasize that system's engineering does not deal with just the integration of multiple *engineering* fields. For instance, cognitive psychology, a decidedly non-engineering discipline is applied to design systems that explicitly accommodate how humans think. Books by Donald Norman (geared toward a general audience) [73, 74] are representative of work that consider the challenge of designing for human users.

If engineering design can work toward integrating cognitive psychology, then integrating economics into network systems engineering should be a relatively simpler challenge. We suspect that economics is inherently more tractable and observable than the vagaries of human cognition. Yet, engineering of the economics of systems remains less developed as a field.

9.5 Internet measurements

CAIDA[22], and other excellent measurement infrastructures such as Paxson[86] or Scastroute[98], have set the standard by which measurement projects are judged. While not supported by a measurement infrastructure, Odlyzko[81] aggregates and collects interesting data on how traffic patterns are evolving.

The U.S. FCC [54] and the OECD[83] both collect and make available limited data about broadband access networks. The FCC data is focused primarily on broadband availability and the levels of competition. They do collect information on the access technology with a limited view of the peak access speeds. The OECD publishes more extensive information about the broadband service plans available in a country. The data includes information about the number of plans available from a provider, the upload and download speeds, usage limitations and any accompanying usage charges, and finally information about the installation fees.

Our own work with Beverly has sought to measure the extent and nature of source

address filtering on the Internet[11]. Our data is currently being used by RIPE to encourage best common practices among their member networks. This demonstrates the power of measurement data to shape operational behavior.

Chapter 10

Conclusions and future work

The challenge that this thesis addresses – one that has fascinated us for a long time – is how to achieve economically desirable behaviors in the real world systems that arise from network engineering designs. Part of the answer, we believe, lies in leveraging the insights, theory, models, and tools of economics as a standard part of the normal engineering process of designing network protocols and architectures. Our thesis is just a small contribution toward showing how this can be done.

To make progress on these engineering challenges, a systematic ability to reason about and analyze the economics of packet networks and protocols is required. The engineering community needs to learn how to design, implement, and analyze network architectures and protocols in ways that integrate both the technical considerations that are historically familiar as well as economic and incentive criteria that are equally important to the success and long-term viability of a system.

Ultimately, this will be a long-term community effort incorporating both economic theory and experience-backed practices of what techno-economic questions are important to consider. Just as thinking about the security of a system has become ingrained in the engineering community's way of thinking, designing to accommodate economic incentives will also become ingrained.

This challenge is not unique to engineering. A quote from a non-engineering community illustrates a similar quest to elevate instinctively learned experiences into clear principles and systematic methodologies [84]:

Admittedly, it's only over time that these ideas have crystallized into their present form. But the clearer and more systematic they became, the more useful they became. That is after all what craft is all about. It is the process of raising instinctive knowledge and the lessons of experience to a conscious level so that they can be applied more efficiently and systematically.

In the engineering community, the challenge is not simply a matter of engineers learning game theory or mechanism design (though these will be important), rather what is required is a way of systematically integrating economic thought into the various stages of the systems and research processes of networking. We see this as a challenge of networking research over the next couple of decades so we readily acknowledge that this thesis does not solve this challenge. Rather we see our work (along with other contemporaries) as making in-roads into this space.

10.1 Future directions

There is a great deal of work that can be done in this space. We see a need for future work in the following categories:

1. Metrics

- (a) Qualitative and quantitative measures that express an engineered system's performance relative to economic objectives. Economic theory provides many evaluation strategies already, the challenge is in applying those strategies to the complexities of real-world networked systems.
- (b) Agreement within the community about weightings that should be given to different economic objectives. As with any engineering systems, tradeoffs are often required between the techno-economic objectives. What makes for a "good" system design in such an imperfect world then is a manner for thoughtful community deliberation.

