Abstract

Many of the most common protocols at layers two (media access), three (network), and above expect to operate over broadcast media such as IEEE 802.3/Ethernet or IEEE 802.5/token ring. LAN emulation provides a widely applicable means for transitioning these protocols to the connection-oriented ATM environment. A combination of centralized and distributed intelligence allows a star-plus-mesh topology of ATM virtual circuits to emulate broadcast media, while providing most of the advantages of connection-oriented media.

ATM LAN Emulation

Norman Finn, Cisco Systems, Inc.

Tony Mason, OSR Open Systems Resources, Inc.

Asynchronous transfer mode (ATM) is an emerging network technology that promises to solve various problems currently suffered by today's networks. Critical to the success of ATM is the question of how ATM will interoperate with existing local area network (LAN) technologies. The better this fit, the more rapidly ATM can be deployed and the more quickly ATM's benefits will outweigh its costs.

Key among the efforts to ease this transition is in the area of *LAN emulation*, a technique of emulating Ethernet and token ring LANs over an ATM infrastructure. This article describes the motivation of LAN emulation, how the ATM Forum version works, and the status of this ongoing work.

Motivation

Three things motivate LAN emulation versus other possible solutions (e.g., "classical IP over ATM," Internet Engineering Task Force RFC 1577). First, the vast majority of networks today are mixed-protocol environments, using Internet Protocol (IP), IPX, NetBEUI, AppleTalk, DECNet, and so on; hence, an IP-only solution is not generally usable in production environments. Very few of these protocols have defined mappings to ATM, which means that either a mapping must be defined for each such protocol, or a more common solution must be found. Second, ATM must be integrated with existing networks using bridges and routers. Third, there is a huge base of existing applications that are not ATM-aware, and do not wish to be ATM-aware, but work with Ethernet or token ring.

With these in mind, LAN emulation provides the basic functionality of Ethernet and token ring, while sacrificing some ATM advantages and a few LAN features. In return, LAN emulation provides a general solution that allows any protocol defined to run over Ethernet or token ring to work transparently in an ATM environment.

Existing LANs

Ethernet and token ring networking technologies use *broadcast media,* while ATM uses a *nonbroadcast multiple access* (NBMA) medium. In the broadcast model, every message sent is available for reception by every station on the network segment. In the NBMA model, messages are visible only to the sender and the recipient (for point-to-point connections) or to the sender and the recipient list (for point-to-multipoint connections). Broadcast LAN segments are connected together with bridges and routers. ATM networks are joined together using switches.

LAN emulation works by simulating the broadcast nature of Ethernet and token ring as necessary. LAN Emulation eases the requirement that each station's hardware filter all incoming traffic to find frames destined for itself because most inbound traffic is only for that particular station. Since the principal goal of LAN emulation is to enable basic data communications, there are no provisions for guaranteed bandwidth, quality of service, and so forth. Those choices are made by the LAN emulation system.

Trivial LAN Emulation

There are several possible models for LAN emulation. The simplest is a "flooding forwarder." The forwarder maintains a point-to-multipoint connection to every member of the emulated LAN (ELAN). When a member of the LAN wishes to send a message, it sends it to the forwarder who in turn sends it to every member of the ELAN.

While this model works, it merely transforms ATM into a fast broadcast LAN. It imposes substantial overhead on the end systems by failing to utilize the ability of ATM to channel traffic directly from sender to receiver.

The actual implementation of ATM Forum LAN emulation provides a general solution that allows for bridging and routing, and attempts to keep the vast majority of network traffic on point-to-point connections between sender and recipient.

Mechanics of ATM Forum LAN Emulation

ATM LAN emulation mimics the behavior either of an IEEE 802.3/Ethernet LAN segment or an IEEE 802.5/token ring LAN segment. Emulation of a broadcast medium such as Ethernet over the connection-oriented medium of ATM requires the cooperation of each end station on the ELAN as well as the implementation of more or less centralized services. The LAN emulation component that replaces the conventional Ethernet or token ring driver in each end station is called the LAN emulation cient (LE client or LEC). The three centralized components are the broadcast and unknown server (BUS), the LAN emulation server (LES), and the LAN emulation configuration server (LECS).

