Survivable ATM Group Communications Using Disjoint Meshes, Trees, and Rings

William Yurcik

Department of Applied Computer Science
Illinois State University
Normal, IL. 61790 USA
wjyurci@ilstu.edu

Abstract. This paper examines the potential benefits of providing survivability to ATM group communications based on the extension of circuit-switched restoration techniques. Specifically I compare the feasibility and cost of using disjoint dedicated backup route sets established by the VC Mesh model, shared multicast trees, and rings. Results show promise for further development of self-healing survivable rings.

1 Introduction

The search for a native ATM group communications solution began with the introduction of multicast capability (point-to-multipoint circuits) in the ATM Forum's User Network Interface (UNI) specification 3.1. The ATM Forum MPOA (Multi-Protocol Over ATM) subworking group proposes two approaches for ATM intracluster group communications: (1) the VC Mesh Model and (2) the Multicast Server Model (MCS). As a result of the perceived complexity and inefficiency of these two approaches, other techniques have been proposed such as shared trees and rings.

System survivability is important due to the increasing number of information systems and society's increasing dependence of these systems for dependable service. The majority of work on providing survivability for ATM networks focuses on extending circuit-switched facility restoration techniques to ATM using Virtual Paths (VPs). VPs can be viewed as semi-permanent routes through the network with dedicated bandwidth onto which Virtual Circuits (VCs) are grouped. Different algorithms for ATM network survivability have been proposed using either a preplanned backup VP for each working VP and/or dynamically searching for new paths(s) after a single link failure [9, 8].

While ATM network survivability has focused on point-to-point VP/VC restoration, there has been very little work on the survivability of ATM group communications. The same types of vulnerabilities present in ATM unicast are also present in ATM group communications but the potential risk of these vulnerabilities is considerably greater. For instance, while an ATM unicast path is a collection of links and nodes between one source and one destination, ATM group communications can have multiple such paths. To exacerbate this vulnerability, it is often more complex to determine and control which links are in use for a given group routing scheme as compared to the equivalent point-to-point communication. Additionally, users may require an all-or-nothing restoration where if all group members can not be restored simultaneously then the session is abandoned.

In this paper the focus of investigation is the extension of circuit-switched restoration techniques to the survivability of connection-oriented ATM group communications. The basic concept is each group communication session requiring survivability will establish a disjoint backup set of routes with reserved bandwidth for restoration. When a single link or node fault occurs within a "working" set of routes, traffic flow is rerouted to the corresponding disjoint "backup" set of routes. The remainder of the paper is organized as follows: Section 2 briefly overviews some of the techniques proposed for ATM group communications. Section 3 summarizes survivability issues and tradeoffs specific to these proposed schemes. Section 4 reports comparative experimental results and I conclude with a summary in Section 5.

2 ATM Group Communication Schemes

2.1 VC Mesh Model

Figure 1A shows an example of group communications between three nodes using the VC Mesh Model. For a single group containing multiple senders, a unidirectional VC for each sender is established on the network with no coordination between VCs. This criss-crossing of VCs across a network gives rise to the name "VC Mesh" [1]. If a member joins or leaves the group, each point-to-multipoint VC for each sender in the group must be modified. In the ATM Forum UNI 3.1 specification, the sender keeps track of all members in the connection in all cases, only the sender can add a new group member and either the sender can delete a member or the member can delete itself from the point-to-multipoint VC [3]. The ATM Forum UNI 4.0 specification allows a new member to join a point-to-multipoint VC without the intervention of the sender by using a Leaf-Initiated Join (LIJ) capability [2].

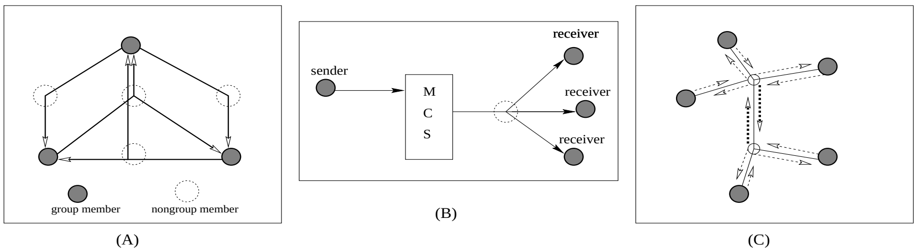

Fig. 1. *(A) VC Mesh Model; (B) Multicast Server Model; (C) Shared Multicast Tree*

