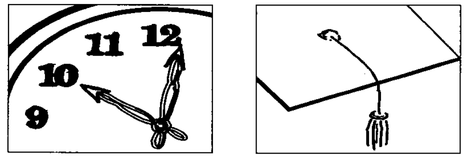


Data Communication in ATM Networks

An increasing number of customers require LAN access with high bandwidth and low delay over long distances. To satisfy these needs, several high-speed network techniques have been developed recently. ATM is superior compared to other networking technologies, as it offers high bandwidth and is scalable in the sense that the bandwidth capacity of an ATM system is not fundamentally limited to the technology itself.

Nail Kavak

Although emerging high bandwidth multimedia applications are one of the major driving forces behind the B-ISDN initiative, its success or failure largely depends on how well currently used data communication applications and protocols can be supported in these networks. This article analyzes architectures and protocols specifically designed for providing data communication services in ATM LANs and ATM wide area networks. Current trends and activities within standard bodies and forums are outlined, the merits and weaknesses of various solutions are given, and the issues that still need to be addressed are discussed.

An increasing number of customers require LAN access with high bandwidth and low delay over long distances. To satisfy these needs, several high-speed network techniques have been developed recently, offering data rates up to hundreds of Mb/s. Typical examples are FDDI, Frame relay, Fast Ethernet, Ether switch, and, most recently, Asynchronous Transfer Mode (ATM), all of which are designed to satisfy connectivity requirements emerging from current LAN, as well as high-speed workstations in order to run powerful distributed multimedia applications.

Much interest has been expressed in ATM, due to its flexibility and support of multimedia traffic. Although ATM initially targeted wide area networks, interest today is growing from LAN networking vendors. ATM is superior compared to other networking technologies, as it offers high bandwidth and it is scalable in the sense that the bandwidth capacity of an ATM system is not fundamentally limited to the technology itself. Furthermore, it can support multimedia traffic offering seamless integration with wide area ATM networks, both public and private.

Initial ATM installations will operate as subnetworks of existing networks and MAC layer protocols. One of the main challenges in ATM is the transparent support of existing connectionless LAN services. Recently, several activities have been launched within international standard bodies and forums to specify ways of providing data communication services over ATM. Most notable examples are Switched Multimegabit Data Service (SMDS) [1] and the similar Connectionless Broadband Data Service (CBDS) supported mostly by public network service providers [2]. But also other approaches such as IP over ATM, and LAN emulation [3, 4] that show more adherence to the existing local and campus area networking paradigms.

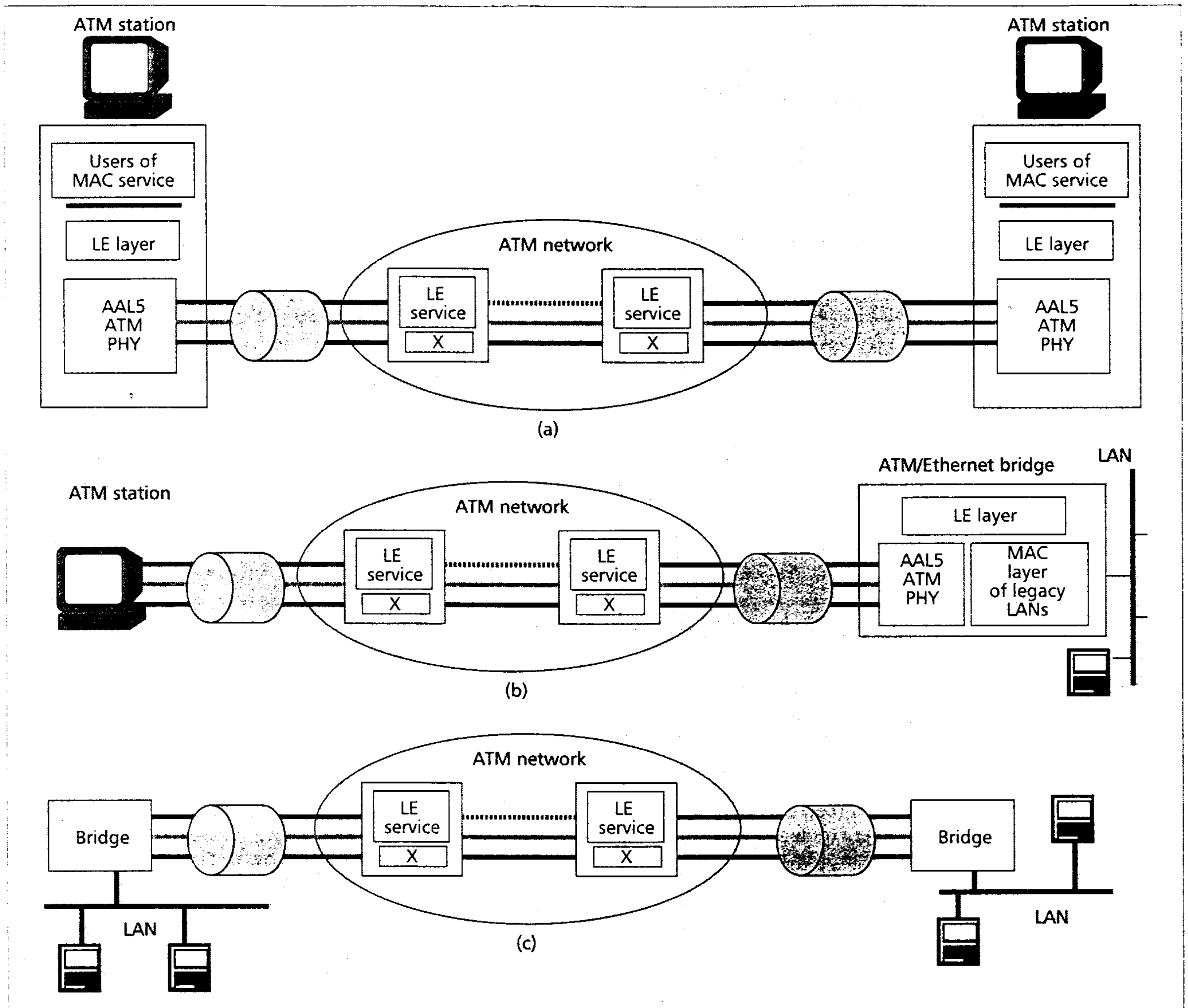
The organization of this article is as follows: the first section presents the requirements and architecture of the LAN emulation service. The following section describes the alternative methods for carrying IP packets over ATM. Public broadband service architecture and CBDS are the subject of the next section. The traffic management aspects of the data communication services are discussed following that. At the end of each section, the strength and weakness of each described solution is given. Conclusions are drawn in the final section.