LAN emulation chooses to solve the problem of distributing broadcast and other frames to all clients over the connection-oriented ATM medium by means of the BUS. In the simplest implementation, there is one BUS per ELAN. Every LEC has a point-to-point multicast send virtual channel connection (VCC) to the BUS. There is one point-to-multipoint multicast forward VCC from the BUS to all LECs on the ELAN. Any frame sent by an LEC down its multicast send VCC is retransmitted by the BUS to all clients via the multicast forward VCC.

The specification does not limit implementations to this simple model of one BUS and one point-to-multipoint VCC,

0163-6804/96/$05.00

but the differences between this simple model and the allowed alternatives are not apparent to an LEC.

Most unicast frames are sent via a mesh of point-to-point connections, called "data direct VCCs," established between pairs of LECs. Each LEC is required to discover the other LECs and make the required connections to them. LE clients use the BUS only when necessary, preferring direct connections.

The LE server (LES) facilitates the protocols that enable the LECs to make these discoveries and decisions. Every LEC has a point-to-point VCC, the control direct VCC, to an LES. The LES may also, at its own option, establish a control distribute VCC to some or all clients. Typically, this is a point-to-multipoint connection for distributing control frames, but other configurations are allowed. Each LEC sees exactly one control distribute VCC and at most one control direct VCC (Fig. 1).

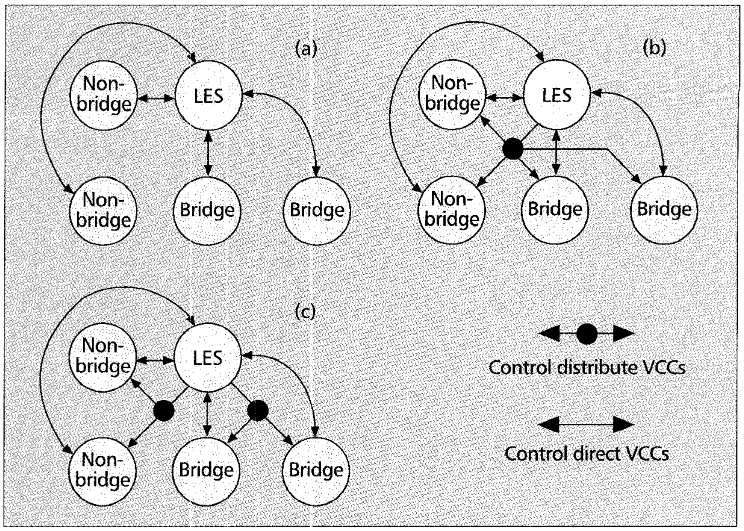

Figure 1. *Possible control plane arrangements.*

The LECS enables LECs to configure themselves and join networks without requiring manual intervention. Its point-to-point connections to LECs are called "configure direct VCCs."

The data direct, multicast send, and multicast forward VCCs connect the LECs and the BUS in a network of data plane connections. The configure direct, control direct, and control distribute VCCs connect the LECS, LES, and LECs in a network of control plane connections. These are illustrated in Fig. 2.

Configuring the LE Client

In order to achieve "plug-and-play" operation for LECs, LAN emulation requires a cnfiguration server (LECS). An LEC discovers the ATM address of the LECS and establishes a configuration direct VCC to the LECS. There are three ways for the LEC to find the LECS's ATM address: the address can be obtained from the ATM switch via the interim link management interface (ILMI), the standard defines a fixed LECS ATM address, and it also defines a well-known permanent virtual circuit (PVC).

Having established the configure direct VCC, the LEC sends a Configure Request control frame to the LECS specifying at least the LEC's ATM address. Optionally, the LEC may include a 48-bit medium access control (MAC) address, a preferred ELAN name, and/or choices for maximum frame size and LAN type (IEEE 802.3/Ethernet or IEEE 802.5/token ring). In addition, provision is made for supplying user-specific optional parameters in the Configure Request. The LECS uses any or all of this identifying information and a database configured by the network administrator to decide which ELAN, if any, the LEC should join, and returns the result in a Configure Response.