2.2 Multicast Server Model

Figure 1B shows an example of ATM group communications between three nodes using the MCS Model (this figure needs to duplicated for each active sender). In the MCS Model, a server is chosen within each cluster to serve as a proxy for all senders in order to relieve senders from VC connection setup and release operations resulting from group dynamics. Conceptually the MCS serves as a centralized connection manager mediating and disassociating senders and receivers within a group [12]. All senders establish a unidirectional point-to-point VC to the MCS. The MCS establishes a unidirectional point-to-multipoint VC to the rest of the multicast group. During transmission, the MCS reassembles cells into packets arriving on all incoming VCs (avoiding interleaving cells from different packets) and queues cells belonging to particular packets for transmission on an outgoing point-to-multipoint VC. All requests to create a group, delete a group, to add a member to a group, or delete a member from a group are sent to the MCS which maintains all state information.

2.3 Shared ATM Multicast Tree Schemes

Figure 1C shows a generic shared tree approach connecting a group of three nodes. I consider three specific shared ATM multicast tree schemes: SMART, SEAM, and SPAM [5–7]. These schemes are similar in that they aim to provide a general-purpose control architecture by modifying in-band control mechanisms of ATM switches [11]. Resources are reserved in both directions on all of the VC links of a shared multicast tree until the connections are released[5]. Each of these schemes individually address two inherent problems of a shared tree architecture: (1) cell interleaving because cells from different sources may arrive interleaved at one destination and (2) resource management because resources allocated to a connection are shared between a number of different sources.

SMART (Shared Many-to-many ATM ReservaTions) proposes a shared tree of one to several associated many-to-many VCs [6]. The same service class and traffic parameters are given to all connections of the tree at call setup. By granting access to a shared tree through the use of resource management (RM) cells, SMART ensures that only one source is active on a tree at any given instant. SEAM (Scalable and Efficient ATM Multicast) relies on per-VC queueing and per-packet forwarding to avoid interleaving of cells from different packets by requiring ATM switches to detect the first and last cells in a packet. SEAM introduces two concepts: (1) "cut-through forwarding" which is a mechanism to forward all cells from a packet together while buffering cells from later packets until the last cell is forwarded, and (2) "short-cutting" which is a signaling mechanism allowing cells to follow the shortest path along a shared tree instead of having to go all the way to the "root" and be forwarded back to receivers.

SPAM (a Simple Protocol for ATM Multicast) is a third attempt to support multipoint-to-multipoint VCs in a shared tree approach by incrementally improving SPAM. In contrast to the forwarding model of SEAM where there is no multiplexing of cells from different packets, SPAM multiplexes cells from different packets through the use of a proposed AAL-SPAM and native "cut-through forwarding". As a compromise between AAL5 and AAL3/4, AAL-SPAM is AAL5 plus an added 16 bit MID (Multiplexing IDentifier) field. Unique MIDs are assigned to senders within a multicast group such that receivers can distinguish cells from different senders. "Native cut-through forwarding" refers to cells from different packets being switched immediately and multiplexed over the shared many-to-many VC connecting the group. Results comparing SPAM and SEAM report the following: (1) SPAM buffer requirements are independent of packet size while SEAM buffer requirements increase significantly with increase in packet size and (2) SPAM delay is the expected propagation and transmission delay while certain SEAM switches in a network will experience maximum delay due to increased blocking [7]. The major shortcomings of SPAM are: (1) the practicality of modifying AAL5 to AAL-SPAM and (2) the global allocation of unique MIDs.

2.4 Virtual Ring Multicasting

Ofek and Yener have proposed the Virtual Ring [10, 13] for window-based packet group communications. The Virtual Ring is a unidirectional circular overlay of routes that join all desired members for a group communication. Using Virtual Rings is straightforward since a message sent from one member of the group travels around the ring and back to the sender.

In the application of Virtual Ring multicasting to ATM networks, each unicast connection between adjacent nodes on the Virtual Ring can be implemented as a separate point-to-point VC. The Virtual Ring provides a global sense of direction for concatenated VCs from switch to switch around an enclosed ring.

The main advantage of the Virtual Ring is that it extends multicast to large groups since it incorporates feedback from receiver to sender to confirm cells have been received correctly. Through a combined strategy of feedback mechanisms referred to as "Implicit ACK" and "Explicit NACK", the Virtual Ring minimizes the feedback needed for reliability while overcoming the "ACK implosion effect". The Implicit ACK procedure considers each cell that circulates the Virtual Ring and returns to its source as an implicit acknowledgment. An Explicit NACK is a unicast retransmission request initiated by a receiver (at a higher layer) when ordered cells are missing.

