IP Transport Over Switched ATM Networks

Harris & Jeffries™

Ultimate Networking Software™

Important Notice

Trademark Acknowledgements

The Trademark™ (or Registered Trademark)	belongs to
Harris & Jeffries	Harris & Jeffries, Inc.
High-Performance Frame Relay™	Harris & Jeffries, Inc.
Soft-ATM™	Harris & Jeffries, Inc.
Microsoft (C, Visual C/C++, NMAKE, MS-DOS)	Microsoft Corporation.
WATCOM, WMAKE	WATCOM Intl. Corporation.
Microtec®, MCC68K v4.5T	Microtec Research, Inc.
SunOS	Sun Microsystems, Inc.
UNIX	UNIX System Laboratories, Inc.
VxWorks	Wind River Systems, Inc.
Windows NT	Microsoft Corporation.

Harris & Jeffries, Inc.

888 Washington Street
Dedham MA 02026
Phone: (781) 329-3200
Fax: (781) 329-4148
http: www.hjinc.com

Introduction

In recent years, the Internet Protocol (IP) has become the dominant layer 3 transport mechanism throughout Enterprises and in the Internet. During this period, Asynchronous Transfer Mode (ATM) has emerged as a powerful and viable carrier based service with its ability to deliver Quality of Service (QoS). Concurrently, bandwidth requirements from all sectors, ranging from business to residential use, are exploding, as increasingly powerful PC's seek to use more sophisticated applications. The possibility to use ATM based switched services combined with different QoS levels in a standards based manner is a paradigm shift in the way that IP networks are engineered and used. This is the potential for increased performance, on demand services, and QoS tailored to requirements of a particular application.

LAN Emulation, RFC 2225 (formerly RFC 1577), referred to here as Internet Protocol Over ATM (IPOA), Multi-Protocol Over ATM (MPOA) and Multicast Address Resolution Server (MARS) form the set of ratified standards that enable these advanced networks to be built today. Although these standards can be deployed with Permanent Virtual Circuits (PVCs), which are analogous to leased line or Frame Relay PVC networks, it is Switched Virtual Circuit (SVC) based networking that is the focus of this paper, offering the potential to tailor network requirements in a dynamic manner.

Whether it is enterprises seeking enhanced IP support for new applications, or carriers/Internet Service Providers (ISPs) seeking to offer differentiated services, such as outsourced intranets, these standards based mechanisms are being employed to create sophisticated networks, capable of delivering differentiated services, such a Voice Over IP (VoIP), which requires a high QoS, and email which can be supported with "best efforts" traffic characteristics.

This paper explores the standards that enable advanced IP networks to be built over switched ATM networks, suggests alternative topologies that may be employed and applications that can effectively be deployed over those networks.

Harris & Jeffries is a leading supplier, worldwide, of software protocol stacks for the communications industry, specializing in ATM, Frame Relay and IP transport applications related to these technologies. Software components, meeting the standards required to develop IP services over switched ATM networks, are core products of the Harris & Jeffries portfolio and can be pre-integrated with commercially available operating systems and SAR chipsets.

Technology Background

ATM Standards

Signalling

All the standards used to support an IP routed network over a switched ATM network specify the use of ATM Forum signalling specification (User Network Interface) UNI 3.0 or UNI 3.1. MPOA is the only specification which allows the use of UNI 4.0. However, since MPOA is not specific in its use of class of service, other than UBR (Unspecified Bit Rate), this is of little significance. As UNI 4.0 becomes a widely deployed standard, particularly for the use of ABR (Available Bit Rate), it is reasonable to assume that it will be used as the signalling mechanism for the all the standards discussed below. Initially, use may be restricted to the same parameters as specified in UNI 3.0 or UNI 3.1, but as QoS becomes widely deployed over these networks, the full capabilities of UNI 4.0 will be utilized over time.

Class of Service

ATM traffic is generally divided in a number of “classes of service”, each of which is defined by the general requirements of type of traffic to be transported through the network. IP traffic is generally transported using Unspecified Bit Rate (UBR) class of service, (although there is no restriction on this) which is defined in all the ATM Forum UNI specifications (3.0, 3.1 and 4.0). UNI 4.0 also introduces an additional class of service, Available Bit Rate (ABR), which operates with many of the characteristics of UBR, but is designed to provide superior performance for packet traffic.

UBR provides a Virtual Channel Connection (VCC) through an ATM network, but does not reserve any resources within the network. This provides a true “best efforts” type of service. Traffic using a UBR class of service will be indiscriminately discarded in the event of network congestion. This is very similar to the operation of frame relay networks with a zero Committed Information Rate (CIR). Further UBR traffic can be transmitted into an ATM network at up to the line rate and is not subject to either traffic shaping or traffic policing, since there is no established traffic contract. In lightly loaded networks, UBR can offer excellent performance, but as a network utilization increases and traffic has to be discarded to meet traffic contracts, ATM cells are randomly discarded. Although many IP protocols, such as the Transmission Control Protocol (TCP) provide error recovery in the event data loss (random ATM cell discard in this case), performance can severely degrade in a heavily congested network, and TCP re-transmission can actually make the situation worse.

ABR

ABR expands on the concept of a “best efforts” type of service by providing a feedback mechanism and a minimum level of throughput for a VCC, described by a Minimum Cell Rate (MCR). This allows a connection to consume a small amount of network resource. Additionally, ABR operates with a feedback mechanism which allows connections to indicate how much traffic can be injected into the network depending on resource availability. This mechanism increases and decreases bandwidth dynamically, between the MCR and the Peal Cell Rate (PCR), according to network resource availability. For correct operation, ABR implicitly requires that each connection has per Virtual Circuit (VC) queuing to enable adjustment on a per VCC basis.

With this mechanism, arbitrary cell loss is greatly reduced, providing better overall throughput for each IP traffic flow, often described as “goodput”. To enhance this mechanism further, the concepts of Early Packet Discard (EPD) and Partial Packet Discard (PPD) can be employed. Both mechanisms recognize that in the event of ATM cell discard, that particular cell represents information lost from an (IP) packet. Transmitting the remainder of the packet through the network represents a waste of resources, since the packet will be discarded at the destination. PPD and EPD work by discarding the remainder of a packet, or preferably whole packets to more effectively utilize network resources.