A successful Configure Response tells the LEC at least the ATM address of an LES serving the ELAN the LEC should join. The response may also include additional configuration information for the LEC.

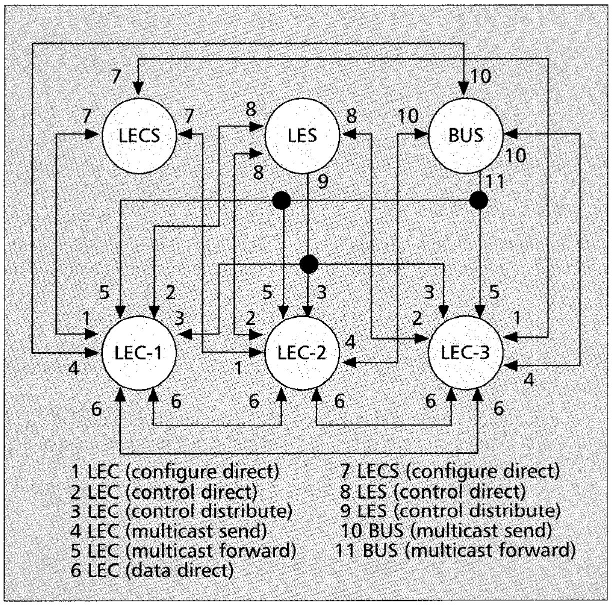

Figure 2. *VCCs in a fully connected three-client ATM emulated LAN.*

Joining an Emulated LAN

The LEC uses the ATM address of its LES to establish a control direct VCC to that LES and transmits a Join Request control frame. The Join Request contains an ATM address for the LEC and a flag indicating whether the LEC will ever act as a proxy for non-registered MAC addresses (e.g., act as an IEEE 802.1D transparent bridge). The LEC may also include a MAC address or preferences for LAN type and maximum frame size.

The LES decides if the LEC may join its ELAN. If the request is to be granted, the LES may choose to establish a control distribute VCC to the LEC (Fig. 1). Once the control distribute VCC, if any, is established, the LES sends a Join Response to the LEC.

The LES has a great deal of flexibility in constructing the control distribute VCC(s). The requirements are that every LEC has exactly one control direct VCC, and no more than one control distribute VCC. Plan (a) might be implemented inside an ATM switch. It has no control distribute VCC; the LES must copy control frames it wishes to broadcast. Plan (b) works just like the BUS; all frames sent by the LES are delivered to all LECs. Plan (c) allows the LES to deliver MAC address resolution (LE-ARP — LAN Emulation Address Resolution Protocol) requests for unregistered MAC addresses only to the proxy LECs, with a copy required to transmit a control frame to all LECs. With any of these plans, the LES may optimize the flow of LE-ARP requests and responses by directing them to specific LECs via the clients' control direct VCCs.

> *A BUS implementation is allowed to be "smart"; that is, it may choose to forward unicast and multicast frames to some subset of the LECs in the ELAN.*

A successful Join Response tells the LEC the correct values for the LAN type, maximum frame size, and ELAN name. Also included in the Join Response is a 16-bit LAN emulation client identifier (LECID). This number is unique among all LECs belonging to a given ELAN. There is a place for this identifier in all data frames and all request-type control frames issued by the LEC after joining the ELAN.

CONNECTING TO THE BUS

To find the BUS, the LEC sends an LE-ARP request to the LES. The LE-ARP request contains the broadcast MAC address (0xffffffffffff) and the LECID and ATM address of the issuing LEC. The LE service provider (presumably the LES or the BUS itself) must respond to this request by returning an LE-ARP Response to the LEC containing the ATM address of the BUS.

Armed with the BUS's ATM address, the LEC can establish the multicast send VCC to the BUS. The BUS responds to this VCC creation by establishing the multicast forward VCC back to the LEC. Parameters passed in the process of establishing these connections ensure that the LEC knows which VCC comes from the BUS, so it can be sure to accept the call and not release it until it wishes to leave the ELAN.