3 Survivability Issues and Tradeoffs

3.1 Relevant Circuit-Switched Restoration

The most relevant survivability issues that translate from circuit-switched restoration to multipoint restoration are: (1) the location to initiate restoration and (2) the method of identifying a backup route for restoration. Considering tradeoffs briefly outlined, end-to-end preplanned restoration is the only option that simultaneously provides both optimal and guaranteed restoration.

There are two locations for initiating restoration: end-to-end restoration initiated between sender and receivers and local restoration initiated between the upstream and downstream nodes adjacent to the detected fault. While local restoration offers greater speed and ease of implementation, it may also result in blocking and "backhaul" situations where the backup route is not disjoint from the remaining working route. In general, the further restoration is initiated away from a fault results in slower restoration but a globally optimal solution, less blocking, and avoidance of backhaul situations.

There are two methods of identifying a backup route for restoration: a preplanned backup route can be computed in advance or a dynamic search for a backup route can be computed in real-time. In the preplanned method, each working route has a prearranged backup route when initially assigned. In the dynamic search method, senders and receivers broadcast signaling messages after being notified of a network failure and backup routes are selected via a specified algorithm. In the preplanned method, survivability is guaranteed because the backup route has reserved bandwidth. In dynamic search, no bandwidth is reserved for backup routes so restoration is not guaranteed (i.e., residual bandwidth may not be available).

3.2 Survivability of Proposed Models: VC Mesh, Shared Multicast Trees, Multicast Server, and Virtual Ring

In the VC Mesh Model, group dynamics drive the establishment of a new VC for each sender in a group in order to reflect updated membership after each join or leave operation. Also in the VC Mesh Model, multiple senders transmitting to the same group members use separate VCs while in the MCS Model multiple senders each establish point-to-point VCs to the MCS and then the MCS integrates point-to-multipoint transmissions to all group members. For both of these reasons, the number of VCs utilized in the VC Mesh Model increases with the number of senders at a faster rate than the MCS Model. As a result, there is potentially a larger number of routes over which a failure may occur for the VC Mesh which means an increased exposure to link and node failures.

The MCS constrains the number of potential paths over which a failure may occur by concentrating traffic on a server which is both a performance bottleneck and single point-of-failure. If an MCS is disabled, all group sessions using that MCS terminate.

For the ATM Shared Multicast Tree, the dominant survivability issue is handling link or node failures in the "trunk" of the shared tree which will interrupt all group communications. On the other hand, sharing a common tree makes the protection of links and nodes an easier task. A general technique to provide survivability to tree-based approaches is to restore faults on the "working" tree by rerouting to a preplanned disjoint "backup" shared multicast tree as depicted in Figure 2A.

I propose extending Virtual Ring multicasting to ATM group communications in the form of Self-Healing Survivable Rings (SHSR). A SHSR consists of a working and backup set of routes both in the form of rings. In the case of a

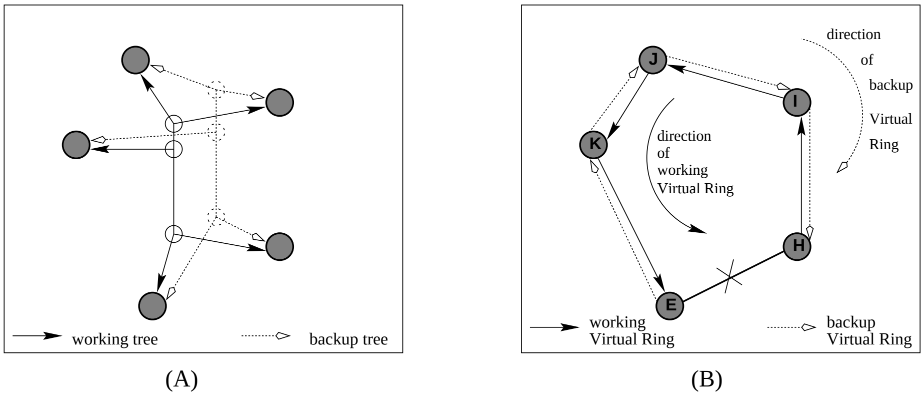

Fig. 2. *Restoration Techniques to Provide Survivability to (A) Tree-based or (B) Ring-based ATM Group Communications*

failure in the working ring, traffic is rerouted to the backup ring by switches adjacent to the failure. To use SHSR to provide survivability, two contributions are necessary: (1) a novel formulation of a *least cost* Virtual Ring connecting all desired group members, that is *link and node disjoint*, such that a single fault will only disconnect a ring in one place, and (2) a real-time ring reconfiguration mechanism to provide restoration of single faults while maintaining the established ring.