LAN Emulation

While it may be foreseen that ATM (and multimedia technology in general) will drive the development of new networking interfaces, it will still be required to support the existing networking interfaces. The majority of installed protocol stacks rely on facilities provided by today's LANs. In order to use the current base of existing LAN applications, it is necessary to define an ATM service, which emulates services of existing LANs on an ATM network without the need of any change in the ATM terminal equipment's interface to the MAC layer. The LAN

NAIL KAVAK is with Telia/Telia Research.

0890-8044/95/$04.00 1995 © IEEE



■ Figure 1. *a) ATM-ATM interworking; b) ATM-LAN interworking; c) LAN-LAN interconnection.*

Emulation (LE) service specified by the LE-SWG of the ATM Forum is exactly designed to meet this requirement.

Providing the LE service at the MAC layer facilitates transparent support of many protocol stacks including IP, SNA/APPN, IPX, NetBios, and AppleTalk. Alternative approaches, e.g., IP layer emulation, restrict support to only the IP protocol. Nevertheless, existing LANs differ radically from ATM in several important aspects [1]. For example, ATM networks are connection-oriented in the sense that data transfer is preceeded by a setup phase and succeeded by a release phase. LANs are connectionless, i.e., data is transferred at any time with no prior explicit warning.

Also, in LANs broadcast and multicast are easily achieved through the shared medium. Every packet, whether unicast, multicast, or broadcast, is broadcast to all stations which reside on the shared medium, and each station filters out the packets it wants to receive. In ATM, such a mechanism would lead to inefficient use of network resources. For example, multicast packets can be exchanged directly between communicating parties, without relying on broadcast filtering.

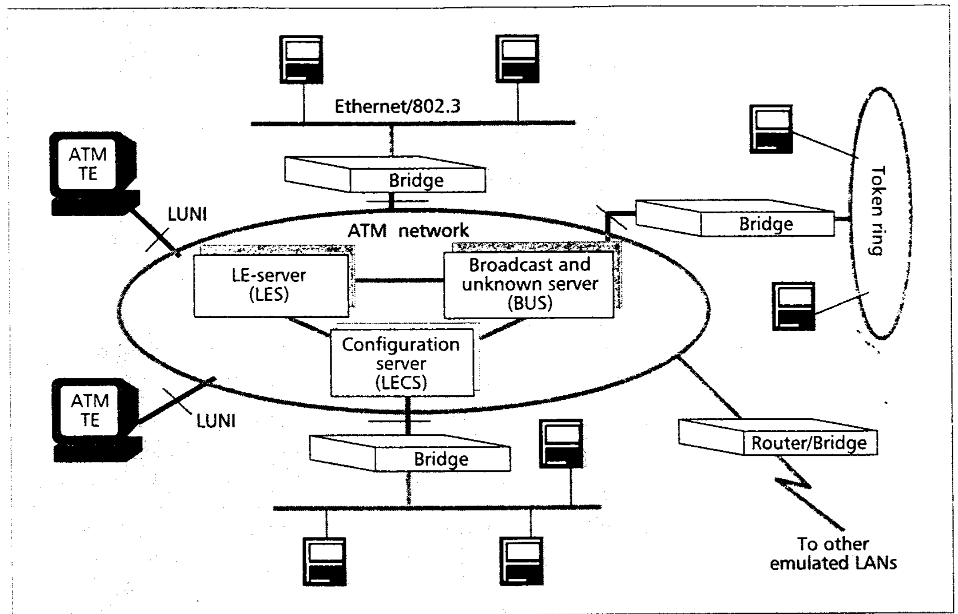
Furthermore, the address format supported by current LANs is based on manufacturing serial numbers and does not reflect the network topology, whereas ATM addresses are hierarchical. Since LAN and ATM addresses have different formats, address resolution functionality is required to associate LAN addresses with ATM addresses.

Another major problem associated with the provision of LAN services is that although the "normal" ATM connection set-up procedure requires the specification of user traffic characteristics, the vast majority of LAN applications are incapable of predicting their own bandwidth requirements in advance. Hence, an explicit guarantee of service cannot be given. In essence, the LE service needs to perform the functions that are required by traditional LANs but not directly supported by the ATM network.

Architectural Issues

The LE service [5] is exclusively designed to support three configuration scenarios; ATM-ATM

To avoid clients abusing the broadcast channel, a client is allowed to send only a limited number of broadcast frames within a given time.



■ **Figure 2.** *An example configuration of emulated LANs.*

interworking (Fig. 1a), ATM-LAN interworking (Fig. 1b), and LAN-LAN interconnection (Fig. 1c). Hence, the interconnection of existing LAN applications across an ATM backbone to other end-user systems, both those which are ATM-attached (servers, high-end systems) and those which are on legacy LANs, is possible by means of bridging methods. The aim is to enable ATM to be used seamlessly as a backbone technology for existing legacy LAN technologies. This will enable the end users to take advantage of the features that ATM offers as backbone technology, while offering them a migration path, to take advantage of native ATM facilities in the future.

