MASTER'S THESIS

IP Multicasting in Hybrid Networks

by Isatou Secka
Advisor: Dr. John Baras

CSHCN M.S. 97-2
(ISR M.S. 97-5)

The Center for Satellite and Hybrid Communication Networks is a NASA-sponsored Commercial Space Center also supported by the Department of Defense (DOD), industry, the State of Maryland, the University of Maryland and the Institute for Systems Research. This document is a technical report in the CSHCN series originating at the University of Maryland.

Web site http://www.isr.umd.edu/CSHCN/

Abstract

Title of Thesis:	IP MULTICASTING IN HYBRID NETWORKS
Degree candidate:	Isatou Secka
Degree and year :	Master of Science, 1997
Thesis directed by:	Professor John S. Baras Electrical Engineering Department and Institute for Systems Research

The asymmetric nature of traffic in most networks, as evident in the Internet, is shifting current networking technology trends more towards the development of hybrid networks. Multimedia traffic with its inherent variability in Quality of Service (QoS) requirements further reinforces this trend. Technologies such as DirecPC which allow users to send traffic terrestrially and receive traffic through satellite have demonstrated the efficiency of the broadcast nature of satellite communications as a means of delivering high bandwidth traffic to end users. Even though the majority of Internet applications rely on point-to-point transmission (unicast), emerging applications such as teleconferencing and information distribution have necessitated the development of an overlay multicast backbone network in the Internet (MBONE) for point/multipoint-to-multipoint data transmission. A major hurdle in multicasting over the Internet

is the potential for high bandwidth traffic to cause congestion in the terrestrial backbone. Introducing hybrid terminals within corporate LANs for incoming multicast streams thus would provide an effective means of preserving gateway bandwidth for other outgoing traffic.

IP MULTICASTING IN HYBRID NETWORKS

by

Isatou Secka

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland at College Park in partial fulfillment
of the requirements for the degree of
Master of Science
1997

Advisory Committee:

Professor John S. Baras, Chairman/Advisor
Professor M. Scott Corson
Professor Mark Austin

Dedication

To my best friend and husband, Muhammed Jah, for giving me the strength and always being there for me.

Acknowledgments

I wish to express my thanks to some of the many individuals who made this work possible. Foremost among these is my advisor, Professor John Baras, who gave me the opportunity to work on a very interesting and practical problem as part of the research for my thesis.

I am grateful to Professors Mark Austin and Scott Corson for kindly consenting to join the defense committee and review this thesis. I am also grateful to the Center for Satellite and Hybrid Communication Networks for the support of my research provided by NASA under contract NAGW-2777, Hughes Network Systems and the State of Maryland under a cooperative industry-university contract from the Maryland Institute Partnerships Program (MIPS).

I wish to thank my friends and co-workers for their support and encouragement, especially Manish and Spyro.

Finally, I would like to thank my family for their love and support throughout my studies, particularly my mum and dad.

Table of Contents

List of Tables

List of Figures

IP MULTICASTING IN HYBRID NETWORKS

Isatou Secka

July 23, 1997

This comment page is not part of the dissertation.

Typeset by LaTeX using the `dissertation` class by Pablo A. Straub, University of Maryland.

Chapter 1

Introduction to Multicasting

Multicasting allows us to send a data packet to multiple sites at the same time. The key here is the ability to send one message to one or more nodes in a single operation. This provides a tremendous amount of savings in bandwidth when compared to traditional unicast transmission which sends messages to multiple nodes through replication of the message to each node. Besides the performance improvement over unicast transmission, multicast allows the construction of truly distributed applications.

There are several new and exiting applications such as real-time audio and video conferencing which make good use of multicast services. Because of the real-time constraints on these services, there is a constant data flow requirement and a very low tolerance to transmission delay jitters, hence multicast routing protocols should satisfy these constraints. Multicasting is also often used for synchronization, duplication, and coherency of data in Distributed and Database Systems. For the implementation of coherency one needs to use atomic operations among different machines. This atomicity can be achieved by using multicasting. The same can be said for synchronization in Distributed Systems

especially when the system is used to implement parallel processing algorithms. Another aspect of distributed systems is the duplication of data in a bit to provide some form of Fault Tolerance. A direct application of this would be for updating a file server with multiple and distributed copies of data in one operation through multicasting. This would also ease the work of coherency between copies of the data. Network resource allocation can also be eased by the use of multicasting

1.1 IP Multicast

Internet Protocol (IP) multicasting allows an IP datagram to be be delivered using "best-effort" to a host group consisting of one or more hosts identified by a single IP destination address. The membership in a host group is dynamic and there are no restrictions on the location or number of members in it. Also, a host may be a member of more than one group at a time and multicast sources need not be be members of the group.

Multicast routers,which may be implemented in an Internet Gateway, are designated the role of forwarding IP multicast packets. A multicast source transmits an IP multicast datagram using a Time_to_live (TTL) of 1 to a local network which reaches all immediately-neighboring members of the destination group. If the source wishes the packet to traverse outside of the local subnetwork, a TTL greater that 1 is used. Then, the multicast router(s) attached to the local network takes up the responsibility of forwarding the packet to all other networks that have members of the destination group. An attached multicast router completes delivery by transmitting the datagram as a local multicast.

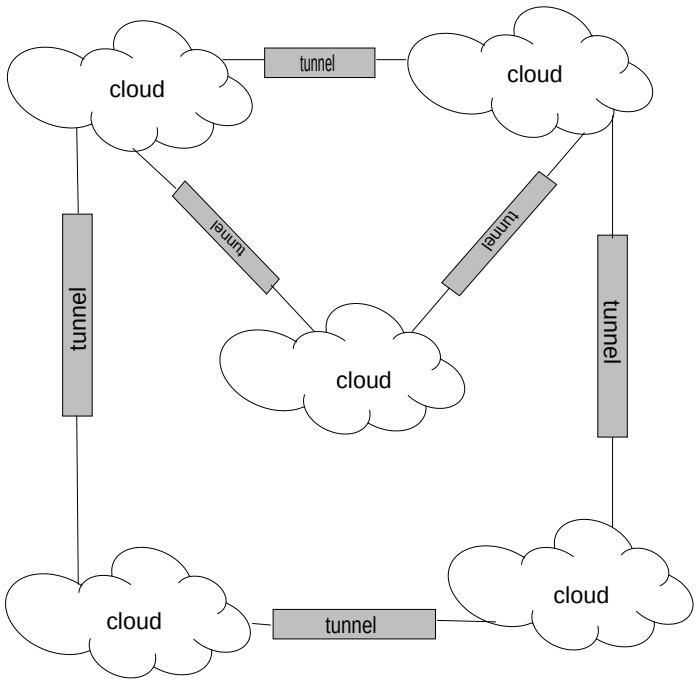

Figure 1.1: The architecture of the MBONE

IP provides an unreliable transmission of data packets from a single source host to a single destination host i.e. unicast transmission service. However, research has shown that only minor modifications are required to add multicast routing support to IP. The resulting IP Multicast routing protocol provides efficient delivery of packets from one source to an arbitrary number of destinations throughout a large heterogeneous network such as the Internet.

