Position Paper on Networks with Aggregate Quality-of-Service

Stephen D. Patek[a] and Jörg Liebeherr[b]

[a] Department of Systems and Information Engineering, University of Virginia
[b] Department of Computer Science, University of Virginia

ABSTRACT

As the Internet continues to expand into new application domains, there is growing demand for differentiated services that provide varying degrees of Quality of Service (QoS). Until recently, the approach for service differentiation in the Internet has focused on providing QoS guarantees to individual traffic flows. This *per-flow* model has not been widely embraced, largely due to the vast amount of state information that needs to be maintained in the network. As a consequence, the network community is currently redefining the notion of QoS for the Internet, leading to a new service model where guarantees are made to *aggregate* flows, rather than to individual flows.

The notions of aggregate QoS which are being defined, e.g., by the Differentiated Services (DiffServ) working group in the IETF, are very different from traditional networking concepts and require novel algorithms for traffic control and provisioning. In our work in this area we seek to derive fundamental insights into the nature and delivery of QoS to traffic aggregates, i.e. *aggregate QoS*. A central feature of service models for aggregate QoS is that there is no requirement for users to specify the composition or the destination of traffic. This significantly complicates the provisioning of network resources and suggests that new, feedback-intensive traffic control algorithms must be developed to mitigate uncertainty about network loading. At present there is little or no theory available to guide the development of such algorithms.

In this paper we lay out a research agenda for illuminating the fundamental issues associated with QoS for aggregate traffic. By providing a theoretical reference frame for aggregate QoS, one which provides tools and techniques for congestion control, capacity provisioning, and admission control, we complement ongoing standardization efforts for differentiated services in the Internet.

Keywords: Aggregate QoS, Differentiated Services, Traffic Control

1. INTRODUCTION

To guarantee the viability of the Internet as an ubiquitous information infrastructure in the next decade, it is essential that the Internet be able to offer a wide range of different services to applications and users. Prior attempts to introduce differentiated services in the Internet have focused on supporting varying levels of Quality-of-Service (QoS) for each individual traffic flow. In this *per-flow* model, network resources are reserved separately to support the desired QoS level for each individual flow. The *per-flow* model has been extensively studied by many researchers throughout the 1990's.[1–5,4,6] In the Internet Engineering Task Force (IETF), the Integrated Services Working Group (IntServ WG) has devised a *per-flow* QoS service model.[7,8] However, Internet service providers have not embraced the *per-flow* model. A primary reason for this is that maintaining state information for each flow at each router on its path introduces a scalability problem, exacerbated by the fact that many applications, especially Web-oriented applications, require an establishment of many short-lived flows. (The time to establish per-flow QoS may exceed the entire lifetime of the flow.) Other reasons for not adapting the IntServ model include the fact that per-flow QoS across multiple networks requires a tight coordination between service providers and the fact that most users are unable to precisely define their QoS requirements on a per-flow basis.

The gap between the growing need for service differentiation and the inability of the existing *per-flow* QoS model to serve this need has triggered a rethinking of the basic tenets of QoS in the Internet and has led to a major revision of the approach to implement QoS. Starting as early as 1995,[9] a revised QoS notion has emerged,[10–12] and, since November 1997, is being made precise by the *Differentiated Services Working Group (DiffServ WG)* group in the IETF.[13,14] A main characteristic of the new QoS model is that service guarantees are given to aggregate flows, rather

This work is supported in part by the NSF through grants ANI-9903001, ECS-9875688 (CAREER), ANI-9730103, NCR-9624106 (CAREER), and ANI-0085955. Please send correspondence to Stephen D. Patek, `patek@virginia.edu` than on a per-flow basis. While proposals vary widely in their specifics, they all share the following characteristics. (1) Service providers and users agree upon a hierarchy of service classes defined with respect to a generalized notion of bandwidth consumption. (2) The service agreements are enforced at the network boundaries, through a combination of marking, dropping, or shaping of incoming packets. (3) Network elements in the core of a network process packets based exclusively on the marking that packets received at the network border; elements in the core of the network do not have any notion of end-to-end flows. (4) Service agreements are made for traffic aggregates as opposed to single traffic flows.

QoS for aggregate traffic is fundamentally different from per-flow QoS. For example, the geographical scope[11] of the QoS guarantees can be quite diverse, e.g. guarantees between a specific source/destination pair, from a specific source to a set of destinations, or from a source to any destination. Clearly, without a clear specification of the composition and destination of traffic, provisioning network capacity for aggregate QoS guarantees is a formidable task.

The point of departure in this paper is the recognition that the concept of QoS for aggregate traffic requires fundamentally different algorithms for traffic control and capacity provisioning from those used in traditional best effort networks or per-flow QoS networks. Our goal is to gain an understanding of the fundamentals of QoS for aggregate traffic and to make progress towards a theory of aggregate QoS traffic. In our research, we attempt to find answers to the following questions.

