[Hein93] B. Heinrichs, K. Jakobs, A. Carone, “High Performance Transfer Services to Support Multimedia Group Communications,” *Computer Communications*, Vol. 16, No. 9, Sept. 1993.

[Huan94] C. Huang, P. K. McKinley, “Communication Issues in Parallel Computing Across ATM Networks,” *IEEE Parallel & Distributed Technology*, Winter 1994, pp. 73-86.

[Jian92] X. Jiang, “Routing Broadband Multicast Streams,” *Computer Communications*, Vol. 15, no. 1, 1992.

[Kaas89] M. F. Kaashoek, A. S. Tanenbaum, S. F. Hummel, H. E. Bal, “An Efficient Reliable Broadcast Protocol,” *Operating System Review*, Vol. 23, No. 4, Oct. 1989, pp. 5-19.

[Kadi94] J. Kadirire, “Minimising Packet Copies in Multicast Routing by Exploiting Geographic Spread,” *Computer Communication Review*, 1994, pp. 47-62.

[Kimb95] B. G. Kim, P. Wang, “ATM Network: Goals and Challenges,” *Communications of ACM*, Vol. 38, No. 2, Feb, 1995, pp. 39-44.

[Komp93] V. P. Kompella, J. C. Pasquale, G. C. Polyzos, “Multicast Routing for Multimedia Communication,” *IEEE Trans. on Network*, Vol. 1, No. 3, June 1993.

[Laub94] M. Laubach, “Classical IP and ARP over ATM,” *RFC1577*, HP Lab., Jan. 1994.

[Math94] L. Mathy, O. Bonaventure, “QoS negotiation for Multicast Communications,” Proceedings on Multimedia Transport and Teleservices, Vienna, Austria, Nov. 1994.

[Maye92] E. Mayer, “An Evaluation Framework for Multicast Ordering Protocols,” *COMM’92*, MD. USA, Aug. 92, pp.177-187.

[Mcki90] P. K. McKinley, W. S. Liu “Multicast Tree Construction in Bus Based Network,” *Communications of the ACM*, Vol. 33, No. 1, Jan. 1990, pp.29-42.

[Moy94] J. Moy, “Multicast Routing Extensions for OSPF,” *Communications of ACM*, Vol. 37, Aug. 1994, pp. 61-66.

[Nava88] S. Navaratnam, S. Chanson, and G. Neufeld, “Reliable group communication in distributed systems,” *Proceedings of the 8th Int. Conf. on Dist. Comp. Sys*. San Jose, June 1988, pp. 439-446.

[Ngoh89] L. H. Ngoh, T. P. Hopkins, “Transport Protocol Requirements for Distributed Multimedia Information Systems,” *Comput. J*. Vol. 32, No. 3, 1989, pp. 252-261.

[Ngoh92] L. H. Ngoh, “Multicast support for group communication,” *Computer Network and ISDN Systems*, Vol 22, 1991, pp. 165-178.

[Raja92] B. Rajagopalan, “Reliability and Scaling Issues in Multicast Communication,” *COMM’92*, MD. USA, Aug. 92, pp.188-198.

[Shac92] N. Shacham, “Multipoint Communication by Hierarchically Encoded Data,” *IEEE Proceedings of INFOCOM ‘92*, 1992, pp. 2107-2114.

[Shri90] D. H. Shrimpton, C. S. Cooper, “Multicast communication on the Unison network,” *Computer Communication*, Vol.13, No.8, Oct. 1990, pp. 460-468.

[Stra92] W. Strayer, B. Dempsey, A. Weaver, “XTP: The Xpress Transfer Protocol,” Addison-Wesley, 1992.

[Topo90] C. Topolcic, “Experimental Internet Stream Protocol, Version 2 (ST-II),” *RFC 1190*, CIP WG, Oct. 1990.

[Verm93] D. C. Verma, P. M. Gopal, “Routing Reserved Bandwidth Multi-Point Connections,” *Sigcomm’93*, Ithaca, N.Y, USA, Sept. 93, pp. 96-105.

[Whet94] B. Whetten, S. Kaplan, T. Montgomery, “A High Performance Total Ordered Multicast Protocol,” *Draft copy for INFOCOM 1994*.

[Wall82] D. W. Wall, “Selected Broadcast in Packet-Switched Networks,” *Proceedings of the Sixth Berkeley Workshop on Distributed Data Management and Computer Network*, Feb. 1982, pp. 239-258.

[Zhan93] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, “RSVP: A New Resource ReSerVation Protocol,” IEEE Network, Sept. 1993.

[Aiel93] R. Aiello, E. Pagani, G. P. Rossi, "Causal Ordering in Reliable Group Communications," *Sigcomm'93*, Ithaca, N.Y, USA, Sept. 93, pp. 106-115.

[Amma93] M. H. Ammar, S. Y. Cheung, C. M. Scoglio, "Routing Multipoint Connection Using Virtual Paths in an ATM Network," *IEEE INFOCOM '93*, pp. 98-105.

[Armi95] G. Armitage, "Support for Multicast over UNI 3.1 Based ATM Networks," *Internet Draft*, Jan. 1995.

[Arms92] S. Armstrong, A.Freier, K. Marzullo, "Multicast Transport Protocol," *RFC1301*, Feb. 1992.

[Atm94] ATM Forum, "ATM User-Network Interface Specification Version 3.0," Pretice Hall, 1993.

[Ball93] T. Ballardie, P. Francis, J. Crowcroft, " Core Based Tree - An Architecture for Scalable Inter-Domain Multicast Routing," *Sigcomm'93*, Ithaca, N.Y, USA, Sept. 93, pp. 85-95.

[Baej91] J. J. Bae, T. Suda, " Survey of Traffic Control Schemes and Protocols in ATM networks," *Proceedings of the IEEE*, Vol. 79, No.2, Feb. 1991, pp.170-189.

[Belk89] N. E. Belkier, M. Ahamad, "Low Cost Algorithms for Message Delivery in Dynamic Multicast Groups," *Proceedings of IEEE 9th Int. Conf. on Distributed Computing*, June, 1989, pp. 110-117.

