WASHINGTON UNIVERSITY

SEVER INSTITUTE OF TECHNOLOGY

ERROR CONTROL FOR CONTINUOUS MEDIA AND LARGE SCALE MULTICAST APPLICATIONS

by

Christos Papadopoulos

Prepared under the direction of Professor Gurudatta M. Parulkar

A dissertation presented to the Sever Institute of

Washington University in partial fulfillment

of the requirements for the degree of

DOCTOR OF SCIENCE

August, 1999

Saint Louis, Missouri

WASHINGTON UNIVERSITY

SEVER INSTITUTE OF TECHNOLOGY

ABSTRACT

ERROR CONTROL FOR CONTINUOUS MEDIA AND LARGE SCALE MULTICAST APPLICATIONS

by Christos Papadopoulos

ADVISOR: Professor Gurudatta M. Parulkar

August, 1999

Saint Louis, Missouri

One of the key elements that has contributed to the runaway success of the Internet today is its ability to support a wide variety of applications with often conflicting requirements. The Internet, however, offers only a best-effort service model which means that packets can be lost. Applications requiring reliable transmission usually employ a transport protocol to provide reliability. The most common reliable transport protocol today is TCP which provides total reliability at the expense of latency, a trade-off that works well for applications like file transfer.

File transfer, however, is just one class of applications. Other classes such as continuous media, transaction-oriented and multicast applications have different, sometimes opposing requirements than file transfer, and thus need different error control mechanisms. In this thesis, we present error control mechanisms for two important classes, namely (a) *interactive continuous media* (CM), and (b) *large-scale multicast* applications. The former include applications like teleconferencing, video on demand and visualization; the latter include Internet radio and movies, web and software updates and distributed interactive simulation.

Retransmission-based error recovery has been, in general, considered inappropriate for continuous media (CM) applications because of its latency. However, retransmission is still attractive because it requires minimal network bandwidth, (less than either peak-bandwidth allocation or FEC for bursty streams) and processing cost. In this thesis we have explored retransmission as a possibility for providing error control for interactive CM applications. We have designed, implemented, and evaluated a retransmission-based error control scheme for CM applications, which aims to provide the best possible reliability at the lowest cost, without violating the application's timing constraints. We have enhanced selective-repeat retransmission with: (1) playout buffering to increase the time available for recovery, (2) gap-based rather than timer-based loss detection to minimize loss detection latency, (3) implicit expiration of sender retransmission buffers to eliminate acknowledgments, (4) conditional retransmission requests to avoid triggering late, unnecessary retransmissions, and (5) delivery of data integrity information to the application to aid in concealment. Our experimental results show that the mechanism achieves reductions in observed loss compared to loss without error control ranging from one to several orders of magnitude, without violating the application's delay constraints. Our work was one of the first to demonstrate the feasibility of retransmission for CM applications. Our conclusions were confirmed later by other researchers who have reached similar conclusions.

Multicast is becoming very important for emerging applications and services on the Internet. Multicast is already in use for audio and video conferencing, stock quote distribution, collaborative workgroup computing, push technologies, news distribution, shared whiteboard applications, broadcasting of live concerts, and web site updates. The size of multicast groups may range in size from a handful to millions. Control mechanisms must be very scalable to handle such groups. The greatest challenges are implosion from receiver feedback, exposure from reception of superfluous packets, and latency due to lack of local recovery. These problems are hard to solve because the existing IP Multicast provides a shared access channel which makes it hard to perform localized actions, e.g., local recovery, ACK and NACK fusion, etc.

We made the observation that forwarding and error control are two clearly separable components, and great benefits can be realized by decoupling and placing each one where it is more beneficial: the forwarding component at the routers, and the error control component at the receivers.

The separation is very clean, and does not violate any layering principles. The result of this separation is a set of simple router forwarding services called *Light-weight Multicast Services (LMS)*. These services do not require examining or storing packets, but greatly simplify multicast control protocols. LMS is highly suited for reliable multicast, where it achieves near-optimal performance. Existing schemes like SRM do not require router assistance, but their performance is significantly less than optimal. The advantages of router assistance have not gone unnoticed: the router giant Cisco has begun marketing PGM, a reliable multicast scheme that postdates LMS, but is less efficient.

LMS is very light because it does not require per-packet state; it is highly scalable because it shields receivers from topology, and adapts almost instantly to any dynamic group changes. LMS has been extensively studied using simulation, and has been implemented in NetBSD Unix. In this thesis we present results from simulations that compare the performance of LMS with PGM, and SRM, which is a non-assisted scheme. We show that LMS outperforms the other schemes by three to ten times in terms of latency; in terms of exposure (reception of unwanted packets), LMS lags slightly behind PGM, which completely eliminates exposure. SRM with no local recovery suffers dramatically from exposure compared to LMS, which offers orders of magnitude lower exposure than SRM. We have also implemented LMS, demonstrating its feasibility, and evaluated its processing overhead, showing that it is on par with normal multicast forwarding.

to my parents and all who worked hard to educate me