- Which traffic control algorithms, implemented at the border routers, can guarantee high resource utilization of network resources for aggregate QoS? Among the vast possibilities for implementing traffic control, which have the most impact?
- What are the time scales over which service guarantees can be offered?
- How do we measure the performance of a network with aggregate QoS? Which metrics are appropriate? How do these performance metrics relate to the provisioning (scale and structure) of the network?
- What are the important considerations in deciding on new service contracts? In other words, how do we do admission control? For a given a set of traffic control algorithms, how do we determine the statistical multiplexing gain?
- What do we need to know to do on-line control and optimization of a network with aggregate QoS? What feedback information is appropriate? Given this feedback, what can be inferred about congestion within the network?

The goal of this paper is to describe a research agenda for answering many of the questions above. In Section 2, we set the context of our discussion by describing a revised view of the Internet. In Section 3, we outline several of the major thrusts of our research on aggregate QoS. The first main component of the research is to develop fluid-based control models for enhancing aggregate QoS. The second main component of the research is to quantify the statistical multiplexing gain available to large aggregates of traffic in a DiffServ context. The third and final component of the research is to evaluate and improve our control models through simulation-based experimentation. In Section 4, we conclude with a brief discussion of our results to date and upcoming research highlights.

2. NETWORK DESCRIPTION

We claim that the notions of aggregate QoS and differentiated services will lead to a new picture of how the Internet will work. As in the traditional Internet, there will be no notion of per-flow resource reservation. Moreover, the directionality and volume of the packet streams, will be unknown *a priori*. Consequently, *feedback* about the state of the network will play a key role in any technique for on-line traffic control and resource allocation.

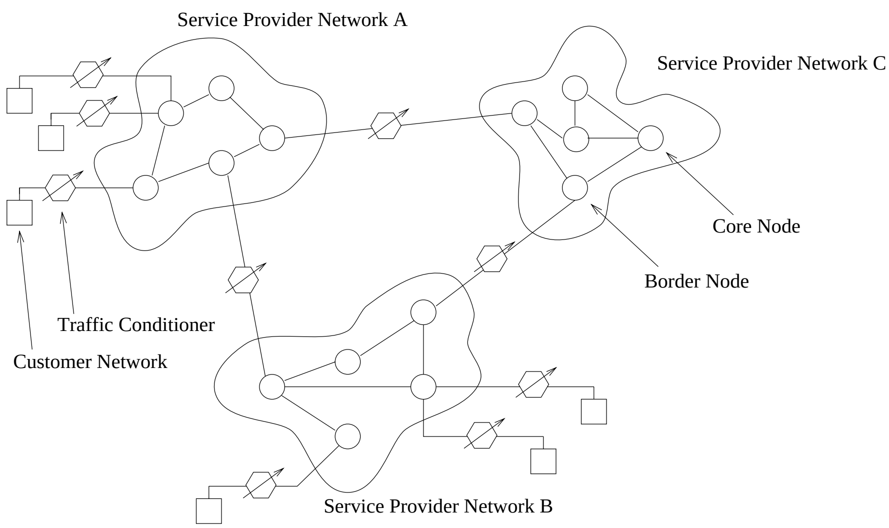

Figure 1. View of the Internet.

2.1. A Revised View of the Internet

The network abstraction used throughout this paper is given in Figure 1. The network is composed of customer networks and service provider networks. Each customer network has access to at least one service provider network, and customer networks are the sources and sinks of traffic. Each service provider network is connected to at least two other networks and consists of a set of interconnected routers. Routers which connect to another network are called *border nodes*; the remaining routers are called *core nodes*. Each link between either (1) a customer network and a service provider network or (2) between two service provider networks is equipped with a traffic conditioner which monitors the level of service being provided, and drops, marks, or delays packets which exceed the nominal levels of service.

In this network view, network service providers can offer customer networks a wide range of network services. Customers and service providers negotiate a traffic profile which specifies the traffic rate which can be submitted to the network for a given service. The traffic profiles are manifested in a so-called service level agreement (SLA), and the traffic conditioners between networks are derived from the negotiated SLAs. The ability of core nodes to prioritize traffic is assumed to be very limited. Core routers give differential treatment to packets based exclusively on the marking of packets. In particular, core routers are unable to differentiate between packets based on their origin, their destination, or the traversed route.

We focus on traffic control algorithms for a single service provider network. The particular service we investigate is the *assured service* as proposed by Clark,[10] wherein traffic which complies to the negotiated profile is unlikely to be dropped in the network. We consider a service with *coarse spatial granularity*,[10] in which the service profile is applied to any possible destination in the Internet. We believe that this type of service, which gives bounded loss probability without requiring users to specify the direction of traffic, poses the most challenging problems and is the least understood.