QoS

Quality if service represents the ability to meet a requested traffic contract for an ATM VCC through a network. It consists of requesting network resources during the signalling/network establishment phase of circuit, using a Connection Admission Control (CAC) algorithm inside ATM switches, or other devices, to ensure that resources are reserved appropriately, and policing traffic through the network to ensure that the traffic contract negotiated by the signalling mechanism is met. The power of ATM is the ability to request different Quality of Service (QoS) levels and have them guaranteed by the network.

Addressing

ATM networks supporting IP functionality are deployed in a variety of topologies. In a purely private campus environment, ATM End System Addresses (ASEA’s) will most likely be used. In a public carrier environment over the Wide Area Network (WAN), E.164 addresses will be used. The standards must therefore support a flexible addressing scheme which may employ one or both types of capabilities.

Routing

The discussions below assume that mechanisms exist to establish connections through an ATM network with the appropriate QoS characteristics that are requested by the ATM signalling. The path through the network may be computed using proprietary mechanims, which limits the network to devices of a single vendor, or it may use standards based mechanisms such as Private Network to Network Interface (PNNI) from the ATM Forum, which provides an integrated signalling/routing mechanism between ATM devices at the core of a network.

Adaptation

All the standards used in developing switched IP networks over ATM use ATM Adaptation Layer 5 (AAL5). AAL 5 provides transport for variable bit rate, delay tolerant, connection oriented traffic for both the signalling (call setup and teardown) and as the transport mechanism for the data. This is well suited to the “best effort” delivery mechanism described by UBR and ABR, relying on the higher layer protocols to provide recovery in the event of lost or corrupt data. Additionally, the data is expected to be insensitive to some measure of delay and delay variation, largely because PDUs transported across a given Virtual Channel Connection must be transmitted contiguously, since each ATM cell is not explictly identified as belonging to a particular PDU. This means that only whole packets of data can be transmitted at once over AAL5, and the variable length of these packets introduces both delay and delay variation for a given VCC.

The choice of AAL5 as an adaptation mechanism does not prevent different QoS levels being associated with any VCC, except that delay and delay variation cannot be tightly constrained, because that cannot be guaranteed.

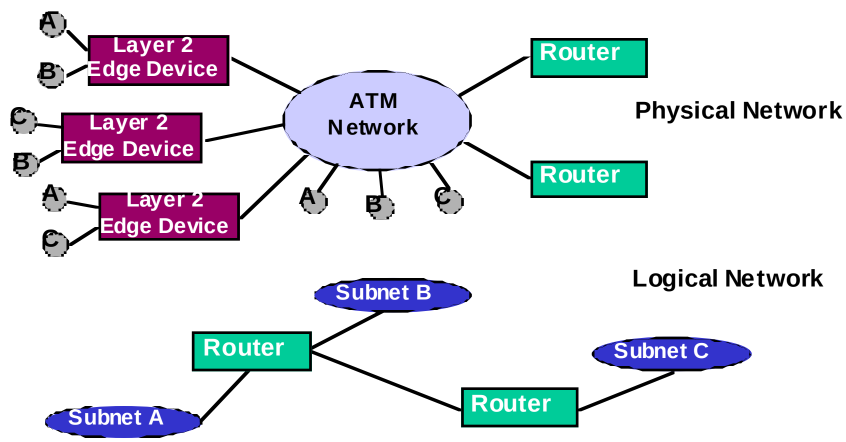
RFC 1483

RFC 1483 provides the basic definition of how multi-protocol traffic, including IP, is encapsulated and transported over ATM AAL5 virtual channel connections. It provides the underlying mechanism used by all the standards discussed below. Key in this standard is the definition of how multiple protocols can be transported over the same ATM VCC using Logical Link Control (LLC) multiplexing.

LAN Emulation

LAN Emulation extends the functions of legacy Ethernet (IEEE 802.3) or Token Ring (IEEE 802.5) LANs over an ATM network. As with legacy networks, LAN Emulation is not limited to the transport of IP since it operates at the MAC layer. Using ATM to emulate Ethernet or Token Ring networks does remove the restrictions imposed by those physical media. The physical location of a device is independent of its membership in a particular LAN Emulation group. This allows the creation of Emulated LANs (ELANs), where the members can be in physically disparate locations. Extending this concept allows the implementation of virtual LANs (VLANS) where membership can be organized by function rather than physical location (Figure 1).

Figure 1 - LAN Emulation - ELANs/VLANs



The basic operation of LAN emulation is to resolve MAC layer addresses to ATM addresses in order to establish a shortcut (data direct Virtual Channel Connection (VCC)) across the ATM network for efficient traffic flow. The component parts of LAN emulation are the LAN Emulation Client, the LAN Emulation Server (LES), the LAN Emulation Configuration Server (LECS) and the Broadcast and Unknown Server (BUS). The LEC is the element that performs the forwarding function in LAN emulation with the other elements providing necessary functions to emulate a legacy LAN, such as broadcasting or sending unicast data prior to determining the destination MAC address.

LAN Emulation Operation

Figure 2 illustrates the steps required to establish connection between two LECs for Unicast data. LECs must first obtain their configuration from the LECS. The ATM address of the LECS is identifed by the LEC through the Interim Link Management Protocol (ILMI), where it is identified in an SNMP MIB object. This enables the LEC to connect to the LECS. The configuration information from the LECS enables the LEC to determine it's ELAN membership and provides the LEC with the address of the LAN Emulation Server(s) (LES) it should use.

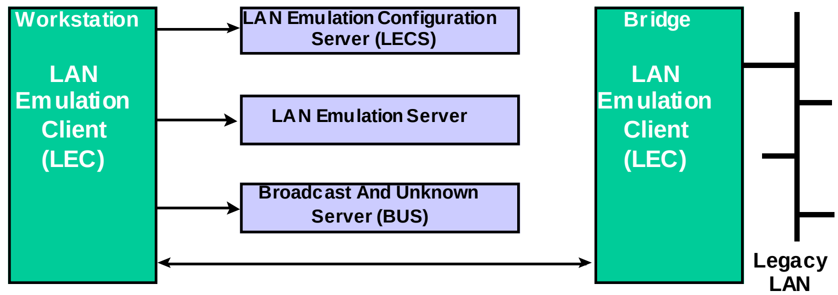


Figure 2 - LAN Emulation Components And Operation

When the LEC has data to send it, it can either use an internal cache it has established, based on previous traffic patterns, or it establishes a direct "control" VCC to the LAN Emulation Server (LES), which enables the LEC to request the ATM address for the destination MAC address. If the destination ATM address is not held in the LEC's cache, it sends a LAN Emulation Address Resolution Protocol (ARP) request to the LAN Emulation Server. If the LES can respond to the ARP it returns the ATM address associated with the destination MAC address, which allows a data direct VCC to be established between the LECs. Until the LAN Emulation ARP can be answered, data is forwarded to the BUS, which sends the data to all the LECs in the emulated LAN, ensuring that it reaches the destination. As this process illustrates, use of the internal cache greatly reduces the overhead on the system, particularly in the overhead of broadcasting Unicast data, which is only destined for one location. Support for true broadcast and multicast data is performed via the BUS, which is beyond the scope of this paper.