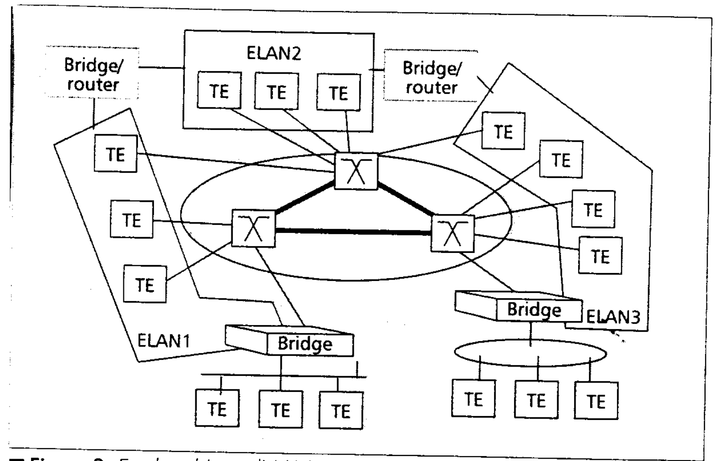
The LE service architecture is based on a client-server (query-response) model. The components of an emulated LAN include (Fig. 2) ATM workstations and ATM/LAN bridges (i.e., clients). The components of the LE service include a LE server (LES), a LE configuration server (LECS), and a broadcast and unknown server (BUS). However, the LE architecture does not imply any particular implementation. In fact, any of the LE service components can be implemented distributed (for reliability or performance reasons) or centralized. All components may even be collapsed into a single physical entity (e.g., for economical reasons). To enable such an implementation flexibility, a number of virtual channels (VCCs) are defined for the communication between LE clients and LE components. Normally, clients use control channels (e.g., configuration-direct VCC, control-direct VCC, etc.) for sending/receiving control messages to/from LE components, and data channels (data-direct VCC, multicast-send VCC) for sending/receiving user data only.

The LES provides a facility for registration and resolving MAC addresses into ATM addresses. The LECS is used for locating the LES and obtaining configuration information for each ATM segment. The BUS is mainly used for forwarding multicast/broadcast frames, but also for delivering unicast frames targeted to the unregistered LAN stations, for which addresses cannot be resolved yet.

The integration of LANs with ATM technology offers significant benefits of virtual networking. Virtual networking means complete separation of the physical and logical network infrastructure. ATM LANs can be virtually segmented into multiple segments which can be organized along administrative boundaries providing increased security and scalability (Fig. 3).

Connectivity between partitioned virtual ATM segments (emulated LANs) can be provided by means of bridging or routing methods. Unlike traditional LANs, membership in an emulated LAN is characterized logically rather than physically. This offers increased flexibility in terms of terminal mobility and network management [6]. For example, a client may be a member of more than one emulated LAN via a single physical attachment, and it can remain a member of the same emulated LAN, even if it moves from one physical location to another in the ATM network. Logically, clients belonging to the same emulated LAN would be connected to the same logical LES and BUS pair and each emulated LAN with its own LES and BUS would provide service of either Ethernet/802.3 or Token ring segment. It is allowed to have multiple physical LES/BUS pairs within a single emulated LAN (e.g., one LES/BUS pair per switch), in which case interactions between LES/BUS pairs would need further specification, if multivendor interoperability is required. It is important to note, though, that broadcast frames stay within the boundaries of an emulated LAN, except when the emulated LANs are interconnected via bridges or routers. The current ATM Forum specification covers only the interfaces and functions of a single emulated LAN.

Within ATM stations, the LE service is provided by an LE layer. The LE layer shields the higher layer protocol stacks from the characteristics of the ATM network and gives them the illusion of being attached directly to a traditional LAN. The LE service provides functions related to initialization, registration, address resolution, and forwarding of unicast or multicast frames. When designing the overall architecture, attempts have been made to cleanly separate control and data transfer paths. The primary goal has been to use the LE service mainly for control functions using the control channels whereas, the data would be transferred via data channels transparent to the LE service. This goal, as will be seen in the following sections, is sometimes violated in order to achieve better throughput.



■ **Figure 3.** *Emulated (virtual) LANs.*

LAN Emulation Service Functions

Initialization — When an ATM station is plugged into a switch port, it executes initialization, joining and registration procedures in order to have access to the LE service. At the end of this phase, the LE client should have initialized some parameters (e.g., the emulated LAN type, client id, connection time-out period, aging time, forward delay time, etc.) and configured its own MAC and ATM address. Bridges can choose to register all LAN stations behind them (non-proxy bridge) or flood all frames it receives to the LAN segment to which it is connected (proxy bridge).

Successive to the initialization phase, a client will contact the LECS to locate the LES. Assuming an ATM connection is established to the LES (if it is not already preconfigured), clients need to go through join, initial registration procedures and connection establishment to the BUS phase before they can perform LE service specific functions (i.e., they can ARP or send multicast frames to the BUS).

Address Resolution and Unicast — In order to set up ATM connections, the clients need to maintain information about MAC to ATM address mappings. In the absence of such information, the address resolution function provides a method that allows a client to resolve a MAC address into an ATM address using LES. The LES constructs a MAC to ATM lookup table upon the registration of the clients, and this table is updated every time a client joins or leaves the network.

In order to send a MAC frame, the client needs to find out which VPI/VCI to send the frame on. If such an entry does not exist, the client forwards a request to the LES to get the ATM address for a given MAC address, since a client may initially not have any information about the MAC to ATM address mapping. The LES may reply directly to this request, on behalf of the registered clients, or it may choose to broadcast the request to other clients via point-to-multipoint connection, with the LES as root and the clients as leaves. The reply is also sent to the client via the LES.

While the address resolution progresses, the client may also send data frames to the BUS, which will flood "unknown" unicast frames to all clients. The rationale behind performing data forwarding in parallel to address resolution is the follwoing: first, without such a mechanism, unregistered LAN stations would never be reachable; second, some LAN applications are delay sensitive and may not operate properly, due to delays caused by address resolution and subsequent connection set-up procedures.

To avoid clients abusing the broadcast channel, a client is allowed to send only a limited number of broadcast frames within a given time. While such a limitation may be useful, the current ATM Forum specification does not make it clear as to how the BUS polices "unknown" frames.

When forwarding unknown frames, the BUS can forward an unknown frame selectively to the target client (assuming the client is registered) instead of flooding to all clients. Selective forwarding would thus reduce the total amount of traffic within the network and the load on the clients. The necessity of intelligent BUS behavior is still disputed. The argument against introducing such intelligence is that this would complicate and make the implementation of the BUS and the clients more expensive. Besides, the intelligent BUS provides negligible improvements in efficiency because the unknown traffic sent by a client to the BUS is insignificant, and most traffic flows over the direct data path between the clients.