2. Methods and Tools

- (a) Analytic methods and toolsets for easily evaluating basic economic properties of network systems. The challenge here is to support and communicate basic economic analysis of a networked system by and to individuals that are not necessarily economic experts.
- (b) Simulation tools to facilitate experiential learning and provide evidence about the economic behavior of networked systems. While simulation may be increasingly discounted in networking if the simulation is not paired with supported “real-world” results, we believe that simulation tools will be essential in initially building more economic knowledge within the systems-oriented networking community.

3. Community practices

- (a) Publication reviews that support and require consideration of the economic behaviors of systems. Just as system’s papers are rejected today if they fail to consider important security properties, papers should perhaps be rejected if they fail to consider important economic properties.
- (b) Integration of economic engineering practices into the education system. Learning to evaluate and design a networked system to achieve economically desirable behaviors is an acquired skill that will have to be taught to students.
- (c) Funding bodies that promote and understand economic engineering work. We are optimistic on this front given the NSF’s announced directs in the GENI and FIND programs. The commitment though will have to be a long-term.
- (d) Tenure review processes are capable of understanding and evaluating contributions in this inherently cross-disciplinary areas. We see this area as being a unique mixture of economic theory and engineering practice that

requires a different understanding of what constitutes intellectual contributions.

These categories illustrate that the challenge for the future is not purely technical or analytic. The challenges is also to build up the community's institutional capability for doing this type of work. We plan to be a part of that community, working on both the methodological as well as core system building challenges.

Bibliography

- [1] Mike Afergan. *Applying the Repeated Game Framework to Multiparty Networked Applications*. PhD thesis, Massachusetts Institute of Technology, August 2005.
- [2] Mike Afergan. Using repeated games to design incentive-based routing systems. In *Proceedings of the 2006 IEEE Infocomm Conference*. IEEE, April 2006.
- [3] Mike Afergan and Rahul Sami. Repeated-game modeling of multicast overlays. In *Proceedings of the 2006 IEEE Infocomm Conference*. IEEE, April 2006.
- [4] Luzi Anderegg and Stephan Eidenbenz. Ad hoc-VCG: a truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 245–259, New York, NY, USA, 2003. ACM Press.
- [5] Grenville J. Armitage. Revisiting IP QoS: Why do we care, what have we learned? ACM SIGCOMM 2003 RIPQOS workshop report. *SIGCOMM Comput. Commun. Rev.*, 33(5):81–88, 2003.
- [6] H. Balakrishnan and S. Seshan. The Congestion Manager. RFC 3124 (Proposed Standard), June 2001.
- [7] Steven Bauer, Peyman Faratin, and Robert Beverly. Assessing the assumptions underlying mechanism design for the internet. In *Economics of Networked Systems (NetEcon06)*, June 2006.

- [8] Gregory Bell. Failure to thrive: Qos and the culture of operational networking. In *RIPQoS '03: Proceedings of the ACM SIGCOMM workshop on Revisiting IP QoS*, pages 115–120, New York, NY, USA, 2003. ACM Press.
- [9] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.
- [10] Robert Beverly and Steven Bauer. The Spoofer Project. <http://spoofer.csail.mit.edu>.
- [11] Robert Beverly and Steven Bauer. The spoofer project: Inferring the extent of source address filtering on the internet. In *SRUTI*, pages 53–59, July 2005.
- [12] Robert Beverly, Steven Bauer, and Arthur Berger. Toward quantifying network neutrality. In *Proceedings of the 8th Passive and Active Measurement (PAM) Workshop*, pages 135–144, April 2007.
- [13] Boudreaux, Donald J. and Holcombe, Randall G. The coasian and knightian theories of the firm. *Managerial and Decision Economics*, 10(2):147–154, jun 1989.
- [14] B. Briscoe, A. Jacquet, A. Salvatori, and M. Koyabe. Re-ecn: Adding accountability for causing congestion to tcp/ip. IETF Internet draft draft-briscoe-tsvwg-re-ecn-tcp-03, October 2006.
- [15] B. Briscoe and S. Rudkin. Commercial models for ip quality of service interconnect. *BT Technology Journal*, 23(2):171–195, 2005.
- [16] Bob Briscoe. Flow rate fairness: Dismantling a religion. *ACM Computer Communication Review*, Feb 2007.
- [17] Bob Briscoe, Arnaud Jacquet, Carla Di Cairano-Gilfedder, Alessandro Salvatori, Andrea Soppera, and Martin Koyabe. Policing congestion response in an internetwork using re-feedback. In *SIGCOMM '05: Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 277–288, New York, NY, USA, 2005. ACM Press.