[Berr90] L. Berry, "Graph Theoretic Models for Multicast Communication," *Computer Networks and ISDN Systems*, Vol. 20, 1990, pp 95-99.

[Birm87] K. Birman, T. Joseph, "Relable communication in the presence of failures," *ACM Trans. on Computer Systems*, Vol5, No.1, 1987, pp. 47-76.

[Birm91] K. Birman, A. Schiper, P. Stephenson, "Lightweight Causal and Automic Group Multicast," *ACM Transactions on Computer Systems*, Vol. 9, No. 3, Aug. 1991, pp. 272-314.

[Brau93] R. Bruade, S. Zabele, "Requirements for Multicast Protocols," *RFC1548*. May, 1993.

[Camp94] A. Campbell, G. Coulson, D. Hutchison, "A Quality of Service Architecture," *ACM SIGCOMM Computer Communication Review*, 1994, pp. 6-27.

[Chan84] J. M. Chang, N. F. Maxemchuk, "Reliable Broadcast Procotols," *ACM Transactions on Computer Systems*, Vol. 2, No. 3, Aug. 1984, pp. 223-233.

[Cher85] D. R. Cheriton, W. Zwaenepoel, "Distributed Process Groups in the V Kernel," *ACM Trans. on Computer Systems*, Vol. 3, No. 2, 1985, pp. 77-107.

[Come90] D. E. Comer, " Internetworking with TCP/IP," Vol. I, Prentice-Hall.

[Chua94] S. C. Chuang, "A Flexible and Secure Multicast Architecture for ATM Network," Internet version, Dec, 1994

[Deer88] S. E. Deering, "Multicast Routing in Internetworks and Extended LANs," *ACM Computer Communications Review*, Vol.18 No.4 1988 pp. 55-64.

[Deer90] S. E. Deering, D. R. Cheriton, "Multicast Routing in Datagram Internetworks and Extended LANs," *ACM Transactions on Computer Systems*, Vol. 8, No. 2, May 1990, pp. 85-110.

[Deer94] S. E. Deering, D. Estrin, D. Farinacci, VanJacobson, C. G. Liu, L. Wei, " An Architecture for Wide-Area Multicast Routing," *ACM SIGCOMM 94*, London, Aug. 1994.

[Doar93] M. Doar, I. Leslie, "How Bad is Naive Multicast Routing?" *IEEE Proceedings of INFOCOM '93*, 1993, pp. 82-89.

[Erik94] H. Eriksson, "MBONE: The Multicast Backbone," *Communications of ACM*, Vol. 37, Aug. 1994, pp. 54-60.

[Fran85] A. J. Frank, L. D. Wittie, A. J. Bernstein, "Multicast Communication on Network Computers," *IEEE Software*, Vol. 2, No. 3, 1985, pp. 77-107.

[Garc91] H. Garcia-Molina, A. Spauster, "Ordered and Reliable Multicast Communication," *ACM Transactions on Computer Systems*, Vol. 9. No. 3, Aug. 1991, pp. 242-271.

[Hand94] R. Handel, M. N. Huber, S. Schroder, ATM Networks: Concepts, Protocols, Applications, Addison Wesley, 1994

VI. Summary and Future Research Problems

Multicast communication has been received more and more attention in recent years because it is a basic support protocol for many multiparty interactive multimedia applications. In this paper, a number of important multicast protocols and their significant features are surveyed. The important issues related to multicast are identified and examined. These issues cover multicast routing topologies, multicast group management schemes, and QoS guarantee techniques. The following conclusions are made.

- The development of multicast protocols started from supplying basic service, i.e. best effort to guarantee enhanced higher quality of services to applications. The multicast over ATM is still in its infant stage. Since the multimedia data and ATM architecture are different from traditional data and network respectively, new multicast protocols that support multimedia communication over ATM network have to be developed and the process for development will follow the similar development process as in other multicast protocols.

- Many multicast routing topologies have been used in different protocols and each of them has its merits in specific situation. The flexibility of ATM VC's could be explored to support various multicast routing topologies. The ephemeral routing solutions which provide optimum solution for a short life time are not practical in dynamic real world applications.

- Based on the characteristic of ATM network, centralized group management schemes are preferred.

- Certain level of QoS should be guaranteed for different multimedia applications. Normally, the QoS guarantee is on the expense of the efficiency.

Finally, it should be noted that there are still some research areas in multicast communication requiring further investigation.

- In ATM network, only point-to-multipoint unidirectional VC's can be established and the VC's do not provide reverse communication. Therefore, methods for sending back the acknowledgment to maintain certain level QoS have to be investigated.

- The ATM allows the integration of various multimedia traffic data. To support the transmission of selective data to receivers of different capabilities is a new topic for multicast communication.

- In most of the past multicast protocols, N-way multicast is not supported. Practical solutions for N-way multicast is still an open research area.

In future, we intend to propose a suitable ATM multicast service that supports real-time multimedia traffic (i.e. video and audio).

References

[Adie93] Adie, "Distributed Multimedia Information Systems," *Computer Networks and ISDN Systems*, Vol. 25, 1993, pp.49-57.

[Agui84] L. Aguilar, "Datagram Routing for Internet Multicasting," *ACM Computer Communications Review*, Vol.14, No.2, 1984, pp. 58-63.

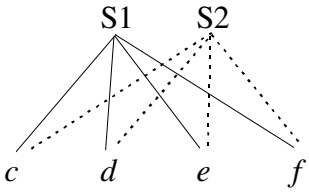

Fig.4 Two sources send to group G = {c, d, e, f}

Birman *et.al*. have proposed another solution based on the network wide priority number [Birm87]. Each receiver maintains a priority queue. The sender sends the packet to destinations which assign their own priority numbers to the packet. This priority number is a system wide unique number higher than any given so far for that receiver. The packet is marked undeliverable and put in the queue. All the receivers return their priority number for the packet to the sender. The sender selects the highest priority number and sends them back to the receivers who replace their original number with the new one and mark the packet as deliverable. Each receiver reorders its queue and whenever the packet at the front of the queue is deliverable, it is delivered. This solution needs intensive exchanging information between sources and receivers that create high protocol overhead.

