IEEE
COMMUNICATIONS
SURVEYS
The Electronic Magazine of
Original Peer-Reviewed Survey Articles

MULTICAST FORWARDING OVER ATM: NATIVE APPROACHES

JOSEP MANGUES-BAFALLUY AND JORDI DOMINGO-PASCUAL,
TECHNICAL UNIVERSITY OF CATALONIA (UPC)

ABSTRACT

Multicasting is growing in importance as new multimedia applications are devised. Throughout this article, multicasting is understood as the efficient multipoint-to-multipoint transmission of information (in terms of network resource consumption) between the members of a group. Most multicast services have been designed up to now to work over connectionless environments. The approach adopted by connection-oriented networks has been to try to imitate these connectionless multicast schemes with the aim of supporting IP multicast or network-layer broadcast. However, these solutions present drawbacks in terms of delay or signaling overhead. The goal of native ATM multicasting is to provide multicast communications support by taking into account the characteristics of ATM. Therefore, the design philosophy of multicast must be rethought by making it more suitable for connection-oriented networks. Native ATM multicasting is based on mechanisms implemented at the switches to allow the correct ATM-layer multicast forwarding of information. These mechanisms seek to avoid the delay and signaling problems of current solutions, e.g., LAN emulation and IP multicast over ATM. This article provides a survey of the literature on the strategies that offer multicast communications in ATM environments, with special stress on native ATM multicast forwarding mechanisms. Other aspects, such as signaling, quality of service, traffic control, and routing, are not addressed in detail in this article.

Multicasting is growing in importance as new applications are devised. Throughout this article, multicasting is understood as the efficient multipoint-to-multipoint transmission of information (in terms of network resource consumption) between the members of a group. In particular, the focus is on the problems of deploying the multicast service in an ATM environment.

ATM has been widely deployed in current networks. However, such networks do not allow the efficient provisioning of multipoint-to-multipoint communications. This lack of efficiency occurs because current solutions seek to reproduce a connectionless environment over ATM, which is connection-oriented.

For connectionless environments, the applications developed using the TCP/IP protocol suite are mostly oriented to work over broadcast networks. The connectionless nature of IP makes broadcast networks a natural choice. In these networks, multicast forwarding is carried out very easily. A multicast packet with a group address is sent over the broadcast network and is only delivered to those hosts that have joined the group. This characteristic is not present in connection-oriented networks (e.g., ATM), where the sender must establish a connection with all group members before packets can be forwarded.

On the other hand, native approaches offer the multicast facility at the ATM level, and they therefore take its connection-oriented nature into account. Apart from this conceptual difference, ATM introduces further challenges to the multicasting problem due to quality of service (QoS) provisioning issues. All these unsolved problems make ATM multicasting a challenging field. These aspects are discussed in the following subsections.

MULTICAST PROTOCOL REQUIREMENTS

We could use the requirements of a generic multicast protocol ([1, 2]) to classify the various problems to be solved when offering multicast communications in general, and in particular, making use of ATM. The issues to be addressed may be divided into a control part, which is usually software, and a forwarding part, which is usually hardware. The separation of the two functions allows both areas to improve in parallel

This research work has been funded by the Spanish Ministry of Education under contract CICYT TEL99-1117-C03-03 (Project SABA-2: New Services for the Spanish Broadband Academic Network-2).

without depending on one another. In particular, ATM has proved to be a good switching technology in terms of scalability and switching speed, but some problems arise when attempting to offer multicast forwarding of cells through an ATM network.

The most important requirements for a multicast protocol, as stated in [1] and [2], are given below. Before transmitting information to a group, there must be mechanisms that allow some hosts to be grouped under the same group identifier without conflict. This is the function of the "multicast group address assignment" mechanisms.

Concerning group set-up, apart from the unambiguous identification of a group, there must be other protocols allowing the allocation of an address for setting up a group. They should also allow the members of a group to know the address of the group they wish to join.

Once the group has been initially established, membership management will allow hosts wishing to be part of a given group to join it, hosts wishing to end its membership to leave it, or to switch from one group to another.

Once the communication has been finished, some protocol must be defined to end the group communication, i.e., to carry out group tear-down.

The above points are related to group establishment and management and may be classified as connection establishment and maintenance, i.e., signaling, in a connection-oriented environment, or group establishment and maintenance in a connectionless environment. The following issues are more concerned with allowing the information transmitted among members of a group to reach the recipients.

Transport reliability is one of these points. Depending on the characteristics of the information transmitted to the group, reliability will take the form of error recovery at the receivers or retransmissions in case of losses. When retransmission is not possible due to time constraints, error recovery mechanisms will use redundant information to recover some losses. On the other hand, retransmission will be used when the focus is on the correctness of the information transmitted regardless of the time it takes to be transmitted.

Another aspect to consider is flow control. Its goal is to adapt to network load. It controls the information placed in the network per unit time and may therefore increase network efficiency by reducing the number of losses and consequent retransmissions.

Support for network heterogeneity is also required. Present networks are characterized by heterogeneous equipment at network nodes and end-systems, different network technologies, and different user requirements. This scenario makes adaptability a complex but necessary issue.

And last but not least, efficient packet forwarding strategies at network nodes must be developed to allow all the above requirements to be fulfilled. This article focuses on this issue.

Problems of Multicasting in ATM

Most present applications and technologies are designed to work over broadcast networks in order to fully exploit their potential. This is mainly due to the development of the Internet and its lateral technologies. However, the inherent connection-oriented nature of ATM introduces new challenges to the multicasting problem with respect to IP multicasting over broadcast networks. ATM is a non-broadcast multiple access (NBMA) technology. Consequently, multicast mechanisms cannot benefit from inherent broadcast facilities offered by a broadcast medium.

ATM introduces further challenges to the multicasting problem due to its inherent QoS provisioning and its connection-oriented nature. When establishing group communications, the routes to the members must be computed according to a requested QoS. Therefore, the signaling and routing protocol responsible for this (e.g., Private Network-to-Network Interface (PNNI)) will be more complex than those found in IP multicast networks, which are based on best-effort service. Furthermore, QoS must be enforced during connection establishment (through Connection Admission Control (CAC)) and during data forwarding (through Usage Parameter Control (UPC)). The problem is further complicated by the heterogeneous and dynamic nature of groups. These characteristics make it difficult to enforce QoS.

Focusing on forwarding, the mechanisms and protocols designed up to now, such as VC mesh and multicast server (see below), tend to imitate broadcast network characteristics over connection-oriented networks. However, such solutions present important drawbacks, e.g., with regard to signaling overhead and scalability.