- [18] Broadband Working Group Communications Futures Program (CFP). The broadband incentive problem, whitepaper, September 2005.
- [19] Dennis M. Buede. *The Engineering Design of Systems: Models and Methods*. John Wiley & Sons, Inc., New York, NY, USA, 1999.
- [20] Kenjiro Cho, Kensuke Fukuda, Hiroshi Esaki, and Akira Kato. The impact and implications of the growth in residential user-to-user traffic. In *SIGCOMM '06: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 207–218, New York, NY, USA, 2006. ACM Press.
- [21] Cisco. Global IP Traffic Forecast and Methodology, 2007.
- [22] K. C. Claffy. Caida: Visualizing the internet. *IEEE Internet Computing*, 5(1):88, 2001.
- [23] David Clark, William Lehr, Steven Bauer, Peyman Faratin, Rahul Sami, and John Wroclawski. Overlay Networks and Future of the Internet. *Journal of Communications and Strategies*, 3(63):1–21, 2006.
- [24] David D. Clark, John Wroclawski, Karen R. Sollins, and Robert Braden. Tussle in cyberspace: defining tomorrow’s internet. *IEEE/ACM Trans. Netw.*, 13(3):462–475, 2005.
- [25] Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson. Ignoring the great firewall of china. In George Danezis and Philippe Golle, editors, *Privacy Enhancing Technologies*, volume 4258 of *Lecture Notes in Computer Science*, pages 20–35. Springer, 2006.
- [26] Elise A. Couper, John P. Hejkal, and Alexander L. Wolman. Boom and bust in telecommunications. *Economic Quarterly*, (Fall):1–24, 2003. available at <http://ideas.repec.org/a/fip/fedreq/y2003ifallp1-24nv.89no.4.html>.

- [27] C. Courcoubetis, V.A. Siris, and G.D. Stamoulis. Application and evaluation of large deviation techniques for traffic engineering in broadband networks. In *In Proc. of ACM SIGMETRICS '98*, June 1998.
- [28] C. Courcoubetis and R. Weber. Buffer overflow asymptotics for a buffer handling many traffic sources. *j. app. prob.*, 1995., 1995.
- [29] Costas Courcoubetis and Richard Weber. *Pricing Communication Networks: Economics, Technology and Modelling (Wiley Interscience Series in Systems and Optimization)*. John Wiley & Sons, 2003.
- [30] Edward Crawley, Olivier de Weck, Steven Eppinger, Christopher Magee, Joel Moses, Warren Seering, Joel Schindall, David Wallace, and Daniel Whitney. The influence of architecture in engineering systems. MIT Engineering Systems Symposium, 29-31 March, at MIT, Cambridge, Mass., Mar 2004.
- [31] I. Liu D. Clark, W. Lehr. Provisioning for bursty internet traffic: Implications for industry and internet structure. MIT ITC Workshop on Internet Quality of Service, November 1999.
- [32] Peter Eckersley, Fred von Lohmann, and Seth Schoen. Packet Forgery By ISPs: A Report On The Comcast Affair, 2007. http://www.eff.org/files/eff_comcast_report.pdf.
- [33] Paul Joskow (ed.). *Economic Regulation*. Edward Elgar Publishers, 2000.
- [34] Casserly et. al. Comments of comcast corporation. http://fjallfoss.fcc.gov/prod/ecfs/retrieve.cgi?native_or_pdf=pdf&id_document=6519840991, Feb 2008.
- [35] Peyman Faratin, David Clark, Patrick Gilmore, Steven Bauer, Arthur Berger, and William Lehr. Complexity of internet interconnections: Technology, incentives and implications for policy. In *Proceedings of Telecommunications Policy Research Conference (TPRC-07)*, 2007.