One of the early centralized multicast protocol that guarantees ordering QoS can be found in Navaratnam's work which uses a single token site to support total ordering and reliability [Nava88]. It requires that each site send back a positive acknowledgment before the next packet can be sent. This solution is limited in scalability by the center site and is vulnerable to failure of the center site.

Chang and Maxemchuk [Chan84] proposed a more burden-shared approach in order to reduce the synchronization cost. Here all the sources send to a central site which assigns sequence number to the packets and then forwards them to the receivers. This center site is identified by a token and the token circulates through the receivers. This strategy reduces the burden on the single site that guarantees message ordering. Similar approaches can be found in the Reliable Multicast Protocol RMP [Whet95] and Aiello's work [Aiel93] which are based on a modified version of Chang's token ring protocol. The RMP provides an N-way virtual circuits, namely token ring, between groups of processes connected by a multicast medium. It lets the reliability control to be shared by all the processes so that each process has the same role in the communication. The RMP uses the negative acknowledgment for error detection and retransmmission while limits the necessary buffer space by passing a token around the members of a token ring.

In summary, QoS's are important service guarantee for various applications. Normally, certain level QoS can only be supported in the expense of efficiency because extra control messages have to be transmitted. The approaches based on the retransmission may degrade the performance of the protocol and give the high burden to receivers for dropping the redundant packets. The centralized token site approach has limited scalability and is vulnerable to token site failure. The multimedia applications have specific QoS requirements for different types of data stream. For instance, it is not necessary to have the fully reliable video stream for a videoconference, while the reliable is obliged for text transmission in update the distributed databases. Therefore, a multicast protocol should have the capability for the application to select appropriate QoS levels.

Source ordering: If message m_1 and m_2 are sent from the same source site and they are destined to the same multicast group then all the processes in this group should receive them in the same relative order and each message is delivered once to the members in the group.

Causal ordering: It guarantees that each message, sent to a group *G*, is delivered to all active, i.e. both correct and faulty, receivers in *G* or to none of them, and are processed according to their causal order.

Total ordering: If two messages m_1 and m_2 are delivered to the same multicast group, then all the addressed processes receive them in the same order even if they come from different sources.

Tracing back to the early multicast implementations such as the V system [Cher85] and IP multicast [Deer88], they implemented only the "best effort" delivery of multicast packets without further guarantee of message reliability and packets ordering. Many multicast applications have been implemented based on the IP multicast that support only the best effort delivery. For instance, the MBONE videoconference system that based on IP multicast allows hosts to participate in the conference taking place in remote location [Erik94]. The quality of these applications is not satisfy both in the image quality and the data transmission speed. Obviously, the throughput is one of the major reason for degrading the quality of MBONE. However, the unreliable and disordered packet delivery schemes does play an important role in degrading the QoS of multicast applications. Generally speak, when the throughput of the network system remains the same, higher QoS levels increase the latency of the data delivery [Whet94]. The ISIS system is one of the earliest protocol that provides causal ordering and total ordering [Birm91]. It is implemented on top of a reliable multicast protocol which requires separate acknowledgments from each destination that limits the performance. Almost at the same time, many other multicast protocols emerged that provide various high level QoS's for multicasting [Kaas89, Pete89, Arms92, Birm93, Whet95]. The techniques for QoS guarantee are discussed in following paragraphs.

To guarantee single source ordering of multicast packets is relatively straightforward and sometime it is done by the underlying network protocols. The basic idea is to number the packets at the source and to have destination sites to order the incoming packets in a buffer. This method allows the receivers to detect missing packets.

To guarantee total ordering needs more effort. Many solutions, range from distributed to centralized controlled, have been proposed. One solution is to stamp each packet with an timestamp that records the sending time, and then deliver packets in the timestamp order. As illustrated in Fig. 4, if sources S_1 and S_2 send packets to group G = {*c, d, e, f*}. Assume that S_1 send to *G* packet P_1 with timestamp T_1. When the destination *c* receives P_1, it can not forward P_1 immediately to the destination process. It must check for all the potential sources if there are other packets with smaller timestamps. When *c* is certain that a packet has minimal timestamp among all the undelivered packets does it deliver it.

system provides a convenient means to identify the multicast services and to locate the second level group management entity namely the Multicast Coordination Center (MCC). The MCC is an enforcer of group policies and manages the group access control.

Centralized group management has two important weak points that have to be solved. Firstly, all the group management activities are conducted by the single group management entity which may cause significant communication delay due to the heavy burden on it. Secondly, the single group management entity is vulnerable to the failure of the hosting site. An obvious solution to this problem is to create several backup group management entities.

V. QoS Definition and Implementation Techniques

As we mentioned before, the B-ISDN has to supply wide variety of services in dealing with different types of traffic. Some new applications, such as real time distributed control, multimedia collaborative work and video conferencing, require timely delivery of data while maintaining the causal relationship among the messages. Reliability and ordering are therefore important QoS (Quality of Service) to be supported by the multicast protocols. The multicast QoS can roughly be classified into data link level QoS (low level) and group level QoS (high level). The data link level QoS concerns the parameters that should be guaranteed by the communication channels. These parameters include minimal value, peak value, average values of throughput, maximum delay, and etc. Many solutions of link level QoS are for one-to-one communication but there are limited work done for the multicast communication [Topo90, Shac92, Math94]. ST-II is one of the pioneer protocol that supports link level QoS for multicast Internet environment. Mathy also proposed a link level QoS negotiation scheme for multicast over ATM [Math94]. Since guarantee link level QoS is a new research topic and few results are presented, this paper will not investigate any further in this topic. The group level QoS defines reliability and ordering requirements for the multicast communication. This survey emphasizes on the guaranteeing the group level QoS schemes

A. Group Level Multicast QoS

This section will first give the definitions of different QoS requirements and then examine several important solutions in guarantee these QoS in the multicast communication environment.