Unicast/Multicast/Broadcast — Once the MAC address is resolved into an ATM address, the client initiates the connection set-up procedure. Upon the completion of this phase, unicast frames can be transferred directly to the receiver client, transparent to the server(s). Connections are torn down after a preconfigured idle time. Clients can cache the ATM connections in use, assuming that future communications to previous destinations are likely to arise.

The data frame format for an Ethernet/802.3 emulated LAN differs from a Token-ring emulated LAN. A common feature is the absence of checksum fields. The advantage of having two different frame formats for each emulated LAN is that it simplifies bridge implementation. This is because, due to logical differentiation between emulated LANs, bridges within a single emulated LAN are not required to perform

While LAN emulation may suffice for workgroups, the interconnection of large scale LAN and WAN networks across ATM will require the development of native mode network layer protocols.

conversion between Token ring and Ethernet/802.3 frames. A conversion mechanism would be necessary if two different emulated LANs are interconnected via a bridge. Such a mechanism is currently beyond the scope of the LAN emulation service.

The procedure for sending and receiving multicast frames is slightly different from unicast. The client will initially request the LES to return the ATM address of the BUS. The client will then establish multicast connection to the BUS in the sending direction. The BUS will, in turn, set up a multicast connection to the client on the return path to enable the client to receive multicast traffic. The BUS may forward the multicast frames either over point-to-point or point-to-multipoint connections.[1] The point-to-point connections will be used mostly in environments where the ATM network does not support a point-to-multipoint connection capability.

In any event, the BUS has to reassemble and sequence multicast frames before forwarding them. A client can identify and filter broadcast/multicast frames by comparing with its own LE client-id which is obtained during the join or registration phase.

Frame Sequence Integrity — Most LAN applications rely on in-sequence delivery of frames. In an emulated LAN, clients are allowed to transfer frames through the BUS prior to establishing a direct data path, and when such a direct connection is established, the client may end up having two data paths, one via the BUS and one direct path. Switching between these paths introduces the possibility that frames may be delivered out of order. The problem of preserving the order of transmission of unicast frames across an emulated LAN has been solved by the flush protocol. The algorithm applied for the flush protocol is rather simple. When switching between the two paths, the sender client transmits a flush message down the old path and holds the frames for the given LAN destination until the receiving client transmits an acknowledgment that the old path is cleared and the new one is ready to use. Upon receiving the acknowledgment, the sending client dumps all frames that were held down to the new path.

There are some other alternative methods to deal with the out-of delivery of frames. One is to let the receiving client accept frames only from the direct path and not from the BUS, which will result loss of initial frames. Another way is to let the sending client wait for some period of time for address resolution to work before utilizing the BUS. In this way, out-of-order frames can be minimized at the cost of some delay. As a matter of fact, the flush protocol itself introduces some delay due to buffering of frames while waiting for acknowledgment from the receiving client.

Discussion

The current LAN emulation architecture lacks robustness. In particular, using a single BUS for the distribution of all broadcast and multicast frames presents a single point of failure. A failure in the operation of the BUS will lead to interruption of the service for all clients belonging to an emulated LAN. Future work needs to investigate possibilities for introducing redundancy and hence improved reliability in BUS implementations.

Furthermore, the current LE service does not specify interaction between LE service components (LNNI). The standardization of LNNI is needed to realize LE service distributed across mulivendor platforms. LAN emulation is well suited for small workgroups, due to its protocol independence and high speed. Multiple emulated LANs can be interconnected through traditional bridging, routing or brouting techniques. However, the inherent and well known limitations of bridging suggest that it would be impractical and imprudent to extend bridged virtual LANs beyond a local area or small number of workgroups. Interconnecting multiple emulated LANs through a wide area ATM network by means of traditional bridging techniques would also lead to performance limitations due to increasing broadcast traffic [7]. In order to reduce the broadcast traffic, a router can be used instead to filter out unnecessary traffic. Alternatively, multiple emulated LANs can be interconnected by means of direct ATM connections. The main advantage would then be the reduction of interconnection devices (e.g., bridges, routers) and thereby improved performance.

IP over ATM

While LAN emulation may suffice for workgroups, the interconnection of large scale LAN and WAN networks across ATM will require the development of native mode network layer protocols. This is essential, for instance, for the operation of ATM routers that will be used as the primary mechanism for interconnecting current LANs and WANs across ATM backbones, as customer premises networks evolve toward the widespread use of ATM. The early work on IP over ATM standards was largely done by the IP over ATM working group of the IETF. So far, this work comprises encapsulation methods [8], default maximum transmission unit, and an address resolution method within a logical subnet [9]. There are also working drafts specifying establishment of on demand connections [10], and IP multicast over User Network Interface (UNI) 3.0 based ATM networks [11]. Most of the work done so far does not attempt to change the fundamental nature of the IP protocol (hence the name "classical IP over ATM") [9]. However, more radical architectures are under discussion which cast out the traditional subnet architectures in favor of the ATM Internet model [12, 13]. The new model is called the peer model which means that ATM network and IP routers are considered as peers (i.e., they exchange routing information between them).

Classical IP over ATM

In the classical model, a logical IP subnetwork (LIS) consists of hosts and routers having the same subnetwork address and netmask. Hosts connected to the same subnet (i.e., LIS) communicate directly. However, communication between two hosts on different LIS is only pos-

[1] *A point-to-multipoint connection has a root and leaf nodes consisting of a tree topology, though the return paths of the VCs have zero bandwidths.*

sible through an IP router, regardless of whether direct ATM connectivity is possible between these two hosts.