Currently, there is an experimental Multicast Backbone (MBONE) which is exploring applications of IP multicast. MBONE, a virtual network that overlays the Internet, allows multicast packets to traverse through routers that are set up to handle only unicast traffic. Datagrams travel through non-multicast capable clouds through tunneling (encapsulating) multicast packets in regular unicast IP packets as indicated in Figure 1.1.

Multicast Addressing

A multicast address is assigned to a group of hosts that form a multicast group. Senders use the multicast address as the destination IP address of a packet that is to be transmitted to all group members.

An IP multicast group is identified by a 32-bit Class D address (IPv4) whose higher order four bits are set to "1110" concatenated with a 28-bit group ID. Hence, multicast group addresses range from 224.0.0.0 to 239.255.255.255 in "dotted-decimal" notation. The Internet Assigned Numbers Authority (IANA) maintains a list of register IP multicast groups. From the range of available addresses, there are reserved blocks:

- 224.0.0.1 to 224.0.0.255 is reserved for the use of routing protocols and other low-level topology discovery or maintenance protocols; 224.0.0.1 is the "all systems on this subnet" address, 224.0.0.2 is the "all routers on this subnet" address;
- 239.0.0.0 to 239.255.255.255 is reserved for site restricted applications;
- the rest are assigned to multicast applications or are unassigned;

Broadcast networks, such as Ethernet, support multicasting whereby multicast packets are sent to a specific group address making it necessary to derive the network-layer group address from the IP class D address. The IANA has been allocated a reserved portion of the IEEE-802-MAC layer multicast address space. The group address is derived from the IP address by placing the low-order 23 bits of the IP address into the low-order 23-bits of the Ethernet multicast address 01-00-5E-00-00-00(hex).

When a source wishes to send a multicast packet to receivers on the same network, the packet is given the IP multicast address destination. The network interface card then maps the address to the corresponding IEEE-802 multicast address. The receivers simply inform their IP layer of their intent to receive packets addressed to the group. In the general case where the sender and receivers lie on different subnetworks, the routers need to learn group membership information so they can forward packets to other routers with attached members. This is discussed further in next paragraph, Section 1.2.

1.2 Issues in Multicasting

To support multicasting several modifications have to be made to unicast transmission protocols because of the additional considerations that have to be taken into account. Not only are routers burdened with the additional task of learning group membership on directly attached subnetworks, but also the construction of a delivery path that enable forwarding of multicast datagrams. As illustrated in Figure 1.2, group membership protocols run between routers and hosts within the same subnetwork whereas routing protocols run between connection routers. Internet Group Management Protocol (IGMP) is used by routers on the MBONE to keep track of group members and join appropriate multicast delivery paths using routing protocols such as DVMRP, MOSPF, and PIM.

The delivery path constructed by a routing protocol is referred to as a multicast tree. Multicast trees can be either source-based or shared as illustrated in Figure 1.3. Source-based trees have uni-directional links and are rooted at the

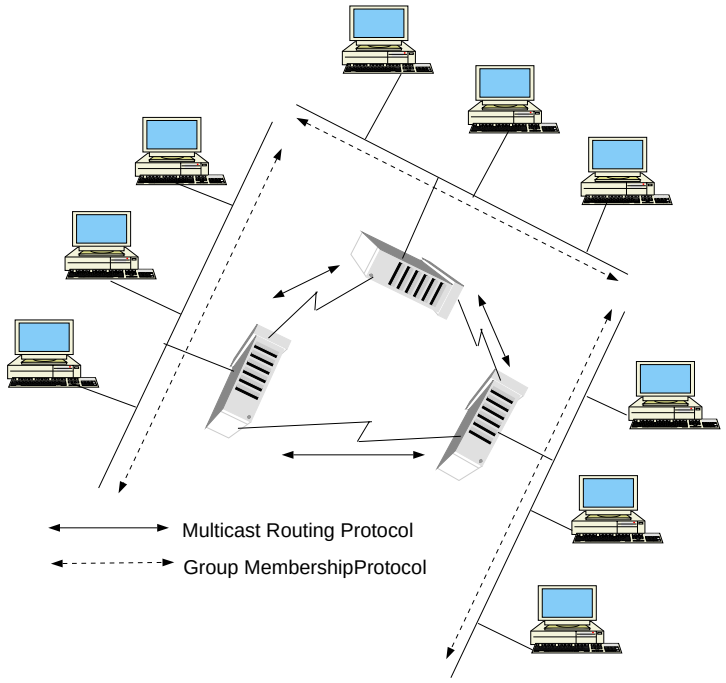

Figure 1.2: Multicasting in Wired Networks

source of multicast packets and therefore a different tree has to be constructed for each member of the multicast group. On the other hand, shared trees consists of bidirectional links and hence the same shared path can be used by each member of the group.

1.2.1 Internet Group Membership Protocol (IGMP)

IGMP is an integral part of IP that is used by IP hosts to report their host group memberships to any immediately-neighboring multicast router. Hosts inform their local router of their intent to receive transmissions attached to a specific multicast group. The router would then periodically query the LAN to determine if group members are still active. Based on the group membership information learned from IGMP, the router joins a multicast delivery tree for

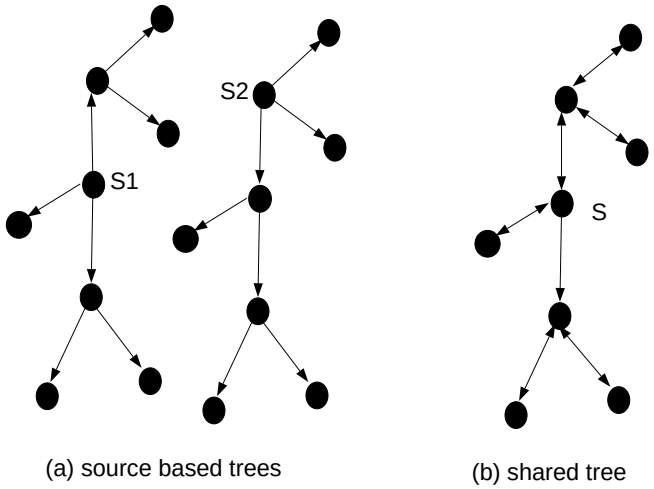

Figure 1.3: Multicast Delivery Trees

each group that determines routes where multicast traffic has to forwarded over.

IGMP Version 1 is specified in the appendix of RFC-1112 [1]. According to the specification, multicast routers send Host Membership Query messages addressed to the all-hosts group (224.0.0.1) with TTL of 1 to discover which host groups have members on their directly attached subnetworks. Hosts respond to a query with a Host Membership Report for each host group to which they belong on the network interface from which the query was received. To minimize the protocol overhead, when a host receives a query, rather than sending reports immediately, it starts a randomly chosen report delay timer for each of its group memberships. If during the delay period, another report is heard from the same group, the local host resets its timer to a new random value. When a timer expires, a report is generated for the corresponding host group.