2.2. Components of Traffic Control for Aggregate QoS

Without the ability to establish per-flow state in the network, and with the limited complexity at core nodes, traffic control algorithms which enable or support differentiated services will be heavily based on algorithms implemented at border nodes. In this section we discuss our assumptions on the entities and algorithms which participate in traffic control for aggregate QoS.

Alternate Routing: An important assumption for the considered control algorithms is that network routes can change dynamically, but do not change frequently. This assumption is justifiable in today's networking environment.[15] We assume that the network uses an existing, distributed routing algorithm.[16] Thus, we do not require or assume that the routing algorithm cooperates with the traffic control algorithms which implement aggregate QoS. Since processing source-routed packets in current routers is slow, we will employ a new mechanism, called *alternate routing*, in which we assume that the network has the ability to implement "IP tunneling" between border nodes, i.e., the network has the ability to perform IP-in-IP encapsulation. With this mechanism, we can devise traffic control algorithms which can send traffic to specific egress points. Note, however, that the path of tunneled packets is not assumed to be controllable.

Traffic Conditioning: Traffic conditioning at network boundaries is a central feature of all proposals for differentiated services in the Internet.[11,13,14,17–19] Traffic conditioning includes metering, marking, dropping and shaping of traffic. A simple way to condition traffic, is to mark packets which comply to the negotiated traffic profile as 'in-profile', and to mark all other packets as 'out-of-profile', implying that out-of-profile traffic has a higher drop priority. Each traffic conditioner is responsible for maintaining state information for the flows it monitors. The conditioning of a packet can be different at each network boundary traversed by the packet, based on the SLAs between adjacent networks.

Core Nodes: Core nodes give differential treatment to packets based only on how they are marked. Core nodes do not carry flow-specific information. We do not make assumptions about the scheduling algorithms implemented at core nodes, other than that they can differentiate packets based on the marking performed at the network boundary. The type of scheduling algorithms which are most suitable for core routers is an open issue. Fair queuing algorithms are best suited if the aggregate QoS should satisfy fairness guarantees,[20,21] otherwise simple priority queuing may be the preferred algorithm.[12]

Probing and Reporting: Since the precise volume and the routes of the traffic are not known in advance in networks with aggregate QoS and coarse spatial granularity, traffic control algorithms must rely heavily on feedback from the network. Feedback can be obtained through probe packets which measure the available bandwidth on a path.[22,23] Alternatively, feedback information can be obtained via explicit congestion notification or other information from the network. Other than the collection of information, the dissemination (reporting) of measurements and the frequency of making measurements are important issues. One of our objectives here is to gain insight into the type of feedback needed for an effective implementation of aggregate QoS.

Provisioning and Admission Control: Since customer networks effectively only specify the access bandwidth that they desire, without saying where their traffic will be flowing, it is extremely difficult to provision aggregate QoS networks and to make the decision whether to admit new customers. As with other measurement-based approaches,[24–30] admission control algorithms for aggregate QoS will rely on feedback information about the state of the network.

3. RESEARCH AGENDA

Our main objective is to gain fundamental insights into the nature and delivery of aggregate QoS. We seek to provide an analytical infrastructure in an area where currently few theoretical insights exist. In Section 3.1 below, we discuss fluid-based models for the routing of traffic, where feedback mechanisms (probing and reporting) serve to inform alternate routing decisions for congestion avoidance. In Section 3.2, we present methods for determining the statistical multiplexing gain of aggregate flows. These first two sections are fundamentally theoretical/analytical. In Section 3.3, we seek to employ these insights in developing prototype traffic control algorithms which we shall test in simulation.

3.1. Fluid-based Traffic Control Models

The primary obstacle to delivering aggregate QoS is uncertainty about the traffic patterns that will emerge. In general, there is no notion of connection establishment in aggregate QoS networks, and resources are not reserved on a per-flow basis. Rather, service providers and (aggregate) users only agree upon a generalized measure of bandwidth to be delivered at their access points. It is unknown ahead of time where traffic will go or where it will come from. One way to address uncertainty about traffic patterns is to adopt a control-theoretic perspective, wherein a network service provider (1) uses probing and reporting to gather "soft-state" information about how its resources are being used and (2) uses alternate routing to explicitly redirect appropriate amounts of traffic in response to perturbations in flow patterns. Our hope is that feedback control will allow us to make QoS guarantees without explicitly reserving resources on a per-flow basis.

A key idea that we wish to exploit is that probing and reporting yield information about network usage that is consistent with the "path flow" formulation of multicommodity network optimization models.[31,32] Recall from Section 2.2 that, at the boundaries of a service provider network, there are traffic conditioners that mark packets according to contracts made with customer networks. The process of marking packets corresponds to classifying packets as being of one "commodity" type or another. We assume that the algorithms by which packets are marked are fixed and that marking is in no way modified by feedback control. Also, we treat alternate routing simply as a means of specifying the egress points from which packets depart the network. Alternate routing commands *do not* specify the routes that packets take through the network; routing *within* the network is automatic, as implemented by protocols within the routers. To complete the feedback loop, ingress points periodically probe the network to determine (1) the prevailing routes taken by packets of various classes in making their way to egress points and (2) the amount of congestion in the links that comprise the prevailing routes. This information is then forwarded to control authorities within the network that use network optimization techniques to determine (1) how much traffic from each ingress point should be alternately routed and (2) the *distribution* of alternately routed flow to various egress points. In what follows, we show how the network-flow perspective (above) translates into algorithms for traffic control in aggregate QoS networks.