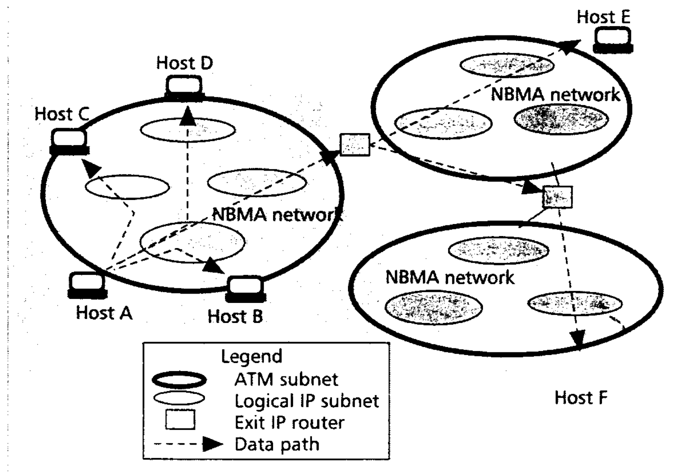
Implementing IP over ATM will require mapping between IP and ATM addresses. In the classical model, IP addresses are resolved to ATM via ARP and vice versa via InARP, within a subnet. Initially, hosts are required to register their own addresses to an ARP server which is connected on a well known channel in the ATM subnetwork. The ARP server uses InARP to determine the IP and corresponding ATM address of the connected host. Thereafter, hosts can query the ARP server to get the ATM address of a given IP address. An ARP server only resolves IP host addresses connected to the same subnet.

In order to encapsulate different network layer protocols, hosts can apply two different encapsulation techniques [8]. The first method allows multiplexing of multiple protocols over a single ATM channel. The carried protocol is identified by prefixing a LLC/SNAP header. The second method does higher layer multiplexing implicitly by assigning different channels for different protocols. Both methods are functionally equivalent. The LLC/SNAP-based encapsulation is suitable if the ATM network only supports semi-permanent ATM connections. Multiplexing per channel would be more suitable in those environments where dynamic creation of large numbers of ATM channels is fast and economical.

Non-Broadcast Multi-Access (NBMA) Networks

Although the classical IP over ATM is conceptually very simple and does not require any changes to existing systems, it is very limited, since communication between different subnets must occur through a router. This is a significant limitation particularly for an ATM based network in which many subnets can be defined. In ATM, hosts connected to the same network are capable of communicating directly, without IP layer switching by routers. This presents an opportunity to optimize performance and perhaps lower costs by eliminating unnecessary hops through the medium.

In order to overcome this limitation, the Routing Over Larger Clouds (ROLC) working group at the IETF has been investigating the possibility to set up direct connections across non-broadcast multi access (NBMA) networks such as ATM [14]. With this aim, the ROLC working group is working on a new protocol named "Next Hop Resolution Protocol" (NHRP) which relies on the use of "super" ARP servers. The ultimate goal of the NHRP protocol is to enable a host to bypass some or all of the routers between it and a given destination host by establishing a direct connection through the ATM fabric. In this protocol, disjoint IP subnets are treated as one logical network, called NBMA network. For every NBMA, there is at least one NBMA-server which resolves IP addresses to NBMA addresses. A NBMA-server constructs the IP-NBMA bindings based on the terminal's NHRP registration packets or through dynamic address learning mechanisms. If there are multiple NBMA-servers within one NBMA network, the servers resolve the addresses cooperatively using, for example, the traditional server to server protocols (e.g., OSPF, GGP). NBMA networks are, in turn, interconnected via IP routers (Fig. 4).



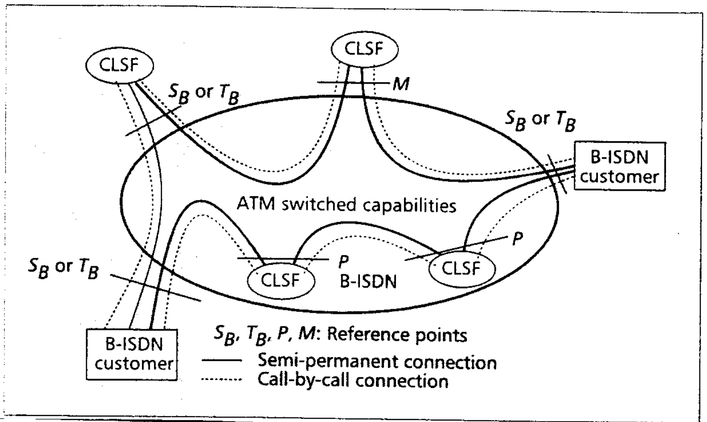
■ **Figure 4.** *NBMA network architecture.*

The address resolution protocol proceeds in the following way. If the source and destination host are connected to the same subnet, address resolution is performed by conventional means using ARP or pre-configured tables. If the destination host belongs to another subnetwork, an ARP request is sent to the NBMA-server. If the address entry is found, the server returns the ATM address of the destination host. Otherwise, the server forwards the query to the next hop towards the destination. The next hop may itself be the exit router through which packets for destination is forwarded. A negative reply is sent back to the source, if the destination cannot be found within the NBMA network. It is important to emphasize that NHRP requests never cross the border of a logical NBMA network.

In order to increase the performance of the protocol, responses are also cached in each node on the return path. Hosts may even transfer data while address resolution progresses. Once a NBMA address is resolved, hosts can initiate the connection establishment phase indicating the required quality of service (QoS).

Discussion

The Internet protocol family is one of the most important network protocols for ATM to support. The classical IP over ATM, being designed to allow rapid implementations, has a number of weaknesses. First, end-to-end QoS cannot be guaranteed as data traverses through several IP routers which do not allocate resources. Second, it does not scale well to the large size of the eventual ATM network due to large number of routers. Although NBMA server(s) reduces the number of hops, it may increase response times up to a round-trip time, which may be critical for some applications. Besides, NBMA servers add complexity for network management, and



■ **Figure 5.** *Indirect and direct conectionless service reference model.*

introduce more network failure points. In fact, NHRP protocol may even cause undetected loops, since the path from which the routing information is received is separated from the path the data is transferred over. As of the time of this writing, this problem has not been fully resolved.

In general, current IP over ATM solutions obscure the full power of ATM, including scalability, performance, and guaranteed QoS. Ideally, the Internet and ATM technology should be more tightly coupled so that direct ATM connectivity can be provided between two endpoints without intervening external servers or routers. To this end, the future version of ATM Internets need to address appropriate (or even the elimination of) address resolution techniques, mapping of Internet multicasting with ATM multicasting, security issues, QoS connections, and coherent routing policy in both ATM and Internet segments. In fact, the success of Internet in ATM depends largely on how well these two networking services can be integrated.