Multicast router interfaces are configured to receive all multicast IP traffic. It is not necessary for a router to keep track of all hosts that are group members. In fact, the router only needs to know that at least one group member is present

on a network interface.

If no reports are received for a particular group after some number of Queries, the router assumes that there are no group members for that group and prunes itself from the delivery tree of the group. To guarantee that a host will receive multicast traffic if it is the first to join a group, rather than wait for a query, it immediately transmits a report for that group when it joins a new group.

IGMP Version 2 [2] enhances and adds additional features to Version 1. It specifically defines a procedure for the election of a multicast querier in a LAN with more than one multicast router. By default, the router with the lowest IP address on the LAN is elected the multicast querier. IGMP Version 2 also defines a Group-Specific Query message that allows a router to specify a group that is being queried. Also, it defines a Leave-Group message used by hosts to inform routers that they are the last to leave a group. This triggers the querier to send Group-Specific Queries on the interface that the Leave-Group message was received.

A preliminary draft for IGMP Version 3 [3] has been submitted to the IETF. The major addition to this version is the support of Group-Source report messages so that a host can elect to receive traffic from specific sources of a multicast group. Group-Source Report messages can either specify sources that it does not want to receive from (exclusion) or sources it wants to receive from (inclusion). Routers will be able to use this additional information to conserve bandwidth when constructing the branches of their multicast delivery tress. Version 3 further enhances the Leave-Group messages introduced in Version 2, allowing a host to leave an entire group or to specify the specific IP address of the <source, group> pair it wishes to leave.

1.3 Overview of Multicast Routing Protocols

1.3.1 Distance Vector Multicast Routing Protocol (DVMRP)

DVMRP, the most predominant routing protocol on the MBONE specified in [4], builds source-based multicast delivery trees dynamically using a variant of the Reverse Path Forwarding algorithm. When a packet arrives on an interface, the reverse path to the source of the datagram is determined by examining a unicast routing table of known source networks. If the packet arrives on an interface that would be used to transmit unicast packets back to the the source, then it is forwarded out of all interfaces that are part of tree. Otherwise, it is considered not to be on the optimal delivery tree and the packet is discarded. To minimize the number of branches necessary to reach all group members, outgoing interfaces are pruned from a tree if they have no members directly attached to it by sending a <source, group> pair Prune message. Tree branches are added dynamically as new members join the multicast group by grafting the new sections onto the delivery trees using a Graft message.

DVMRP uses IP-IP encapsulation to traverse regions (tunnels) that do not support native multicast routing. Tunneling is done by encapsulating IP multicast packets in unicast IP packets and addressing them to routers that support native multicast routing. Neighbor DVMRP routers are discovered dynamically by periodically sending Neighbor Probe messages on local multicast capable network interfaces and tunnel pseudo interfaces. To prevent these messages from propagating beyond a subnetwork, they are sent to the All-DVMRP-Routers IP multicast address. Each probe message contains a list of Neighbor DVMRP

routers for which the probe message has been received so as to ensure that routers know of each others existence.

Furthermore, to ensure a consistent view of the unicast path back to a source, a unicast routing table is propagated to all DVMRP routers as an integral part of the protocol. Although this introduces additional overhead, it removes the burden of synchronization from the network manager and places it on the protocol thereby reducing the risk of creating routing loops or black hosts due to disagreement between neighbor routers on the upstream interface.

A major disadvantage of this type of protocol is that it does not scale well since multicast routers must maintain state per group per active source. Moreover, because prune messages have to be sent for leaf routers with no attached group members, this algorithm is not suitable for sparsely populated group members typical of most wide area networks, and would saturate links with control messages.

1.3.2 Multicast Open Shortest Path First (MOSPF)

MOSPF specified in [5] is built on top of OSPF [6], a unicast link state routing protocol, to provide multicast routing capability. Routers running MOSPF periodically collect reachability and group membership information and flood it in link state packets, to compute the delivery tree. On receiving a multicast packet, each router uses membership and topology information to calculate the shortest path tree routed at the next hop router of the source of the packet, hence it is source-based. If a router falls within a computed tree, it forwards the packet over the interfaces defined by the calculation. Otherwise, packet is dropped.

MOSPF routers maintain a current image of the network topology through

the unicast OSPF routing tables. Within a subnetwork, a single MOSPF router, denoted the Designated Router (DR), is assigned the responsibility of maintaining a list of directly attached group members and communicating it to all other routers in the OSPF area using Group-Membership Link State Advertisements (LSAs). The DR sends a separate Group-Membership LSA for each multicast group having one or more entities in the DR's local group database which is flooded only within a single area.

The shortest path tree is built on demand when a router receives the first multicast packet for a particular <source, group> pair by using the Routers-LSAs and Network-LSAs (see [6]) in the MOSPF link state database to construct a source-rooted shortest-path tree using Dijkstra's algorithm. Group-Membership LSAs are then used to prune those branches that do not lead to subnetworks containing individual group members. Each MOSPF router that is in the delivery path determines its position within the tree and creates a forwarding cache entry containing the <source, group> pair, the upstream node, and the downstream interfaces. The forwarding cache entry is then used to forward all subsequent packets for the <source, group> pair and is updated only if the topology of the OSPF internetwork changes or if there is a change in Group-Membership LSAs indicating that distribution of individual groups has changed.

Unlike DVMRP, MOSPF does not provide support tunnels. In addition to the scalability problems due to its source-based nature, flooding of group membership and reachability information may cause a considerable increase in link traffic. The computation cost of the shortest path tree for each source using methods such as Dijkstra's calculation may also be too high.

1.3.3 Core-based Trees (CBT) Protocol

CBT protocol uses a set of pre-nominated routers called cores to establish a shared multicast delivery tree through an explicit message protocol specified in [7] and [8]. Multicast trees for each group consist of a primary core, secondary cores, and non-core routers. Tree construction is triggered by the receipt of an IGMP report by a CBT capable router, which then sends a join message towards a target core using the next hop address from the unicast routing table. The join request is processed by all intermediate routers that mark the interface on which the join was received as belonging to the group's delivery tree. On receipt of a join message, the core replies with an acknowledgment (ACK) message which traverses the reverse path of the corresponding join to the sending router. On a subnet with multiple multicast routers, the subnet's IGMP querier is designated the CBT-DR for joining trees on behalf of member hosts.

If before reaching the core the message comes across a router which is already on the tree, that router takes up the responsibility of acknowledging the message. When the source router of the join message receives an ACK message, it creates a CBT Forwarding Information Base (FIB) entry, listing the interfaces corresponding to a particular group over which multicast packets should be forwarded. Thus when a packet is received, it is forwarded out of all interfaces dictated by the FIB. Figure 1.4 illustrates how an incoming packet traverses a CBT multicast delivery tree.

