Report and Discussion on the IEEE ComSoc TCGN Gigabit Networking Workshop 1995

James P. G. Sterbenz, Henning G. Schulzrinne, and Joseph D. Touch

Abstract

The strong message provided by the contributions to this year's workshop is that a number of researchers are concerned with the practical implications of running TCP/IP over ATM, and most of those dealing with higher layer issues use TCP/IP over ATM out of expediency.

(This an excerpt from the complete paper.)

4 Discussion Session – ATM vs. IP?

This discussion session was intended to consider not only the relative merits of ATM/B-ISDN vs. IP as a network (layer 3) and internetwork (layer 3.5) infrastructure, but also since there were so many ATM contributions in the workshop, we decided to have Henning Schulzrinne presenting an IP-centric critique of ATM to initiate the discussion.

4.1 ATM: Dangerous at any speed? (Position Paper) Henning Schulzrinne

Since ATM has been the object of much hyperbole as ”the” or even ”the one-and-only” future network technology, it is tempting to summarize some of the open issues and, more importantly, principal limitations of the technology. Surprisingly, there has been little published in the technical literature concerning the demerits and problems of ATM. Some of the problem areas are shared with its two older siblings: X.25 and ISDN (Q.931) signaling.

In the following, we summarize some of the issues that may interfere with widespread deployment, particularly in heterogeneous nets owned by different service providers. Using ATM permanent virtual circuits as links between routers is proven technology and poses fewer challenges, except in the area of traffic integration. As with any network technology, ATM is evolving, so that some of the reservations may be moot in future revisions and capability sets. Indeed, important changes in ATM perception and technology have already taken place. To name but a few, the reduction from four to two adapation layers (AAL1 and AAL5), the use of ATM over ever lower bit rates, the lack of use for the generic flow control (GFC) header field, the rare use of SMDS (switched multimegabit data service) over DQDB (distributed queue dual bus) multiple access channels and the abandonment of service classes for more closely defined QOS parameters. But the evolvability of ATM certainly also applies to missing features of IP and its implementation in high-speed networks.

4.1.1 Architectural Issues for ATM "internetworks"

ATM contributes to a continuing confusion over the network model. Together with the AAL, it is not simply a data link layer, and offers important network layer functionality like routing and global addressing, but it is far from clear whether it can be carried efficiently over as wide a range of media as "traditional" network layer protocols like IP, without sacrificing important data link-layer functionality such as "anonymous" multicast. Running ATM cells over important local area networks like Ethernet is feasible, but the architectural and efficiency implications must be carefully examined.

The division of "public" networks and customer premise equipment (CPE) reflects traditional telecom views and is often artificial. For example, what may appear as a UNI interface to the carrier is actually a network-to-network interface to the local ATM cloud. Is the distinction really necessary (the Internet lives without it)?

The idea of connectionless service as espoused by the ITU-T (International Telecommunication Union - Telecommunication Standardization Sector) seems to merely duplicate the functionality of a classical IP network. Since applications are unlikely to use the ATM connectionless service directly, adding it on top of ATM merely adds it and an encapsulation layer between the network and the application, reducing efficiency and making management more difficult. Thus, carriers might be better off offering private IP networks to their customers. In the local area, the benefits of LAN emulation over, say, switched 10 or 100 Mbps Ethernet seem minor, but the complexity in specification and implementation need to be carefully weighed.

4.1.2 Cell Format

The earliest objections to ATM (Cidon 1992) centered around the small cell size. Overhead is a particular concern for low-speed links and small IP packets; for example, packet voice over IP over ATM wastes about 40% on packet headers, not counting SONET overhead. Furthermore (Sterbenz 1994) points out that the small cell size not only limits speed from below, but also from above, as cell switching times approach nanoseconds. Chopping packets into cells also causes loss of a whole packet for each cell loss, potentially leading to a form of congestion collapse due to retransmissions. The partial packet discard and early packet dropping mechanisms suggested by Romanow and Floyd (Romanow 1994) cure this problem, but imply that cell switching has to be aware of the AAL type. Also, packet-aware mechanisms are feasible only if the number of different AALs remains very small. This, and the desirability to implement packet reassembly in network adapter hardware, is likely to limit ATM AALs to AAL5, with niche uses for AAL 3/4 and AAL1. The original intention for ATM was to have extremely low packet loss due to both the use of fiber and careful traffic control. Both sources of loss would appear to be higher for ATM as a universal packet layer, running over less reliabale media and handling more bursty data. It is unclear wheter the cell size issue can be reconsidered, since existing switches are hard wired for the current size (but if new switches were to also support larger cells connections using only these switches could use larger cells).