To start with, we may naively assume the existence of a single controller that receives reports from all ingress points, processes the data, and issues alternate routing commands. We consider a service provider network with links $(i,j) \in A$, ingress points $o \in O$, egress points $d \in D$, and traffic classes $m \in M$. (Again, each traffic class corresponds to a marking-type enforced by the traffic conditioners.) Let

$r_{o,m}^{i,j}$ be the amount of class-m flow on link (i,j) that results from a unit of conventionally-routed, class-m flow pushed into the network at ingress point o,

$s_{o,m,d}^{i,j}$ be the amount of class-m flow on link (i,j) that results from a unit of alternately-routed, class-m flow pushed into the network at ingress point o that is destined for egress point d,

$b_{o,m}$ be the amount of class-m flow that originates at ingress point o.

The network controller estimates these parameters from the probe-reports that arrive from ingress points. Let

$u_{o,m}$ be the fraction of class-m flow that is *conventionally*-routed from ingress point o,

$v_{o,m,d}$ be the fraction of *alternately*-routed, class-m flow being sent to egress point d from ingress point o.

These are the alternate routing control variables that are adjusted by the centralized controller of this framework. Let u be the vector of conventional flow fractions $u_{o,m}$. Let v be the vector of the alternate flow fractions $v_{o,m,d}$. Let

$$x_m^{i,j}(u,v) = \sum_{o \in O} b_{o,m}[u_{o,m} r_{o,m}^{i,j} + (1 - u_{o,m}) \sum_{d \in D} v_{o,m,d} s_{o,m,d}^{i,j}] \tag{1}$$

be the amount of class-m flow on the link (i,j) that results from the flow allocations (u,v). Let $x^{i,j}(u,v)$ be the vector of flows associated with the traffic classes m. Let $y^{i,j}(u,v) = \sum_{m \in M} x_m^{i,j}(u,v)$ be the *total* flow on (i,j) associated with (u,v). The control variables should be adjusted in response to feedback information to achieve a fair and efficient allocation of resources.

One problem inherent in the centralized control models of the preceding paragraphs is that it may be necessary to wait until reports arrive from *all* ingress points before an informed adjustment to the control parameters may be made. A fully centralized control scheme may be unable to deal effectively with abrupt, drastic changes in the way that the network is being utilized. If each ingress point were allowed to make routing decisions in response to local changes in traffic patterns as they arise, then the system as a whole would be much more responsive. This motivates a *decentralized* control scheme whereby each ingress point acts of its own accord in rapid response to (1) its own observations and (2) the reports from other ingress points (which arrive relatively infrequently).

The main advantage of decentralized control is that individual controllers at ingress points may respond quickly to local changes in traffic flow. However, an important drawback of the decentralized approach is that *an acceptable routing pattern may not be achievable when only one ingress point is rerouting its flow at a time.* Substantial gains in throughput may be had when two or more ingress points coordinate their routing activity. As an example, suppose that traffic entering at ingress point o is congested regardless of the routing decisions that can be made there. Given that the network is adequately sized, the congestion must be due to routing decisions being made independently at other ingress points. Suppose that (1) the flow at another ingress point o' could be redirected without seriously compromising the QoS offered to its customer networks and (2) this change would alleviate the routing problem at ingress point o. In the absence of a higher-level control authority, o' would have no incentive to make this redirection of flow. The higher level control authority could create this incentive by issuing commands to o' which effectively make its original routing decision more expensive than the alternative.

3.2. Statistical Multiplexing Gain in Networks with Aggregate QoS

A key characteristic of packet switching networks is their ability to exploit statistical multiplexing of traffic sources, and therefore, achieve a high utilization of the available resources. A packet network with QoS requirements can exploit statistical multiplexing by taking advantage of (1) the statistical properties of traffic sources, and (2) the statistical independence of flows, and (3) different QoS requirements of traffic sources. In exploiting the statistical properties of traffic sources, we are led to make a distinction between deterministic and statistical network services.[3] *Deterministic services*, which guarantee worst-case end-to-end delay bounds for traffic,[33–36] are known to lead to an inefficient use of network resources.[37] On the other hand, *statistical services*, which make guarantees of the form

$$Pr[Delay > X] < \epsilon \ , \tag{2}$$

i.e. services which allow a small fraction of traffic to violate QoS specifications, can significantly increase the achievable utilization of network resources. A statistical service can exploit *statistical multiplexing gain*, expressed as

$$\left(\begin{array}{l} \text{Resources needed to} \\ \text{support statistical} \\ \text{QoS of } N \text{ flows} \end{array} \right) \ll N \times \left(\begin{array}{l} \text{Resources needed to} \\ \text{support statistical} \\ \text{QoS of 1 flow} \end{array} \right) .$$

Ideally, the statistical multiplexing gain of a statistical service increases with the volume of traffic so that with a high enough level of aggregation the amount of resource allocated per flow is nearly equal to the average resource requirements for a single flow.