Best effort: is similar to UDP traffic in which a data packet can be delivered 0, 1, or more times to a destination without ordering guarantees on delivery.

Reliable: A reliable QoS guarantees that packets are delivered correctly to the intended receivers.

Due to the varying network delay in the multicast routes, multicast messages usually arrive at different points in time at the receiver site [Maye92]. Without synchronization, individual receivers may receive messages ordered differently. Therefore, in some applications, the multicast protocol must provide guarantees of the order in which data packets are sent to the destinations. For instance, consider a deposit and a withdraw transactions for the same bank account in a computerized banking system, if the withdraw is performed first an overdraft occurs and a penalty is charged. With the deposit first, no penalty is incurred and the balance of account is different. Similar examples can be found in the videoconference where ridiculous reactions may occur if a destination receives information in a reversed order.

REQUEST is normally acknowledged by a JOIN_ACK. All the intermediate routers traversed by the JOIN_ACK change their state to CBT-non-core routers for the group identified by the group-ID. It is the JOIN-ACK's that actually creating the span tree branches. Each router records its parent and child interface with respect to a particular tree. The parent interface in each CBT tree is where the JOIN-ACK was received while the child interface is the one over which a JOIN-ACK has been forwarded with respect to a particular group.

The distributed group management schemes shows a greater flexibility in managing the group membership. In IGMP for instance, it facilitates the implementation of a flexible multicast addressing scheme that allows any host to send to a group without knowing all the receivers. But it makes the security checking of group members more difficulty to implement. The routers of the individual subnetwork can decide whether to grant a host's joining-request for a group even when the sender does not wish to. Furthermore, IGMP has no address allocation mechanism where addresses are assigned either by an outside authority or by each application. This may lead to address contention among multiple applications.

B. Centralized Multicast Group Management for ATM

The centralized membership management schemes have been widely adapted by many multicast protocols, such as in the Deering's PIM system [Deer93], where the *Rendezvous Points* (RP's) are used for senders to announce their existence and by receivers to learn about new senders of a group. Recently, many researchers also work on implementing multicast protocols on ATM based networks [Chua94, Armi95, Whet95]. Several typical centralized membership management schemes for ATM network are discussed in following paragraphs.

In ATM, since the sender must have prior knowledge of each intended receiver and explicitly establish a VC with itself as the root node and the receivers as leaf nodes, centralized group management becomes more pragmatic. Armitage has described a protocol that supports IP multicast over ATM [Armi95]. In his approach, a Multicast Address Resolution Server (MARS) acts as a registry of multicast group membership. The host address resolution entities query the MARS when a multicast group address needs to be resolved. It provides asynchronous notification of group membership changes by operating a point-to-multipoint VC to all the hosts that require multicast support. After receiving membership information from MARS, the sender can establish a point-to-multipoint VC to all the members of the group or use the multicast servers to distribute multicast packets to group members. The MARS may reside within any ATM host that is directly accessible by the hosts it is serving. It keeps an extended table of (Multicast address; ATM.1, ATM.2, ..., ATM.n) mappings. When a source has packets for transmission, and there is no outgoing VC established for the packets, the MARS is queried for the set of hosts currently constituting the group. The MARS will return the address ATM.1, ATM.2, ..., ATM.n to the source. Then, a point-to-multipoint VC can be established by the sender. Two messages MARS_JOIN and MARS_LEAVE are used to manage the host join and leave actions. The MARS_JOIN carries a multicast group address and the unicast address of itself. When MARS_JOIN is received by the MARS, it adds the specified ATM address to the table entry for the specified multicast group. The MARS_LEAVE message is processed by removing the specified ATM address from the table for a specified group. The MARS_JOIN and MARS_LEAVE messages are retransmitted to all the members of the group to ensure the membership changes are distributed timely.

Another similar centralized group management scheme for ATM can be found in Chuang's system which uses a hierarchical centralized group management architecture [Chua94]. In the highest level, a directory system enables the groups to be registered. An entry in the directory

The multicast over ATM is still in its infant stage and it is a fast evolving research area. Currently, ATM signaling specification [Atm93] supports only point-to-multipoint VC's but does not support multicast address abstraction. A key of the multicast solution may lie on the ATM signaling which supports the group addressing. The merits of various multicast routing approaches for ATM have to be evaluated before any concrete conclusion can be drawn.

IV. Multicast Group Management

The group management is an important function of a multicast protocol that manages the membership status and member joining or member leaving activities. There are two basic group management schemes: The distributed group management that allows local routers to manage the group membership in the directly attached subnetworks. While in the other extreme, the centralized group management scheme uses a group agent to coordinate all the group management activities. The distributed group management can be found in the early protocols such as Internet Group Management Protocol (IGMP) [Deer88] and Process Group Management Protocol (PGMP) [Ngoh92]. As the usage of computer network grows, network security becomes an important issue. The group members are distributed sparsely across a wider area that makes the traditional receiver initiated and broadcast-based multicast membership schemes inefficient. The trend of multicast membership management tends to be more and more centralized as described in several recently works [Shri90, Deer93, Chua94, Armi95, Whet95]. This section examines the different schemes for distributed and centralized multicast group management.

A. Distributed Group Management Schemes

The multicast address is assigned as a logical group ID to a collection of users in the IGMP. Senders, without knowing all the receivers, simply use the multicast address as the destination address of the data packets for delivery to all the members of the group. Hosts can join and leave a multicast group in a transparent manner. They perform the joining and leaving action by notifying the attached routers, which in turn uses this information to prune or add branches to the multicast tree. The Internet Group Management Protocol (IGMP) is used by the routers to learn the membership of various groups in their directly attached subnetwork.

