INTERNETWORK TRAFFIC MANAGEMENT ENHANCEMENTS: EVALUATION AND EXPERIMENTAL RESULTS

Joseph Macker and Vincent Park
Information Technology Division
Naval Research Laboratory
Washington, D.C.

ABSTRACT

In recent years, there has been an increasing reliance on internetworking technology for military communication networks. As more diverse and mission critical applications begin to coexist within this shared infrastructure the need for improved network traffic and bandwidth management becomes apparent. Recent developments within the research community and within commercial products provide numerous candidate technologies that may be applied as solutions. This paper explores a set of such techniques while evaluating their relative merits and performance tradeoffs. A subset of empirical results from an actual testbed environment is also provided and discussed.

INTRODUCTION

Resource allocation is at the root of many management problems being faced today in and beyond the information processing industry. Within communication system design, we have seen two basic system paradigms emerge.

Stovepipe Environments: Dedicated resources ensuring end user satisfaction through a set of private, separate information delivery systems.

Shared Use Environments: A common pool of resources shared by many users and end systems. With the Internet Protocol (IP) suite, this also provides interoperability across potentially large scale heterogeneous information delivery systems.

To improve operational cost, flexibility, and interoperability military communication networks are largely converting to shared use environments. While this trend is likely to continue and provide numerous benefits, overall traffic management strategies remain an ongoing technical issue. A main concern is that increasing distributed, shared access increases the probability of problematic resource contention situations. To address this issue, we explore the ability of different emerging bandwidth management strategies and techniques to support robust critical mission services and improve cost effectiveness within military internetworks. There is extensive literature on this growing interest area; for some additional technical background see [Braden 94, Clark 92, Floyd 95, Macker 96].

One apparent unique design factor of military internetworks is the extensive reliance on wireless, bandwidth-constrained, Wide Area Network (WAN) links (e.g., satellite circuits) to connect force components. These WAN links are often used for access to and from high bandwidth local environments (e.g., perhaps a local high speed fiber plant on a Naval platform). The wireless WAN issue is an important distinction since system cost and performance tradeoffs between enhanced queueing complexity and packet forwarding requirements are interrelated.

Enhanced datagram traffic control and quality-of-service (QoS) networking are evolving technologies in which there has been an explosive amount of excellent technical work produced in recent years. Despite some of this technical progress, traffic management and network QoS issues remain hotly debated topics with numerous divergent opinions and approaches being espoused. It is our hope to demystify several issues relating to managing network traffic over bottleneck links. While many issues remain open for future exploration, we discuss "how, why, and where" certain traffic management components can be effective solutions.

The goal of improving network traffic management within future military internetworks is motivated by the resultant operational benefits.

- More cost effective, shared allocation of precious network resources
- Increased protection for mission critical communications within a shared environment
- Improved network integration of multimedia and mission planning services
- Increased assurance of maintaining appropriate shared bandwidth policies under stressed conditions and through known system bottlenecks

Traffic service improvements are critical to the distributed operation of applications such as the Joint Maritime Communications and Information System (JMCIS). For example, the time sensitive nature of JMCIS track updates imposes certain performance requirements on the data delivery service provided by the underlying network. In addition to specific application support, enforceable traffic management policies allow diverse warfighting areas to gain confidence in a shared infrastructure that provides stable, predictable performance under stressed and dynamic network conditions.

Due to the uniqueness of military WAN communications, protecting mission critical traffic at the ingestion points of bottleneck backbone links (e.g., SHF satellite) through enhanced servicing is essential. Significant performance improvements may be realized by enhancing traffic management only at the bottleneck link(s) when a large part of the network architecture is essentially overprovisioned (i.e. in processing power and available capacity) relative to a bottleneck link. This may often be true where interconnected high-speed shipboard local area networks (LANs) converge to a moderate rate (e.g., T1) wireless interface for external communications. We will term this type of an architectural approach a *partially deployed* QoS. The initial evaluation of network traffic management techniques for *partially deployed* QoS architectures will be a primary focus of further discussion in this paper.

TESTBED OVERVIEW

A testbed has been established at NRL for the experimentation and evaluation of IP traffic management technology components. The testbed facilitates rapid reconfiguration and allows for future growth.

Initial Testbed Configuration

The testing configuration, in Figure 1, consisted of multiple distributed Ethernet segments separated by commercial routers with adjustable low-to-moderate rate network bottleneck connections (i.e., 2.4kbps-4Mbps). Components used in the testbed included a variety of Cisco router models. Component selection was not intended as an endorsement of any one particular vendor, and our choice was based upon the fact that the Cisco routers allowed us to experiment with a number of alternate queueing and QoS strategies.