4.1.3 Connections and Signaling

Due to the small cell size, ATM can only operate in a connection-oriented mode, where signaling establishes the path for data cells switched by short VC (virtual channel) identifiers. The performance and features of the network are largely determined by those of the signaling protocol.

The currently standardized signaling protocol (Q.2931) is extremely heavy weight, with numerous timers, protocol states and a difficult-to-parse packet structure featuring BCD numbers and 7-bit bytes. The several round-trip times and the processing costs impose a heavy delay burden on short-lived associations like RPC (remote procedure call) and query-response applications such as HTTP (hypertext transfer protocol). The overhead gets worse if one or more name resolutions (URN to URL) are needed in the context of the World-Wide-Web and other applications that we expect to inherent this naming and location model.

Currently, there is no security on call establishment, unless you trust all network operators along the path. There is also no provision for per-AAL-packet authentication, thus "stealing" VCs is possible even with authenticated call setup.

Just like AAL5, as a "simple and efficient adaptation layer" (SEAL), more or less replaced the more cumbersome AAL3/4, a lightweight signaling protocol, unifying UNI and NNI protocols and operating within a single round-trip time is called for. A processing-efficient signaling protocol is also important if end systems dumber than workstations are to be directly connected to ATM networks, e.g., as suggested by the concept of desk-area networks.

The limited number of VCs available (and thus a likely charge for holding VC even without generating traffic), both due to cell header and ATM switch implementation restrictions, force relatively frequent signaling within a single upper-layer (e.g., TCP) connection. Some such connections can last days, and can be inactive for most of the time. Even with a timeout of 64 seconds, a router sees 2000 flows. Typical current ATM switches have a capacity of 4096 VCs (virtual connections), which seems inadequate. With IPng (IP next generation) flow labels, similar flows can be aggregated; there is no mechanism yet to aggregate VCs into VPs (virtual paths). If VCs are charged for setup and simply for being there, applications will have to learn subtle optimization strategies, depending on the current tariff structure, such as trying to guess when it is better to drop a VC and reconnect later. In contrast, with TCP, transient network failures can be hidden as long as the end system stays up. With ATM, applications may have to reconnect on VC failure. After a backhoe fade, this would yield massive signaling load with all applications trying to reestablish connectivity ("redialing"). Probing packets (a la TCP) is a lot simpler than trying to decide whether to abandon a connection. (To avoid ATM layer signaling for local failures, lower layers like SONET or fiber switching are likely to handle fiber cuts.)

It is also not clear what happens if an end system or application crashes without cleaning up connections. Will this have the same undesirable consequences as leaving the phone off the hook when calling New Zealand from Europe? What happens to resource reservations (and charges) if the access link is down, so that the end system cannot tear down a connection?

Unlike other technologies such as Ethernet, Tokenring or FDDI, ATM is point-to-point only and requires additional signaling support for multicast. Currently, only root-initiated joins are possible; even with leaf-initiated joins in the ATM Forum UNI 4.0 specification, explicit knowledge and management of source and sinks is still required, with the attendant management and rendezvous problems. In ATM, joining a 1,000-member group takes 1,000 ADD PARTY messages and responses. For example, for a distant learning application, the teacher needs to find all student receivers, where the number of students is dynamic. Having a student ask a question could require all other students setting up connections to the interlocutor. Contrast that with IP multicast, where senders do not need to know the identity of potential receivers and receivers specify a single group address to receive packets from any number of senders. Also, for IP multicast, routers only need to maintain state per group, while ATM requires per-sender state.

4.1.4 ATM as part of the Internet

Pure ATM solutions come in two degrees of purity: applications directly on top of the adaptation layer, or end-to-end ATM with a traditional stack (e.g., TCP/IP) on top. Using ATM within an Internet (today's and particularly a more ATM dominated one), poses additional problems, beyond the current packet loss delaying deployment of the ATM-based NAPs (network access points) in the Internet. For example, there now is a separate address space, that of E.164 numbers embedded in NSAPs (network service access points). Previous experience with NSAP assignment has shown severe scaling difficulties with the purely organizational rather than topological assignment of NSAPs. It would be helpful if the management of a separate namespace, above IP numbers, could be avoided by appropriate management of the NSAP space. Otherwise, another layer of global address lookup becomes necessary, yielding a host name - IP address - ATM NSAP chain. If large numbers of local ATM systems are assigned numbers that are not globally and topologically meaningfull, the problems faced by the interconnection of IPX and AppleTalk networks will be repeated.

There are a number of operational and network management considerations, for example, these are no ATM tools like ping or traceroute, and the necessary protocol and switch support for such tools is missing.