The focus of this article is on forwarding. In this case, the main issue to be addressed is the cell-interleaving problem (Fig. 1). This problem appears at merge points of a shared tree through which multiple senders simultaneously transmit information packed into ATM Adaptation Layer 5 (AAL5) packets (or any other adaptation layer that does not have a multiplexing identifier inside the cell). These packets are usually longer than the payload of an ATM cell. Therefore they must be fragmented into cells if they are to be transmitted through an ATM network. The adaptation layer most widely used to carry out this process is AAL5. However, AAL5 does not have any multiplexing ID inside each cell indicating to which packet the cell belongs. This is not a problem for point-to-multipoint (or point-to-point) connections, but it is a problem for multipoint-to-multipoint (or multipoint-to-point) connections. The problem only appears at merge points. At a merge point, two or more input virtual circuits (VC) belonging to the same multicast connection are forwarded through the same output VC. Therefore, the virtual circuit identifiers (VCI) of the input cells are mapped to the same output VCI. In this way, cells belonging to different packets may become interleaved at the output VC. Since AAL5 places no identifier (ID) inside each cell, the end-system will not be able to tell if a cell belongs to one packet or the other.

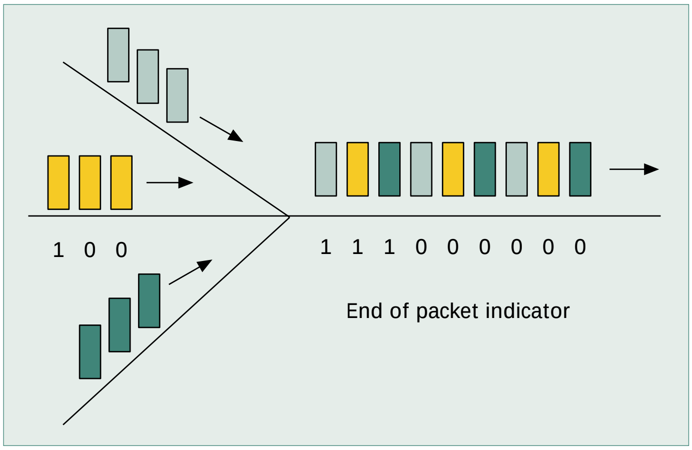


FIGURE 1. *The cell-interleaving problem.*

Figure 1 is an example of a merge point where three input VCs are forwarded through the same output VC. In AAL5, one bit in the header of each cell is used to indicate whether this is the last cell of the packet (=1) or not (=0). If cells with the same output VCI became interleaved, the end-system would consider that the first seven cells in the figure form a packet and the other two form a packet each.

The following sections of this article explain the diverse solutions that have been proposed to solve the cell-interleaving problem. The article is divided into two main parts. The first deals with solutions that seek to integrate IP and ATM in order to offer a multicast service. The second (and largest) part focuses on native ATM multicast proposals. A classification of these mechanisms, as well as a description of each of the proposals, is presented.

REVIEW OF CURRENT TECHNIQUES FOR MULTICASTING IN ATM: INTEGRATION OF IP AND ATM

This section provides a brief summary of current techniques being used to offer multicasting over ATM in the context of IP and ATM integration. This will allow their drawbacks to be considered and to justify the interest in exploring native forwarding approaches that will improve performance.

The problem of offering multicasting over ATM while attempting to maintain the interoperability with connectionless environments has been studied by many organizations. The IETF has studied this issue in the Internetworking over NBMA (ION) working group and the Integrated Services over Specific Link Layers (ISSLL) working group. At the ATM Forum, the LAN Emulation (LANE) and MultiProtocol over ATM (MPOA) working group and the Multiway BOF also studied the problem, as did the study groups SG-11 and SG-13 at ITU-T. This section provides a summary of this work.

There are two main models, namely, the *overlay model* and the *peer model*. Key elements about the overlay model are:

- Separate addressing schemes, i.e., one entity has an IP and an ATM address, which are not algorithmically related. Therefore
- IP to ATM address resolution is performed manually or dynamically via Address Resolution Protocols (ARP).
- Separate IP and ATM routing protocols and topologies.

A second model, the peer model, presumes that all devices support a single address space, maintain a single topology, and run a single routing protocol, all based on IP. This model has been realized through the development of MPLS mechanisms within the IETF. Earlier, the ATM Forum had worked on several initiatives to have homogeneous routing in heterogeneous environments, such as PNNI Augmented Routing (PAR) [3] and Integrated PNNI (I-PNNI) [4].

When offering the multicast service over ATM using the overlay model, there are two main options (IP multicasting over ATM [5] and LAN emulation [6, 7]). Interoperability with current equipment is possible without much effort, but the price paid is added overhead and, consequently, less efficiency. A brief description of both options follows.

IP MULTICASTING OVER ATM

IP multicasting over ATM [5] proposes a solution to support multicast communications in an ATM environment by using user-network interface (UNI) signaling version 3.1. Specifically, this makes use of point-to-multipoint connections to offer the multicast service. The first additional component is the Multicast Address Resolution Server (MARS). The MARS approach is built on the classical IP and ARP over ATM model, which provides an ARP and IP unicast service over ATM. MARS offers a layer 3 multicast service with the signaling of UNI 3.1. The main function of MARS is to map IP multicast addresses to a list of ATM addresses.

There are two options for forwarding data to the members of a multicast group when using the MARS model. These are the *VC mesh* strategy and the *multicast server* (*MCS*) strategy. The VC mesh strategy establishes a point-to-multipoint tree from each sender to the rest of the members of a multicast group. In this case, a point-to-multipoint control VC from MARS to the users of a group is required to notify membership changes. When such an event occurs, each point-to-multipoint connection from each sender to its receivers needs to be modified to add or remove the member. Therefore, the signaling load at end-systems is very high, and the time it takes for all the connections to be modified could be considerable. This strategy suffers from scalability problems, e.g., group stability latency is high. The main advantages of this approach are the distribution of the load across all the switches of the network, and the optimization of the paths, minimizing end-to-end latencies as a result.

In the multicast server strategy, multicast data distribution is centralized in a server. All the members of a group send data to this server. When a member wishes to transmit to the group, the packet is sent to the MCS by means of a unicast connection. All members may transmit at the same time; therefore, if AAL5 is used, the MCS must reassemble all cells belonging to the same incoming AAL5 packet data unit (AAL5 PDU) previous to its distribution through a point-to-multipoint connection. Cells belonging to the same packet must be transmitted together through the point-to-multipoint VC in order to avoid cell interleaving. The control connections between each member and the MARS and between the MARS and all the members are the same as in the previous option. The main advantages of this approach are: less resource consumption at end-systems is required, less signaling load between all entities involved in the multicast service, and less group stability latency. All these points are derived from the use of just one point-to-multipoint connection for data forwarding. However, it also presents some drawbacks: higher end-to-end latency due to buffering and non-optimal routes, and load concentration at the switches near the MCS.

LAN EMULATION (LANE)

IP multicasting over ATM solves the problem for IP, but the rest of the network-level protocols (IPX, IPv6, NetBIOS, DECNet, AppleTalk, etc.) are not considered. LAN emulation ([6, 7]) was designed to work in a multiprotocol environment, so that every application developed to work over any of the existing network-level protocols could benefit from the capabilities offered by ATM, and uses AAL5.