In the IP multicast, group management involves hosts issuing IGMP report messages either when these hosts perform a *JoinLocalGroup* or in response to an router's IGMP query. By periodically transmitting queries, IP multicast routers are able to identify which IP multicast groups have non-zero membership on a given subnetwork. All IP multicast hosts must issue *JoinLocalGroup* for the address during their initialization. Each host keeps a list of IP multicast groups it has joined to. When a router issues a IGMP query on the address, each host begins to send IGMP reports for each group it belongs to. IGMP reports are sent to the group address so that other members of the same group on the same network can overhear the report. IP multicast routers conclude that a group has no members on the subnetwork when IGMP query no longer triggers associated replies.

Group management in CBT is a receiver initiated scheme in which the router uses acknowledgment for a joining host to establish a path from the joining host to the core. In CBT, whenever a router receives a group membership report from a host in the directly attached subnetwork, it will proceed the membership request by sending a JOIN-REQUEST toward the core. This JOIN-REQUEST is then repeated to the next-hop router on the path toward the core of the specified group. The JOIN-REQUEST traverses until it reaches either a core or a router that is already part of the tree as identified by the group-ID. The JOIN-

point and hence lack the low level multicast ability as in the bus or ring-based LANs. Although the existing ATM switches are able to transmit copies of an incoming cell to different outlet, there are several fundamental limitations in this multicast supporting mechanism. Firstly, only point-to-multipoint unidirectional VC can be established and the sender must have prior knowledge of each intended receiver and explicitly establish a VC with itself as root node and the receivers as leaf nodes. Secondly, only the root node of a given multicast VC may add or remove the VC's for the group members.

Recently, the extension of IP multicast over ATM has received a lot of attention in the Internet community. The draft version for implementing IP multicast over ATM is under intensive discussion by IETF. In IP multicast over ATM, two proposed routing topologies namely multicast mesh and multicast server have been implemented [Armi95]. In the multicast mesh scheme, each sender is the root of a point-to-multipoint VC that has every other host in the group as a leaf; while in the multicast sever scheme, all senders send their packets directly to a server that is located somewhere in the ATM cloud, which then retransmits copies to all group members. Another solution, an approach independent from IP multicast proposed by Chuang [Chua94], tries to construct SBT and CBT over ATM network. Ammar uses another approach which tries to develop a new type of virtual paths in ATM [Amma92], namely the VP with intermediate exits, where a node that performs VP switching can copy switched packets to the local destinations. Based on the new type of VP, a SBT can be established for transmitting multicast packets.

A matrix topology, inspired from the parallel processing [Fran85, Huan94], shows great promising to support the N-way multicast communication in ATM network. The ATM virtual channels, both unicast and multicast, can be used to support 2-D matrix virtual topologies in following ways.

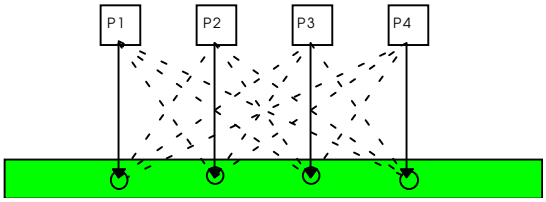

Fig. 3a Four-way multicast connection

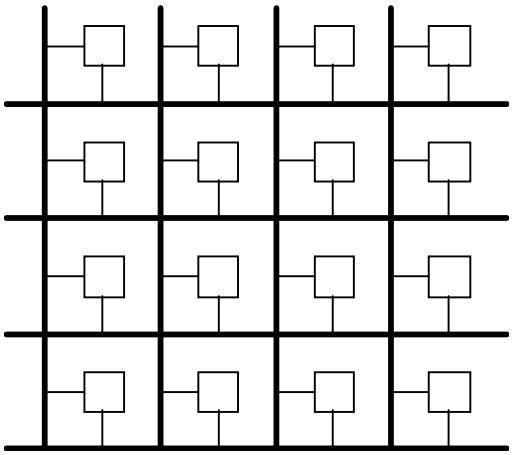

Fig. 3b 4 x 4 matrix connection

Fig. 3a shows four processes, p1 through p4, are interconnected with a virtual bus by using a four-way complete multicast connection where one VC for each process to be used for sending and three VC channels for receiving. Fig. 3b shows how 16 processes are connected by a matrix of virtual buses. Established multicast connections in such a grid lets each process multicast to all other processes in maximum two-step message passing; first it sends to all the processes on the same row and then all receiving processes send to their respective columns. In such a topology, group members are identical that makes the possibility of the N-way multicast over ATM.

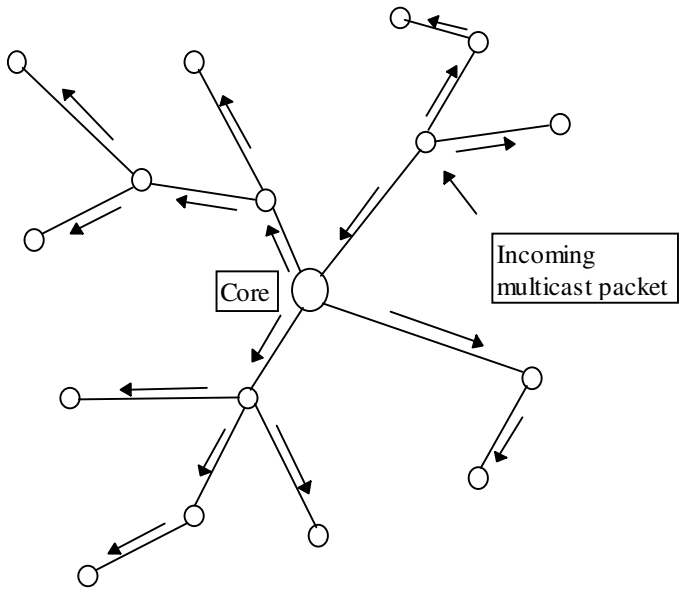

Fig. 2 Core based tree

The previous described multicast routing schemes have their advantages in different network configurations (e.g. DVMRP in regions where a group is widely represented, or CBT in network where bandwidth is universally plentiful). Recently, Deering *et. al*. developed a Protocol Independent Multicast (PIM) that tries to constrain the routing data so that a minimal number of routers in the network can receive it. The PIM is unique in the sense that it supports both types of trees i.e. SBT and CBT. It might be desirable to support both types of the trees such that the selection of tree types becomes a configuration decision within a multicast protocol [Deer93].

