# An Update on ATM Traffic Management

Shivkumar Kalyanaraman, Rensselaer Polytechnic Institute

*It has been a year since the first article of this column was written. That article discussed the GFR/UBR+ service which was then under hot debate in the ATM Forum. In the current article, Shivkumar Kalyanaraman gives us an update on the status of the GFR service definition, as well as other topics discussed in the ATM Forum.*

The Traffic Management (TM) group at the ATM Forum has been fairly focused on developing a new service called *guaranteed frame rate* (GFR) service [1]. This service is expected to be positioned as an intermediate alternative between available bit rate (ABR) service, which uses a closed-loop feedback control framework, and unspecified bit rate (UBR) service, which provides no assurances whatsoever. The specification is expected to go to a straw vote in December 1998.

The motivation for GFR service comes from the perceived need for frame-based minimum rate guarantees instead of cell-based guarantees. This is beacuse cell-based guarantees do not easily translate to well understood frame-based guarantees. In GFR service, complete frames are accepted or discarded, and traffic shaping/marking is frame-based. Moreover, the goal is to minimize policing and signaling functions, avoid feedback, and achieve the minimum rate guarantees through simple buffer management and scheduling at switches. This way, an abstraction of a "leased-line-plus" or "frame relay CIR-like" service can quickly be offered to data traffic customers.

Recent GFR work has focused on the definition of the conformance tests. The conformance consists of three parts — one for maximum frame size (MFS), one for peak cell rate (PCR), and one for cell loss priority (CLP) marking (CLP bits in a frame are all marked zero or all marked one). The debate was whether to require all three tests and whether to execute them serially or allow parallel execution. The agreement was to require the PCR test, and allow MFS and CLP tests to be optional. In addition, these tests can execute in parallel. Also, a motion to position GFR as an ITU-T variable bit rate (VBR) conformance definition (VBR.4) was defeated in February. GFR will be defined as a new service.

Another examined issue was how to accumulate the peak-to-peak cell delay variation (CDV) of a path based on the CDV at each switch in the path. TM specification 4.0 [2] mentions two methods of doing this. The first method is a simple one where the CDV of the path is simply the sum of the CDVs at each switch. This method gives an overestimate and would lead to underutilization and increased blocking. The second, more complex asymptotic accumulation method involves the use of mean, variance, and a discrepancy term at each switch, requiring the switch to measure these quantities — an expensive task. An intermediate approach has been

(Continued on page 7)

the physical medium attachment (PMA) sublayer. Commercially available Serializer/Deserializer (SerDes) components first developed for Fibre Channel can be used for gigabit Ethernet. Only modest additional testing requirements are needed, mainly due to the fact that gigabit Ethernet operates at a clock rate of 1.25 Gb/s (1 Gb/s data rate) while Fibre Channel operates at a clock rate of 1 Gb/s (800 Mb/s data rate). This reuse of off-the-shelf components has contributed to the rapid development and deployment of gigabit Ethernet systems. The PMA sublayer is also responsible for recovering a clock reference from the received data stream. The expansion which results from encoding 8-bit octets into 10-bit symbols requires a signaling rate of 1.25 Gbaud at the serial interface to the medium.

The PCS includes a function referred to as *auto negotiation*, which is a link startup and initialization procedure first defined in IEEE Std 802.3u for 100BASE-T. Within IEEE Std 802.3z, auto negotiation is used to select between the CSMA/CD and full-duplex operating modes, and to select whether the Pause flow control mechanism is enabled or disabled on a link-by-link basis.

At the bottom of the diagram are the 1000BASE-SX and 1000BASE-LX fiber optic transceiver specifications. The 1000BASE-SX specification for short-wavelength laser transceivers supports multi-mode fiber optic links at distances up to 275 m using 62.5 μ fiber, and 550 m using 50 μ fiber. Once again, Fibre Channel provided the starting point for the 1000BASE-SX specification, but in addition, the 1000BASE-SX specification embraces VCSELs as well as the older CD style of laser developed for the Fibre Channel market.

1000BASE-LX supports longer distances using higher-cost components, spanning 550 m on 62.5 μ or 50 μ fiber, and up to 5 km on single-mode fiber. The 1000BASE-LX laser transmitter is optimized for single-mode fiber, and requires a mode-conditioning patch cord to support multimode fiber optic cable. The patch cord mitigates an effect known as differential mode delay by altering the launch characteristics of a 1000BASE-LX laser transmitter so that the resulting optical beam more closely resembles the overfilled launch pattern produced by a light-emitting diode (LED). Multimode fiber modal bandwidth is specified under the launch conditions produced by an LED. Through use of the mode-conditioning patch cord, a 1000BASE-LX link can be characterized on the basis of the rated modal bandwidth of a multimode fiber. Both 1000BASE-SX and 1000BASE-LX specify the familiar duplex SC optical connector, eliminating the most common installation problem encountered in fiber optic networks, the misconnection of the transmitting and receiving fibers.