LEC performance is critical to efficient network operation and this is best accomplished by providing an optional hardware forwarding function, or "hardware assist" that can be utilized in conjunction with the software MAC to ATM address resolution once a shortcut has been established through the ELAN in the ATM network.

Standards Development

The initial version of LAN Emulation, version 1.0 from the ATM Forum provided the basic mechanism for transporting LAN traffic over an ATM network. The objective was to hide the complexities of ATM from the user application. However, it rapidly became apparent that this specification had some weaknesses, particularly restricting the use of ATM QoS characteristics and the lack of support for RFC 1483 LLC multiplexing, requiring a separate data direct VCC for each traffic flow. Additionally, the inability to support multiple and/or redundant LAN Emulation Servers and BUS's were seen as long term limiting factors to overall network reliability and performance, although these issues have been addressed by manufacturers in a proprietary manner.

The current standards for LAN Emulation V2.0 have addressed the shortcomings of the initial specification. They are now split into main areas, the LAN Emulation User Network Interface (LUNI), which provides the communication standard for LEC to LEC, and the LAN Network to Network Interface (LNNI) which details the operation of multiple server and BUS's in a network (LNNI is ongoing work in the ATM Forum). From a LEC perspective, LUNI provides for the multiplexing of traffic on a single ATM connection for efficient transport, as defined in RFC 1483 and it provides the ability to associate quality of service levels with each data direct VCC.

QoS

Version 2 of the LAN Emulation specification provides the ability for higher layer applications to request a Quality of Service associated with a given data flow. Rather than exposing all the complexities of ATM signalling, the association is made indirectly via a "handle". The handle is associated with a specific set of QoS parameters which include all the static information associated with a call. Applications may then select the appropriate handle that best suites their requirements.

Applications

LAN Emulation can be deployed in a pure ATM environment, where ATM transport is used directly to the desktop, but it is commonly used in conjunction with traditional networking devices, such as bridges and routers. This allows existing shared media Ethernet and Token Ring LANs to be extended into the ATM environment in a seamless manner. From the IP perspective, a single emulated LAN supports an IP subnet in an equivalent manner to a traditional Ethernet or Token Ring.

ELANs can be interconnected using traditional devices such as bridges or routers, in the same manner as legacy LANs. Bridging technology allows the creation of Virtual LANs (VLANS), where the underlying requirement is connectivity at the MAC layer.

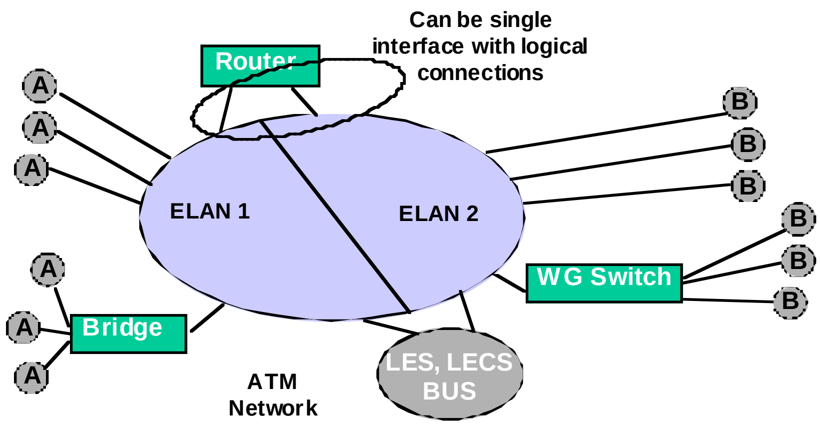


Figure 3 - IP Transport Across Multiple ELANs

Interconnection of IP subnets, which are overlaid in ELANs is accomplished via routers, as with traditional LANs, providing the same level as functionality as to-day (figure 3). This makes IP networks in an ATM environment subject to the same constraints as any other underlying media, such as frame relay or leased lines. The technologies discussed below demonstrate how these restrictions can be removed when a complete ATM infrastructure is used.

IPOA (RFC 2225)

IPOA provides a mechanism for transporting IP data over an ATM network. This introduces the concept of a Logical IP Subnet (LIS) which maps traditional IP subnets onto an ATM network, removing the physical relationships imposed by traditional mechanisms, in a similar manner to LAN Emulation. For the correct operation of IPOA, all devices in the LIS must be directly connected over the ATM network.

The main difference between ELANs and LISs is that ELANs operate at the MAC layer, and as such ELANs are insensitive to the layer 3 protocol they transport, which could be IP, IPX or any of the other layer 3 protocols in use. LISs are specific to IP, and operate under the same rules as traditional IP subnets. LISs do not require an underlying MAC layer transport, the equivalent is provided by ATM, and this removes much of the complexity required by LAN Emulation, which must simulate the operation of legacy LAN broadcast operation in a connection oriented environment.

The base standard for the definition is RFC 1577- Classical IP and ARP over ATM from the IETF, but there have been updates to consolidate this specification and related specifications into new standards, RFC 2225.

Operation

Conceptually IPOA is similar to LAN Emulation, but it is much simpler in operation. It is specific to the IP layer, does not use MAC headers, and leaves the issue of support for broadcast, or multicast IP data open. QoS characteristics are also ignored by the IPOA standards and are left at the discretion of the user to implement.

The role of the LAN Emulation Server is replaced by an ATM Address Resolution Server (ATMARP). The ATMARP server performs the function of resolving IP addresses to ATM addresses in order to establish a connection across the ATM network between the two points. Additionally, it also supports an Inverse ATMARP capability, InATMARP, which enables known ATM addresses to be associated with IP addresses. This is primarily used when Permanent Virtual Circuits (PVCs) are established between locations.