CBT operates under two forwarding modes. In native mode, when a CBT router receives a data packet, the packet may only be forwarded over outgoing tree interfaces if and only if it has been received via a valid on-tree interface or the packet has arrived encapsulated from a non-member. On the other hand in CBT

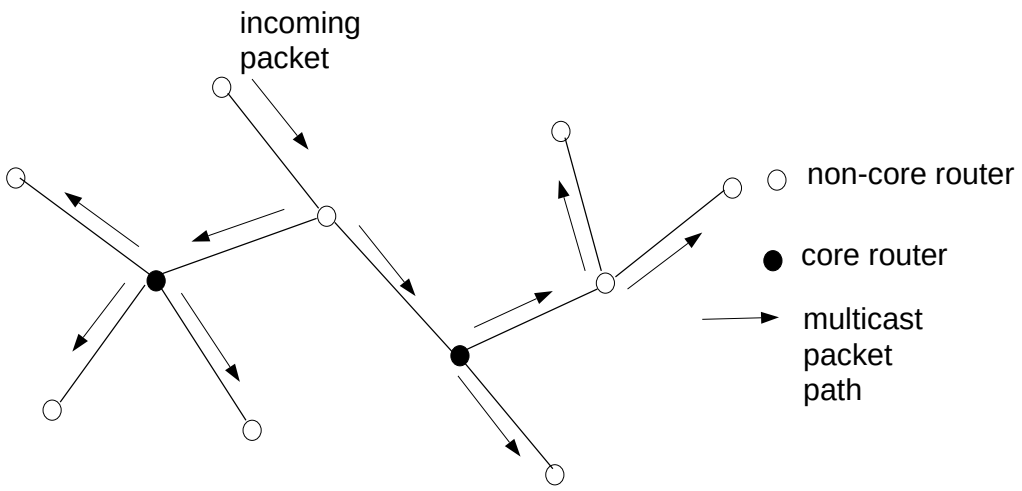

Figure 1.4: CBT packet forwarding

mode, routers ignore all non-locally originated multicast data packets. Locally-originated packets are forwarded native mode by the DR , TTL 1, over outgoing member subnets for which that router is DR. Additionally, the DR encapsulates the packets and then forwards them over all tree interfaces specified in the CBT FIB entry.

Certainly, a major disadvantage of this protocol would be the traffic concentration on the shared path since all packets for that group traverse the same link. However, a great advantage of CBT is that it totally supports non-member sending of multicast packets. Sources that are not members of a multicast group encapsulate packets and then send them towards the core of the tree. If the encapsulated packet hits the tree at an on-tree router, the packet is forwarded as dictated by the FIB entries.

1.3.4 Protocol Independent Multicast (PIM)

PIM, as the name suggests, builds a multicast routing tree that uses unicast routing information independent of the particular unicast routing protocol deployed. PIM operates in two modes, Dense-mode and Sparse-mode described in [9], [10] and [11]. PIM Dense-mode (DM) is designed to work in environments where multicast group members are densely populated and bandwidth is abundant. PIM Sparse-mode (SM) is designed to support multicast groups with members that are sparsely distributed across many regions and bandwidth is not necessarily widely available. The motivation for developing PIM was that existing multicast protocols are specifically developed for either densely populated regions (e.g. DVMRP or MOSPF) or sparsely populated regions (e.g. CBT), but not both.

PIM-DM uses the Reverse Path Multicasting (RPM) algorithm, but unlike DVMRP, multicast packets are forwarded downstream until explicit prune or truncation messages are received. The designers traded-off packet duplication for routing protocol independence and less overhead in building a parent/child database as is done in DVMRP.

PIM-SM requires routers with directly attached downstream members to join a sparse-mode distribution tree by transmitting explicit join messages to the group's primary Rendezvous Point (RP) which acts as the root of the tree. PIM-SM operates very much like CBT where a Designated Router (DR) upon receipt of an IGMP group report, sends a join/prune message towards the designated RP for the group. Each router along the path toward the RP builds a <anysource, group> state for the group before forwarding the request. This state creates a shared, RP-centered distribution tree that reaches all group members.

A major advantage of PIM is the option it provides routers to switch from an RP-shared tree to a Shortest-Path-Tree (SPT) as soon as they start receiving data packets from the source station. For high data rates, it makes sense for routers with local receivers to join source-specific trees and prune the source's packets off the shared RP-centered tree. Because PIM is still an Internet draft, there are several important issues to be resolved, for example, minimizing state information, reducing implementation complexity, and defining interfaces with other multicast protocols.

1.4 Goal: Multicasting in Hybrid Networks

Traditional multicasting on the MBONE has been used for exchanging information between a group of users in applications such as video or audio conferencing but a major hurdle in multicasting over the Internet is the potential for high bandwidth traffic to cause congestion in the terrestrial backbone. For groups with many members that are sparsely distributed over a wide area, the multicast packets would have to traverse several links before reaching all group members, hence the potential for causing congestion. Some companies may wish to engage in multicast conferencing applications but may have limited gateway bandwidth to the Internet. For such users, introducing hybrid terminals within their corporate LAN to route incoming traffic through a satellite link would be a way of preserving the corporate wireline gateway bandwidth for other outgoing traffic. Another motivation of multicasting in hybrid networks is its use in military or medical applications, where individuals in remote areas equipped with hybrid terminals would be able to receive critical high data rate packets.

There are several issues to consider when extending multicast over hybrid networks. First and foremost, a group membership protocol has to defined for keeping track of group membership information in the hybrid network but this will only be covered briefly in Section 3.3.1. The work described in this thesis is mainly directed towards developing asymmetric multicast routing techniques for constructing multicast trees at remote LANs, so that all outgoing traffic is directed toward the corporate wireline gateway while incoming multicast traffic comes through a satellite link. The protocol established for this special case (satellite-terrestrial) could then be extended to other hybrid networks.

Construction of a multicast tree gives the ability to both send and receive multicast packets. The motivation for multicasting is to support high_data_rate applications such as video conferencing. In hybrid networks where there is limited bandwidth on the uplink, it is impossible to support such applications. Hence, use of the asymmetric nature of hybrid networks for multicasting data makes sense only on the receiving end. Thus what we are doing, in effect, is **constrained** multicasting where hybrid hosts take advantage of the high bandwidth downlink to receive packets, but are restricted to sending only low_data_rate voice and data packets which can tolerate the degradation of quality.

One of the biggest challenges faced is that the asymmetric nature of traffic, out through the Corporate LAN and in through a satellite receiver, creates the potential for the formation of loops, breaking the concept of tree construction completely. Further complications could arise at a multi-homed (multiple routers) local LAN with a hybrid host particularly when more than one router are multicast capable because this would make construction of an internal delivery tree difficult. Generally, Internet routing protocols were developed assuming bi-

directional and symmetric links. and may no longer work in the uni-directional environment. For example, routers on the receiving end of a uni-directional link have no means of announcing routes to feeds at the source of link because they cannot communicate directly with them. A subcommittee, the Uni-Directional Link Routing (UDLR) working group, has been formed at the Internet Engineering Task Force (IEFT) to find solutions for dynamic routing problems caused by uni-directional links. The UDLR working group currently focuses on support of alternative uni-directional links on top of a bi-directional internetwork. There are currently two proposed approaches that address this problem. One is based on the modification of the common routing protocols to support uni-directional links. The other one proposes adding a layer between the network interface and the routing software to emulate bi-directional links through tunnels. Both approaches are being studied in order to come up with a solution for dynamic routing in the presence of uni-directional links.