This technology is named after its main characteristic: it emulates a legacy LAN over a NBMA ATM network. In LANE, the ATM network is divided into different subnets called emulated LANs (ELANs). Intra-ELAN unicast traffic is forwarded by means of point-to-point connections and broadcast traffic is served by means of the broadcast and unknown server (BUS), which forwards multicast traffic to all the members belonging to the same ELAN as the sender. However, Inter-ELAN traffic is still forwarded through a router.

Focusing on multicast forwarding, the main components to consider are the broadcast and unknown server (BUS) that already existed in LANE 1.0 and the selective multicast server (SMS), which first appeared in the LANE Network-to-Network Interface (LNNI) specification of LANE 2.0. The forwarding procedure of both functional components is the same. The only difference is whether the message reaches all the members in an ELAN (as with BUS) or whether there are procedures to selectively choose a group of receivers in an ELAN (in SMSs).

Avoid cell-interleaving		VP switching		Allow multiplexing inside a VC			Multiple VC switching	
Buffering	Token control	Source ID	Packet ID	Added overhead	GFC	Signaling	Group ID	Packet ID
Cut-through (SEAM)		VP switching		SPAM				DMVC
CT-NC CT-T	SMART	VP-VC switching	DIDA	CRAM	Subchannel (WUGS)	VC-merge scheduler	FMVC	SMVC
Store-and-forward		VP switched CLIMAX		AAL5+ based CLIMAX				CVC

Table 1. *Classification of native ATM multicasting mechanisms.*

BUSs and SMSs establish point-to-multipoint connections with the server itself as a root and the receivers as leaves. The sender wishing to send multicast information sends it through a point-to-point connection to the server. When this information arrives at the server, it is stored in reassembly buffers and reassembled. Once the destination is obtained, each packet is segmented into cells and forwarded, avoiding cell interleaving.

In summary, the philosophy of multicast forwarding in LANE is the same as in MCS. As a consequence, the same advantages and drawbacks appear.

From all that we have seen in this section we may conclude that IP multicasting over ATM (overlay model) is problematic in terms of complexity, signaling overhead, and delay. On the other hand, a peer approach scales better and is less complex, but the cell-interleaving problem must still be addressed. Furthermore, as ATM networks evolve and incorporate MultiProtocol Label Switching (MPLS), it is likely that MPLS ATM label switch routers (LSR) will employ some of the mechanisms described below to address the cell-interleaving problem. All these points justify research into native ATM multicasting mechanisms. The following section describes the research that has been carried out in this field.

MULTICASTING IN ATM: NATIVE ATM MULTICASTING

In the context of this article, native ATM multicasting refers to the mechanisms implemented at the switches to allow the correct ATM-level forwarding of the information being interchanged by the members of a multicast group. That is, the cell-interleaving problem is solved without having to reassemble cells into AAL5 PDUs inside the network, unlike the multicast server (MCS) case [5].

The mechanisms presented in the above section solve the problem for legacy protocols by allowing them to interoperate in an ATM environment. Their technical approach is the imitation of a broadcast medium over a non-broadcast medium. As the focus is on interoperability, these mechanisms do not take full advantage of ATM characteristics, e.g., QoS provisioning and forwarding speed. If these characteristics are to be exploited, a different philosophy must be conceived.

Native ATM multicasting mechanisms aim to design a multicast strategy based on the ATM technology with no more restrictions, except the requirements all multicast protocols have. In this way, the capabilities of ATM could be fully exploited and efficiency in terms of resource consumption would increase.

Native ATM multicasting techniques provide solutions for offering true multipoint-to-multipoint connections by solving the cell-interleaving problem at the ATM level, i.e., without any reassembly inside the network. True multipoint-to-multipoint refers to those group connections using a unique shared tree for all the members in the group. A classification of these mechanisms is included in [8]. However, VC merging is used there as a particular case of the more generic native ATM multicasting techniques. (Sometimes, VC merging is also used in the literature as a synonym for native ATM multicasting.) Consequently, Table 1 presents this classification with the new notation and some other mechanisms that have recently appeared in the literature.

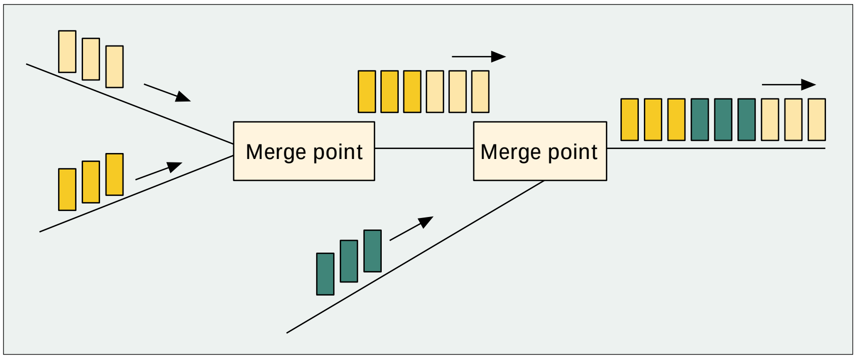
Techniques belonging to the first type solve the cell-interleaving problem by avoiding cells from different packets to be interleaved. They are generically referred to as strategies that *avoid cell-interleaving*. Buffering techniques reassemble all the cells of each PDU in separate buffers and forward them without mixing cells belonging to different buffers (or PDUs) ([9–11]). Shared Many-to-many ATM ReservaTions (SMART) uses a token-passing scheme to allow just one sender to put data in the multicast tree at any instant [12]. In the second type (*VP switching*), the virtual path identifier (VPI) identifies the connection, and the virtual circuit identifier (VCI) of the ATM cell header is used as the multiplexing ID (identifying the PDU [13] or the source [14, 15]). The third type of strategies allow *multiplexing inside the same VC*. This is done either by adding overhead in the transmitted data ([16, 17, 15]), by using the generic flow control (GFC) field in the header of the ATM cell [18], or by negotiating, at connection establishment, the sequence with which cells are going to be transmitted to the downstream node [19]. Finally, in *multiple VC techniques*, two or more VCIs are used as PDU IDs ([20, 8]) or group IDs [20] for the same compound multicast connection.

A more detailed description of these mechanisms may be found in the following sections.

AVOIDING CELL-INTERLEAVING

The main advantage of these approaches is their scalability in terms of number of groups. There can be as many multicast connections as there are VCs available, because only one VC is used for all the traffic of the group.

Buffering — The first type of technique solves the cell-interleaving problem by avoiding cells from different packets to be interleaved. This is referred to as cut-through (CT) forwarding. Simple and Efficient ATM Multicast (SEAM) [9] is one example of these techniques. It buffers the cells of a packet until no other cells are being forwarded to the same output VC. This buffering, when carried out at each switch in the path, presents the additional effect of increasing burstiness, latency, and cell-delay variation (CDV); as a result, traffic characteristics may be violated. The term cut-through forwarding means that the forwarding of a PDU starts when the first cell of the PDU arrives if the outgoing VC is idle and continues until the last cell of the PDU arrives. Therefore, a



■ **FIGURE 2.** *The store-and-forward strategy avoids cell-interleaving by buffering all the cells of a packet prior to its transmission.*

slow source could block the rest of the sources for significant time intervals. The store-and-forward (SF) proposal follows the same idea, but in this case, the first cell of a packet is not forwarded until all the cells of that packet have been buffered (Fig. 2). Once the last cell arrives, all the cells of that packet are buffered together in the output buffer where they wait to be transmitted consecutively. This latter approach seems to be the most likely implementation when ATM is used in Multi-Protocol Label Switching (MPLS) environments [10]. In this article, we will jointly refer to SEAM and store-and-forward as buffering techniques.