- [36] J. Feigenbaum, C. Papadimitriou, R. Sami, and S. Shenker. A BGP-based mechanism for lowest-cost routing. In *Proceedings of the 2002 ACM Symposium on Principles of Distributed Computing.*, 2002.
- [37] J. Feigenbaum and S. Shenker. Distributed algorithmic mechanism design: Recent results and future directions. In *Proceedings of the 6th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, pages 1–13. ACM Press, New York, 2002.
- [38] Joan Feigenbaum, Christos H. Papadimitriou, and Scott Shenker. Sharing the cost of multicast transmissions. *Journal of Computer and System Sciences*, 63(1):21–41, 2001.
- [39] Joan Feigenbaum, Christos H. Papadimitriou, and Scott Shenker. Sharing the cost of multicast transmissions. *J. Comput. Syst. Sci.*, 63(1):21–41, 2001.
- [40] Bob Fernandez. Comcast agreement in dispute with bittorrent. http://www.philly.com/philly/business/20080327_Comcast_agreement_in_dispute_with_BitTorrent.html, Mar 2008.
- [41] S. Floyd. Congestion Control Principles. RFC 2914 (Best Current Practice), September 2000.
- [42] S. Floyd and J. Kempf. IAB Concerns Regarding Congestion Control for Voice Traffic in the Internet. RFC 3714 (Informational), March 2004.
- [43] M. Fischer T. Fowler. Fractals, heavy tails and the internet. Published in Sigma, Mitretek Systems, Fall 2001.
- [44] Milton Friedman. *Essays In Positive Economics*, chapter The Methodology of Positive Economics. Univ. of Chicago Press, Chigaco, IL, 1966.
- [45] Drew Fudenberg and John Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.

- [46] Kensuke Fukuda, Kenjiro Cho, and Hiroshi Esaki. The impact of residential broadband traffic on japanese isp backbones. *SIGCOMM Comput. Commun. Rev.*, 35(1):15–22, 2005.
- [47] R. Gibbens and F. Kelly. Resource pricing and the evolution of congestion control automatica 35(12)1969– 1985, december 1999., 1999.
- [48] G. Gilder. Fiber keeps its promise: Get ready. bandwidth will triple each year for the next 25.
- [49] Jim Gray and Prashant J. Shenoy. Rules of thumb in data engineering. In *ICDE '00: Proceedings of the 16th International Conference on Data Engineering*, page 3, Washington, DC, USA, 2000. IEEE Computer Society.
- [50] M. Handley, S. Floyd, J. Padhye, and J. Widmer. TCP Friendly Rate Control (TFRC): Protocol Specification. RFC 3448 (Proposed Standard), January 2003.
- [51] Shai Herzog, Scott Shenker, and Deborah Estrin. Sharing the “cost” of multicast trees: an axiomatic analysis. *IEEE/ACM Transactions on Networking*, 5(6):847–860, 1997.
- [52] Elgan Huang, Jon Crowcroft, and Ian Wassell. Rethinking incentives for mobile ad hoc networks. In *PINS '04: Proceedings of the ACM SIGCOMM workshop on Practice and theory of incentives in networked systems*, pages 191–196, New York, NY, USA, 2004. ACM Press.
- [53] Hughesnet. Fair access policy. <http://gsa.hughesnet.com/fairaccess.cfm>
Accessed on June 2, 2007.
- [54] Industry Analysis and Technology Division Wireline Competition Bureau. High-Speed Services for Internet Access: Status as of December 31, 2005, July 2006.
- [55] Matthew Jackson. *Optimization an Operations Research*, chapter Mechanism Theory. EOLSS, Oxford, UK, 2003.