The main objective of this thesis is to develop a system-level design of a multicast routing protocol that would allow hybrid hosts in hybrid satellite-terrestrial networks to dynamically receive multicast packets. The rest of the work is structured as follows: chapter 2 describes the Systems Design Process employed in developing the protocol including the requirements engineering and the preliminary design; chapter 3 takes a closer look at the protocol design; chapter 4 analyses and evaluates the performance of the proposed protocols using mathematical and simulation techniques; and chapter 5 summarizes work done in this thesis and suggests further studies to be done.

Chapter 2

Systems Design Process

2.1 Systems Engineering Approach

In developing a multicast protocol design for hybrid networks, the systems engineering approach, which emphasizes the translation of a system's needs to a set of formally written requirements and specifications for system performance and configuration, was taken. With the requirements that glue the system in place, the next step would be the use of systems analysis techniques to understand the structural, dynamic and functional relationships within the system's domain. The final and crucial step is to identify a high level systems engineering development model which describes the expected evolution and management of the system.

The system engineering life cycle outlines six phases to be followed in any system design process.

Phase 1, Requirements Engineering: involves identifying the requirements that must be met to achieve the goals of the system. The challenge is to identify the requirement drivers that are important in the final design of the system and

focus on those first.

Phase 2, System Design: identifies the functions that implement the system and come up with an architecture to develop the system. This includes the design of subsystems and the relationship among them as well as systems specification and modeling

Phase 3, Detailed Design: designs the individual components and modules that will implement the top-level-specifications. Each module should have a well-defined purpose and meaning, and should be weakly coupled with other modules.

Phase 4, System Integration: involves assembling the modules of the system in a fashion that ensures that the design requirements are met.

Phase 5, System Verification and Optimization: testing to ensure that the system is performing well. Optimization tools can be used to enhance system performance.

Phase 6, System Validation: verifies that the final system is working according to the initial design and requirements specifications defined in phase 1.

2.2 Protocol Design Process

The design goal is to come up with a multicast protocol that can be implemented to allow hybrid hosts on hybrid satellite-wireline networks to receive multicast packets. One of the major design constraints is that the hybrid system

architecture is already in place and hence the design should eliminate or at least minimize changes to the current architecture. Furthermore, since hybrid hosts should be able to send or receive multicast packets to or from other hosts on the MBONE, the routing protocol developed should be adherent to Internet routing standards. It is therefore crucial that we reuse as much as possible existing routing protocol modules.

The reuse and modification of existing modules and components readily lends itself to the bottoms-up approach in system design and object-oriented life cycle modeling. Unlike the traditional top-down model which begins with a high-level design and works its way down to subsystems and modules, the bottoms-up approach begins with the low level modules and subsystems and tries to combine them into higher level entities. The object-oriented approach aims to provide a seamless process between different stages of the life cycle by delaying component implementation and specification until a much later stage of the development process. The key goal of object-oriented modeling is to develop a knowledge-based library containing reusable and pluggable components using an iterative approach as illustrated in Figure 2.1. This would be very relevant in our design since the protocol developed would go through many fine tuning and enhancement stages.

2.3 Requirements Engineering

The most important phase of the systems engineering life cycle is the requirements engineering phase. The term "requirements driven development" is generally used to highlight the central role requirements play in all design activities

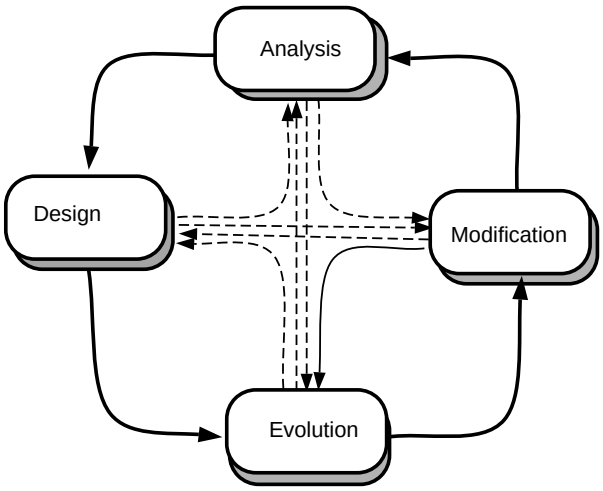

Figure 2.1: Object Oriented Life Cycle Approach.

including development, production, and testing. More formally, requirements engineering has been defined in the following ways:

> A process in which "what is to be done" is elicited, modeled, and communicated. The process has to deal with the different view points, and it uses combination of methods and tools. The product of this process is a model from which a document, usually a requirement definition, is produced.
>
> A discipline for development of a complete, consistent, unambiguous specification - among all parties concerned - describing what the system or product will do.

In accordance with the requirements engineering process, all the requirements important in the design of a multicast protocol for hybrid networks are identified so that they can be traced down in the final system using validation techniques. The protocol should satisfy the following requirements:

- enable "at least once" delivery of a multicast packet to all hybrid hosts that are group members
- allow both hybrid and terrestrial hosts on the hybrid network to be sources or recipients of multicast packets
- allow dynamic adds and joins of hybrid hosts to a multicast group
- construct a true multicast delivery tree that is free from loops
- not introduce significant additional delays to the routing of packets (unicast) to hybrid hosts
- introduce significant bandwidth savings in corporate wireline gateway
- be scalable as the number of hybrid hosts increases

Before proceeding to the design phase, it is essential that the design constraints be identified. Because a system architecture is already in place for hybrid network under consideration, any implementation of the protocol should require minimal or no changes to the current system and should not affect performance of other functions. Furthermore, since source of multicast packets may be any host on the Internet, which may support other multicast routing protocol, the protocol developed should require changes to only routers responsible for direct delivery of multicast packets to hybrid hosts and not to other routers or hosts.

2.4 Preliminary Design

In this stage of the design process, the current hybrid system architecture designed for Internet Access in satellite-terrestrial hybrid networks defined in [12]

and [13], was carefully studied in order to identify the design drivers of our system. In the former, all packets from the hybrid hosts are tunneled using IP-IP encapsulation through a SLIP provider to a Hybrid Gateway where packets are decapsulated and routed to their final destination. Packets destined for Hybrid Hosts (HHs) are intercepted by the Hybrid Gateway (HGW) and encapsulated in a special packet format and sent over Ethernet to the Satellite Gateway for broadcasting over the satellite link. The driver in the HH scans all packets broadcast over satellite for packets addressed to it, removes the satellite header and sends them to the TCP/IP stack through a SLIP driver.