An extension of buffering techniques for providing some kind of quality of service (QoS) classes is briefly discussed in [21]. The proposal of the authors is to use different output buffers for traffic requiring different QoS. Each output buffer is assigned a different VC, and a cell-level scheduling mechanism is responsible for interleaving cells of different classes in order to minimize traffic distortion.

A stability study of CT strategies may be found in [11]. The authors argue that CT techniques using a round-robin scheduling discipline to assign the buffers are not stable when the packet rates of all input VCs are not the same. Instability means that the content of the buffers grows without any bound. Some variations are proposed to solve this problem. Cut-through-on-no-contention (CT-NC) does not forward a partially arrived packet if there is a complete one in another queue. CT-NC combines the delay reduction of traditional cut-through and the stability of store-and-forward.

Cut-through-with-thresholds (CT-T) allows its queues to dynamically change from CT to SF. For any given queue, there are three working options. First, the queue performs CT forwarding while the number of complete packets in the other queues of the same connection is below threshold H1. Second, if the number of packets is above H1, it performs SF forwarding for this queue. Third, it returns to CT operation when the number of packets in the other queues is below another threshold, H2.

The results obtained in their simulations show a significant improvement in delay for CT-NC with respect to CT. They also show, in general, a slight improvement with respect to SF, though in some cases this improvement could be greater.

The main advantage of buffering strategies is their simplicity. Though these techniques may allow easier implementation when compared to the other types, their main drawback is the buffering requirements at input queues of the switch. This buffering, when carried out at each switch in the path, presents the additional effect of increasing burstiness, latency, and CDV with the result that traffic contract may be violated. As a consequence, their main application would be data transmissions, and not real-time transmissions. The solution proposed in [21] to offer some kind of QoS to store-and-forward does not entirely solve this issue. Problems derived from buffering will remain for traffic belonging to the same class. Therefore, if the class granularity is not very small, i.e., if there are not many different buffers, one could expect that all the multimedia traffic will pass through the same buffer and its traffic parameters will therefore be distorted. Furthermore, if a large number of classes are defined to allow cell interleaving between classes, VC consumption will increase. Therefore, the scalability advantage claimed by buffering strategies over other strategies is diminished.

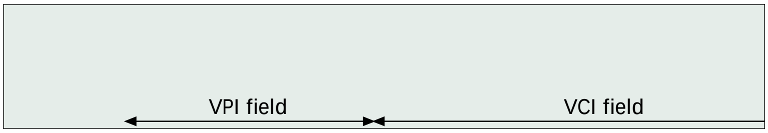
Token Control — In Shared Many-to-many ATM Reservations (SMART) [12] cell-interleaving is avoided by means of a token-passing protocol that allows only one sender to put information in the shared tree at any given time. In this case, the shared tree is accessed as if it were a shared medium. This mechanism allows the enforcement and accomplishment of the traffic contract, because enforcing the contract of the group at any given time corresponds to the enforcement of the sending end-system.

The main advantage of SMART is that the management of the QoS offered to the whole group is reduced to solving the problem for the sender that is transmitting in a given instant.

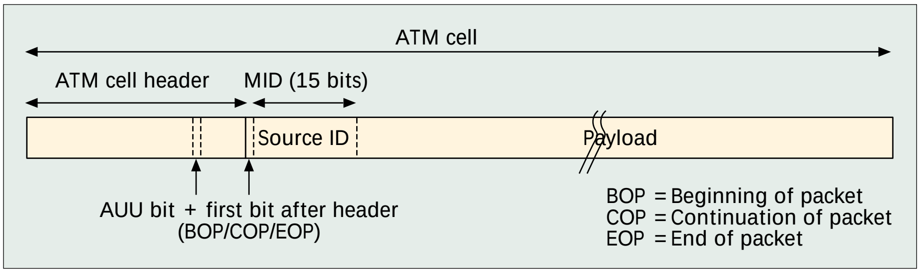
But SMART is a complex protocol because all switches must interchange resource management cells in order to allow the token to move from sender to sender, which imposes a considerable overhead. The mechanism is further complicated if more than one tree has to be managed simultaneously in order to allow some senders to send to the group at the same time. Therefore, complexity in group management leads to scalability problems.

VP SWITCHING

In VP switching techniques, the VPI identifies the group and the VCI is used as the multiplexing ID. A further subdivision differentiates between the VCI that identifies just



■ **FIGURE 3.** *VP switching.*



■ **FIGURE 4.** *Simple Protocol for ATM Multicasting (SPAM): cell format.*

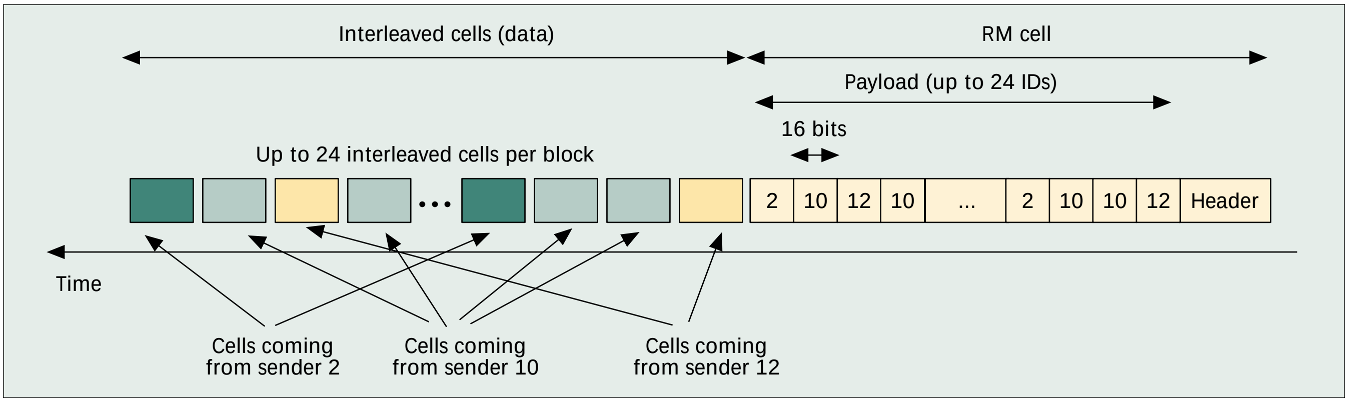


FIGURE 5. *Cell Re-labeling At Merge Points (CRAM): block format.*

the packet and the VCI identifying the source. Dynamic IDentifier Assignment (DIDA) [13] follows the former scheme, while [14] deals with the latter and proposes a modification, which is named VP-VC switching. The modification seeks to combine the advantages of what the authors call VP switching and VC switching. VP switching in [14] imposes a globally unique VCI identifying the sender. In addition, what the authors call VC switching corresponds to a strategy with the VCI value being changed at each switch. Therefore, additional mechanisms are required to identify the sender. The VCI mapping is static and once established, it lasts until there are no more cells coming with a given input VCI. VP-switched CLIMAX (CelL-Interleaved Merged ATM conneXions) [15] is another VP switching mechanism using source ID in the same way as in [14].