MAC ADDRESSES AND ATM ADDRESS REGISTRATION

The maintenance of the tables relating MAC addresses and ATM addresses may take a considerable effort by an LEC. In order to make it possible for a "smart" LES to relieve LECs of some of that burden, the specification provides LECs the ability to register {MAC address, ATM address} and {token ring route descriptor, ATM address} pairs with the LES.

One such registration pair (typically the LEC's own MAC address and ATM address) can be piggybacked on the Join Request sent by the LEC when it first contacts the LES. This capability suffices for most LEC implementations. In addition, Register Request, Register Response, Unregister Request, and Unregister Response frames are all defined in the specification to allow an LEC to register and unregister further {MAC address or route descriptor, ATM address} pairs. This capability might be useful, for example, to a data concentrator or legacy LAN switch whose list of served end stations was essentially static.

ADDRESS RESOLUTION

The LEC is fully operational once it has established its VCCs to the LES and BUS and has registered its MAC addresses. In order to transfer a data frame, the LEC must discover to which LEC it should send the frame. This is the function of the LE-ARP protocol.

Each LEC maintains a table, the LE-ARP cache, relating MAC addresses to ATM addresses. When an LEC gets a data frame for transmission whose destination MAC address is not in its LE-ARP cache, it sends an LE-ARP Request to the LES. This LE-ARP Request includes the requested MAC address and the LECID and ATM address of the requesting LEC.

In the simplest LES implementation, it relays the request to all of the LECs. The LEC responsible for that MAC address responds to the LES with an LE-ARP response containing its own ATM address. The LES relays the response to at least the originating LEC (Fig. 1).

The LEC knows the ATM address of the ATM end station at the other end of each of its VCCs. If it has no VCC to that ATM address, it signals the ATM switch to create one. The LEC associates the requested MAC address with the specific VCC, called a "data direct VCC." This information is placed in the LE-ARP cache.

This procedure is symmetrical, and can result in two data direct VCCs being established simultaneously between a pair of LECs. When this occurs, a simple algorithm is used so that both LECs know which VCC to use. The unused VCC is released due to disuse after a time-out period expires.

Token ring route descriptors, required by LECs that perform source routing, are associated with LECs' ATM addresses using the same LE-ARP mechanism as used for MAC addresses. Every LEC must register all its route descriptors with the LES.

An optional feature allows an LEC to transmit a Negative LE-ARP request to all other LECs. This control frame is sent when one LEC (say LEC A) detects that another LEC's LE-ARP reply (say, LEC B's) is in error. It cancels the other client's (B's) reply and supplies replacement information.

FRAME FORMATS

ATM LAN emulation uses ATM adaptation layer type 5 (AAL5) frames. AAL5 frames occupy an integral number of ATM cells, 48 data bytes per cell, with the last cell containing a trailer that includes the number of octets in the frame and a four-octet end-to-end cyclic redundancy check (CRC) to ensure data integrity.

There are two kinds of frames used in LAN emulation: data frames and control frames.

Data frames have either the value "0" or the LECID of the LEC in the first two octets. Following this marker is a normal IEEE 802.3/Ethernet or IEEE 802.5/Token-Ring frame. Token ring frames also include a pad byte and a frame control (FC) byte. No frame check sequence (FCS) is included in a data frame; LAN emulation relies on the AAL5 checksum. Each ELAN has a maximum frame size, whose value is distributed to each LEC as it joins the ELAN. There are four values for frame size, 1516, 4544, 9234, and 18,190 octets. The choice of maximum frame size is independent of the choice of ELAN Type; an "Ethernet" with an 18,190-octet maximum frame size is allowable.

Control frames have a common format (Fig. 3). The first two octets contain the value X"FF00," an illegal value for an LECID, to distinguish them from data frames.

DATA TRANSFER

All broadcast and multicast frames are sent to the BUS for retransmission to all LECs. All unicast frames for which a {MAC address, data direct VCC} or {route descriptor, data direct VCC} association has been established are sent down that data direct VCC. Unfortunately, there is a category of IEEE 802.3/Ethernet unicast frames that present a problem: those whose destination MAC address is not in the LEC's LE-ARP cache.