In Switched Virtual Circuit (SVC) operation, as the IPOA devices come on line, they register their ATM and IP addresses with the ATMARP server, using the same mechanism as when requesting a connection to a destination IP address, except using their own IP address. When an IPOA device has data to send it establishes a VC to the ATMARP server and requests the destination ATM address associated with the destination IP address. The returned ATM address is then used to establish a connection to the destination. This connection is a point to point bi-directional VCC which is used by both ends to transmit and receive data. The destination ATM addresses have an associated aging process which requires re-validation every 15 minutes for active VCs. Inactive circuits are removed after this time.

The establishment of point to point connections between IPOA enabled devices leads to the creation of a mesh of VCs. The number of VCs required is described by (n*(n-1))/2. As the network size grows, the VC count can become large, but should should not be a problem in a properly engineered network, where LISs are of a reasonable size.

Building Networks with IPOA

As with traditional IP subnets, LIS's are interconnected via routers, as shown in Figure 4. Since IPOA only operates within a LIS, devices directly connected to the LIS must be able to determine the IP address of the appropriate router in order for traffic to transit between LIS's.

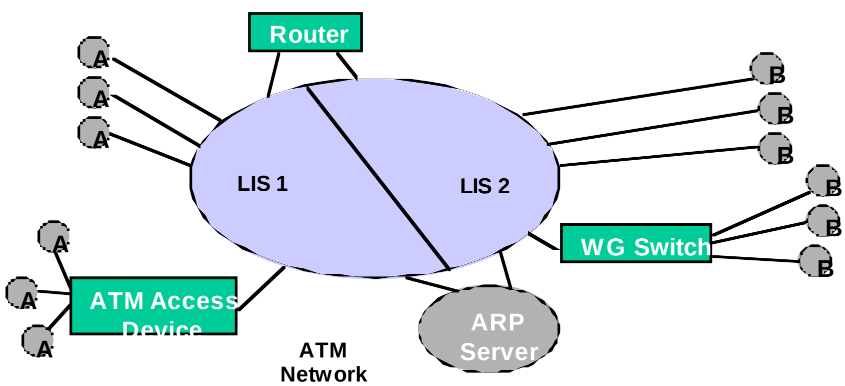


Figure 4 - IPOA - Classical IP and ARP Over ATM

One of the key issues faced in configuration of the routers to enable the devices in one LIS to communicate with those in another. One option is manual configuration, but this is time consuming and error prone. It does not allow router discovery mechanisms, developed over many years, to operate effectively. Automatic router discovery and configuration mechanisms can be enabled over LISs by the use of additional protocols such as Multicast Address Resolution Server (MARS). This is discussed later in this paper.

Multi-Protocol Over ATM (MPOA)

From an IP perspective, LAN Emulation and IPOA both deal with datagram transport within a single IP subnet. In typical networks, IP datagrams need to traverse multiple subnets between source and destination. Although the power of ATM can be used inside the subnets, advanced applications are limited by the internal capabilities of the routers to cross between subnets. MPOA effectively addresses these restrictions and substantially enhances the limiting capabilities of many IP routing devices by enabling the establishment of ATM VCC's, or "shortcuts" directly between IP endpoints on different subnets.

Traditional routers combine routing calculation and IP datagram forwarding in a single device. MPOA splits these functions into an MPOA Client (MPC) which performs traffic forwarding, and an MPOA Server (MPS) which is responsible for route determination (figure 5). This introduces the concept of a virtual router and greatly reduces the reliance on pre-determined paths through the network, established by traditional routers, which are insensitive to different QoS requirements. MPOA allows individual "traffic flows" to take the appropriate path through an ATM network based on their resource requirements, and the current resource allocation in the ATM network.

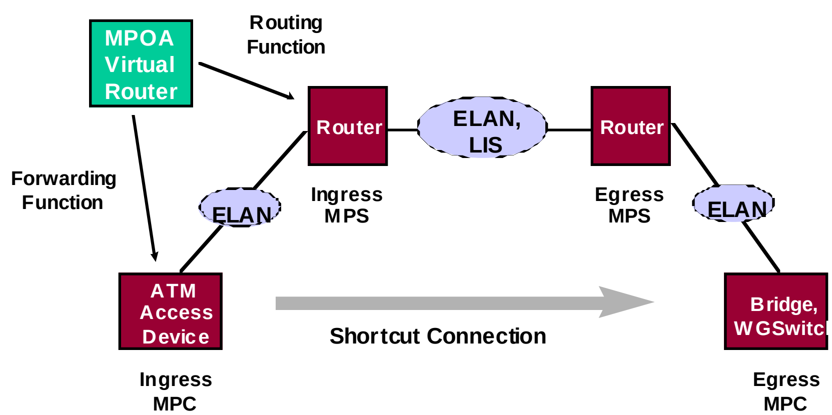


Figure 5 - MPOA Functions

MPOA builds on the previously discussed standards of LAN Emulation, but also draws from standards being developed by the Internet Engineering Task Force (IETF) by using the Next Hop Resolution Protocol (NHRP). The component parts of the MPOA Client and Server are shown in figure 6.

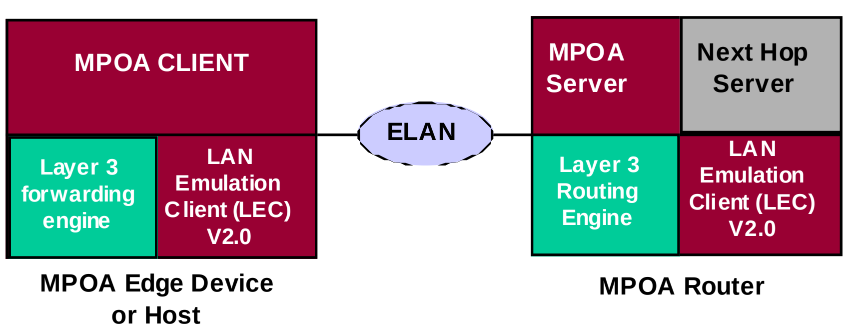


Figure 6 - MPOA Components

MPOA Operation

The MPOA Client is responsible for detecting the traffic flows and establishing the shortcut through the network once the destination IP address has been resolved to an ATM address using the MPOA/NHRP protocol. Prior to this process, packets are forwarded through the local IP subnet using LAN Emulation. The use of LAN Emulation allows existing traditional LAN devices to be integrated into an ATM network without the MPOA edge device being a fully functional router. This does not preclude a router at the edge of the network containing both an MPC and an MPS.