Connectionless Broadband Data Service (CBDS)

Unlike the previous architectures, ITU's connectionless service is provided at the user-network interface. The ITU-TSS draft recommendation I.364 identifies two different configurations [15] as depicted in Fig. 5, indirect and direct.

Architecture

In the indirect method, the Connectionless (CL) service together with associated adaptation layer (AAL) entities are installed outside B-ISDN. A transparent ATM connection, either permanent, reserved, or on demand, is used between B-ISDN interfaces. All connectionless-service-related functions are performed outside B-ISDN. This operation mode may be suitable in case the network does not support signaling and when the number of connection end points are small enough to allow for a fully interconnected mesh [16].

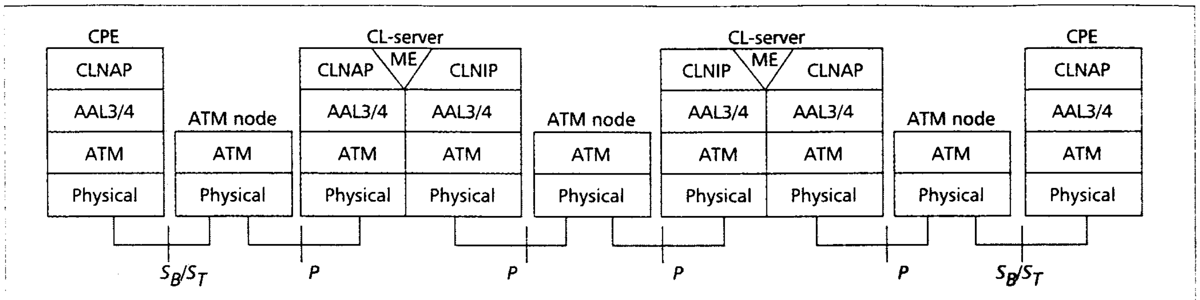
In the direct method, a Connectionless Service Function (CLSF) is installed within B-ISDN. The CLSF terminates CL protocols and routes CL-packets to their destinations according to routing information included in CL-PDUs. The CL-servers can be installed at various ATM switches as well as ATM cross-connect nodes. The connectionless user sends data to a well known CL-server, e.g., the closest one, to which it either signals or has a semi-permanent connection already established. The CL-server then forwards the data to the destination user possibly via a route of other CL-servers. The CL-servers may be interconnected by a "virtual overlay network" consisting of several Virtual Paths (VP) with preallocated bandwidth resources. The choice of PVC or SVC depends on the user traffic characteristics and QoS requirements. The use of CL-servers within the ATM network will lead to a reduction of the number of VPs needed (as compared to the full VP mesh) and thus to a concentration of connectionless traffic on fewer VPs. By statistically multiplexing several sources on the same VP, burstiness can be reduced. Furthermore, the number of connections each end point needs to setup is reduced to only one.

Figure 6 depicts the protocol architecture of the CL-servers at the UNI as well as between the network nodes (NNI).

The CLSF terminates the B-ISDN Connectionless Network Access Protocol (CLNAP), which includes functions for the mapping of connectionless protocol onto the connection-oriented ATM service by means of AAL3/4 entities. The CLNAP layer includes, among other things, functions for the routing [17] and addressing of variable length CBDS packets transferred between one source to one or more destinations without the explicit establishment of any ATM connection by the user. It also supports multiprotocol encapsulation and the following QoS parameters: transit delay, cost, and residual error probability. Routing is performed based on the E.164 address information contained in the CLNAP-PDU header. To achieve higher traffic concentration at CLSs, VPs between CLSs can be configured not to form full mesh connectivity. Alternatively, the CL-servers can be interconnected arbitrarily or by means of other topological schemes such as hierarchical tree, bus or ring [1].

The Connectionless Network Interface Protocol (CLNIP) supports the CL-service between CL-servers inside a network operator domain as well as between two network operator domains. CLNIP provides for the transport of both encapsulated and non-encapsulated data units. Interworking functions between the CLNAP and the CLNIP are provided by a Mapping Entity (ME) in CL-servers.

The CLSs can operate in message mode or streaming mode. In the first case, each CBDS packet received is completely reassembled by AAL3/4 and then passed to the CLNAP layer which performs the routing function based on the E.164 address and QoS requirements. The same packet is fragmented again to cells before being forwarded to the next hop. In streaming mode, the AAL3/4 will pass only the first cell of the packet to the CLNAP layer which then establishes an association between destination E.164 address, incoming VPI/VCI, and Multi-



■ **Figure 6.** *Protocol architecture of the connectionless service using CL-servers.*

plexing Identifier (MID) to the outgoing VPI/VCI/MID. All subsequent cells of the same packet will follow the same route.

Obviously, the message mode causes processing delays and requires large reassembly buffers. On the other hand, erroneous packets can be detected when reassembled so that if a single cell is lost, the whole frame can be discarded.

There are two alternatives regarding the placement of CLSF in B-ISDN. One is to implement the CLSF integrated with the ATM node [18]. The other is to implement the CLSF in an unit externally attached to the ATM node [19, 20]. In the integrated approach, incoming connectionless cells which are identified by a special tag (e.g., the CLP bit or a special VPI) will be sent to a switching unit, which transfers cells directly to the outgoing link. The integrated method performs better than the stand-alone method at the cost of less flexibility. The major drawback of having CLSF in an external unit is that every cell has to be switched twice. All connectionless cells will be switched to the CLS, in which the routing decision for the outgoing link will be taken based on destination addresses along with the QoS requirements. The main advantage of the stand-alone approach is that the network designer can manage the CLSs, without intervening in the underlying ATM network. In this way, addition of new CL-service related functions will not cause disruption to "pure" ATM switching services.

Discussion

As noted above, the ITU/ETSI architecture for the direct support of connectionless service makes use of CLSs. The key distinction between LAN emulation and ITU connectionless service architecture is that in the latter every cell has to go through the CLS. This requires the CLSs having high-speed packet (or cell) forwarding capability. Particularly, if the CLS is implemented in an external unit, the CLS may become the performance bottleneck.