For this system to support multicasting, additional modules need to be added to both the HGW and HH. These modules would be directly responsible for administering multicast related functions such as keeping track of group membership and routing of packets to group members. Since one of the design requirements is scalability, the protocol should be able to support large number of HH group members in different multicast groups. A major concern in protocol development is reducing traffic overhead when the size of network increases. Hence, our design should minimize as much as possible, the messaging traffic so as to avoid congestion or overflow at the HGW where all traffic is routed through. At the same time, the protocol should maintain enough state information to guarantee "at least once delivery" of multicast packets to every HH group member. Another design parameter is the traffic patterns of data. This kind of network is expected to support multicast sessions that generate both "bulky traffic" and "short-length" patterns which may have a time constraint or may not tolerate packet losses.

Chapter 3

Protocol Design

3.1 Protocol Architecture

During the protocol design process, we emphasized the need to adopt the object-oriented systems engineering approach which would allow us to reuse as much as possible existing designs. In the preliminary design, the two major funtional modules, group membership and routing modules, were identified. Therefore, it makes sense to study the terrestrial conterparts of these modules to see if they can be modified to suit our hybrid design. The system definition proposed in this thesis for extending multicast protocols hybrid networks uses a modified version of CBT, hereafter referred to as Hybrid Core-based Trees (HCBT), and assumes the architecture model described in [12] and [13] for hybrid Internet Access. In addition, HCBT architecture assumes the scenario illustrated in Figure 3.1 where we have:

- N users that want to form a multicast group
- Out of these N users, H are static HHs and (N-H) are terrestrial users on the MBONE

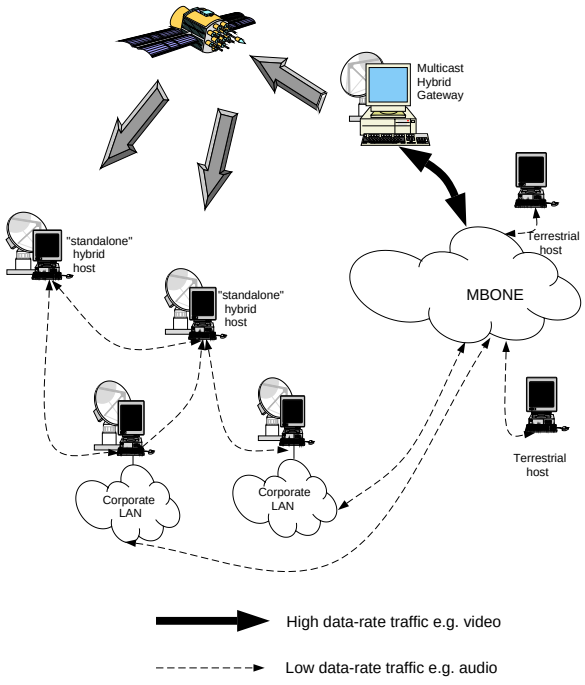

Figure 3.1: Diagram illustrating the HCBT architecture.

- Out of the H Hybrid Hosts, L are attached on LANs and (H-L) are "standalone" hybrid hosts. Note that the LANs also have terrestrial (wireline) access to the MBONE.
- The HHs attached on LANs may be responsible for forwarding packets to other users on the LAN.
- A modified version of IGMP described in Section 3.3.1 is running between the HHs and the gateway

For our system, we define two types of traffic, low-data rate or "short length" traffic (e.g. audio, web browsing), and high-data rate or "bulky" traffic (e.g. video, images, books). All traffic below a certain threshold, T (bits/sec), is considered low-data rate traffic and all traffic with above T rate is considered high-

data rate. Likewise, all traffic beyond size, S (bits), is considered "bulky", otherwise it considered "short length". An intelligient routing scheme will be deployed that routes high_data_rate or "bulky" traffic through satellite and low_data_rate or "short length" traffic through the terrestrial network.

We are proposing that all HHs be required to join multicast trees through a Multicast Hybrid Gateway (MHGW) which is analogous to the Hybrid Gateway in hybrid Internet access architecture. It is assumed that the MHGW would be the IGMP querier for all HHs and is thus aware of group membership information of HHs. Necessarily, all multicast traffic to and from the HHs is routed through MHGW. When packets are multicast to a group with HH members, the MHGW would observe the data rate to determine whether to send them terrestrially or via satellite to HHs. If the latter is required, the packets are put on the satellite interface for broadcasting to the HHs.

Since packets put on satellite are broadcast and would be available to everyone, some authentication mechanism need to be established to allow only HHs that are members of the group to receive multicast packets. Therefore some "key sending process" needs to be included in the IGMP version for Hybrid Networks so that when a HH registers with the MHGW to be a member of a multicast group, the MHGW sends it a "special key" to be used for receiving messages. The alternative to this is for the MHGW to keep track of all group members and unicast a copy of the message to each of them, which obviously wastes satellite downlink bandwidth.

The HHs that are attached on LANs would have an extra responsibility of forwarding multicast packets to and from other hosts attached on the LAN. Therefore, in addition, these HHs would run a proxy to enable them to act as a

multicast router for the LAN.

The system architecture defined raises a lot of interesting issues to be addressed. Let us suppose that a group of multicast users are having an audio conference terrestrially (wireline) and in the middle of the conference, a user decides to multicast an image to others. The users equipped with hybrid terminals would be receiving this image through their satellite link instead. Therefore, it is important that certain performance issues, such as which link will act as a bottle neck to the conferencing, be carefully studied. Another interesting question is determining how many HHs can be served by a MHGW with minimal delay because there is the potential for congestion since all multicast traffic is routed through the MHGW.

3.2 Motivation for using CBT

In considering a routing protocol to be used for multicasting in hybrid networks, one has to carefully look at the issues unique to this type of network and make use of its asymmetric nature to minimize the overhead introduced by routing. The best approach would be to modify an existing routing protocol to accommodate hybrid networks since this would ensure changes are only made on gateways to HHs. As previously mentioned, the most predominant multicast routing protocol is DVMRP. However, the assymetric nature of traffic in hybrid networks almost eliminates using any distance-vector-based protocol which only forwards multicast packets if they arrived over interfaces used to reach the source of a packet. Thus, if a HH is the source of a packet, the hybrid gateway would not forward it to other hosts since the packet arrived on a different interface

(e.g. terrestrial) from the one used to reach the source (e.g.satellite interface). MOSPF was also eliminated since it uses a flooding based scheme and has high SPT computational costs, thus limiting its use on the Internet. PIM was not considered as an option because of the implementation complexities involved in switching between its two modes of operation. Even though implementation of CBT has not been completed, ongoing work shows that its merits, outlined below make it well-suited for hybrid networks.

Non-Member Multicast Source: One of CBTs' attractive features is support of non-member sending, which makes it the best choice for resource discovery applications. Data driven protocols such as DVMRP and PIM dense mode are less suitable for such applications since a group forwarding state is established as data flows in all routers from point of source. On the other hand, routers in between a non-member sender and the corresponding CBT delivery tree incur no group-specific overhead for forwarding that sender's multicast data packets; these are encapsulated by the sender's local CBT router and unicast to one of the group's core routers. The core would then decapsulate packets and distribute them over the corresponding delivery tree.