IEEE Std 802.3z also includes a specification for a transceiver technology referred to as 1000BASE-CX, which supports shielded copper cables links spanning 25 m. The SerDes component which makes up the PMA sublayer is designed to drive this cable directly, which makes 1000BASE-CX an economically attractive choice for short-distance interconnections, for instance, between devices located within the same rack or within a computer room or telephone closet.

A new project within the IEEE 802.3 Working Group, referred to as 1000BASE-T, is chartered to develop a PHY specification which will support 1000 Mb/s operation on four pairs of category 5 UTP cabling, at a maximum link distance of 100 m. This project is being conducted in the P802.3ab Task Force. 1000BASE-T will take advantage of recent advances in silicon process technology which permit complex high-speed digital signal processing algorithms to be implemented cost effectively. A 1000BASE-T PHY will transmit its signal on all four pairs of wire simultaneously, thus reducing the data rate on each pair to 250 Mb/s. The use of a five-level pulse amplitude modulation scheme further reduces the signaling rate on each pair. Hybrids and digital echo cancellation are used to achieve full-duplex communication.

Network administrators will actively embrace gigabit Ethernet because the familiar management attributes that have been in use in 10 and 100 Mb/s Ethernet networks have been retained. The attribute definitions have been updated to reflect the fact that statistics counters tick 10 times faster for gigabit Ethernet, and certain attributes have been extended to embrace the new physical layers, but the "look and feel" of Ethernet from the network manager's point of view has been preserved.

### Biography

HOWARD FRAZIER (hfrazier@cisco.com) is employed by Cisco Systems, Inc. within the Workgroup Business Unit. He is chair of the IEEE 802.3z Gigabit Task Force, which developed the Gigabit Ethernet standard. Previously, he was chair of the IEEE 802.3u 100BASE-T Task Force, which developed the standard for Fast Ethernet. Prior to joining Cisco he was employed by Sun Microsystems, Inc. He graduated from Carnegie-Mellon University with a B.S.E.E. in 1983.

ATM FORUM/from page 7

proposed which uses the Chernoff method (which in turn uses the Markov, or Chebyshev, inequality) and assumes that local delays at switches are independent and gamma distributed [1].

Other brief discussions centered around the issues at a VC-to-VP aggregation point. One issue was how to determine the QoS of the VP relative to the QoS of the VCs. It has been documented that no general solution is known to this problem. Another issue was how VP sources should handle explicit forward congestion indication (EFCI); the problem is that source behavior 12 requires sources to reset the EFCI state. At a VC-VP boundary, this means that EFCI information from the VC control loop would be lost. The solution proposed for this is to reflect the EFCI state of the incoming data cells back to the VC source, stopping short of a full virtual source/virtual destination implementation.

A number of joint sessions have been held with network management, SAA, and RBB groups. The work with network management centered on how to count valid and invalid RM cells, and how to log traffic descriptors. The work with SAA centered around ABR API issues to provide an interface to query and set ABR parameters. Another issue was how to request SCR and MBS parameters given the mean and PCR of a video stream. The work with RBB centered around simplification of traffic parameters for residential users, and issues of shared access over asymmetric links such as cable.

Several performance contributions on GFR, IP over ATM, and VS/VD were presented in recent meetings.

### References

[1] J. Kenney, Ed., "ATM Traffic Management Living List," ATM Forum LTD-TM-01.07, Apr. 1998.
[2] ATM Forum Traffic Management, *The ATM Forum Traffic Management Specification Version 4.0*, Apr. 1996; available as ftp://ftp.atmforum.com/pub/approved-specs/af-tm-0056.000.ps.

### Biography

SHIVKUMAR KALYANARAMAN (shivkuma@ecse.rpi.edu) is an assistant professor at the Department of Electrical, Computer and Systems Engineering at Rensselaer Polytechnic Institute in Troy, NY. He received a B.Tech degree from the Indian Institute of Technology, Madras, India in July 1993, followed by M.S. and and Ph.D. degrees in Computer and Information Sciences at the Ohio State University in 1994 and 1997, respectively. His research interests include multimedia networking, traffic management in ATM networks and Internet, Internet pricing, and performance analysis of distributed systems. He is a co-inventor in two patents (the ERICA and OSU schemes for ATM traffic management), and has co-authored several papers and ATM forum contributions in the field of ATM traffic management. He is a member of IEEE-CS and ACM. His homepage is: http://www.ecse.rpi.edu/Homepages/shivkuma