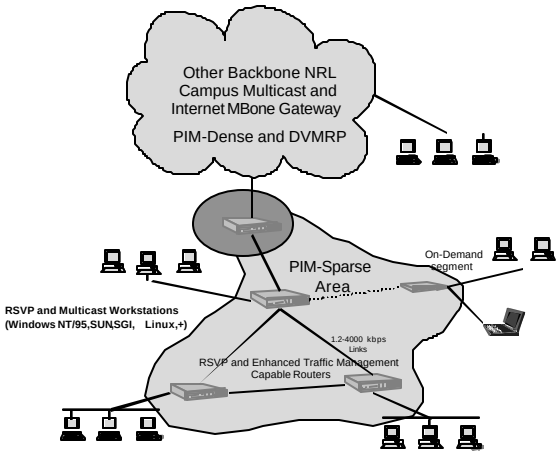

Figure 1: NRL Traffic Management Testbed

A variety of end systems (i.e., variety of UNIX and Windows NT workstations) were used to generate and log network traffic. The NRL-developed MGEN software toolkit[1] was employed for accurate generation and logging of both unicast and multicast traffic. During specific test runs the modeled sources injected traffic levels well above the capacity of the bottleneck links to produce heavily stressed traffic loads within the testbed. The goal was to verify the performance of enhanced packet forwarding and queueing techniques under varying congestion conditions with a variety of datagram source modeling. During testing, the end system clocks were synchronized using the Network Time Protocol (NTP) which allowed for accurate measurement of end-to-end delivery delay

Test Tools

Multicast and unicast UDP traffic are anticipated to be important military IP network transport mechanisms. The reason behind this is the increasing demand for group data dissemination/collaboration tools, multimedia applications, and support for relatively short self-contained messages (JMCIS track updates) more suitable to UDP datagram delivery. Preliminary study results mainly focused on the effects of the various queueing mechanisms on UDP data flows, including extensive support for multicast network traffic and varied datagram message sizes. In addition, some tests were conducted using TCP data flows to provide insight into the behavioral differences and coexistent interactions of data flows using different transport mechanisms.

The MGEN toolkit was used to generate network data flows and log end-to-end statistics. The MGEN toolkit provides the ability to produce accurate time scripted traffic loading from multiple traffic sources and log data from the multiple flows at the receivers. MGEN also provides a Resource Reservation Protocol (RSVP) [Braden 97, Zhang 93] application-layer interface for interoperability with RSVP-capable network components. IVOX[2], an NRL-developed Internet voice application which also provides an RSVP application-layer interface, was used in the testbed to create more tangible demonstrations.

ENHANCED TRAFFIC MANAGEMENT IN PACKET FORWARDING MECHANISMS

Before discussing results from actual testing we first review some basic queueing mechanisms and consider how they can be used to enhance QoS. This is not an exhaustive list of technical approaches but is representative of techniques available within commercial routers. While a variety of enhanced queueing mechanisms have been implemented in commercial router products, many have not been verified in light of potential military usage and the highly stressed performance requirements anticipated. To apply available mechanisms to specific QoS problems, it is essential to understand respective capabilities and limitations.

In general, a queueing mechanism, which can comprise multiple internal queues, can be functionally separated into two distinct parts.

> *A Sorting Algorithm (Packet Classifier)*: identifies, separates and inserts packets into separate internal queues.
>
> *A Forwarding Algorithm (Packet Scheduler)*: determines the order and manner in which the separate queues are serviced (i.e. which packet is forwarded next).

FIFO Queueing

First-in-first-out (FIFO) queueing[3] is perhaps the simplest mechanism conceived and implemented. Packets are forwarded in the same order they arrive. FIFO queueing uses a single internal queue, and thus requires no classification/sorting of packets. Packets are simply inserted at the tail of the queue upon arrival, and removed from the head of the queue upon availability of the transmitter. Upon congestion, packets are typically dropped from the tail of the queue. This classic algorithm has been widely used in networking products, partly

[1] ftp://manimac.itd.nrl.navy.mil/ManiMac/Pub/MGEN/

[2] ftp://manimac.itd.nrl.navy.mil/ManiMac/Pub/IVOX/

[3] This type of queueing is also commonly referred to as first-come-first-serve (FCFS).

due to its simplicity and low processing overhead. However, it provides no outgoing service isolation (i.e. it offers no variability in QoS for individual data flows or packet types). Furthermore, some undesirable behaviors are exhibited when packet trains from different data flows are inserted into the same FIFO queue. This will be further discussed in the section on fair queueing.