3.2.1. Effective Envelopes

One way to quantify the statistical multiplexing gain for traffic aggregates is through the use of *effective envelopes*, as introduced by Boorstyn, Burchard, Liebeherr, and Oottamakorn.[38] Within this framework, we assume that the arrivals $A_j(t_1, t_2)$ of traffic to a core router from source network j which receives a specific (preferential) service in the time interval $[t_1, t_2]$ are described by a random variable which is constrained by a deterministic envelope A_j^*, in the sense that $A_j(t, t+\tau) \le A_j^*(\tau)$ for all $t \ge 0, \forall \tau \ge 0$. Further, we assume that A_j is *stationary*, i.e., $P[A_j(t, t+\tau) \le x] = P[A_j(t+y, t+\tau+y) \le x]$ for all $t, \tau, y \ge 0$. The assumption of stationarity says that all time shifts of A_j are equally probable. Recent measurements (of best effort traffic) indicate that traffic in a core network is stationary over 5-minute long intervals[39] justify the stationarity assumption. The modeling assumptions listed above are not new[38,40–49] and are sometimes collectively referred to as 'regulated adversarial traffic'.

The total amount of traffic to arrive at the core router in the interval from t to $t+\tau$ from a set $\mathcal{N}$ of flows is denoted by $A_{\mathcal{N}}(t, t+\tau) = \sum_{j \in \mathcal{N}} A_j(t, t+\tau)$. We are interested in characterizing the aggregate of all of the traffic arriving from the source networks $j \in \mathcal{N}$. The *effective envelope* for this collection of flows is defined as a probabilistic subadditive bound on the amount of traffic that can arrive in any interval τ.

DEFINITION 3.1. *Given a set of flows $\mathcal{N}$, a* **local effective envelope** $\mathcal{H}_{\mathcal{N}}$ *satisfies for all $t, \tau \ge 0$*

$$P\Big[A_{\mathcal{N}}(t, t+\tau) \le \mathcal{H}_{\mathcal{N}}(\tau; \varepsilon)\Big] \ge 1 - \varepsilon. \tag{3}$$

Using a bound on the moment generating function for the traffic from any individual flow,[38] it is possible to construct the local effective envelope for any aggregate of flows all of which are constrained by the same subadditive upper bound A^*. For example, if all of the sources in $\mathcal{N}$ are policed by leaky bucket regulators with peak rate enforcement, so that $A_j^*(\tau) = A^*(\tau) = \min(P\,\tau,\, \sigma + \rho\,\tau)$, then

$$\mathcal{H}_{\mathcal{N}}(\tau;\varepsilon) \;=\; N \min(x, A^*(\tau)) \tag{4}$$

with N being the number of sources in $\mathcal{N}$ and x being the smallest real number such that

$$\left(\frac{\rho\tau}{x}\right)^{\frac{x}{\rho\,\tau}} \left(\frac{\frac{P}{\rho}-1}{\frac{P}{\rho}-\frac{x}{\rho\,\tau}}\right)^{\frac{P}{\rho}-\frac{x}{\rho\tau}} \;\leq\; \varepsilon^{\frac{P}{\rho\,N}}. \tag{5}$$

It is noteworthy that our bound $H_{\mathcal{N}}(.;\varepsilon)$ is a line through the origin, with some slope between $N\rho$ and NP.

3.2.2. Statistical QoS Guarantees for Traffic Aggregates

Effective envelopes provide a convenient way of describing, in a probabilistic sense, the maximum amount of traffic that can arrive in any interval of length τ. Boorstyn, Burchard, Liebeherr, and Oottamakorn[38] show how effective envelopes can be used to derive schedulability conditions guaranteeing statistical QoS for the traffic from individual networks arriving at a single core node. The challenge remains to provide new tools for provisioning end-to-end statistical QoS by extending the theory of effective envelopes to handle the network case.