- [56] Ehud Kalai and John O. Ledyard. Repeated implementation. *Journal of Economic Theory*, 83(2):308–317, April 1998.
- [57] D. Katabi, M. Handley, and C. Rohrs. Internet congestion control for future high bandwidth-delay product environments. *acm sigcomm 2002*, 2002.
- [58] F. Kelly. Notes on effective bandwidths. *In Stochastic Networks: Theory and Applications*, pages 141–168, 1996.
- [59] F. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *In Journal of the Operational Research Society*, volume 49, 1998.
- [60] Kevin Kelly. When bandwidth is free: The dark fiber interview with George Gilder, *Wired*, 1: 4, pp. 38-41., 1993.
- [61] E. Knightly. Second moment resource allocation in multi-service networks. *In Proc. of ACM Sigmetrics'97*, 1997.
- [62] Mark Kortekaas. Keynote Presentation: Video over the Internet: Can we break the Net?, Feb 2007. NANOG.
- [63] Jean-Jacques Laffont and Jean Tirole. *Competition in Telecommunications*. MIT Press, Cambridge, MA, 2002.
- [64] Dave Levin. Punishment in selfish wireless networks: A game theoretic analysis. *In Economics of Networked Systems (NetEcon06)*, June 2006.
- [65] J. MacKie-Mason and H. Varian. Pricing congestible network resources. *IEEE Journal on Selected Areas in Communications*, 13(7):1141–1149, 1995.
- [66] Jeffrey K. MacKie-Mason and Hal R. Varian. Pricing the Internet. *In Public Access to the Internet, JFK School of Government, May 26–27, 1993*, page 37, 1993.

- [67] R. Mahajan, M. Rodrig, D. Wetherall, and J. Zahorjan. Sustaining cooperation in multi-hop wireless networks. In *Proceedings of 2nd Symposium on Networked Systems Design and Implementation*, 2005.
- [68] Ratul Mahajan, Maya Rodrig, David Wetherall, and John Zahorjan. Experiences applying game theory to system design. In *PINS '04: Proceedings of the ACM SIGCOMM workshop on Practice and theory of incentives in networked systems*, pages 183–190, New York, NY, USA, 2004. ACM Press.
- [69] A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, New York, NY, 1995.
- [70] Lee W. McKnight and Joseph P. Bailey. Internet economics: When constituencies collide in cyberspace. *IEEE Internet Computing*, 1(6):30–37, 1997.
- [71] Jungwon Min. *Cross-user subsidy in residential broadband service*. PhD thesis, Massachusetts Institute of Technology, September 2005.
- [72] Herve Moulin. Axiomatic cost and surplus sharing. In K. J. Arrow, A. K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*.
- [73] Donald A. Norman. *The Design of Everyday Things*. Basic Books, September 2002.
- [74] Donald A. Norman. Emotional design. *Ubiquity*, 4(45):1–1, 2004.
- [75] Bill Norton. Internet Video: The Next Wave of Massive Disruption to the US Peering Ecosystem (v1.2), Jan 2007.
- [76] W.B. Norton. The art of peering: The peering playbook. Technical report, Equinix.com, 2001.
- [77] A. Odlyzko. The economics of the internet: Utility, utilization, pricing, and quality of service, <http://www.research.att.com/amo/doc/internet.economics.pdf>, october 14, 1999., 1999.

- [78] A. Odlyzko. Internet pricing and the history of communications. *Computer Networks (Amsterdam, Netherlands: 1999)*, 36(5–6):493–517, 2001.
- [79] A. Odlyzko. Internet traffic growth: Sources and implications, 2003, 2003.
- [80] Andrew Odlyzko. Minnesota Internet Traffic Studies (MINTS). <http://www.dtc.umn.edu/mints/home.html>.
- [81] Andrew Odlyzko. Minnesota Internet Traffic Studies (MINTS). <http://www.dtc.umn.edu/mints/growth.html>.
- [82] Andrew Odlyzko. Minnesota Internet Traffic Studies (MINTS). <http://www.dtc.umn.edu/mints/2002-2008/analysis-2002-2008.html>.
- [83] OECD. OECD Broadband Statistics, available at <http://www.oecd.org/sti/ict/broadband>.
- [84] Darwin Ortiz. *Designing Miracles: Creating the Illusion of Impossibility*. A1-MagicalMedia, 2007.
- [85] IC Park and W. Willinger. *Self Similar Network Traffic and Performance Evaluation*, chapter Self-similar network traffic: an overview. Wiley Interscience, 2000.
- [86] Vern E. Paxson. *Measurements and Analysis of End-to-End Internet Dynamics*. PhD dissertation, University of California, Lawrence Berkeley National Laboratory, April 1997.
- [87] D. Payne and R. Davey. Avoiding the bandwidth price dilemma using optical networking, 2004. BTextact.
- [88] Henry Petroski. *To Engineer Is Human: The Role of Failure in Successful Design (Vintage)*. Vintage, March 1992.
- [89] J. Postel. Transmission Control Protocol. RFC 793 (Standard), September 1981. Updated by RFC 3168.