The CBDS is based on an overlay architecture. In general, the main drawbacks with the overlay architectures, which both the classical IP over ATM and to some extent the LAN emulation are based on, are the increased network equipment costs and increased complexity in terms of network management, security, and routing.

Furthermore, the CBDS is the only ATM service which is specified on top of AAL3/4. It has been shown, though, that CBDS can be designed more efficiently [19, 20] making use of AAL5 instead. An AAL5-based CBDS service would be interoperable (at least with respect to the AAL layer) with other data communication services such as Frame relay, IP over ATM, or LAN emulation, which are all designed on top of AAL5.

Traffic Management Aspects

The diversity of datacom applications results in a wide range of traffic characteristics. While some applications may be more sensitive to cell delay and cell loss (e.g., video conference) and thus require performance guarantees, others may have loose delay constraints and strict error requirements (e.g., file transfer). The applications requiring QoS guarantees can only be handled with reservation of bandwidth and/or buffer resources, which requires a preliminary declaration of the profile of the source, described in terms of, e.g., peak rate, mean rate, burst size, and negotiation with the Call Admission Control entity. However, most LAN applications generate bursty traffic on every time scale [21], and the application users are incapable of predicting their own bandwidth requirements. For bursty data applications, allocating "fixed" bandwidth in advance might not be the best way to utilize network resources. If fixed bandwidth is allocated, this may cause either underutilization of the allocated capacity, or limitation in the throughput of the LAN traffic by the allocated capacity. An alternative strategy would be to send traffic on a "best effort" basis in which users can transmit data "without" reserving bandwidth, where all users compete on equal terms for the available bandwidth [22-27]. The assumption made here is that the best effort service should not affect the traffic with reserved bandwidth. However, the amount of best effort traffic sent to the network may momentarily exceed the available capacity, which may cause loss of cells. The unique characteristic of the LAN traffic is that large packets (e.g., 9 kB) have to be fragmented into small ATM cells (53 bytes) and the loss of any single cell will cause retransmission of the whole packet (192 cells). The loss of cells and the corollary loss of packets is bad enough, but a more significant impact is that these lost packets will be retransmitted into the congested network, causing more congestion and even more

It is expected that the UBR and ABR service definitions and exact details of the rate based scheme will be consolidated quite soon.

cell loss. In order to reduce the effect of losses, there is a need for the network to indicate the congestion status via feedback mechanism to the source.

For bursty data type traffic, two different ATM service categories are under consideration in the ATM Forum [28], Unspecified Bit Rate (UBR), and Available Bit Rate (ABR) service. Both UBR and ABR service classes are intended for delay-tolerant applications (i.e., those which do not have stringent delay and delay variation requirements). Unlike UBR, the ABR service explicitly supports a feedback mechanism that allows the source end system to adapt (i.e., increase or decrease) its transfer rate to a time varying bandwidth. In that sense, the ABR service is more reliable than UBR. But, on the other hand, it requires more detailed specification of service specific QoS parameters (e.g., cell loss ratio and minimum cell rate).

Originally, two feedback mechanisms were under consideration in the ATM Forum: a rate-based scheme[28], and a credit-based scheme [29]. The Forum has recently adopted the rate-based scheme. The basic idea with this scheme is that the destination and/or the network periodically inform the source of the rate with which it is allowed to send so that the ATM-level of the source adjusts its sending rate in response to congestion indications from the network. The exact details of the rate-based scheme are still under discussion. The common problem with feedback-based schemes is that, on the long distance links, long buffers may be required to absorb bursts. Furthermore, also due to propagation delays, end systems may not receive the congestion indications in time to react.

An alternative to the feedback scheme would be to use adaptive window flow control mechanisms at the transport layer, such as in TCP or TP4, with appropriate packet and window-size settings, supported by sufficiently large buffers in the switches to hold a reasonable number of packets. The main problem with this scheme is correct estimation of the system configuration parameters (i.e., switch buffer size, user packet size, transport protocol window size, etc.). For example, large buffers contribute to end-to-end packet delay which in turn may trigger the retransmission of the whole packet into the congested network. On the other hand, small buffers may seriously penalize because of buffer overflows. Besides, the effects of small buffers and small windows are minimal if the number of contending connections increases. The answer to the optimal buffer size depends on tradeoffs between a number of parameters such as throughput, delay, and cost.

In the event of congestion, the ATM network can apply intelligent cell dropping mechanism to improve the overall throughput. In its simplest form, a selective cell discarding mechanism could be applied such that when a cell is lost, all other cells belonging to the same packet may be discarded as well. This mechanism, however, has a limited effect, since there may be other cells of the same (long) packet already transmitted down the link. An alternative strategy, the so called Early Packet Discard [30] proposes a mechanism whereby the switch drops entire packets whenever proportion of the buffer in use exceeds a fixed threshold (i.e., prior to buffer overflow). In a sense, the ATM switch in this case behaves as a conventional router. This scheme prevents transmission of useless cells on the congested link and reduces the total network load. The difficulty is the determination of the optimal value for the fixed threshold. Also, the fairness is not guaranteed since cells may be discarded that do not belong to the connection that caused the congestion.

It is expected that the UBR and ABR service definitions and exact details of the rate-based scheme will be consolidated quite soon. However, further investigations of cell dropping schemes in different traffic environments and with different feedback-based congestion control mechanisms would be useful. Care should be taken to study the behavior of the end user transport layer protocols under different congestion control (feedback and/or cell dropping) mechanisms.

Conclusions

This article presents data communication service architectures designed for both private and public ATM networks. It is clear from the discussions that different solutions will apply for different networking environments, which will inevitably lead to interoperability problems. Up to now, such problems have not even been addressed in standards bodies. For the architectures described in this article, even though some interim results exist that may lead to short term products, work still needs to be conducted with respect to robustness, architectural scalability, multicasting, optimized routing, and resource and congestion management.

One important design policy in both LAN emulation and IP over ATM was to demonstrate that the introduction of ATM technology into LANs must be an evolutionary process that respects past, current, and future investments. While this may be good first step in ATM deployment, today both of these approaches lack functionalities and means to fully utilize the unique properties of ATM. In the long term, it is essential to develop more powerful architectures which expose ATM to higher layers instead of hiding them. This will also require careful design of migration strategies and graceful transition from the current models.