The MPOA Server is a component of a router. It requires a LAN Emulation Client to forward and receive traffic from MPC devices prior to the establishment of an MPOA shortcut ATM VCC. It takes no part in flow detection and is responsible for determining the destination ATM address for the given internetwork layer (IP) address that is requested by the MPC.

MPOA traffic flows are determined by the MPC, which monitors traffic between end points in the network at the IP layer. The MPOA protocol is engaged once a pre-determined number of packets have been sent between the end points in a specified period of time.

NHRP/MPOA Protocols

The Next Hop Resolution Protocol (NHRP), RFC 2332, developed by the IETF is the basis for the mechanism used to determine the destination ATM address that is used by an MPC to establish a shortcut connection between the source and destination IP addresses. The use of NHRP is not restricted to MPOA, but this is the first practical application for the protocol. In general NHRP can be used with Non-Broadcast Multi-Access (NBMA) networks. Another possible application of NHRP is in SVC based frame relay networks, either in a pure frame relay environment, or when interworking with ATM, when the standards are ratified.

The MPOA/NHRP protocol allows the resolution of the destination IP address to the destination ATM address by following the path through the network at the IP layer from the source to the destination. The basic message flow is illustrated in figure 7. Message flow is regarded from the perspective of MPOA as a closed entity with messages entering the system at the ingress, for transport across the network, and exiting the system at the egress.

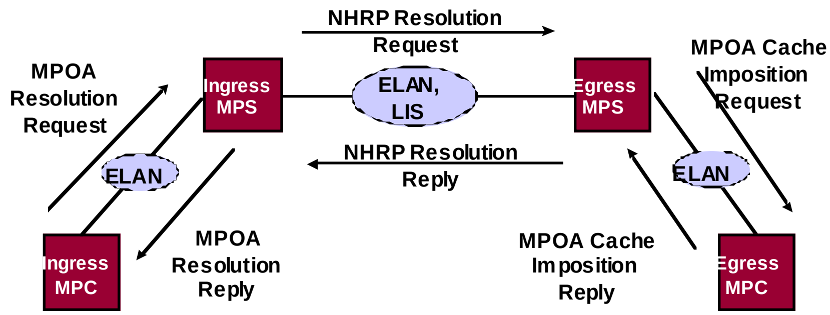


Figure 7 - MPOA Message Flow

NHRP works purely at the internetwork layer and requires access to routing tables in order to determine the next hop for a given destination IP address, in the case of TCP/IP. The MPC also works at the internetwork layer but since there is no requirement for the MPC to have an IP address, it cannot use NHRP directly because it cannot necessarily create the correct message, and it may not know the IP address of the local MPS where the message should be forwarded as the next hop. LAN emulation does know the MAC address of the MPS router where it is forwarding frames that are not destined for its local ELAN. The MPC uses this information to send an MPOA resolution request message to the MPS. The MPOA message set is therefore a slightly modified form of the NHRP message set which takes advantage of the ability of LAN Emulation to correctly forward requests and relaxes the stringent requirements of NHRP at the IP layer.

Once a flow has been detected, the MPOA message flow is as depicted in figure 7. When the ingress MPC receives the MPOA resolution reply, it can establish a direct ATM connection through the network to the destination, removing the need to traverse the routed network. During the MPOA/NHRP resolution process, the IP datagrams continue to flow along the routed path.

In addition to the basic message flow described above, the MPOA/NHRP protocols must account for routing changes in the IP networks, due to congestion, network/equipment failure, or re-configuration such as adding MPSs, and the failure/addition of MPCs and IP hosts to networks. This provides MPSs with the ability for example, to purge MPC's of existing information and the re-establishment of connections to ensure that destinations can be reached.

QoS and MPOA

The MPOA specification does not discuss the use of QoS parameters beyond UBR, which inherently is a best effort service. However, it is easy to see that similar mechanisms to those employed by LAN Emulation could be used to extend a QoS capability to MPOA.

Multicast Address Resolution Server (MARS)

The final element that provides fully functional IP transport over switched ATM networks is RFC 2022 Multicast Address Resolution Server (MARS). MARS and extension RFCs provide the capability to support multicast and broadcast IP traffic using a client/server approach, in an extended manner to IPOA.

MARS operates on a Logical IP Subnet (LIS) basis and expects that all the routers for a particular LIS are directly connected to an ATM network. As with IPOA, the MARS is expected to support multiple LIS's and the routers and the MARS may use the same physical connection to provide multiple logical ATM VCC's to different LIS's.

Routers typically operate as MARS clients and are configured, or learn via extensions to ILMI, the address of the IP Multicast group(s) of which they are members. The MARS clients register their IP addresses for the multicast groups within the particular LIS with the MARS server (figure 8). This process is not limited to multicast IP addresses an can easily be extended for use with the IP broadcast address, as is described in RFC 2226.

It should be noted that LIS membership is not restricted to routers, but can consist of a wide variety of devices, such as ATM equipped workgroup switches, ATM access devices and ATM hosts. IP Multicast membership is overlaid on top of the LIS, which for example, allows all the routers to be placed in the same multicast group, enabling them to easily discover each other, while not involving non-routing devices that are members of the LIS.

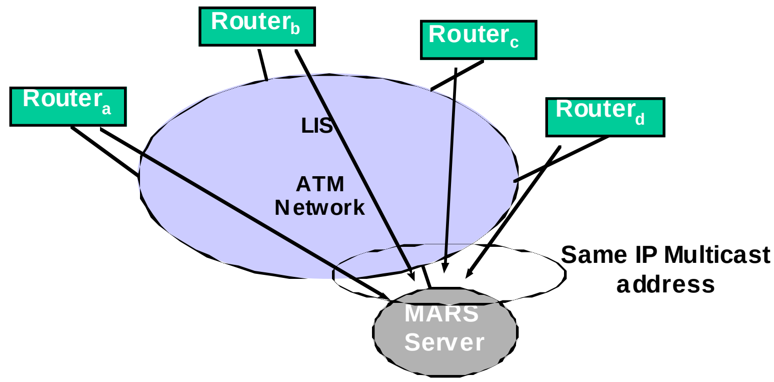


Figure 8 - MARS Registration

MARS has two modes of operation for enabling multicast communication; establishing a mesh of point to multipoint VCs from each device in the multicast group, or using a Multicast Connection Server (MCS) to perform the multicast function from one device to all the others in the group. Both mechanisms have their advantages and drawbacks and the selection of which method to use is very much dependent on network size, topology and performance requirements. In either case, the broadcast mechanism is transparent to the MARS Clients.