In Liebeherr, Patek, and Yilmaz,[50] we consider the problem of guaranteeing end-to-end statistical QoS for aggregate traffic. By making QoS guarantees to the aggregate and not for specific flows within, the design of the core network can be greatly simplified since no per-flow information is required inside the network. The main difficulty of provisioning statistical QoS for a multi-node network lies in addressing the complex correlation of traffic at downstream multiplexing points. One way of addressing the "downstream problem" is to reconstruct traffic characteristics inside the network so that arrivals to core nodes satisfy the same properties as the arrivals to an edge node, perhaps through the use of delay-jitter control. Another approach is to allocate network capacity for each path or 'pipe' between a source-destination pair in the network, and only exploit the multiplexing gain between the flows on the same path. The former approach can achieve a very high level of aggregation at each node within the network, resulting in a high level of statistical multiplexing gain; however, this comes at the expense of having to restore the statistical independence of the flows at each node. In the latter approach, there is no requirement to reconstruct the statistical properties of the flows since the statistical multiplexing gain is extracted only at network's border nodes.

Aside from our earlier work,[50] there are only a few previous studies which apply the traffic model of adversarial regulated arrivals to multiple node networks. First, the lossless multiplexer of Reisslein, Ross, and Rajagopal[48] assumes that routes are such that traffic arrivals at core nodes from different flows are always independent. Subsequently, other researchers[51,52] have derived probabilistic bounds for end-to-end delays in networks with so-called coordinated schedulers. This class of scheduling algorithms addresses end-to-end provisioning of QoS, by taking into consideration the time that a packet has already spent in the network at previous nodes.

3.3. Experimentation with Aggregate QoS

Simulation experiments play an important role in developing new analytical and procedural devices for aggregate QoS. In our own work, the simulations we have developed model aggregate QoS networks at a relatively high level of abstraction. By sacrificing details of transport-level or session-level protocol functions, we are able to simulate networks with a relatively large number of nodes. Typically, the components of the packet-level network simulator include traffic conditioning and packet forwarding mechanisms, routing algorithms,[16] and traffic profiles.

Simulation Experiments

1. *Studies in Fluid-based Traffic Control* - Simulation allows us to determine experimentally the roles that alternate routing and probing play in delivering aggregate QoS, as discussed in Section 3.1. We can explore the tradeoffs between accuracy in network parameter estimation and performance of the overall system. Also, we can identify the measurements that need to be made in probing the network, and we will determine the amount of information that needs to be reported to the various control authorities within the network. Finally, we can determine the role and value of coordination through hierarchical control, especially in light of potential instabilities.

2. *Studies in Admission Control via Effective Envelopes* - We can develop and test algorithms for determining (1) whether to accept new aggregate users, (2) what the terms should be for the next contract with a given user, (3) what types of services are or will be required from neighboring networks in delivering aggregate QoS to existing or potential customers. In general, we hope to determine the practical importance of traffic modeling (e.g. characterizing burstiness) in satisfying QoS guarantees.

3. *Studies in Conditioning* - We can use simulation to determine experimentally the performance sensitivity of alternative methods of traffic conditioning (marking, dropping, shaping packet streams) at ingress points. We can investigate the effect of shaping at egress points.

4. CONCLUSIONS

Emerging notions of aggregate QoS are very different from traditional networking concepts and require novel algorithms for traffic control. In order to better understand the fundamental nature of aggregate QoS, we have laid out a three-pronged research agenda involving research into (1) fluid-based control models for alternate/multipath routing, (2) effective envelopes for characterizing the statistical multiplexing gain available to traffic aggregates, and (3) simulation experimentation for verifying our theoretical insights.

Our preliminary results are encouraging. In Patek, Venkateswaran, Liebeherr,[53] we describe a very simple alternate routing scheme where, based on feedback information from probe packets, we attempt to reroute high-priority traffic around points of congestion. Our initial, packet-level simulation results indicate that the scheme can have a positive impact on the performance of aggregate QoS networks. We have tested this alternate routing mechanism for a representative network topology under a wide variety of perturbations, and we have observed an improvement in the performance of the network at least with regard to high priority packet loss. The experimental results have encouraged us to study the theoretical properties of alternate/multipath routing.

With regard to effective envelopes and the statistical multiplexing gain associated with traffic aggregates, our initial results[50] suggest that there is a tradeoff between (1) achieving a high level of multiplexing gain throughout the network and (2) the requirement for extra machinery for reconstructing the statistical properties of flows along their end-to-end paths. Effective envelopes seem to provide a powerful characterization of aggregate traffic.