Though these strategies could be implemented with small or no modifications to current switches, their main drawback is a lack of scalability in terms of the number of groups that can be established. The VPI field has just 8 bits at the user-network interface (UNI) and 12 at the network-network interface (NNI). Lack of group size flexibility is another disadvantage. For instance, in DIDA all the VCIs of a VP are assigned to a group even if the group is small, because the VP identifies the connection. Thus, having 2^{16} identifiers in a group is the smallest granularity. Moreover, the advantages of using packet IDs as multiplexing IDs is not fully exploited in DIDA, because the identifiers are of a fixed and large size. As a consequence, the efficiency in ID consumption due to using packet IDs is lost due to a large number of IDs being unused for a given connection.

The utilization of the VPI field of the cell header as group ID may also represent a problem if carriers attempt to use it for other purposes.

Another drawback for VP switching source ID strategies is that the implementation of Early Packet Discard (EPD) is difficult because switching is carried out using the VPI, without keeping any state information for the VCIs inside the VPI.

ALLOWING CELL MULTIPLEXING INSIDE A VC

The underlying philosophy of these strategies is to use just one VC for each multicast connection. There are three main ways to do this: by adding extra overhead to the cell to carry the multiplexing ID; by using the GFC field to carry the multiplexing ID; or by negotiating the multiplexing order of cells through signaling. The following subsections describe each of these strategies in more detail.

Added Overhead — In this group, we classify those techniques that propose a modification of AAL5, and particularly its segmentation and reassembly PDU (SAR-PDU), by adding an extra field that carries multiplexing information for each cell. Again, this field could be used to identify the packet or the sender. In the latter case, a global ID assignment is needed. Simple Protocol for ATM Multicasting (SPAM) [16] is an example of these techniques that uses per source IDs (Fig. 4). The same approach is followed by AAL5+-based CLIMAX [15].

We have grouped together several techniques as *added overhead* to make a more generic definition. Therefore, we include Cell Re-labeling At Merge Points (CRAM) [17] as one particular case. In CRAM, the multiplexing IDs for each cell are somehow related with the source that transmitted the cell and these IDs are locally remapped at each merge point. However, unlike in SPAM the multiplexing IDs are not carried inside each cell. In this case, they are carried in resource management (RM) cells, which precede a block of interleaved cells (Fig. 5). At the first merge point, an RM cell (and its corresponding block) is built by assigning source IDs to all the cells from the same input VC that are to be transmitted through the same output VC. When a block arrives at a merge point, the multiplexing IDs are extracted from the cell, they are remapped, and a new block is built by buffering the cells from all the merging blocks.

The problem of flexibility of group size also arises in SPAM and CRAM, where multiplexing identifiers have fixed length. As in these cases the VCI (not the VPI) identifies the connection; the flexibility problem as such is not as important as in the other cases. In addition, these techniques include an overhead that reduces the bandwidth available to user information on a given link.

Furthermore, in SPAM the switch needs to perform AAL processing by looking at the multiplexing ID in the SAR-PDU, a task that, in principle, corresponds to the end-system. In the case of CRAM, RM cells are processed in the switch. This processing should be added to the table look-up operations. Moreover, the process of analyzing and creating a block consisting of the RM cell followed by the cells indexed by it needs some buffering, and it would affect latency, CDV, and burstiness, though this effect is not as harmful as in buffering mechanisms.

Another drawback is that the multiplexing ID in the payload is not protected by the HEC error detection/correction capability, which only covers the cell header. Therefore, an error in a multiplexing ID may affect cells from sources using different IDs.

Generic Flow Control (GFC) used as Multiplexing ID — Another mechanism that cannot be classified within one of the three types above is proposed in [18]. The multiplexing identifier, called the subchannel identifier, is carried in the GFC field of the ATM header. That makes possible 15 simultaneous packets in a switch (one ID is left to indicate an idle subchannel). It identifies a burst or packet formed of data cells delimited by RM cells. The subchannel ID is dynamically changed at each switch for each packet. With this strategy,

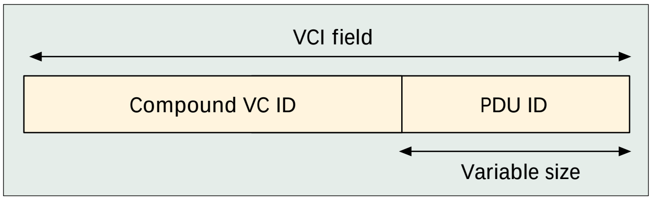


FIGURE 6. *Compound VC (CVC).*

there is no extra overhead due to multiplexing information, but there is some due to the insertion of RM cells, though it could be avoided if no reliability concerns are imposed and the end-of-packet indication of AAL5 is used to differentiate the packets. This mechanism is implemented in the Washington University GigaSwitch (WUGS). Turner also proposes using more than one VC if 15 IDs are not enough to cope with all the traffic, and establishing a block of VCs (one per subchannel ID) that are shared by all senders solves interoperability with non-subchannel switches.

One of the main problems of this mechanism is the utilization of the GFC field, which means that it will not be available at the UNI for user access, e.g., in a passive optical bus, or for other purposes. Moreover, it could limit the potential widespread use in NNI interfaces, as there is no GFC in these interfaces.

Signaling — In this case, the cell-interleaving problem is solved by negotiating at connection establishment the order in which cells from the different queues at an upstream node will be transmitted downstream. In [19] there is presented a VC-merge capable scheduler with a two-level hierarchical structure based on the output module of a per-VC queuing ATM switch. These switches determine the order in which each VC is served by means of a scheduler. This scheduler is preserved in their proposal. The second level is introduced through the concept of the virtual queue. A virtual queue consists of a sequencer and a set of subqueues for different incoming VCs having the same outgoing (merged) VC. Recall that a VC corresponds to a multicast connection in strategies that allow cell-interleaving inside a VC.

Therefore, if both the sequencer at the upstream node and the de-sequencer at the downstream node know this order, cells belonging to different packets may be correctly separated. The authors argue that their proposal allows the provisioning of per-flow QoS to both VC-merge and non-VC-merge connections. Furthermore, their work shows that it is fair and that it presents a bounded delay for the worst case.

However, the operation of this scheduler may not be correct under low traffic conditions. The snooping mechanism used to solve this problem may solve the sporadic lack of a cell but not the continuous lack of cells from different virtual queues. Moreover, the snooping mechanism requires a modification in the utilization of the VPI, as the authors propose the use of the least significant bit of the VPI to mark snooper cells. Finally, the connection establishment procedure becomes more complicated due to the negotiation of the sequencing of cells.