- [90] K. Ramakrishnan, S. Floyd, and D. Black. The Addition of Explicit Congestion Notification (ECN) to IP. RFC 3168 (Proposed Standard), September 2001.
- [91] Peter Reichl, Pascal Kurtansky, and et al. ISP Cost Model (ICOMO) Design ETH Zrich, Computer Engineering and Networks Laboratory TIK, TIK Report Nr. 99, December, 2000.
- [92] Jrg Rieskamp, Jerome R Busemeyer, and Barbara A Mellers. Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature*, 44(3):631–661, sept 2006.
- [93] Jerome H. Saltzer, David P. Reed, and David D. Clark. End-to-end arguments in system design. *ACM Transactions on Computer Systems*, 2(4):277–288, November 1984.
- [94] Sandvine. Policy traffic switch, information available at http://sandvine.com/products/policy_traffic_switch.asp.
- [95] Stefan Savage, Neal Cardwell, David Wetherall, and Tom Anderson. TCP congestion control with a misbehaving receiver. *Computer Communication Review*, 29(5), 1999.
- [96] S. Shenker, D. Clark, D. Estrin, and S. Herzog. Pricing in computer networks: Reshaping the research agenda. *ACM Computer Communication Review*, 26:19–43, April 1996.
- [97] Jeffrey Shneidman and David C. Parkes. Specification faithfulness in networks with rational nodes. In *PODC*, pages 88–97, 2004.
- [98] Neil T. Spring, David Wetherall, and Thomas E. Anderson. Scriptroute: A public internet measurement facility. In *USENIX Symposium on Internet Technologies and Systems*, 2003.
- [99] Brett Swanson and George Gilder. Estimating the exaflood: The impact of video and rich media on the internet, Jan 2008.

- [100] Bernhard Thome, editor. *Systems engineering: principles and practice of computer-based systems engineering*. John Wiley and Sons Ltd., Chichester, UK, UK, 1993.
- [101] S. Uhlig. A review of scaling behaviors in internet traffic. In *CIRM Power-laws workshop*, March 2004.
- [102] Barbara Van Schewick. Towards an Economic Framework for Network Neutrality Regulation. In *Proceedings of the Telecommunications Policy Research Conference*, September 2005.
- [103] Walter Vincenti. *What Engineers Know and How They Know It*. The Johns Hopkins University press, Baltimore MD, USA, 1990.
- [104] W. Willinger. On Internet traffic dynamics and internet topology: Internet model validation. <http://www.ima.umn.edu/talks/workshops/1-7-9.2004/willinger/willinger1.2.pdf>, Jul 2004.
- [105] G. Ramachandran Y. Kogan, G. Maguluri. Satisfying customer bandwidth demand in ip data networks. *Performance Evaluation*, 52(2-3):105–117, April 2003.
- [106] Sheng Zhong, Li (Erran) Li, Yanbin Grace Liu, and Yang (Richard) Yang. On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks: an integrated approach using game theoretical and cryptographic techniques. In *MobiCom '05: Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 117–131, New York, NY, USA, 2005. ACM Press.
- [107] Ziff Davis Media. Comcast Cuts Off Bandwidth Hogs, April 05, 2007 available at <http://http://www.physorg.com/news94981460.html>.