REFERENCES

1. "ATM traffic management specification version 4.0." ATM Forum, <af-tm-0056.000>, April 1996.
2. D. D. Clark, S. Shenker, and L. Zhang, "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanisms," in *Proc. Sigcomm '92*, pp. 14–26, August 1992.
3. D. Ferrari and D. Verma, "A Scheme for Real-Time Channel Establishment in Wide-Area Networks," *IEEE Journal on Selected Areas in Communications* **8**, pp. 368–379, April 1990.
4. J. Liebeherr and D. Wrege and D. Ferrari, "Exact Admission Control For Networks With Bounded Delay Services," *IEEE/ACM Transactions on Networking* **4**, pp. 885–901, December 1996.
5. J. Hyman, A. Lazar, and G. Pacifici, "Real-Time Scheduling with Quality of Service Constraints," *IEEE Journal on Selected Areas in Communications* **9**, pp. 1052–1063, September 1991.
6. D. Yates and J. Kurose and D. Towsley and M. Hluchyi, "On Per-Session End-to-End Delay Distributions and the Call Admission Problem for Real-Time Applications with QOS Requirements," in *Proceeedings of ACM SIGCOMM'93*, pp. 2–12, (San Francisco, CA), September 1993.
7. S. Shenker and C. Partridge, "Specification of Guaranteed Quality of Service." IETF Internet Draft, Integrated Services WG, draft-ietf-intserv-guaranteed-svc-05.txt, July 1995.
8. R. Braden, D. Clark, S. Shenker, "Integrated Services in the Internet Architecture: an Overview." Internet RFC 1633, June 1994.
9. D. D. Clark, "Adding Service Discrimination to the Internet," *Telecommunications Policy* **20**, pp. 169–181, 1996.
10. D. D. Clark and J. Wroclawski, "An Approach to Service Allocation on the Internet." IETF Internet Draft, Diffserv Working Group, draft-clark-diff-svc-alloc-00.txt, October 1997.
11. D. D. Clark and W. Fang, "Explicit Allocaiton of Best Effort Packet Delivery Service," *IEEE/ACM Transactions on Neworking* **6**, pp. 362–373, August 1998.

12. K. Nichols, V. Jacobson, and L. Zhang, "A Two-bit Differentiated Service Architecture." IETF Internet Draft, Diffserv Working Group, <draft-itef-diffserv-arch-02.txt>, October 1998.
13. Y. Bernet *et al.*, "A Conceptual Model for DiffServ Routers." IETF Internet Draft, Diffserv Working Group, <draft-itef-diffserv-model-01.txt>, October 1999.
14. S. Blake *et al.*, "An Architecture for Differentiated Services." IETF Internet Draft, Diffserv Working Group, <draft-itef-diffserv-arch-02.txt>, December 1998.
15. B. Chinoy, "Dynamics of Internet Routing Information," in *Proc. Sigcomm '93*, pp. 45–52, September 1993.
16. J. T. Moy, *OSPF - Anatomy of an Internet Routing Protocol*, Addison-Wesley, 1998.
17. J. Heinanen *et al.*, "Assured Forwarding PHB Group." Internet RFC 2597, June 1999.
18. V. Jacobson *et al.*, "An Expedited Forwarding PHB." Internet RFC 2598, June 1999.
19. I. Stoica and H. Zhang, "LIRA: A Model for Service Differentiation in the Internet," in *Proceedings of NOSS-DAV'98*, 1998.
20. A. Demers, S. Keshav, and S. Shenker, "Analysis and simulation of a fair queueing algorithm," in *Journal of Internetworking Research and Experience*, pp. 3–26, Oct. 1990.
21. J. Nagle, "On packet switches with infinite storage," *IEEE Trans. On Communications* **35**, pp. 435–438, April 1987.
22. R. L. Carter and M. Crovella, "Measuring Bottleneck Link Speed in Packet-Switched Networks," Tech. Rep. BU-CS-96-006, Boston University, Computer Science Department, March 1986.
23. S. Keshav, *Congestion Control in Computer Networks.* PhD dissertation, University of California at Berkeley, September 1992.
24. Z. Dziong, M. Juda, and L. Mason, "A framework for bandwidth management in ATM networks – aggregate equivalent bandwidth estimation approach," *IEEE/ACM Transactions on Networking* **5**, pp. 134–147, Feb. 1997.
25. S. Floyd, "Comments on measurement-based admissions control for controlled-load services," July 1996. Lawrence Berkeley Laboratory Technical Report.
26. R. Gibbens, F. Kelly, and P. Key, "A decision-theoretic approach to call admission control in ATM networks," *IEEE Journal on Selected Areas in Communications* **13**, pp. 1101–1114, Aug. 1995.
27. I. Hsu and J. Walrand, "Dynamic bandwidth allocation for ATM switches," *Journal of Applied Probability* **33**, pp. 758–771, Sept. 1996.
28. S. Jamin, P. Danzig, S. Shenker, and L. Zhang, "A measurement-based admission control algorithm for integrated services packet networks," *IEEE/ACM Transactions on Networking* **5**, pp. 56–70, Feb. 1997.
29. E. W. Knightly and J. Qiu, ""measurement-based admission control with aggregate traffic envelopes"," in *Proceedings of the 10th IEEE Tyrrhenian Workshop on Digital Communications*, September 1998.
30. D. Tse and M. Grossglauser, "Measurement-based call admission control: Analysis and simulation," in *Proceedings of IEEE INFOCOM'97*, (Kobe, Japan), Apr. 1997.
31. D. P. Bertsekas, *Network Optimization: Continuous and Discrete Models*, Athena Scientific, Belmont, MA, 1998.
32. D. P. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, Inc., 1992.
33. R. Cruz, "A calculus for network delay, Part I : Network elements in isolation," *IEEE Transaction of Information Theory* **37**(1), pp. 114–121, 1991.
34. R. L. Cruz, "A Calculus for Network Delay, Part II: Network Analysis," *IEEE Transactions on Information Theory* **37**, pp. 132–141, January 1991.
35. A. Parekh and R. Gallager, "A generalized processor sharing approach to flow control - the single node case," *IEEE/ACM Transactions on Networking* **1**, pp. 344–357, June 1993.
36. A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: The multiple node case," *IEEE/ACM Transactions on Networking* **2**, pp. 137–150, April 1994.
37. D. Wrege, E. Knightly, H. Zhang, and J. Liebeherr, "Deterministic delay bounds for VBR video in packet-switching networks: Fundamental limits and practical tradeoffs," *IEEE/ACM Transactions on Networking* **4**, pp. 352–362, June 1996.
38. R. Boorstyn, A. Burchard, J. Liebeherr, and C. Oottamakorn, "Statistical service assurances for traffic scheduling algorithms," *IEEE Journal on Selected Areas in Communications. Special Issue on Internet QoS* **18**, pp. 2651–2664, December 2000.
39. K. Thompson, G. Miller, and R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics," *IEEE Network Magazine* **6**, December 1997.