Minimal Delay: The asymmetric nature of traffic has been a major motivating factor in the development of hybrid networks as a means of preserving wireline corporate Internet bandwidth for other outgoing traffic. In the case of satellite broadcast for incoming traffic, delay incurred at the satellite link could be significant. The CBT architecture that routes all multicast traffic towards the cores of the distribution tree suggests that by careful selection of cores, we can minimize delay incurred in CBT trees.

Scalability: Current multicast routing schemes such as DVMRP, MOSPF, PIM dense mode employ some sort of source-based routing where a multicast tree is constructed per source per group. This type of architecture works well when multicast traffic is densely populated in a region. However, in hybrid networks that mostly span wide areas sparsely, CBT which was designed to suit low traffic distribution areas would work better since there is less protocol messaging overhead involved. Moreover, since only one shared tree is built per group, the number on entries in the CBT routing table is exactly the same as the number of groups thereby providing a considerable reduction in storage space required. It would also be easier to construct the FIB table since each group's members are attached to the same satellite interface.

Interoperability: The CBT operation mode which assumes a region is heterogeneous with routers using different protocols, as is typical of WANs, makes it possible for multicast packets to traverse regions that are not CBT capable. This facilitates Inter-Area routings and compliments the interoperability with other protocols. Already the interoperability of CBT with DVMRP has been defined in [14].

Routing Protocol Independence: Most of existing multicast routing protocols depend on the underlying unicast routing protocol used. For example, DVMRP is based on RIP while MOSPF only runs on networks running OSPF. Because of the spontainety of applications of multicasting such as conferencing, a server multicasting video packets to hybrid hosts may belong to a network running a different routing protocol. Hence CBT which builds its multicast tree independent of unicast routing protocol would be at an advantage

3.3 Protocol Specifications

For the HCBT architecture proposed in Section 3.1, all routing of multicast packets to and from hybrid terminals is done through the MHGW. To make this possible, modifications would have to be done on both the HHs and the MHGW.

The HHs would run a modified version of IGMP to enable the MHGW to learn group membership. In addition, those HHs that act as routers for members on their LANs would have to run a proxy to enable them to act as a "semi-querier" for the LAN and forward membership information to the MHGW. Furthermore, these special HHs would be responsible for multicasting received packets to member hosts on their LAN (either through broadcasting, say on Ethernet, or some other multicasting scheme).

The MHGW has to be CBT capable in order to join the corresponding multicast trees on behalf of the HHs. As specified by CBT, the group joining process will be triggered by the receipt of an IGMP message for a multicast group. The MHGW would then send a join message towards the target core as specified in [8] for attachment to the multicast tree. After receiving an acknowledgment message, the HCBT module would include in its Forwarding Information Database (FIB), an entry corresponding to the tree joined. Since the IGMP message arrives over a different interface from the one where multicast packets have to be forwarded (the satellite interface), slight modifications have to be made to the way CBT operates to ensure that the correct entry is put in the FIB.

The elegance in the proposed architecture would lie in its capability to do intelligent routing based on traffic type. To support this feature, the MHGW will have to implement a switching mechanism that routes high_data_rate packets through satellite and low_data_rate packets through terrestrial wireline links. In

effect, this would be equivalent to maintaining two separate multicast delivery trees. A simple solution would be to have the MHGW encapsulate all low data rate packets and unicast them to the HHs but obviously this is resource wasteful.

When the MHGW receives a multicast packet, it would consult its multicast routing table to determine the interfaces out of which packets have to be forwarded. If data rate warrants, it would forward packets to the satellite interface for broadcasting. The HHs would receive packets by listening to the channel for multicast packets sent using a scheme similar to Ethernet multicasting where a mapping is defined between an IP multicast address and the HHs' adapter addresses. Because broadcast packets would be available to all HHs, the MHGW would have to run some authentication scheme to allow only registered group members to receive packets. The authentication mechanism could be included in the IGMP messaging process so that once multicast trees are joined, all the necessary information to send and receive packets is available to HHs.

To establish a reliable multicast delivery mechanism that guarantees "at least once" delivery of multicast packets, MHGW would keep a copy of all packets until an acknowledgment is received from all HHs. Hence, the MHGW would have to keep track of all HHs members for each group. However, this deviates from traditional IP multicast schemes (IGMP) where multicast routers only keep track of group membership information on their attached networks and not individual members of each group.

3.3.1 Group Membership Protocol

IGMP (discussed in section 1.2.1) , used by multicast routers to learn about group membership information on their local subnet, is ill-suited for satellite-

terrestrial hybrid network considered in this thesis because some of the assumptions made may no longer hold for this scenario. IGMP specifically assumes that all hosts within a local subnetwork can hear each other and that routers need not keep track of individual members of each group. In our scenario, HHs form a virtual subnetwork with the MHGW as their gateway. However, HHs have no direct link with each other since the satellite link is uni-directional. Therefore, certain modifications have to be done to IGMP before it can be used.

IGMP specifies that a Querier router on the subnetwork periodically (about every 1 second) send a general Query to all hosts on their attached LAN to determine group membership information for a each group with directly attached group members. When a host that is a member of the group hears the Query, it sets a random delay timer for each group of which it is a member. When a group's timer expires before a another host's report is received, the host broadcasts a membership report on the local subnetwork. If a local host receives another host's report while it has a timer running, it stops its timer and suppresses the report it was about to send. In the hybrid network considered, the only logical choice for the Querier is the MHGW. However, if IGMP is used as specified, the HHs within the MHGW's logical subnet would not hear each others' group report since traffic to the MHGW is sent via a terrestrial link and hence would not be able to suppress their own reports. This would lead to an undesirable flooding effect of messages to the MHGW from a HH once a query is issued. The trivial solution would be to have the MHGW broadcast reports received from HHs on the satellite link so that other HHs could hear them. This would involve increasing the random timer delay to account for the time it takes for a report to reach the MHGW and be broadcast.

If reliable multicast delivery is desired and HHs are allowed to suppress their group membership reports, then the MHGW would not have information on the individual membership information of each group, and hence would not be able to guarantee delivery of packets to hosts. In this case, it would be better to remove the query option from IGMP and have all HHs send a membership report to the MHGW when they join or leave a group. To cover the case of lost packets, the report should be duplicated if an acknowledgment is not received within a specified delay timer. This method would cause problems during startup or end of a multicast session when all HHs try to join or leave group because the MHGW would be flooded with group messages. Therefore this technique is only suitable for groups with a small number of HH members.

On the other hand, if reliability is not desired, the MHGW can still forward reports over the satellite link so that other HHs may suppress their reports. Query-Requests need not be sent since Leave-Reports would also be broadcast. Hence if a Leave-R-eport is heard by a HH for a group it is still a member of, it sends another Join-Report to the MHGW after its delay timer for that group expires before it receives a Join-Report from another HH.