Consider an IEEE 802.1D transparent bridge B that is bridging frames between an ATM ELAN and a number of legacy LANs, which in turn may be connected to other bridges. LEC A wishes to send a frame to Ethernet station X residing on an Ethernet segment behind bridge B. Suppose that bridge B is unaware of the existence of station X; MAC address X is not in bridge B's forwarding table.

When LEC A sends the LE-ARP request for MAC address X, it receives no answer. No LEC knows MAC address X. In order for LEC A to reach station X, it must transmit the frame to the BUS, which distributes it to all LECs, including Bridge B. Bridge B, not knowing the whereabouts of station X, floods the frame to all its legacy LANs. The frame eventually reaches station X; X responds, bridge B learns its whereabouts, and answers a subsequent LE-ARP Request from LEC A. Further transmissions traverse the Data Direct VCC from LEC A to Bridge B, rather than using the BUS.

	0	1	2	3
0	MARKER = X"FF00"		PROTOCOL = X"01"	VERSION = X"01"
4	OP-CODE		STATUS	
8	TRANSACTION-ID			
12	REQUESTER-LECID		FLAGS	
16	SOURCE-LAN-DESTINATION			
24	TARGET-LAN-DESTINATION			
32	SOURCE-ATM-ADDRESS			
52	LAN-TYPE	MAXIMUM-FRAME-SIZE	NUM-TLVS	ELAN-NAME-SIZE
56	TARGET-ATM-ADDRESS			
76	ELAN-NAME			
108	Optional TLV parameters			

Figure 3. *Control frame picture.*

The possible presence of IEEE 802.1D bridges requires all LECs to use the BUS as a transmission path for frames whose {MAC address, ATM address} association is unknown. While an LEC could send all frames to the BUS this is undesirable because it would severely impact the performance of the ELAN. In order to accommodate certain special cases and balance conflicting needs, an LEC may violate the normal procedures. It may try for "a while" to get an LE-ARP response before resorting to the BUS path, and may send "a few" frames to the BUS without establishing a data direct VCC. Of course, "a while" and "a few" are well defined in the specification.

A BUS implementation is allowed to be "smart"; that is, it may choose to forward unicast and multicast frames to some subset of the LECs in the ELAN. It may accomplish this by using the bidirectional capability of the multicast send VCCs, or by splitting the multicast forward VCC into multiple segments.

Frame Ordering

Unfortunately, the procedure just described results in having two communications paths between a transmitting LEC and a receiving LEC, one via the BUS and one via the data direct VCC between those clients. Both paths are required, one for connectivity in the presence of transparent bridges, one for efficiency. Having two paths can result in frames being delivered out of order. The first frame sent down a data direct VCC, for example, may arrive sooner than the last frame relayed through a busy BUS. While many protocols can tolerate this situation, others cannot.

LAN segments, and even bridges, are required to preserve the order of transmitted frames, especially frames between a single pair of MAC addresses.

The solution to this problem is yet another protocol. Whenever an LEC must switch between two paths to another LEC, it may send Flush Request control frames down the old path. The receiving LEC is required to return a Flush Response control frame. The originating LEC may hold or discard further frames to the address that is switching paths until the response is returned, thus ensuring that the old path is empty of frames before the new path is put into service.

Permanent Virtual Circuits

The description so far has assumed that all VCCs are created as needed using user-network interface (UNI) signaling or switched virtual circuits (SVCs); PVCs are also supported. This is accomplished by requiring that all ELAN components be configured with (imaginary) ATM addresses for themselves and for the component at the other end of each PVC, as well as the type (control, multicast, etc.) of the PVC.

Alternatives Not Chosen

Several alternatives to the single BUS model were explored. Having each LEC establish a point-to-multipoint connection reaching every other LEC makes large demands on the connection management and traffic management resources expected to be available in ATM networks in the near to mid term. Having multiple multicast servers for different multicast groups proved to complicate the protocols considerably, and promised to offer little benefit in the presence of transparent bridging.

The LES and BUS are not allowed to "learn" associations between MAC addresses and ATM addresses except through the registration protocol. This is because transparent bridge LECs are expected to learn such associations by inspecting traffic carried on data direct VCCs, often enabling them to bypass the LE-ARP protocol. Such traffic is not visible to the LES or BUS. An LES that learned associations only via LE-ARP might give incorrect answers if it made the "optimization" of answering LE-ARP Requests for unregistered MAC addresses.