Priority Queueing

Priority queueing--available in some form in many networking products--provides basic enhancement to traffic management. This mechanism sorts packets based on differences in "relative importance" and inserts them into separate internal queues or shuffles their relative positions within a single queue. The forwarding algorithm generally always transmits packets of the highest priority first. If there are no packets of the highest priority level, the next highest priority queue is serviced, and so on.

Priorities established in this type of queueing are *absolute* (i.e., if there is sufficient high priority traffic to saturate a link, all lower priority traffic will be "locked out"). Particular bandwidth, delay, or packet size considerations are usually not taken into account. This can be a considerable problem, as some protocols (e.g., TCP) seek to detect and use the entire available resource dynamically. Determining how to assign the relative priority levels can also be a considerable problem. In light of the potential for any traffic other than the highest priority to be totally locked out, great care must be taken when using priority methods. Consideration must be given to what this really means in terms of shared use management. Finally, FIFO queueing is still essentially used for all packets of a given priority; thus, the known undesirable behaviors of FIFO queueing still applies to data flows within a common priority level queue.

Fair Queueing

As for priority queueing, fair queueing also involves the use of multiple internal queues with an ability to sort and insert different packets into each queue. The primary difference lies in how the queues are serviced. The objective is to provide "fair" or "equivalent" service to each of the queues. Although the notion of fairness has been defined many ways, perhaps the most widely accepted is that the traffic in each queue should obtain an equal portion of the bandwidth. An extension of this idea is that of Weighted Fair Queueing (WFQ), which adds the capability to divide the bandwidth unequally [Partridge94].

When multiple data flows are inserted into a common internal queue the undesirable behaviors of FIFO queueing still apply. However, if individual data flows are separated into different internal queues, competing packet trains can be interleaved (which tends to alleviate the undesirable behaviors). Obviously, consideration must be given to the impact on processor and memory constraints, as the number of queues being serviced increases.

An important benefit of the WFQ model is that it can be used to provide strong guarantees for a given data flow within a heterogeneous internetwork. Unlike priority queueing, the concept of bandwidth and delay bounds is introduced. Given that WFQ is used at every hop for a particular data flow and the traffic injection source conforms to certain token bucket model assumptions, it has been shown that the worst case queueing delay is bounded within an internetwork [Parekh92].

Other Issues

Due to limited space, we cannot provide a comprehensive discussion of all the related QoS architecture and traffic management for internetworking. There are other areas and techniques not discussed here that will likely be important in future networks. They include, but are not limited to, Class-based Queueing (CBQ) [Floyd 95], TCP shaping methods, Random Early Detection (RED), lower layer ATM interaction, and probabilistic servicing models. These areas are deserving of further analysis and evaluation as potential mechanisms for improving the operation of future military networks.

TEST RESULTS

Here we present a few simple illustrative example test cases from the NRL traffic management testbed. The first test case (Figures 2 and 3) demonstrates a custom queueing operation and the second case (Figures 4 and 5) demonstrates WFQ operation in combination with a dynamic RSVP service.

Test Case 1: Statically Configured Minimum Bandwidth Guarantees Using WFQ Variant

Over thirty separate test scenarios were conducted to evaluate basic custom queueing capabilities under a variety of traffic loading and stressed network conditions. Only the results of one simple test case is presented in Figures 2 and 3 to illustrate the basic functionality of this capability. This test case shows how custom queueing or variants of WFQ provide protection to mission critical data flows for time or bandwidth sensitive traffic even during times of congestion. The MGEN toolkit was used to generate three distinct network traffic flows (i.e. one "mission critical" data flow and two "non-critical" poorly behaving data sources). The queue configuration was configured to allocate a 75% minimum bandwidth guarantee to the mission critical data queue under congestion.