In [14], the author develops a novel "Disjoint Steiner Ring" (DSR) formulation which provides a theoretical basis upon which SHSRs can be implemented. A real-time ring reconfiguration mechanism to provide restoration of single link or single node faults can be achieved by bandwidth provisioning of circuits for counter-rotating rings. When a link or node fault occurs, the adjacent upstream node in the working ring will loop-back onto the counter-rotating backup ring [14] as shown in Figure 2B. This real-time ring reconfiguration mechanism requires no new ATM signaling capabilities.

4 Experimental Results

Here I report the results of experimentation to compare the feasibility and cost of preplanned dedicated backup route restoration techniques to provide survivability for ATM group communications implemented via the VC Mesh Model (VCMESH), the Shared Multicast Tree (SMT), and the Self-Healing Survivable Ring (SHSR). The MCS model was not included because of the dominant single-point-of-failure vulnerability which requires a different set of restoration techniques. To enable comparisons of feasibility and cost I make the following state-of-the-art assumption: only single link/node faults are considered.

I consider two actual network configurations used previously in [14, 15]. For these networks, all groups of varying sizes (3 and 4) are formed. When a link or node fault occurs within a "working" set of routes, traffic flow is automatically rerouted to the corresponding disjoint "backup" set of routes. Minimum cost working and backup rings are calculated by an implementation of the DSR formulation [14]. The SMT approaches leave the issue of how to build the shared tree to an external routing protocol so I have assumed the best case scenario by identifying the minimum cost tree calculated using a implementation variant of a Steiner Tree procedure due to Lawler known as the spanning tree enumeration algorithm [15]. The VCMESH approach specifies that each sender establishes circuit(s) to connect with all other group members but does not specify the mechanism so I have assumed the best case scenario such that each sender optimally selects the set of circuits that minimizes cost.

For the networks studied, the feasibility of restoration using the different techniques are not equal. Figure 3A shows that while the SHSR displays 100% feasibility, the feasibility of the SMT and VCMESH techniques begin at approximately 88% and 58% respectively and rapidly decrease as group size increases.

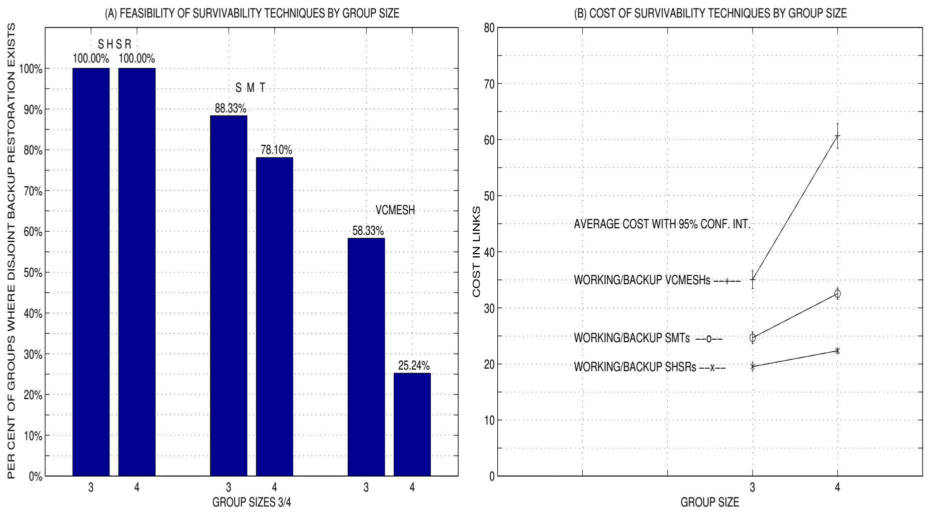

Fig. 3. *Results Comparing Survivability Techniques for ATM Group Communications (for group sizes 3 and 4): (A) feasibility; (B) average cost*

The explanation for the poor feasibility of the SMT and VCMESH techniques is the existence of "multipoint traps". A "trap" is a topology where a corresponding set of backup routes are not available due to the disjointness constraint although disjoint working and backup routes may be available if selected differently [4]. Traps occur because the routing algorithm for the working set of routes optimizes selection according to least cost/minimum hop or similar metric without considering the survivability provided by the selection of a non-optimal set of working routes that can be paired with a disjoint set of backup routes. "Multipoint traps" become more prevalent as the number of group members increases because more links must be used for the "working" SMT/VCMESH leaving less links for forming a disjoint "backup" SMT/VCMESH. "Multipoint traps" also explain why the VCMESH technique has a dramatically lower feasibility than the SMT technique since the VCMESH utilizes more working routes. SHSRs have superior feasibility due to the DSR formulation which solves for both the working and backup routes jointly to eliminate "multipoint traps". Multipoint traps have become the dominant factor for preplanned end-to-end restoration of ATM group communications after being a relatively rare event in point-to-point restoration.