C. Steiner Tree

Another multicast routing approach treats the routing as a Steiner tree problem [Wall82, Jaffe83, Waxm88, Jian91, Komp93]. The problem can be defined formally as following: For a given undirected network $G = (V, E, c)$ and a subset of vertices $X \in V$, find a tree T such that there is a path between every pair of vertices in X, and the cost of T is a minimum. Finding such a tree has been proved as an NP-complete problem [Wall82, Berr90]. Some heuristics can be used in practice for constructing a Steiner tree. One heuristic algorithm is based on joining clusters of small trees, which contain required destination nodes, to build up a Steiner tree. Another heuristic is based on finding minimum span trees [Waxm88]. Both of them have been shown to give near optimum results with respect to tree cost [Jian92]. However, this routing scheme is not suitable for multicasting in a frequently changed group in which members joining and leaving will change the topology of Steiner tree frequently. Doar concluded that future work should focus on the ability to add and remove connections from a multicast tree and attempt to improve the routes in an incremental way, rather than in constructing monolithic solutions that may only be valid for an unprofitable short duration [Doar93].

D. Multicast Routing in ATM Network

ATM has some major advantages in scalability, statistical multiplexing, traffic integration and network simplicity [Kimb95]. However, ATM connection services are inherently point-to-

Another major algorithm for constructing multicast tree is based on the link-state routing. In the link-state routine algorithm, every router monitors the state of each incident links. Whenever a link change, the router broadcasts the new state to all the routers in the network by using a flooding protocol. Consequently every router receives topological information from all the routers in the network. Based on the topology, each router can compute the shortest path span tree rooted at itself using Dijkstra' shortest path algorithm. The link-state routing algorithm can easily be extended to support the shortest path multicast tree. Whenever a new group ID appears or an old group ID disappears on a link, the router attached to that link floods the new state to all other routers. With full information of which groups have members on which links, any router can compute the shortest path multicast tree from any source to any group using the Dijkstra algorithm. If the router doing the computation is on the multicast tree, it can determine which links it must use to forward copies of multicast packets from the given source to its group members.

The algorithms to implement SBT rely on the unicast algorithms which complicates the development of unicast algorithm and limits the flexibility of multicast routing. Another drawback of the SBT is its construction which has a poor scaling property. The DVMRP algorithm requires the routers to store membership information for each source. If the number of active sources is S and G is the multicast groups, it results in a scaling factor of $O(S \times G)$. In the link-state routing, the processing cost of Dijkstra shortest path tree is another major factor preventing the scheme from scaling to large wide-area networks because it requires intensive computing of the shortest path tree for all the active sources in a group. Finally, it is not efficient in the DVMRP that routers, which are not on the multicast delivery tree, still have to process truncated-broadcast packets periodically and perform the pruning of branches for all the active groups.

B. Core Based Tree

A Core Based Tree (CBT) uses one router as the core of the tree from which the branches emanate, as shown in Fig. 2. The nodes on the branches are made up of other routers which form the shortest path between a member host's directly attached router and the core router. A router at the end of a branch is called a leaf router in the tree. The major feature of CBT is that only one multicast tree is needed for each group. The cores can be placed in a heuristic way. For instance, the cores could be statistically configured throughout the backbone network [Ball93]. Alternatively, any router could become a core when a host in one of its attached subnetwork wishes to initiate a group. Two distinctive routing phases can be identified in the CBT data delivery. Firstly, an unicast routing is used to route multicast packet to the core of the specified multicast tree. This is achieved by using the unicast address of the core in the destination field of a multicast packet. Secondly, once a multicast packet is on the tree, it is flooded on the tree corresponding to the packet's group identifier.

The major advantage of CBT is the improvement of scalability which has the scale factor of $O(G)$. Furthermore, only routers in the path between the core and the potential group members are involved in the routing process. In comparison with the DVMRP, CBT needs not to broadcast or floods the truncating packets across the whole network. Finally, the construction of a CBT is independent of the unicast routing. The major disadvantage of the CBT tree is the concentration of all the sources' traffic at the core router which may result in network congestion at the core. Hence the CBT is also vulnerable to core failure which can partition the tree. Another problem in the CBT is the core placement may not lead to the optimal paths between group members.

A. Source Based Tree

The source based tree can be defined as follow: for each multicast group, the network builds the shortest path source-based delivery tree between each sender and corresponding multicast receivers, as shown in Fig. 1. The source based tree is a popular multicasting routing topology adopted in the many computer networks. To construct a SBT, several algorithms have been developed in the last ten years. Among them, the distance-vector multicast routing and the link-state multicast routing have received more attention.

The distance-vector routing has been used for many years in many networks as an unicast routing algorithm [Come90]. Routers that use the distance-vector routing algorithm maintain a routing table containing an entry for every reachable destination in the network. Each router sends a routing packet periodically out of its incident links. On receiving a routing packet from a neighboring router, the receiving router may update its own table if the neighbor offers a new shorter route to a given destination, or if the neighbor has no longer offered the route that the receiving router used. In this way, the routers can maintain the shortest routes to all network destinations.