The MARS clients each must establish a control VCC to the MARS server in order to register themselves and to receive information regarding the other members of the group. Initially, each MARS client establishes a transient bi-dectional VCC to the MARS server to register its membership in the multicast IP group. This is torn down after a period of inactivity, typically 20 minutes. The MARS server registers the ATM address of the client and establishes a point to multipoint circuit, to that client, if this is the first member of the group, otherwise it adds the client as a leaf to an already existing point to multipoint call.

If there is no traffic for the multicast group over the ATM VC's then the routers will invalidate their membership in the group and teardown their leaf on the multipoint connection. The client can then re-register with the MARS server to ensure correct operation.

Mesh of Point to Multipoint VC's

As routers register with the MARS server, they receive information on other members of the IP multicast group. This enables each router to establish a point to multipoint VC to all the other routers in the multicast group (figure 9).

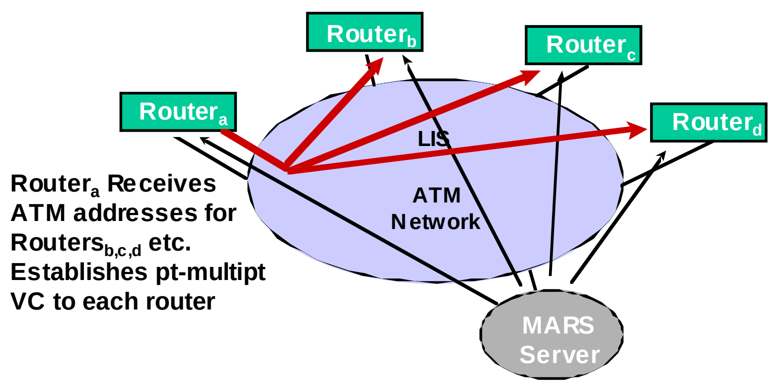


Figure 9 - MARS Meshed VC Operation

Using a mesh of point to multipoint VC's provides high performance since each router in the multicast group is responsible for forwarding traffic to the other members of the group. The number of VC's required in this mode of operation can be substantial and is of the order of n^2, since each router establishes a call. Whether a network can support the required number of VC's, and whether the tariffing aspects are favorable to a public carrier environment are major questions that must be answered prior to deployment. For small networks this is probably not an issue.

MultiCast Server (MCS)

An alternative approach to a mesh of point to multipoint VCs is to use a MultiCast Server. This performs a similar function to the BUS in LAN

Emulation and drastically reduces the number of VC's required at the expense of performing all the multicast aspects from a single point.

To the MARS client, the use of an MCS is transparent, since they register their multicast membership in the same manner as for VC mesh. The MARS server needs to be aware that an MCS is used because it only returns the ATM address of the MCS to each client as it registers. The ATM address of the newly registered client is sent to the MCS for inclusion in the appropriate multicast IP group (figure 10).

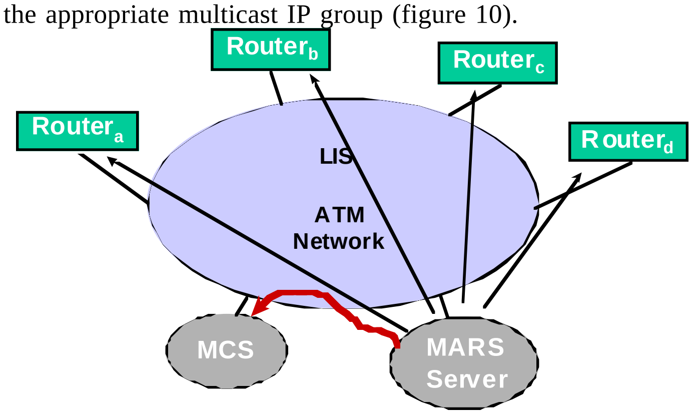


Figure 10 - MultiCast Server (MCS) Registration

The MARS clients "believe" that there is only one other member of their multicast group and establish a VC to the MCS in order to send any traffic for the multicast. The MCS establishes a point to multipoint VC to all the members of the multicast IP group and forwards data it receives from any of the members (figure 11).

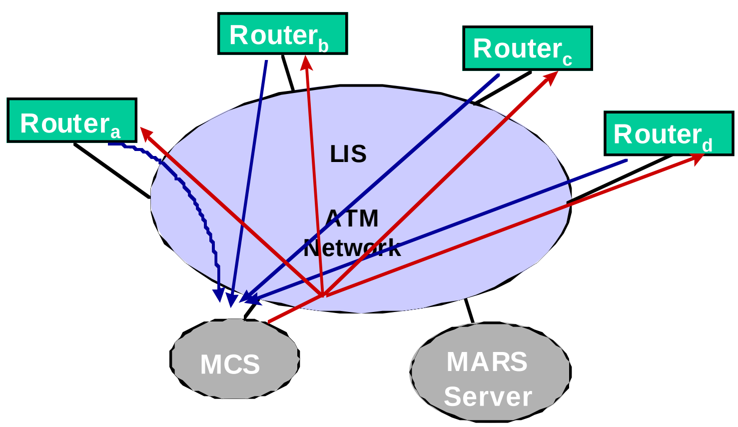


Figure 11 - MultiCast Server (MCS) Operation

Using an MCS to distribute traffic reduces the number of VC's required to the order of 2n at the price of placing the multicast burden on the MCS. In large networks where VC's are a premium this mode of operation delivers substantial savings over the meshed VC approach. In smaller networks, or for performance sensitive applications, the MCS can become a bottleneck in the system unless it very carefully engineered. One additional side effect of the MCS is that the source of the multicast IP datagrams will receive its own information since the MCS is not required to filter traffic before performing the multicast forwarding.

MARS and IPOA

A close relationship exists between MARS and IPOA since they perform similar functions. It is quite possible that both of these capabilities will be implemented in the same device. In this case, the standards provide for the control VC's established between the client and the servers can be shared between the applications.

ATM VCs, QoS and MARS

The MARS specification was developed when the ATM Forum UNI V3.1 specification was the most recently ratified version. Call setup for IP datagrams can only be established using point to multipoint procedures, which requires the originator of the call the add parties as they register with the multicast IP group.

QoS characteristics are unspecified by MARS and are currently at the discretion of the application to select an appropriate characteristics, other than requiring the use of AAL5 and LLC/SNAP encapsulation.