Acknowledgments

The author would like to thank Kim Laraqui, Steve Pink, Gerald Maguire, Gunnar Almgren, Nils Bjorkman, and Alexander Latour-Henner for their helpful comments and suggestions.

References

[1] B. J. Vickers and T. Suda, "Connectionless Service for Public ATM Networks," *IEEE Commun. Mag.*, vol. 32, no. 8, Aug. 1994, pp. 34-42.

[2] W. Fischer *et al.*, "Data Communications Using ATM: Architecture, Protocols and Resource Management," *IEEE Commun. Mag.*, vol. 32, no. 8, Aug. 1994, pp. 24-33.

[3] N. Kavak *et al.*, "Partial Connections Method for the Support of Ethernet Emulation over ATM" 7th IEEE Workshop on Local and Metropolitan Area Networks, Florida, March 26-29, 1995.

[4] P. Newman, "ATM Local Area Networks," *IEEE Commun. Mag.*, vol.

32, no. 3, March 1994, pp. 86-98.
[5] LAN Emulation SWG Drafting Group, "LAN emulation over ATM – Version 0," ATM Forum/94-0035, March 1995.
[6] B. Li and P. Vankwikelberge, "Virtual LAN configuration and address resolution in an ATM network," Interworking symposium 1994, (IOS Press, pp. 179-190).
[7] K. Laraqui and N. Kavak, "Configuration Management Scheme to Support LAN Emulation Service Over a Public ATM Connectionless Broadband Data Network," RACE/LACE Deliverable: Business CPN Communication, Control, and Management Specification-Part B, Dec. 1994.
[8] J. Heinanen, "RFC-1483: Multiprotocol encapsulation over ATM adaptation layer 5," July 1993.
[9] M. Laubach, "Classical IP and ARP over ATM," RFC 1577, April 1994.
[10] M. Perez *et al.*, "ATM Signalling Support for IP Over ATM," Internet draft, April 1994.
[11] G. Armitage, "IP Multicast Over UNI 3.0 Based ATM Networks," Internet draft, Nov. 1994.
[12] B. Braden, J. Postel, and Y. Rekhter, "Internet Architecture Extensions for Shared Media," Internet draft, May 1994.
[13] R. Cole, "IP over ATM: A Framework Document," Internet draft, Jan. 1994.
[14] D. Katz and D. Piscitello, "NBMA Next Hop Resolution Protocol," Internet draft, Nov. 1994.
[15] ITU SG13 Draft Recommendation I.364, "Support of Broadband Connectionless Data Service on B-ISDN," July 1993.
[16] N. Kavak, "Interconnecting 802 LANs through B-ISDN," Interworking '93, Bern Switzerland, ISBN 90 5199 1355, June 1993.
[17] A. Nazari, T. Hellman, and N. Kavak, "Routing and Configuration Management for Broadband Data Service," Telia Research internal report, submitted to the *Computer Commun. Review*, Jan. 1995.
[18] D. Deloddere, P. Reynders, and P. Verbeek, "Architecture and implementation of a connectionless server for B-ISDN," ISS '92 Yokohama, vol. 1 pp. 224-228.
[19] N. Kavak *et al.*, "Experience and Analayzes of a Connectionless Server for Provision of Broadband Data Communication Service," Interworking '94, Sophia Antipolis.
[20] N. Kavak, K. Laraqui, and A. Nazari, "Implementation of a B-ISDN/ATM Connectionless Data Service in an Internet Environment," INET'94/JENC5, Prague.
[21] W. E. Leland *et al.*, "On the Self-Similar Nature of Ethernet Traffic," Proc. SIGCOMM '93 Symposium on Communications Architectures and Protocols, 1993.
[22] P. Newman, "Traffic Management for ATM local Area Networks," *IEEE Commun. Mag.*, vol. 32, no. 8, Aug. 1994,
[23] R. Castro *et al.*, "Support of data communications in an ATM LAN," International conference on Information Networks and Data Communication INDC '94, Madeira, Portugal, April 1994.
[24] Y. Ohba, M. Murata, and H. Miyahara, "Analyses of dynamic bandwidth control for LAN interconnection through ATM networks," *IEICE Trans. Commun.*, vol. E76-B, no. 3, March 1994.
[25] T. Yokotani and T. Shikama, "Bandwidth allocation for connectionless service in private networks based on ATM technology," *IEICE Trans. Commun.*, vol. E-77-B, no. 3, March 1994.
[26] N. Yin and M. G. Hluchyj, "On closed-loop rate control for ATM cell relay networks," *Proc. IEEE INFOCOM*, Toronto, June 1994.
[27] M. Gerla, T. Charles, and G. Galassi, "Internetting LANs and MANs to B-ISDNs for Connectionless Traffic Support," *IEEE JSAC*, vol. II, no. 8, pp. 1145-1159, Oct. 1993.
[28] ATM Forum Traffic Management Specification, working document, atm95-0013, Jan. 1995.
[29] H. T. Kung , T. Blackwell, and A. Chapman, "Credit-based Flow Control for ATM networks: Credit Update Protocol, Adaptive Credit Allocation, and Statistical Multiplexing," *Proc. of SIGCOMM '94*, London, Sep. 1994.
[30] A. Romanow and S. Floyd, "Dynamics of TCP traffic over ATM networks," *Proc. of SIGCOMM '94*, London, Sept. 1994.

Biography

NAIL KAVAK received an M.Sc. in computer science and technology from the Linkoeping University, Sweden, in 1986. In 1994 he received his technical licentiate from the Royal Institute of Technology, Stockholm, where he is currently finishing his doctorate degree. He joined Telia/Telia Research, Haninge, Sweden, in 1987. His interests are network design, internetworking, and high-speed network protocols. He has been involved in the European research program — RACE — for several years, where he has done research and developed data communication architectures for local ATM networks. His e-mail address is: nail.kavak@haninge.trab.se.

It is essential to develop more powerful architectures which expose ATM to higher layers instead of hiding them.