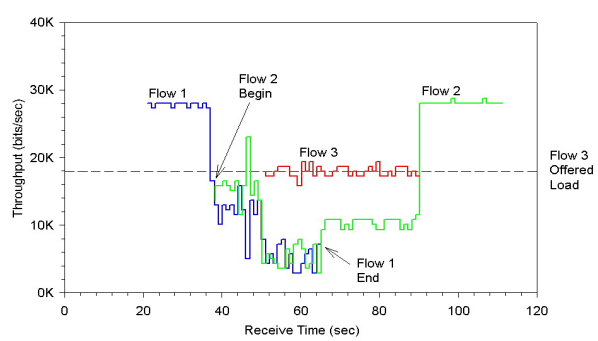

Figure 2: Custom Queueing Throughput

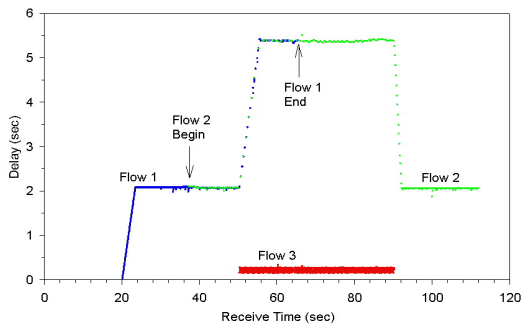

Figure 3: Custom Queueing Delay

Initially, as is shown in Figure 2, a non-critical data flow (Flow 1) was the only source of traffic, and thus dynamically used the full capacity of the link. A significant fraction of Flow 1 packets were lost (i.e., dropped) since the offered load far exceeded the available bandwidth and queue depth of the bottleneck. Once the second non-critical data flow (Flow 2) was introduced, the two data flows subsequently divided the capacity of the link about evenly—each achieving a throughput well below respective offered loads. This response is expected as these two data flows are serviced within the same internal queue. In contrast, when the modeled mission critical data flow (Flow 3) is introduced it essentially maintains a throughput equivalent to its offered load. When the mission critical traffic was present, the non-critical data flow traffic is reduced to a non-interference level and those flows shared the remaining fraction of the capacity allocated to their respective queue. The delay graph for the same test (Figure 3) shows the relative performance gain in end-to-end delay for the mission critical traffic class.

During the more exhaustive series of performance tests within the NRL testbed, design flaws were discovered in a router implementation related to the servicing of variable length datagram sizes. We are working with the manufacturer in reporting and diagnosing such problems leading to software assurance improvements in future product releases that we can subsequently verify through additional testing.

Test Case 2: Dynamically Signaled Minimum Bandwidth Guarantees Using WFQ Variant and RSVP

Test Case 2 demonstrates a scenario similar to Test Case 1 with the introduction of RSVP, and additional random traffic congestion sources. The real-time or mission critical traffic flow is representative of real-time collaborative planning or time sensitive track update information and the congestion sources may represent lower bandwidth precedence e-mail or Web traffic. The throughput and delay curves for the test are depicted in Figures 4 and 5. In Figure 4, the mission critical traffic flow solely occupies the bottleneck link for around 30 seconds. During this period the flow is classified with no special queueing treatment and since there is no competing traffic the throughput is equivalent to the source rate. Next, the introduction of 15 random flows occurs—representative of bursty background traffic conditions. Individual bursty traffic source rates are each within the bottleneck link rate, but in aggregate they produce a congested traffic condition. The results illustrate that during this stressed period, unprotected mission critical traffic is severely affected and significant data is dropped or lost. After an intentional delay of approximately 30 seconds, an RSVP signaled QoS reservation is established across the network for the mission critical flow. We see that the desired throughput of the mission critical flow is largely recovered and maintained irrespective of the continuous traffic congestion condition. Figure 5 shows the improved end-to-end delay effects of the same test scenario.