As a general remark, the main advantage of strategies that allow multiplexing inside a VC is their scalability in terms of number of groups, because just one VC per group is used. However, this scalability comes at the price of extra overhead, limited group size, or extra signaling complexity.

MULTIPLE VC SWITCHING

A new proposal that did not appear in the original classification [15], called multiple VC switching, has been added. Multiple VC switching shares some characteristics with "VP switching" and "allow cell multiplexing."

Multiple VC switching is similar to VP switching in that the multicast connection is not a single VC, but multiple VCIs are assigned to the same connection. Multiple VC switching techniques are similar to SPAM and the like because cell multiplexing is allowed in the connection, but in this case not in a single VC but in a compound connection.

The purpose of these strategies is to reduce the buffer requirements at the cost of increased utilization of VPI/VCI space. This should not involve scalability problems as there are usually far fewer connections than VCIs in any given link.

There is a further subdivision of these kinds of mechanisms. Some mechanisms use the VCIs as packet IDs and others use them as identifiers for a group of senders.

Group ID — In Fixed Multiple VC-merge (FMVC) [20] each sender is assigned a fixed connection ID at a merge point. Some senders share this ID, which will consequently identify a group of senders. Therefore, this mechanism allows the interleaving of cells belonging to different groups.

FMVC uses the multiplexing IDs more efficiently than source ID mechanisms. However, efficiency may be significantly reduced in cases where there is a very active group and many inactive groups, because the active group would run out of IDs while the others are not using theirs.

Packet ID — In this case, the VCIs assigned to the multicast connection are used as packet IDs. In Dynamic Multiple VC-merge (DMVC) [20], each switch maintains a set of unassigned IDs at each outgoing link pertaining to a given connection. When the first cell of a packet arrives, an ID is assigned and is maintained for all the cells of the packet. When there are no free IDs, cells are stored until one ID becomes free.

Selective Multiple VC-merge (SMVC) [20] is an enhancement to FMVC and DMVC for dealing with store-and-forward (SF) VC-merge. When using two IDs, one is used for SF forwarding and the other for cut-through (CT) forwarding in the following way. If an entire packet has been received, it is forwarded through the SF ID. Cells of partially arrived packets use the CT ID.

These mechanisms do not represent an improvement in terms of throughput with respect to SF VC-merge, but they have great impact on those using CT. Furthermore, DMVC obtains a buffer reduction of 80 percent with two connection IDs, when compared to cut-through. Finally, some simulation results also showed that SMVC requires 50 percent less buffer then SF if two IDs are used.

Compound VC (CVC) [8] is another mechanism using packet IDs. The name "compound VC" comes from the utilization of more than one VCI to a single connection. However, the switch treats these VCs as if they were a single compound connection. In CVC, the multicast group is associated with a group of adjacent VCs (or compound VC). The VCI field in the ATM cell header is divided into two parts. The first part corresponds to the CVC ID; the other part carries the PDU ID (Fig. 6). Both parts are locally mapped at the switches. At CVC connection establishment, a static mapping of CVC IDs is assigned and will not be modified during the connection, while the PDU ID is dynamically modified each time the first cell of a new PDU arrives at the switch.

The size of each part is negotiated at connection establishment. This feature allows more flexibility matching different group sizes and traffic characteristics. A unique entry in the switching table is needed to switch the traffic of the entire group using a mask. Some aspects of the negotiation of the size of each part depend on the packet loss probability, and traffic and group parameters are presented in [22] and [23].

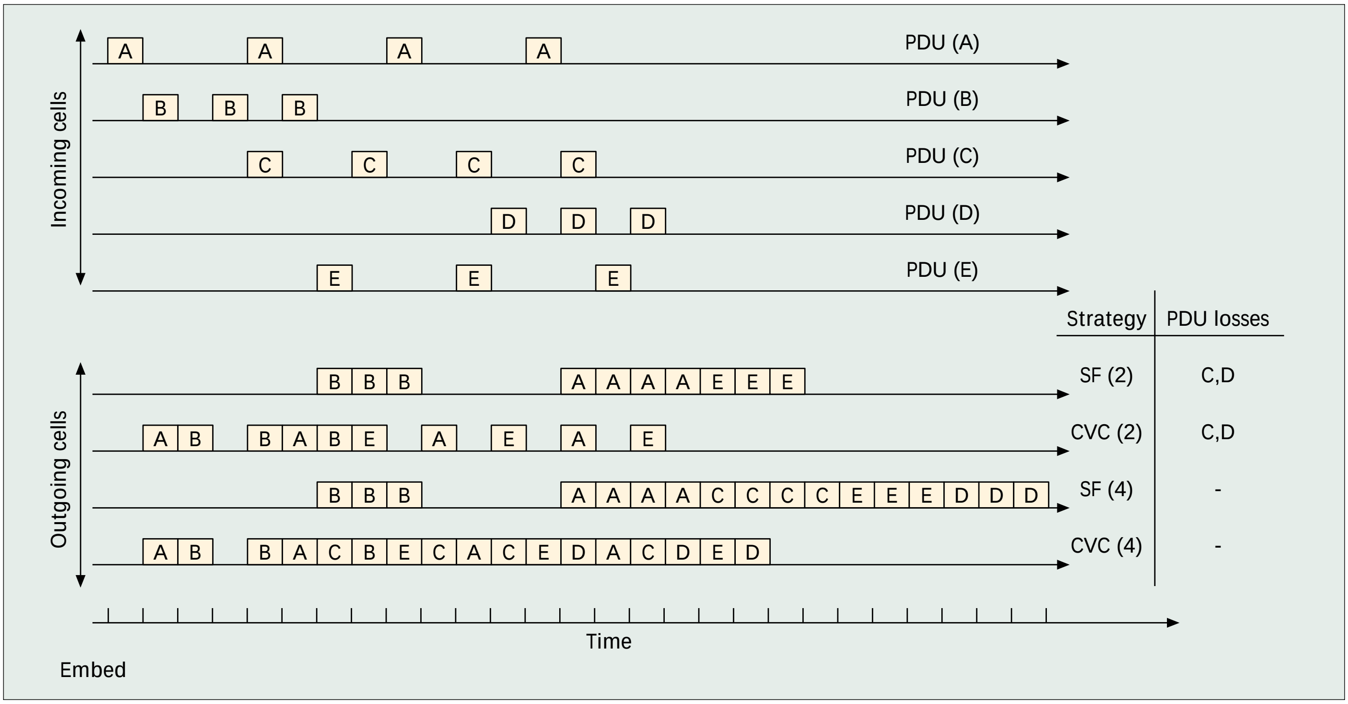


FIGURE 7. *Comparison of some native ATM multicasting mechanisms.*

Unlike for DMVC, in CVC there is no buffering of cells that did not acquire an ID. The authors made this design choice because their aim was to build a mechanism able to deal with real-time connections as efficiently as possible. This is the reason why buffering is avoided as much as possible.