Multicast Work in Progress

Described above is the basic mechanism to support IP multicasting over an ATM network. RFC 2022 describes mechanism to support backup or redundant MARS. Further IETF drafts discuss the use of multiple MARS in a network and how they can utilize related developments such as the Server Cache Synchronization Protocol (SCSP) to maintain coherent data bases. As with IPOA, MARS is currently limited to a single LIS and further work relates to extending MARS capabilities beyond the LIS.

MARS is currently limited to the use of UNI 3.0/3.1 for establishing multipoint connections. This requires that the originator of the call add parties as they join a multicast group. UNI 4.0 adds the capability of leaf initiated join, which, from an ATM signalling perspective, would allow routers to join a multicast group without requiring the MARS to control the call. This is more in line with the IP multicast protocols which allow users to join already active multicast groups.

In addition to work on MARS itself, there is substantial scope for overlaying IP multicast protocols, such as Protocol Independent Multicast (PIM) directly over a MARS capability. As multicast becomes an increasingly important application, where limiting bandwidth use through a network in addition to reaching multiple users is a key requirement, further work can be expected in this area.

IP Transport and Switched ATM Services

Networks based on nailed up bandwidth, such as leased lines and frame relay cannot effectively accommodate the changing traffic flows experienced in today's IP based networks. Frame relay does offer some flexibility, with the ability to burst above the Committed Information Rate (CIR). However, there is no guarantee of delivery and the basic CIR must typically be paid for whether it is used or not. Newer carrier based services with proprietary guaranteed mechanisms, and associated service charges, may provide some measure of flexibility.

Using switched services over an ATM network, IP transport can be made both flexible and QoS aware, providing enhanced services in a standards based manner for a variety of emerging applications, which require guaranteed performance to be effective. MPOA allows the creation of shortcuts between MPC enabled devices at the edge of ATM networks. These devices could take many forms, including ATM host devices, servers, ATM enabled Etherent workgroup switches, ATM access devices, ATM enabled bridges and ATM enabled routers. In the case of routers, the MPC and the MPS may be co-located in the same device.

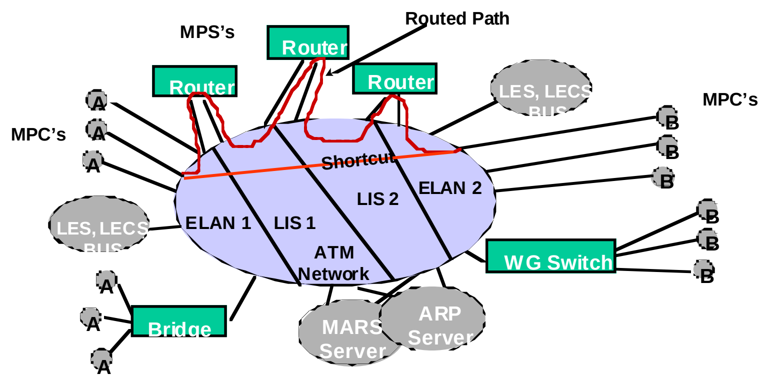


Figure 12 - Practical IP Network Over ATM

To implement MPOA an ELAN must exist between the MPC and the MPS. The MPS's may be connected via ELAN's or more likely in a WAN, using IPOA to create a LIS. Since IPOA substantially reduces traffic over LAN Emulation, it is potentially more popular in a public service or WAN environment where bandwidth costs are at a premium (Figure 12). With this combination of LAN Emulation, IPOA and MPOA, switched ATM networks can effectively be built allowing adaptation to traffic patterns and allowing data flows to utilize shortcuts through the ATM network, with associated QoS parameters.

Throughout the development of IP networking, the IETF has placed great emphasis on simplifying the creation of large routed networks. This means reducing the amount of human intervention for initial configurations, and providing sophisticated dynamic routing algorithms, such as OSPF and BGP4 to interconnection routers and routing domains.

In a switched network, the major issue becomes one of router discovery. Routers initially need to understand how they can establish a connection to each other prior to running protocols that enable them to discover each other at the IP layer, in order to construct the network topology. With leased line networks, this is not an issue, since the router are permanently connected.

Router discovery typically relies on using well known IP addresses, such as the IP broadcast address. Since this is common to all routers, it only has significance within the confines of an IP subnet. Inside ELANs, this is not an issue, since the underlying LAN technologies support broadcasting and the defacto implementation is within an IP subnet.

When interconnecting routers using IPOA, there is no underlying mechanism to support router discovery in the switched environment. However, MARS, extended to include the use of the IP broadcast address, can effectively address this requirement. This allows the routers to establish broadcast or multicast connectivity, at the IP layer, by establishing ATM point to multipoint connections.

The common IP routing protocols all make provision for the use of broadcast or multicast to either exchange initial routing information, or, in the case of hosts, to listen for connectivity to a local router. Routing Information Protocol (RIP) (RFC 1058) uses the IP broadcast address to transmit routing information. RIP V2 (RFC 1388) and Internet Control Message Protocol (ICMP) (RFC 792 and related RFC 1256) support the use of both broadcast and multicast addresses and Open Shortest Path First (OSPF) (RFC 1583 and RFC 2178, OSPF V2) uses multicast addresses.

As routers and hosts learn the IP addresses of their neighbors, or local routers in the case of hosts, they can then use IPOA to resolve the destination IP addresses of their neighbors, to ATM addresses, which enables them to generate point to point VCC’s.

Using MARS as a discovery mechanism in switched ATM networks greatly reduces the manual configuration required. The only information required are ATM addresses of the MARS and IPOA servers. If these servers are combined in a single device, and use shared VCs then configuration is reduced even further. Additionally, extensions to ILMI will allow automatic discovery of the ATM address of the MARS and ATMARP servers.

These components are the equivalent of today’s IP networks and can be easily constructed over a switched ATM network. The real advantage of this new topology is in supporting advanced applications, where traditional mechanisms are in effective.

Enterprise networks are now well positioned to implement IP networks over switched ATM services, either by employing ATM over existing leased lines, or by using increasingly available commercial ATM services. Alternatively, service providers can now leverage their ATM networks and offer flexible, outsourced IP services, hiding ATM complexities from the end user.