The LAN emulation standard provides for only two ELAN types, Ethernet and token ring. It makes no explicit provision for fiber distributed data interface (FDDI) support, in order to minimize the burden on LEC implementations. Well-known techniques for supporting Ethernet–FDDI and token ring–FDDI bridging can be applied to bridging between FDDI and ATM ELANs.

Status of Standards Development and Futures

The ATM Forum is an industry consortium whose explicit purpose is to forge multivendor interoperability agreements to foster the rapid deployment of ATM. The LAN

Emulation Sub-Working Group of the Technical Committee has been charged with responsibility for the LAN emulation standard.

The initial effort of the LAN Emulation SWG was to define the UNI component of LAN emulation (affectionately known as the "LUNI"). This specification was completed in January 1995, followed by a standard for a management information base (MIB) for the LEC in May. Included in the LANE SWG's task list are: MIBs and specifications for intra-ELAN service components; extending an ELAN across a wide area network; reducing the number of VCCs required for large numbers of ELANs; and utilizing ATM's quality of service capabilities in an ELAN.

While ATM-aware applications will appear and take direct advantage of ATM capabilities, LAN emulation provides a migration path for existing LANs into ATM.

A Multi-Protocol Over ATM (MPOA) SWG has also been formed within the ATM Forum. Basing its effort on LAN emulation and ongoing work in the Internet Engineering Task Force, the MPOA SWG is addressing issues above the MAC layer.

The development of this specification and interoperable implementations will greatly ease the transition of ATM network technology into existing networks over the next few years. While ATM-aware applications will appear and take direct advantage of ATM capabilities, LAN emulation provides a migration path for existing LANs into ATM.

Additional Reading

The standard for LAN emulation has been published by the ATM Forum. Paper copies of the specification, the LEC management specification, and the Addendum to the specification can be obtained from the ATM Forum in Mountain View, California, USA. They are available electronically via the ATM Forum's web and FTP sites at "http://www.atmforum.-com" and "ftp://atmforum.com/pub/specs":

- The ATM Forum, "LAN Emulation Over ATM v1.0 Specification," af-lane-0021.000.
- The ATM Forum, "LAN Emulation Over ATM v1.0 Addendum," af-lane-0050.000.
- The ATM Forum, "LAN Emulation Client Management Specification v1.0," af-lane-0038.000.

The ATM UNI specification on which LAN emulation is based has also been published, and the de Prycker book is a good introduction to asynchronous transfer mode technology.

- The ATM Forum, *ATM User-Network Interface Specification, Version 3.0*, Englewood Cliffs, NJ: PTR Prentice Hall, 1993, ISBN 0-13-225863-3.
- The ATM Forum, *ATM User-Network Interface Specification, Version 3.1*, Englewood Cliffs, NJ: PTR Prentice Hall, 1995, ISBN 0-13-393828-X.
- M. de Prycker, *Asynchronous Transfer Mode: Solution for Broadband ISDN*, 2nd Ed., Hertfordshire, U.K.: Ellis Horwood, 1993, ISBN 0-13-178542-7.

Biographies

Norman Finn is a senior MTS with Cisco Systems, San Jose California, specializing in ATM and VLAN architecture. He is currently active in both the ATM Forum LANE and MPOA working groups, and in the IEEE 802.1 VLAN standardization effort. Norm authored a significant portion of the LANE v1.0 specification. Before joining Cisco, he was with Ultra Network Technologies, building gigabit network devices. He holds a B.S. in astronomy from the California Institute of Technology. He may be reached via e-mail at nfinn@ cisco.com.

W. Anthony Mason is a consulting partner with OSR Open Systems Resources, Inc. a consulting firm specializing in the design and development of systems software. Prior to joining OSR, Mr. Mason was an architect and development manager with FORE Systems, as well as a key member of the OSF DCE/DFS design and implementation team at Transarc Corporation. He may be reached via e-mail at mason@osr.com