Assigning PDU IDs instead of sender IDs allows a better usage of the low number of bits in the ATM cell header. Moreover, multimedia communications require traffic characteristics to be respected. This is only attained if cells from different PDUs are interleaved, as in CVC. This strategy is also scalable in terms of the number of groups when compared with VP switching strategies, because more bits could be assigned to the Compound VC ID. Furthermore, the size of each group could be negotiated, which offers added flexibility. Another advantage is that CVC includes some of the previous mechanisms as particular cases depending on the size of the mask being chosen. As a consequence, with CVC one may still apply them in those specific cases for which they are appropriate. Finally, the utilization of part of the VCI field to carry the multiplexing identifier presents the additional advantage of benefiting from the header error correction mechanisms of ATM.

The main drawback of this approach is the switch implementation complexity and the need for a new signaling procedure to establish the CVC.

COMPARISON OF NATIVE ATM MULTICASTING MECHANISMS

There are several approaches to provide native multicasting in ATM. Their main advantages and drawbacks are presented in Table 2.

Another (more graphical) comparison of the behavior of some of these mechanisms is presented in Fig. 7. We have chosen SF and CVC as the most representative mechanisms because avoiding cell interleaving using token control leads to highly complex management and usage of many VCs. The mechanisms based on the assignment of identifiers (VPI or VCI) per source suffer from scalability problems when a large number of groups must be maintained. The mechanisms that allow multiplexing within a VC add extra overhead (modifying AAL5 PDU or using RM cells), or are limited to the UNI (using the GFC field of the ATM cell). For these reasons and in view of the comments in the above section, the most interesting comparison is between SF, because it is the approach most likely to be implemented, and CVC, because it shares some of the characteristics of the other mechanisms.

Figure 7 presents an example of the behavior of SF and CVC. In the example, there are five incoming PDUs (A through E). The upper part — labeled as incoming cells — represents the time arrival instants of the cells belonging to each PDU. For instance, PDU A is composed of four cells arriving with a timing of one every four time units.

The part of the figure labeled as outgoing cells presents the exit instants of the cells above for each mechanism after having passed through processing in the switch. Each horizontal line represents the exit instants for the mechanism specified in the column "Strategy." The number in parentheses represents the number of IDs for CVC or the number of reassembly buffers for store-and-forward (SF) buffering techniques. The column labeled "PDU losses" gives the PDU losses when applying each of the mechanisms.

The following assumptions have been made:

- If the first cell of a PDU is discarded, so are the remaining cells, as in early packet discarding strategies.
- For comparison purposes only, we have chosen two and four reassembly buffers for SF. It does not claim to be a general case in practice.
- The initial state of the system corresponds to empty buffers.
- For the sake of clarity, we have represented a one-cell delay for the processing time of the switch in all cases.

We will focus on the comparison between SF and CVC. Analytical studies and simulation results show that they may behave similarly in terms of throughput [8]. The main difference is the behavior in the pattern of the outgoing traffic. SF techniques include buffering and tend to transmit the cells of the same PDU at the peak rate while CVC simply multiplexes the arriving cells without increasing either the end-to-end

Type	Subtype	Advantages	Drawbacks
Avoid cell-interleaving	Buffering	–Scalability (number of groups) –Easy implementation –No additional overhead for mux IDs	–Buffering requirements –Increased burstiness, latency, and CDV
	Token control	–Scalability (number of groups) –Easy group QoS provisioning	–Difficult connection management –Signaling overhead
VP switching	Source ID	–Easy implementation (no hardware modification) –No additional overhead for mux Ids	–No scalability (number of groups) –No flexible mux ID size –Carriers may use VPI –Difficult EPD implementation
	Packet ID	–Easy implementation (slight hardware modification) –No additional overhead for mux IDs –Efficiency in mux ID usage due to packet ID	–No scalability (number of groups) –No flexible mux ID size –Carriers may use VPI
Multiple VC switching	Group ID	–Better ID usage than strategies with fixed size mux ID –No additional overhead for mux Ids –Group size flexibility	–Increased VPI/VCI space utilization –Inefficient operation under some conditions
	Packet ID	–Better ID usage than strategies with fixed size mux ID –No additional overhead for mux IDs –Efficiency in mux ID usage due to packet ID –Traffic characteristic unchanged (QoS) –Mux ID size flexibility	–Increased VPI/VCI space utilization –Complex implementation
Allow multiplexing inside a VC	Added overhead	–Scalability (number of groups) –Reduced buffer requirements	–High overhead –No ID size flexibility –AAL or RM processing in the switch –Mux IDs not protected by HEC –(CRAM) buffering -> increased latency, CDV, and burstiness –(SPAM) modification of standard AAL5
	GFC	–Scalability (number of groups)	GFC only available at the UNI
	Signaling	–Scalability (number of groups) –Per-flow QoS (bounded delay) –Fairness	–Bad operation in low traffic conditions –Connection establishment complexity –Uses one bit of VPI

Table 2. *Advantages and drawbacks of native ATM multicast forwarding.*

delay or the cell delay variation. This aspect may not be very important for data transmission but may significantly distort the traffic characteristic of multimedia communications.

The price paid by CVC is a higher VC (or ID) consumption than SF (and buffering techniques in general), but this is not a major problem for a local or access network environment.

However, VC consumption is not that high with respect to buffering strategies if they are modified as in [21] to offer some kind of quality of service (QoS). If we wish to attain the QoS granularity as in CVC, we should use many VCs. As a consequence, buffering techniques will end up by having the same scalability concerns as CVC due to VC consumption. Furthermore, QoS is not respected as in CVC because reassembly buffers remain, and that will introduce some delay inherent to buffering techniques that is not present in CVC. In normal scenarios data will pass through many switches, and the delay and CDV will accumulate.

SUMMARY

A review of forwarding strategies offering multicasting over ATM has been presented. Two main groups are established: those strategies that attempt to solve the interoperability between IP and ATM, and native ATM multicasting mechanisms. The main focus of this article is on the latter mechanisms.

The importance of having mechanisms for offering native ATM multicasting becomes clear when we consider what ATM can offer to multicast connections. This technology was designed to provide high switching and transmission capacity, while offering the requested quality of service (QoS) for each connection. Multicast connections will usually carry multimedia information with real-time requirements. Consequently, the provisioning of a mechanism that fully exploits the characteristics of ATM, while offering an efficient multicast service, would be an important step toward widespread multicast deployment. In particular, the emerging MultiProtocol Label Switching (MPLS) implementations of MPLS ATM label switched routers (LSR) might benefit from the multicast forwarding approaches presented in this article.

If we take a look at current ATM networks, AAL5 is the most commonly used adaptation layer, including multimedia and computer supported cooperative work (CSCW) applications. If AAL5 is to be used in multicast connections, the most important issue to solve for multicast forwarding mechanisms is the cell-interleaving problem. Native ATM multicasting includes those mechanisms that solve such problems at the ATM layer, without the need for higher layer processing at the switches or in servers. Four main groups of solutions are explained, namely: avoiding cell interleaving, VP switching, Multiple VC switching, and allowing multiplexing inside a VC.