Given the use of disjoint dedicated backup SMT/VCMESH techniques to provide survivability to ATM group communications is not feasible for groups on arbitrary networks, there still exist densely connected networks where these techniques can provide guaranteed single fault tolerance. For these dense networks only, the question of restoration cost is relevant. Figure 3B shows the cost of the SHSR technique is lower, with statistical significance, than either the SMT and VCMESH techniques with a cost differential that increases as the group size increases. The size of the cost differential is not small, SHSRs are 20% less expensive than SMTs and 56% less expensive than VCMESHes.

5 Conclusion

This research has investigated the extension of circuit-switched techniques to the restoration of ATM group communications implemented via the VC Mesh model, shared multicast trees, and self-healing survivable rings. Experimental results show that for the networks studied the use of disjoint dedicated backup is 100% feasible for the SHSR technique and less feasible for the SMT and VCMESH techniques due to "multipoint trap" topologies. Results also show that the relative ranking order of each technique by cost is consistent across different networks and group sizes: SHSR (lowest cost), SMT (middle cost), and VCMESH (highest cost).

I conclude that providing survivability to ATM group communications via SHSRs is both feasible and the least cost preplanned backup restoration technique. Further research is needed to develop less complex but accurate heuristics that can scale to large networks.

6 Acknowledgments

I thank Luigi Rizzo/University of Pisa and the anonymous referees for comments which have significantly improved this manuscript. This paper is a partial summary of Ph.D. research under the direction of David Tipper/University of Pittsburgh who has made substantial intellectual contributions to the work.

References

1. G. Armitage, "Support for Multicast Over UNI 3.0/3.1 Based ATM Networks," *IETF RFC*, rfc2022.txt, November 1996.
2. ATM Forum, *ATM User-Network Interface (UNI) Signaling Specification Version 4.0*, July 1996.
3. ATM Forum, *ATM User-Network Interface (UNI) Specification Version 3.1.*
4. D. Dunn et. al., "Comparison of k-Shortest Paths and Maximum Flow Routing for Network Facility Restoration," *IEEE J. on Sel. Areas in Comm.*, Vol. 12, No. 1, pp. 88-99.
5. E. Gauthier et. al., "SMART: A Many-To-Many Multicast Protocol for ATM," *IEEE J. on Sel. Areas in Comm.*, Vol. 15, No. 3, pp. 458-472.
6. M. Grossglauser and K.K. Ramakrishnan, "SEAM: Scalable and Efficient ATM Multicast," *IEEE INFOCOM'97*, 1997.
7. S. Komandur and D. Mossé, "SPAM: A Data Forwarding Model for Multipoint-to-Multipoint Connection Support in ATM Networks," *IEEE IC3N'97.*
8. K. Murakami and H.S. Kim, "Virtual Path Routing for Survivable ATM Networks," *IEEE/ACM Trans. on Networking*, Vol. 4, No. 1, pp. 22-39.
9. T. H. Noh et. al., "Reconfiguration For Service and Self-Healing in ATM Networks based on Virtual Paths," *Computer Networks and ISDN Systems*, Vol. 29, pp. 1857-1867.
10. Y. Ofek and B. Yener, "Reliable Concurrent Multicast From Bursty Sources," *IEEE J. on Sel. Areas in Comm.*, Vol. 15, No. 3, pp. 434-444.
11. J.E. van der Merwe and I.M. Leslie, "Service-Specific Control Architectures for ATM," *IEEE J. on Sel. Areas in Comm.*, Vol. 16, No. 3, pp. 424-436.
12. Y. Xie et. al., "Multicasting Over ATM Using Connection Server," *ICC'97.*
13. B. Yener, Y. Ofek, and M. Yung, "Combinatorial Design of Congestion-Free Networks," *IEEE/ACM Trans. on Networking*, Vol. 5, No. 6, pp. 989-1000.
14. W. Yurcik and D. Tipper, "Providing Network Survivability to ATM Group Communications Via Self-Healing Survivable Rings," *7th Intl. Conf. on Telecomm. Systems*, 1999, pp 501-518.
15. W. Yurcik, "Providing ATM Multipoint Survivability Via Disjoint VC Mesh Backup Groups," *7th IEEE Intl. Conf. on Computer Comm. and Networks*, 1998, pp 129-136.