40. B. T. Doshi, "Deterministic rule based traffic descriptors for broadband ISDN: Worst case behavior and connection acceptance control," in *International Teletraffic Congress (ITC)*, pp. 591–600, 1994.
41. A. Elwalid, D. Mitra, and R. Wentworth, "A new approach for allocating buffers and bandwidth to heterogeneous, regulated traffic in an ATM node," *IEEE Journal on Selected Areas in Communications* **13**, pp. 1115–1127, Aug. 1995.
42. G. Kesidis and T. Konstantopoulos, "Extremal shape-controlled traffic patterns in high-speed networks," Tech. Rep. 97–14, ECE Technical Report, University of Waterloo, Dec. 1997.
43. G. Kesidis and T. Konstantopoulos, "Extremal traffic and worst-case performance for queues with shaped arrivals," in *Proceedings of Workshop on Analysis and Simulation of Communication Networks*, (Toronto), Nov. 1998.
44. E. Knightly, "Enforceable quality of service guarantees for bursty traffic streams," in *Proceedings of IEEE INFOCOM'98*, pp. 635–642, (San Francisco), Mar. 1998.
45. F. LoPresti, Z. Zhang, D. Towsley, and J. Kurose, "Source time scale and optimal buffer/bandwidth tradeoff for regulated traffic in an ATM node," in *Proceedings of IEEE INFOCOM'97*, pp. 676–683, (Kobe, Japan), Apr. 1997.
46. P. Oechslin, "Worst Case Arrivals of Leaky Bucket Constrained Sources: The Myth of the On-Off source," in *Proc. IEEE/IFIP Fifth International Workshop on Quality of Service (IWQoS '97)*, pp. 67–76, (New York), May 1997.
47. S. Rajagopal, M. Reisslein, and K. W. Ross, "Packet multiplexers with adversarial regulated traffic," in *Proceedings of IEEE INFOCOM'98*, pp. 347–355, (San Francisco), Mar. 1998.
48. M. Reisslein, K. W. Ross, and S. Rajagopal, "Guaranteeing statistical QoS to regulated traffic: The multiple node case," in *Proceedings of 37th IEEE Conference on Decision and Control (CDC)*, pp. 531–531, (Tampa), Dec. 1998.
49. M. Reisslein, K. W. Ross, and S. Rajagopal, "Guaranteeing statistical QoS to regulated traffic: The single node case," in *Proceedings of IEEE INFOCOM'99*, pp. 1061–1062, (New York), Mar. 1999.
50. J. Liebeherr, S. D. Patek, and E. Yilmaz, "Tradeoffs in Designing Networks with End-to-End Statistical QoS Guarantees," 2001. University of Virginia, Department of Computer Science, Technical Report CS-2001-11. See also the proceedings of the IEEE/IFIP Eighth International Workshop on Quality of Service (IWQoS '2000), pp. 221-230. June 2000.
51. M. Andrews, "Probabilistic end-to-end delay bounds for earliest deadline first scheduling," in *Proc. IEEE Infocom 2000*, pp. 603–612, (Tel Aviv), Mar. 2000.
52. C. Li and E. Knightly, "Coordinated network scheduling: A framework for end-to-end services," in *Proc. of IEEE ICNP 2000, Osaka, Japan*, November 2000.
53. S. D. Patek, R. Venkateswaran, and J. Liebeherr, "Simple Alternate Routing for Differentiated Services Networks," *Communication Networks* , 2001. to appear.