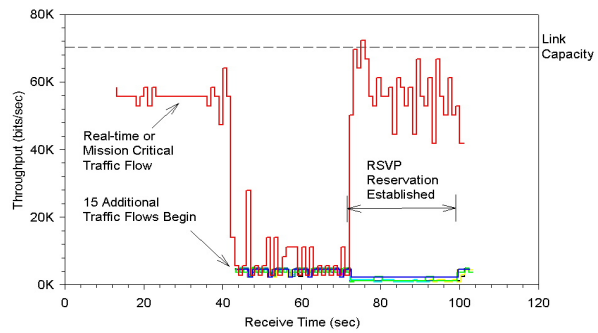

Figure 4: RSVP WFQ Throughput Test

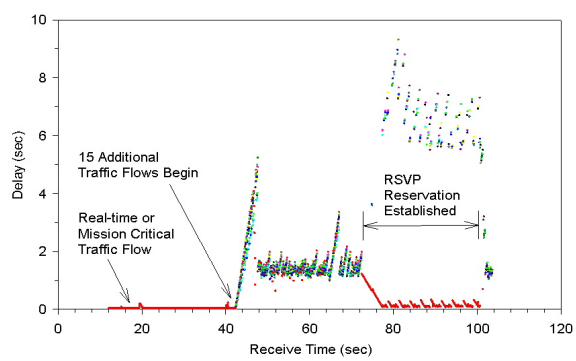

Figure 5: RSVP WFQ Delay Test

While the prospects for RSVP utility and architecture transition are promising, there are a number of issues requiring continued technical exploration over the coming years. First, RSVP was designed to be a flexible QoS signaling format, yet its usage involves additional protocol overhead and scaling issues—an aspect not yet widely investigated in wireless or low bandwidth environments. Second, operational concepts and policies for the use of RSVP need development. To simplify deployment issues, proxy agents can provide QoS signaling *on behalf* of network applications that actually desire QoS but have no RSVP interface capability. Thus, if desireable, RSVP can be deployed with the existing application set intact and with a more

hierarchical security infrastructure. Potential RSVP proxy agent software includes Cisco router-based RSVP agents, NRL MGEN toolkit, or Intel's PC-RSVP toolkit.

SUMMARY AND RECOMMENDATIONS

The test results presented here clearly illustrate the merits of enhanced queueing mechanisms for network traffic management. Preferential treatment, isolation, and/or service guarantees to specific data flows or mission areas within a military network is often desirable. In addition to advantages of different mechanisms, we explored some of the pitfalls.

We discussed priority-based queueing and related issues. Absolute priority-based treatment increases the potential for lower priority traffic to be *locked out* (i.e. receive no service) under certain conditions. This behavior is problematic if priority queueing is coupled with a transport mechanism such as TCP, which attempts to consume available uncongested bandwidth. Priority queueing also cannot be used to provide any deterministic service guarantees (e.g., bandwidth contracts) in the face of arbitrary traffic conditions without additional support mechanisms like traffic shaping.

As an alternative to priority queueing styles, custom queueing (a WFQ style implementation) provides minimum bandwidth guarantees for aggregate traffic serviced within a specific queue. This style service was shown to significantly enhance the QoS of distinct data flows or traffic through bottleneck links under congestion. We discussed how such a technique can be applied to protect *mission critical traffic* or for providing minimum bandwidth requirements for traffic from different *mission areas.* An example application for initial deployment could be the servicing and protection of JMCIS traffic networked within a shipboard architecture and more importantly across bottleneck communication elements.

Despite the encouraging results, there remains an ongoing issue of how to best configure queueing service parameters within a network. Configuration through a system administration or network management function is perhaps the easiest way to envision initially transitioning this technology in a secure and useful fashion. Thus, partial deployment of traffic management at network system bottleneck locations could be securely managed without requiring signaling from the end systems. In the longer term, as shown in the RSVP experiments, QoS signaling allows for more dynamic bandwidth management based upon explicit end system requirements. In conclusion, we feel initial router traffic management capabilities are mature enough and should be considered in present architectures, especially in unique military situations to improve constrained bandwidth bottleneck management (e.g., wireless WAN). We recommend incremental and secure approaches to fielding improved network traffic management, as outlined above.

REFERENCES

[Braden 94] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," Internet RFC 1633, June 1994.

[Braden 97] R. Braden, L. Zhang, S. Berson, Z. Herzog, and S. Jamin, "Resource ReSerVation Protocol (RSVP)—Version 1 Functional Specification," Internet Draft (work in progress), June 1997.

[Clark 92] D. Clark, S. Shenker, and L. Zhang, "Supporting Real-time Applications in an Integrated Services Packet Network: Architecture and Mechanism," *Proc. ACM SIGCOMM*, September 1992.

[Floyd 95] S. Floyd, V. Jacobson, "Link-Sharing and Resource Management Models for Packet Networks," *IEEE/ACM Trans. on Networking*, 3(4): 365-386, August 1995.

[Macker 96] J. Macker, "Controlled Link Sharing and Quality of Service Data Transfer for Military Internetworking," *Proc. MILCOM*, October 1996.

[Parekh 92] A. Parekh, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks", Technical Report LID-TR-2089, Laboratory for Information and Decision Systems, MIT, 1992.

[Partridge 94] C. Partridge, Gigabit Networking, Addison-Wesley Professional Computing Series, 1994.

[Postel81] J. Postel, "Transmission Control Protocol - DARPA Internet Protocol Program Specification," Internet RFC 793, September 1981.

[Zhang 93] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, "RSVP: A New Resource Reservation Protocol," *IEEE Networks Magazine*, September 1993.