In the first solution, there are two alternatives:

- Buffering strategies are simple to implement but modify traffic characteristics, so their application to multimedia could be restricted.
- Token control mechanisms have simple traffic contract management but complex connection management.

The main drawback of VP switching strategies is the lack of scalability in terms of the number of groups, especially when the identifiers are assigned to the sources. Multiple VC switching presents flexibility in group size, no extra overhead, and respects traffic characteristics at the price of a more complex implementation of the switch.

Finally, the last type presents three alternatives:

- Added overhead shows scalability in terms of number of groups, but at the price of extra overhead.
- GFC uses the GFC field, only available at the UNI, which could limit its potential deployment in a wide area.
- Signaling determines the order in which cells from VC-merged connections are going to be transmitted.

REFERENCES

[1] S. Fahmy *et al.*, "A Survey of Protocols and Open Issues in ATM Multipoint Communications," http://www.cis.ohio-state.edu/~jain/ papers/mcast.htm

[2] R. Braudes and S. Zabele, "Requirements for Multicast Protocols," IETF RFC 1458, May 1993.

[3] ATM Forum, "PNNI Augmented Routing (PAR), Version 1.0," document AF-RA-0104.000, Jan. 1999.

[4] ATM Forum, "Integrated PNNI (I-PNNI) v1.0 Specification," document btd-pnni-ipnni-01.00, Dec. 1996.

[5] G. J. Armitage, "IP Multicasting Over ATM Networks," *IEEE JSAC*, vol. 15, no. 3, Apr. 1997, pp. 445–57.

[6] ATM Forum, "LAN Emulation over ATM Version 2 - LUNI Specification," document AF-LANE-0084.000, July 1997.

[7] ATM Forum, "LAN Emulation over ATM Version 2 - LNNI Specification," document AF-LANE-0112.000, Feb. 1999.

[8] J. Mangues-Bafalluy and J. Domingo-Pascual, "Compound VC Mechanism for Native Multicast in ATM Networks," *Proc. 2nd Int'l. Conf. ATM (ICATM '99)*, Colmar (France), June 1999, pp. 115–24.

[9] M. Grossglauser and K. K. Ramakrishnan, "SEAM: Scalable and Efficient ATM Multicast," *Proc. IEEE Infocom '97*, Kobe (Japan), Apr. 1997, pp. 867–75.

[10] E. C. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol Label Switching Architecture," IETF Draft, draft-ietf-mpls-arch-06.txt, Aug. 1999.

[11] A. L. Stolyar and K. K. Ramkrishnan, "The Stability of a Flow Merge Point with Non-Interleaving Cut-Through Scheduling Disciplines," *Proc. INFOCOM '99*, 1999, pp. 1231–38.

[12] E. Gauthier, J. Y. Le Boudec, and P. Oeschlin, "SMART: A Many-to-Many Multicast Protocol for ATM," *IEEE JSAC*, vol. 15, no. 3, Apr. 1997, pp. 458–72.

[13] J. Calvignac *et al.*, "Dynamic IDentifier Assignment (DIDA) for Merged ATM Connections," ATM Forum/97-0504, July 1997.

[14] R. Venkateswaran *et al.*, "Support for Multiway Communications in ATM Networks," ATM Forum/97-0316, Apr. 1997.

[15] M. Baldi *et al.*, "A Comparison of ATM Stream Merging Techniques," *Proc. IFIP High Performance Networking (HPN'98)*, Vienna, Sept. 1998, pp. 212–27.

[16] S. Komandur and D. Mossé, "SPAM: A Data Forwarding Model for Multipoint-to-Multipoint Connection Support in ATM Networks," *Proc. 6th Int'l. Conf. Comp. Commun. Networks (IC3N)*, Las Vegas, Sept. 1997.

[17] S. Komandur, J. Crowcroft, and D. Mossé, "CRAM: Cell Relabeling at Merge Points for ATM Multicast," *Proc. IEEE Int'l .Conf. ATM (ICATM '98)*, Colmar (France), June 1998.

[18] J. Turner, "Extending ATM Networks for Efficient Reliable Multicast," *Proc. Wksp. Communication and Architectural Support for Network-Based Parallel Computing*, Springer Verlag, Feb. 1997.

[19] H. K. Chow and A. Leon-Garcia, "VC-Merge Capable Scheduler Design," *Proc. IEEE ATM Wksp. '99*, 1999, pp. 153–60.

[20] R. Venkateswaran *et al.*, "Improved VC-Merging for Multiway Communications in ATM Networks," *Proc. 7th Int'l. Conf. Comp. Commun. and Networks (IC3N '98)*, 1998, pp. 4-11.

[21] I. Widjaja and A. I. Elwalid, "Performance Issues in VC-Merge Capable Switches for IP over ATM Networks," *Proc. INFOCOM '98*, San Francisco (USA), Mar. 1998, pp. 372–80.

[22] J. Mangues-Bafalluy and J. Domingo-Pascual, "Analysis of the Requirements for ATM Multicasting Based on per PDU-ID Assignment," to appear in *Proc. IFIP-TC6 Networking 2000*, Paris (France), May 2000.

[23] J. Mangues-Bafalluy and J. Domingo-Pascual, "Performance Issues of ATM Multicasting Based on Per-PDU ID Assignment," to appear in *Proc. IEEE Int'l Conf. Commun. (ICC '00)*, New Orleans (USA), June 2000.

BIOGRAPHIES

JOSEP MANGUES-BAFALLUY (jmangues@ac.upc.es) is an assistant professor at the School of Computer Science of Barcelona, the Universitat Politècnica de Catalunya (UPC). He received his degree in telecommunications engineering from UPC in 1996. Since then, he has been working in the broadband communications research group in the computer architecture department of the same university, where he is currently working toward a Ph.D. He is also working as a researcher in the Advanced Broadband Communications Center (CCABA), where he has participated in several Spanish and European projects. He is also involved in the design and deployment of the i2CAT next-generation Internet infrastructure. His main research interests are multicasting over ATM and MPLS.

JORDI DOMINGO-PASCUAL (jordi.domingo@ac.upc.es) is an associate professor of computer science and communications at the Universitat Politècnica de Catalunya (UPC) in Barcelona. There, he received an engineering degree in telecommunication (1982) and a Ph.D. in computer science (1987). Since 1983 he has been working in the computer architecture department. His research topics are broadband communications and applications. Since 1988 he has participated in RACE projects (Technology for ATD and EXPLOIT), several Spanish broadband projects (PLANBA, AFTER, TR1, IRMEM, and SABA), ACTS projects (INFOWIN, MICC, IMMP), and IST projects (LONG). Since 1995 he has been a researcher at the Advanced Broadband Communications Center (CCABA) at UPC, which participated in the Spanish National Host and in the PLANBA demonstrator. He is currently in charge of the i2CAT next-generation Internet infrastructure.