Formerly, X.25 showed that interconnecting connection-oriented networks is more difficult and more fragile than connectionless networks. It will remain to be seen whether ATM improves upon this record.

Overall, from a purely technical standpoint, it appears we could get the (relatively) cheap switching of ATM by simply using application-sized frames at the host and on the wire and cells of whatever size within switches. The queueing delays due to long packets are not significant at optical speeds; at lower speeds, ATM overhead becomes non-competitive. Generally, if bandwidth is cheap, ATM multiplexing is not needed since we can simply use fixed-size pipes; if bandwidth is expensive, we may not be able to afford the ATM overhead.

4.2 Discussion

The reason this session was titled "ATM vs. IP?" was to not only discuss whether ATM or IP was a better solution, but to consider whether ATM and IP could rationally co-exist in the future GII (global information infrastructure) as the successor to the current Internet.

The IP-centric view was presented by Henning. We did not specifically have an ATM-centric presentation and had no defenders of "everything over ATM solves the world's problems" (hopefully they were not too intimidated by the IP-centric presentation). The discussion ranged from the IP-centric view to the idea that the IP and ATM communities would merge and provide a common solution (but not necessarily current IP or IPng over current ATM). The presentation by Gurudatta Parulkar and [4.1.3] also present this later view.

There was general agreement that the problems that are being solved are the same, regardless whether it is the IP or ATM communities providing the solution.

Furthermore, there was significant opinion that ATM and IP can't coexist well as currently implemented.

One problem is that both ATM and IP implement network layer functionality, and in a manner that is not terribly compatible.

Note that we are including the AAL, VC routing, and siognaling along with ATM cell relay when considering ATM network layer functionality.

If multiplexing is done in ATM, it can't be accessed in IP, and vice-versa. The same is true for signalling.

The control mechanisms of ATM have not been designed for interaction with TCP (and of course TCP and most of its optimizations predates ATM). Examples of this problem are given in the presentation by Mishra and in [4.1.2].

The issue of connectionless (IP) vs. connection oriented (ATM) is generally regarded as one of the most fundamental differences. We generally agree that both types of services need to be available to the end user, and is it possible to offer connections over a connectionless infrastructure (e.g. TCP over IP) and datagrams over a connection oriented infrastructure (as ABR traffic will do in ATM).

The real question has to do with performance: How well can datagrams operate over ATM, particularly for transaction type services, such at HTTP in the World-Wide Web (where response to clicking on a link should be sub-second), and of control messages requiring low latency with reliable delivery, such as for session control (higher level than ATM signalling) How well can connection oriented multimedia traffic operate over IP, particularly in the presence of congestion? How well will these two traffic types interact over either infrastructure?

We also generally agreed that state is needed in the network either way. In the case of a connectionless IP network, state can be used to to allow reuse between datagrams. In the case of a connection oriented ATM network, state can be used to cache between short-lived connections to provide a datagram-like service. We note that this soft state caching only helps ATM for long streams (connection-like or connection-oriented) and reuse of host pairs, and will be of limited value for World-Wide Web browsing, particularly where the ratio of traversed links per host is small.

Is is clear that IP provides a useful operational infrastructure for higher layer protocols and applications, whatever its shortcomings may be for emerging connection-oriented applications. Similarly, ATM provides a high data rate switched infrastructure that is being widely deployed, even in the face of technical flaws and details that have not yet been resolved in signaling, traffic management and network management. Thus in the near term, the evolution of the Internet appears to be in the direction of IPng over an increasing number of ATM subnets. There was far less agreement on how this will evolve in the long term, and how this will evolve in the long term, and how the telephone, CATV, and entertainment broadcast networks will be integrated into a single infrastructure.

References

[4.1.1] Cidon, Israel and Derby, Jeff and Gopal, Inder and Kadaba, Bharath, “A critique of ATM from a data communications perspective”, *Journal of High Speed Networks*, 1 (4), 1992, pp. 315–336.

Romanow, A. and Floyd, S., “The Dynamics of TCP Traffic over ATM Networks”, Sigcomm’94, London, p. 79-88.

[4.1.2] Kjersti Moldeklev and Per Gunningberg, “Deadlock Situations in TCP over ATM”, in *Protocols for High Speed Networks IV*, Gerald Neufeld and Mabo Ito (ed.), Chapman & Hall, London, 1994, pp. 243–259.

[4.1.3] James P.G. Sterbenz, ”Protocols for High Speed Networks: Life After ATM?”, in *Protocols for High Speed Networks IV*, Gerald Neufeld and Mabo Ito (ed.), Chapman & Hall, London, 1994, pp. 3–18; also available as http://info.gte.com/jpgs/paper/pfhsnIV.html