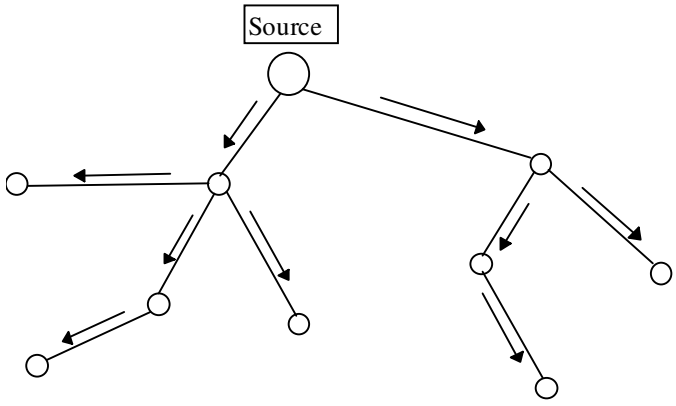

Fig. 1 Source based span tree

Two simple multicast routing algorithms, based on the distance-vector routing, are Reverse Path Forwarding (RPF) and Reverse Path Broadcasting (RPB). These algorithms are implemented by broadcasting along the source based shortest-path broadcast tree and rely on the receivers to select the data packets that are destined to them. These algorithms are not efficient for multicasting in large extended LAN's or WAN's. A more sophisticated approach is the Distance-Vector Multicast Routing Protocol (DVMRP) which uses a modified RPF algorithm to provide on demand pruning of the shortest-path multicast tree. In DVMRP, the first multicast packet is sent by a broadcast along the shortest broadcast tree to all the links except the non-member leaves. When the package reaches a router for whom all of the child links are leaves and non of them has members in the destination group, a NMR (Non Membership Report) is generated and sent back to the router that is one hop toward the source. If the one-hop-back router receives NMRs from all of its child routers and if its child links have no members, it inturn sends an NMR back to its predecessors. Eventually, information about the absence of members propagates up the tree along all branches that do not lead to members. In this way, a prune tree is created and subsequent multicast messages are blocked from traveling down the unnecessary branches by the NMRs sitting in the intermediate routers.

As ATM is becoming the major link layer protocol in the new generation of B-ISDN to support multimedia communication, many researchers have focused on implementing multicast protocols over ATM. The implementation of classical IP and ARP (Address Resolution Protocol) over ATM does not support group addressing and IP multicast as described in RFC1577 [Laub94]. Basically, there are two strategies of implementing multicast over ATM - *extension* of existing multicast protocols or development of *new protocols* from the scratch. The *extension* strategy is to implement existing multicast protocols (e.g., IP multicast) over ATM, then all the existing higher level protocols can be used directly. The *new protocol* strategy follows the development path of existing multicast protocols. It starts from implementing best effort multicast specifically for ATM for construction of higher level protocols that guarantee various QoS's. The *extension* strategy is easy to fit in the existing protocol stack while the *new protocols* strategy has more scope in exploring the potential capabilities of ATM technology. Currently, the extension of IP multicast over ATM becomes an active research area in multimedia communication. The Internet draft for IP multicast over ATM proposed by Armitage [Armi95] is under intensive discussion by IETF. In this draft, a Multicast Address Resolution Server (MARS) is used to manage the group membership. The protocol supports two types of routing topologies, namely multicast meshes and multicast servers. There are also ATM specific approaches that use the flexibility of ATM virtual channels to construct various multicast routing topologies [Chua94].

In summary, multicast over ATM is a new research area at its infant stage. Some important issues have to be investigated further. The issues include: (i) How to guarantee different levels of QoS's; (ii) How to ensure the bandwidth and delay of multicast links which consequently support the selective transmission of multicast data; and (iii) How to support N-way multicast. The last two issues are also the topics that should be investigated for non-ATM based multicast protocols.

III. Multicast Routing

The main objective of multicast communication is to supply various group communication services with required QoS while reduce the cost of data transfer (i.e. minimizing number of data copies are sent for a group). Multicast routing is inherited in some LANs such as the Ethernet that provides efficient broadcast delivery and a large space of multicast addresses [Deer90]. However, two problems have to be solved when implementing multicast over multiple networks: Firstly, an efficient routing mechanism has to be developed; Secondly, the cost of maintaining multicast routing has to be reduced in supporting the scalability of the multicast services. To tackle these issues, several routing topologies and implementation algorithms have been developed. Among them the Source Based Tree (SBT) [Agui84, Deer88, Raja92, Verm93, Moy94] and Core Based Tree (CBT) [Wall82, Ball93, Deer93] have received much attention from the researchers. Recent developments in the Internet multicast have further emphasized the importance of supporting for both tree types [Deer93]. Other types of multicast routing schemes are also being investigated such as the Steiner tree [Wall82, Jaff83, Waxm88, Jian91, Kadi94]. Another interesting routing topology is the matrix based multicast routing that is used in distributed or parallel systems [Fran85, Mcki90, Huan94]. Each of the above mentioned multicast routing schemes has their merits in specific network environments. Therefore, following paragraphs describe briefly the concepts of different routing topologies and the algorithms to implement them, and analyze the pros and cons of these approaches.

negative acknowledgment for error corrections. The Reliable Multicast Protocol (RMP) [Armi95] is a promising multicast protocol that supports a wide range of QoS guarantees from best effort to total ordered data delivery. It is a transport level protocol implemented on top of IP multicast. RMP uses a rotating master approach to guarantee the reliability and ordering. ST-II [Topo90] is one of the pioneer protocols that guarantees end-to-end bandwidth and delay in multicast communication. It requires bandwidth reservations for all links before setting up and guarantees that the requested bandwidth is available through the lifetime of the link. RSVP [Zhan93] is a newly developed resource reservation protocol for supporting real-time multicast applications over Internet. It enables a receiver to reserve resources and the reservation is repeated at regular intervals. We observe that the development of multicast protocols evolves from supporting simple service (i.e. best effort) to enhanced high level services that guarantee the reliability, ordering, delay, and end-to-end bandwidth. The idea is to develop reliable multicast protocols on top of a best effort multicast data delivery protocol (e.g., IP multicast) to support higher level of QoS's,