Applying IP Transport Over ATM

IP networks have traditionally delivered a "best effort" service which makes no distinction between IP datagram delivery irrespective of the application it is transporting. When bandwidth utilization is not an issue, this mechanism can be quite effective. With the rapid growth of the Internet, increasing use of web technology, and the emergence of applications which are delay sensitive, both traffic patterns, which directly affect bandwidth consumption, and QoS requirements on a per application basis, need to be considered.

Voice Over IP (VoIP) and to a slightly lesser extent, FAX transport over IP, are perhaps the most classic and increasingly widespread examples of applications which require enhanced QoS beyond best efforts delivery. Emerging video related applications will also require associated QoS characteristics. Perhaps the widest scope for these technologies is combining their applications to provide interactive capabilities, such as browsing a web site and establishing a voice call to a help desk to answer specific questions.

As the world wide web becomes an increasingly important business tool, company's will seek to provide an enhanced level of performance for improved access over best efforts for sales, marketing and support functions. While these capabilities may not require the same level of QoS amongst themselves, or require the guarantee of performance that VoIP does, superior performance over best efforts will become increasingly necessary.

Switched ATM networks overlaid with LAN Emulation, IPOA, MARS and MPOA can effectively support these applications today in a standards based manner.

Alternative Technologies

If ATM is deployed to the desktop and used as the ubiquitous layer 2 transport technology, then applications can directly request QoS, and indeed, it can be argued that LAN Emulation and IP become superfluous. The current investments in LAN and router technology dictate that the ATM everywhere scenario is an unlikely reality for the foreseeable future. The issue then remains of associating a given application with a particular QoS for LAN Emulation and MPOA. The IETF has already published the Resource Reservation Protocol (RSVP) which is designed to allow applications to request QoS from router networks. The complexity of the protocol and an inability to effectively operate on an end to end basis has restricted the use of RSVP. New and often proprietary mechanisms enable the use of Type Of Service (TOS) bits within the IP header to identify different classes of traffic. Currently limited to 3 bits, by IPv4 it remains to be seen whether this mechanism can be standardized and whether it is sufficient to meet varied demands.

The IEEE is also introducing specifications designed to enhance LAN protocols with traffic prioritization. The IEEE 802.3 p and q specifications provide some basic QoS mechanisms to enhance traffic capabilities in LANs.

Multi-Protocol Label Switching (MPLS) is perhaps the most significant area of new work. This is being driven by the IETF and it aims to enhance existing routing protocols by combining switching and routing capabilities. Routing is distinguished by examining every IP datagram and making a forwarding decision based on the source and destination addresses. Layer 2 switching is characterized by forwarding all traffic for a given destination based on some simple mechanism rather than examining each packet. In ATM relationships are established between incoming and outgoing VPI/VCIs through the network to establish a VCC. In frame relay the relationship is between incoming and outgoing DLCIs. MPLS uses the concept of labels which are established between neighboring routers and used to forward IP datagrams without having to examine the contents. MPLS is designed to run in conjunction with existing IP protocols, such as RIP and OSPF and is designed to run over a variety of wide area technologies, including ATM and frame relay.

Criticism has been leveled at MPOA that it does not scale due to the latency issues in establishing a shortcut connection. This potentially arises from the NHRP query/response if many subnets must be crossed, and from latency in the setup time for the ATM connection. From a practical perspective these issues are very much tied to overall network topology and equipment performance, which make the issues little different from those in MPLS.

Prior to establishing MPOA shortcuts, IP datagrams must traverse the traditional routed path through the network. Since heavily used flows can be assumed to have already established shortcut connections through the ATM network, more bandwidth and processing capability is available for data, where no flow has been established, and for NHRP messages themselves. This effectively removes the first issue of latency across many subnets.

The other aspect is call latency setup time in ATM switches through a network. Early generations of ATM switches had call setup rates of only a few hundred call per second, which is a potential limiting factor. Newer generations of switches have increased hardware capability and use highly optimized call processing stacks, from companies such as Harris & Jeffries to deliver performance in thousands of call setups per second.

MPLS attempts to combine the best attributes of MPOA with the strengths and experience of IP routing in the Internet. The current Internet routing mechanisms are based on identifying the topology of the network and transporting traffic by identifying the best possible path over that topology. One mode of MPLS operation, topology driven, uses the same mechanism and overlays labels on the topology. This provides increased performance by using labels rather than examining every IP packet, but still raises issues in identifying different QoS levels and how traffic paths are adjusted in the event of congestion.

MPOA derives its effectiveness by establishing shortcut connections based on traffic demands - it is a data driven topology. Connections are established through the network based on the currently available resources which offers a guaranteed level of performance. MPLS has a similar mechanism to support data driven applications.

MPLS is clearly targeted at large network applications, such as the Internet. Work is still to be completed in the IETF and the ultimate effectiveness of the protocol is yet to be proven in widespread applications, particularly as it relates to QoS. MPOA is a mature standard that has widespread application in many networks and with new generations of technology can scale to large networks. It can be effectively deployed today and with the inherent QoS capabilities of ATM can support new emerging applications which are reliant on QoS.

Summary

A complete set of standards now exists to deploy advanced IP services, requiring QoS, over switched ATM networks. LAN Emulation, IPOA, MARS and MPOA are the core capabilities that enable applications such as VoIP, and emerging interactive web sites to be effectively deployed.

These network capabilities can be implemented on a wide range of equipment and are being deployed as next generation Enterprise networks as well as being offered as outsourced IP based services by carriers and ISPs. MPOA and its related standards have widespread applicability today. Emerging applications and the fluid nature of traffic patterns demand that congested inflexible IP networks give way to advanced networks that utilize bandwidth based on user requirements, such as those that ATM and MPOA can deliver.

As Enterprises and service providers seek to enhance their IP networks, equipment manufacturers have an increasingly broad scope to implement LAN Emulation, IPOA, MPOA and MARS in their product range to provide a component of the overall standards based application, or a complete end to end range of equipment.

About Harris & Jeffries

Harris & Jeffries, Inc. is the industry's 'ultimate' source for networking software solutions and services. Its *Soft-ATM*™, *High-Performance Frame Relay,*™ and *UltraLinq*™ product lines provide comprehensive "carrier-class" Asynchronous Transfer Mode, Frame Relay, redundant, and interworking software capabilities, respectively. H&J's source code enables networking manufacturers, OEMs and integrators to achieve fast time-to-market with technically advanced products that provide market-leading performance. H&J's products reduce development costs and complexity and are in use by over 120 major networking companies worldwide. H&J can be reached at (781) 329-3200, or at www.hjinc.com.