High Performance Transport for Real-Time and Quasi-Realtime Applications[*]

David Lapsley
MIT Haystack Lab
dlapsley@haystack.mit.edu

Alan Whitney
MIT Haystack Lab
awhitney@haystack.mit.edu

John Wroclawski
MIT CSAIL
jtw@mit.edu

Most protocol development for high bandwidth, large delay networks targets TCP-like reliable communication. In contrast, we are currently developing and testing network protocols and algorithms that take advantage of the special characteristics of *real-time* scientific instrumentation applications to transport their data across fast, long-distance networks. These applications share with classic real-time transport the characteristics of loss tolerance, time sensitivity, and the need for clocking and synchronization, but differ from audio and video by generating several orders of magnitude more traffic. In this paper, we discuss the development and application of several techniques key to the efficient, economical transport of this class of data. These include

- multi-dimensional performance optimization – the ability to dynamically optimize the data transport across multiple concurrent objectives, such as reliability, timeliness, and cost.
- protocols for high speed, quasi-real time transport – extension of RTP-like protocols to the 1-100Gs/s range.
- scavenged bandwidth – the use of background and reserve capacity to minimize the cost of transporting high volumes of loss-tolerant real-time data

One key component of the system being developed is the use of an intelligent end system that uses congestion feedback from the network, a high level transfer "profile" and a multi-parameter optimized algorithm for determining the best transmission strategy. This transmission strategy could include: the use of alternate paths through the network, dynamically varying the transmission rate, marking IP packets with a less than best effort priority, varying the retransmission strategy to allow some (tolerable) packet loss, etc. The intelligent end system optimizes a benefit function based on the constraints specified in the profile and the feedback from the network. In this manner, the end system is able to achieve the transfer of data in a manner that is optimal in some sense.

The second key component of this system is the use of specialized protocols for high speed, quasi-real time transport of scientific data. The Real-time Transport Protocol (RTP) [1] has been chosen for its high degree of suitability for this task. It provides timing synchronization, a simple session management model, statistics and error reporting. It also makes use of the multi-cast model which allows it to make efficient use of network bandwidth for applications that require the distribution of a single set of data to multiple end systems. However, RTP demonstrates several limitations as data rates pass 1GB/s, and as data synchronization requirements drop towards nanoseconds or less. We describe extensions to RTP that address these limitations.

---

[*] Work supported in part by the National Science Foundation under Grant No. ANI-0230759: Development of Shared-Network Protocols and Strategies for e-VLBI.

The third component is the use of scavenged or "less-than-best-effort" bandwidth – bandwidth that is not currently being used by other applications in the network. Our objective is to explore ways to make the maximum possible use of this scavenged bandwidth, to reduce the cost of data transport to the scientific community and to ensure that that the large, long-lived data flows generated by such applications does not interfere with regular Internet traffic. We explore support for this capability in several scenarios. Some networks (e.g. Abilene) offer this bandwidth as an additional service. In this case, IP packets are marked with a special code point that identifies them as being less-than-best-effort and indicates that they should be dropped in preference to regular best effort traffic at any point in the network where there is contention for buffer space/bandwidth. Where network support is lacking, end-to-end resource control algorithms tuned to this purpose provide some of the same capabilities. Our aim is to use scavenged bandwidth as effectively as possible over a wide range of network conditions.

By combining these three components into a single intelligent system, it is possible to transport large volume, high bandwidth, real-time scientific data flows in a manner that is in some sense optimal while minimizing the impact of the scientific data on other users' data flows. A range of scientific uses, such as our motivating application of Very-Long-Baseline Interferometry (VLBI), have characteristics that can be exploited by the intelligent system described above. Our present work is carried out in cooperation with astronomy's VLBI community, and is subject to ongoing development and validation in international testbeds and production high-speed networks available to that community. We describe VLBI briefly below.

VLBI is one of the most powerful techniques available for the high-resolution imaging of distant radio sources in the universe and for making accurate measurements of the motion of the earth in space. Multiple radiotelescopes scattered over the surface of the earth simultaneously record data from a radio source at streaming data rates as high as 1 Gbps for a 24-hour period; the data are then shipped to a central processing site for correlation analysis. Because the signal-to-noise ratio achieved by VLBI increases with the bandwidth of the observations, there has been a continual effort to increase the captured data rate. The advent of modern high-speed networks offers the possibility to transport VLBI data electronically with much higher data rates and lower costs to maximize the scientific potential of the observations. The electronic transmission of VLBI data (dubbed e-VLBI), however, presents a special challenge to the use of high-speed global networks. While e-VLBI requires the transmission of vast amounts of quasi-real-time instrumentation data, latency and data-loss tolerances are less stringent than in more traditional applications. For this reason, e-VLBI and similar instrumentation applications can share global high-bandwidth networks unobtrusively, increasing scientific productivity and drastically lowering costs.

References

[1] H. Schulzrinne, S.Casner, R. Frederick, V. Jacobson, RFC 3550: RTP: A Transport Protocol for Real-Time Applications available: http://www.ietf.org/rfc/rfc3550.txt?number=3550, IETF, 2003.

[2] A. Whitney, J. Wroclawski, Development of Shared-Network Protocols and Strategies for e-VLBI. NSF Grant # ANI-0230759. 2002, MIT: Cambridge, MA.

[3] D. Lapsley, J. Wroclawski, A. Whitney, Development of new IP Protocols for eVLBI. available: http://web.haystack.mit.edu/e-vlbi/lapsley_dwingeloo_03_files/frame.htm, MIT Haystack Observatory, 2003.