Table 1. Important Multicast Protocols in History

1984-	1988-	1990-	Future
V system (Cheriton et. al. 1984) - best effort	IP multicast (Deering, et. al. 1988) -best effort - Flexible address	PIM (Protocol Independent Multicast, Deering, et. al. 1993 - network level protocol - best effort - support two routing topologies - selection of sparse and dense modes XTP (Xpress Transfer Protocol, PEI, 1992) - network & transport level protocol - best effort, semi-reliable - go-back-n error correction ISIS (Birman, et. al. 1991) - transport level protocol on reliable multicast - causal ordering, total ordering - separate acknowledgment from receivers MTP (Multicast Transport Protocol, Armstrong, et. al. 1992) - transport level protocol on top of IP multicast. - best effort, reliable, ordering - multicast master control joining / leaving - master issues token for flow control / ordering - NACK for error correction RMP (Reliable Multicast Protocol, Whetten, et. al. 1995) - transport level protocol on top of IP multicast - from best effort to total ordering - use a rotate master approach ST-II (Experimental Internet Stream Protocol, Topolcic, 1990) - network level protocol - guarantee end-to-end bandwidth, delay - sender and receiver negotiate for joining	<u>Trend</u> - develop multicast protocols on ATM network. - IP over ATM - ATM specific approaches <u>New topics</u> - guarantee different QoS's, best effort, reliable, ordering, etc. - Support real-time traffic. - support selective transmission of partial streams. - support N-way multicast.

establish the data delivery paths from senders to all the active group members in the network. The multicast group management is an important function that manages the membership status and member joining or member leaving activities. The reliability and ordering issues are considered as the QoS (Quality of Service) guarantee that should be supported by the multicast protocols. Research issues of existing techniques for these three functions are also reviewed in this paper.

The structure of this paper is organized as follows: In section II, the important multicast protocols are reviewed. In section III, various multicast routing topologies and implementation algorithms are surveyed. In section IV, various methods for implementing the multicast group management are examined. In section V, different QoS requirements are defined and their implementation methods are examined. Finally, a summary of the paper is given and further research topics are identified.

II. Overview of Existing Multicasting Protocols

Tracing back a decade, the earliest multicast protocol was reported in Cheriton's work of the V system extension which implemented the multicast communication among logical groups of processes [Cher84]. The V system supports only "best efforts" delivery of messages (i.e. the message is delivered zero, one, or more times). By 90's, several multicast protocols were developed for better supporting of wide-area groupware. Among them, the most distinguished multicast protocol should attribute to the IP multicast which implemented the group communication over the Internet with best effort delivery of data [Deer88]. In IP multicast, group addresses are assigned by each application. The sender simply transmits packets to that address with routers determining the paths for multicast data delivery. Deering's work has been used as main driving force behind the formation of the multicast backbone (MBONE) across Internet.

Recently, multicast protocols have received increasing attention by multimedia communication research communities because of the need to support emerging multiparty interactive applications, such as videoconference and distribution of stock market information. These applications require guarantee of higher level QoS and real-time information delivery [Camp94]. To support these types of applications, Ngoh and Hopkins have identified varieties of QoS's (Quality of Services) that are important for a multicast protocol [Ngoh89]. In early 90's, many multicast protocols were developed which guarantee certain levels of QoS. These protocols and their features are listed in Table 1. The significance of these multicast protocols are discussed in following paragraphs.

The Protocol Independent Multicast (PIM) [Deer94] is an extension of the IP multicast. It allows the coexistence of two different multicast routing schemes, namely Source Based Tree (SBT) and Core Based Tree (CBT). PIM is the first protocol that allows the selection between dense or sparse mode routing based on the group member distribution in the network. The multicast of Xpress Transfer Protocol (XTP) [Stra92] is similar to IP multicast but goes one step further by supporting limited reliability, called semi-reliable, which is defined as providing high probability of success in delivery of multicast data [Hein93]. It uses a go-back-n approach for error correction. ISIS [Birm91] is one of the pioneer protocol that provides high level QoS such as causal ordering and total ordering of data delivery. It was developed on top of a reliable multicast protocol. The Multicast Transport Protocol (MTP) [Arms92] is based on the multicast master concept which controls all the activities of group communication. It supports total ordering and the flow control by using the master that issues tokens to the senders in the group which allows data to be sent in a specified rate and be received in order. MTP uses

A Survey of Multicasting Protocols for Multimedia Communication

Hongyi LI[†], Hung Keng PUNG[†] , Lek Heng NGOH[‡]

[†] Dept. of Information Systems and Computer Science
[‡] Institute of Systems Science
National University of Singapore, Singapore 0511

Abstract: Multicast is the transmission of information to a specified group of hosts or processes in the network. Supporting multicast communication becomes more and more important in today's multimedia applications such as video conferencing, stock information distribution. This paper surveys the current state of the arts of the multicast techniques by reviewing various important multicast protocols reported in the literature and then identifies and discusses the important issues of multicast communication in the aspects of multicast routing, group management and guarantee quality of services. These issues are closely related with each other and the existing techniques that deal with these issues are reviewed. The future research areas of multicasting are also suggested in this paper.

Keywords: Multimedia communication, Multicast, Protocols, ATM.

I. Introduction

In recent years, considerable research and development effort has been focused on the development of high capacity and reliable B-ISDN transfer technology, network architecture and different service models. ATM (Asynchronous Transfer Mode) is often considered as the technology that allows total flexibility and efficiency to be achieved in future's high-speed multi-service multimedia networks [Baej91]. Some important multimedia applications such as high quality videoconferencing, video-on-demand service and tele-shopping are becoming realistic over ATM networks. Sophisticated multicast support from the underlying network protocols is essential for these applications. Therefore, multicast is becoming an active research topic.

Multicast refers to the transmission of messages to a group of hosts or processes in a computer network. An efficient multicast protocol should provide data delivery to group of hosts at a lower network and host overhead than broadcasting to all hosts or unicasting to hosts in that group [Deer90]. Generally speaking, there are three types of multicast services, namely one-way multicast, two-way multicast, and N-way multicast. The one-way multicast service requires point to multipoint routes that start from sender to all the members in the group. In the two-way multicast, a sender sends messages to all the members of the multicast group and may also receive replying messages from the group members. In N-way multicast, any message sent by a member is multicasted to every other member of the same multicast group.

This paper surveys the current state of the arts of the multicast techniques by reviewing various important multicast protocols reported and identifying the key research issues. A multicast protocol for group communication provides a set of multicast primitives that is capable of adapting to different application requirements and guarantees the performance of multicast communication. Functionally, a multicast protocol should at least perform *multicast routing*, *group management*, and *guarantee quality of service*. The multicast routing is to