3.3.2 HCBT Subsystems

Before proceeding with our design specifications, several simplifying assumptions are made that introduce some level of abstraction so that details not immediately essential are delayed until needed. As we proceed, our model would be validated to determine how close it is to the design requirements, and new subsystems added so that the whole abstraction process is re-iterated. We consider the special case of the HCBT architecture illustrated in Figure 3.2 where:

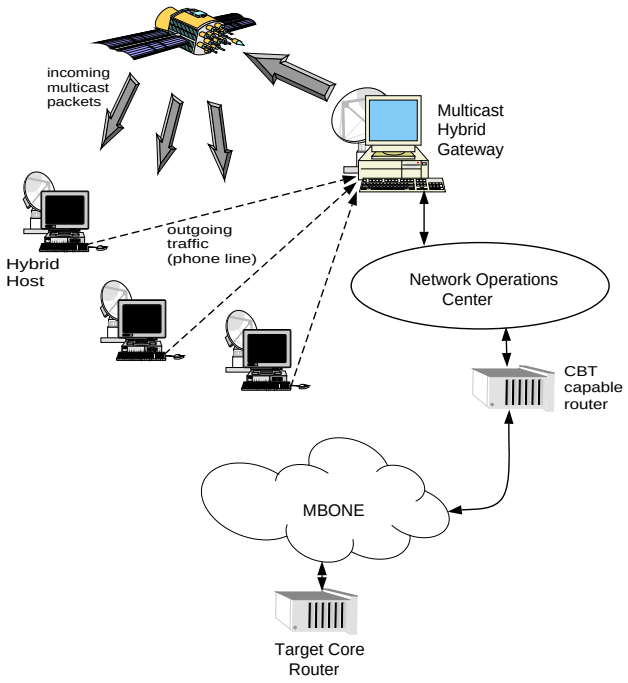

Figure 3.2: Simplified HCBT architecture.

- there are only "stand-alone" HHs, i.e. HHs are not attached on LANs where they are responsible for routing multicast packets to other terrestrial members.
- there is no intelligent routing at each HH, i.e. there is no differentiation among the different traffic types. Hence, all traffic from HHs goes out terrestrially and all incoming traffic is routed on satellite.
- all HHs that are multicast sources only send low data rate traffic.
- there is only one hybrid gateway serving all HHs.

Multicast Hybrid Gateway Subsystem.

Supporting multicasting in the architecture shown above requires implementing three new modules at the hybrid gateway; an IGMP module, a Multicast Database (MDB) module, and a Hybrid CBT router module (HCBT).

The IGMP module would run a modified version of IGMP and would be responsible for keeping track of group membership information of the HHs. It would query the HHs to determine which HHs are members of multicast groups. When it receives a group membership report from a HH, it would query the MDB to determine whether it has already joined the corresponding tree for that group. Once the corresponding tree has been joined, it would run an authentication process to authorize HHs to receive multicast packets.

The MDB module would maintain and manage a local database of trees joined by the MHGW. It would consist of entries denoting which multicast trees have been joined. Furthermore, for reliable "at least once" delivery of packets, this table will keep track of all hosts that are members of each group. The MHGW will keep a copy of all packets until they are acknowledged by all HHs in the group. It is necessary to separate this module from the HCBT module which contains a FIB with the same information because as we drop some of the assumptions made, it may be necessary to maintain more state information.

The HCBT module will run a CBT router function that enables the MHGW to join multicast trees on behalf of the HHs. It will be responsible for sending join messages towards the core of the tree and routing multicast packets to and from HHs.

On receipt of a IGMP report, the IGMP module will consult the MDB module to determine if it has already joined the corresponding tree of that group for the

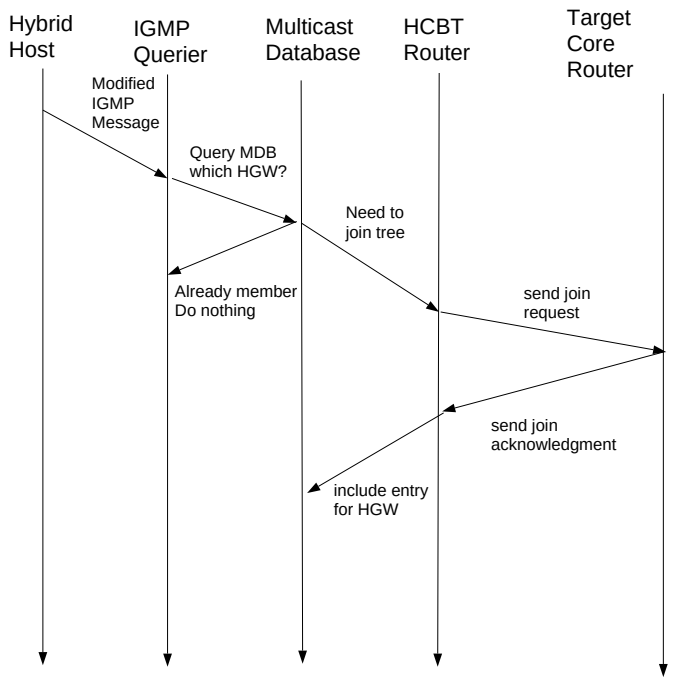

Figure 3.3: HCBT Tree Joining Process.

HGW responsible for the HH. If not, it will inform the HCBT router module of its intention to join the tree. The HCBT module would then send a join message towards the target core as specified in [8] for attachment to the multicast tree. After receiving an acknowledgment message, the HCBT module would include in its Forwarding Information Database (FIB), an entry corresponding to the tree joined and an entry will be added to the MDB specifying the HH as belonging to that group. It should be noted that the IGMP report arrives over a different interface than one where multicast packets are to be forwarded. Therefore, it would be necessary to modify CBT to include the correct interface to which the packet has to be sent. A timing diagram for the tree joining process is shown in Figure 3.3.

When the HCBT module receives a multicast packet, it would use the FIB

information to forward it over the satellite interface if there are HH group members. It would also encapsulate a copy of the packet and send it CBT mode to other interfaces as specified in the FIB since other CBT capable routers on the MBONE could join the delivery tree through it. Because the MHGW will need to keep a copy of all multicast packets until acknowledgment is received from all HH group members, a good buffering management scheme has to be devised.

Security has been of growing concern especially for multicast applications because it becomes relatively more difficult to distribute group keys to each of the group's receivers than to authenticate a session of a single source and destination. A scalable multicast distribution key has been described in [15] which uses CBT to establish secure multicast groups. The solution allows multicast routers to become Group Key Distribution Centers (GKDCs) after receiving a CBT Join ACK to become part of a multicast tree. Thereafter, the GKDCs are responsible for distributing group keys and key encrypting keys to group members on attached subnetworks. Therefore, we could have the MHGW act as the GKDCS for all HH group members and provide them with authentication keys. Because the keys would be broadcast on satellite, maintaining confidentiality would be difficult and extra precautions such as encryption techniques would need to be taken to ensure that only HH members receive packets.

Hybrid Host Subsystem

There are several functions that need to be implemented in the HH for it to support multicasting. The HH must run an IGMP module that allows it to listen for IGMP queries on its satellite interface and respond (send group reports) using its terrestrial interface. The HH has to be level 2 compliant with

IGMP to be able to both send and receive multicast packets. A mapping has to be defined between its satellite IP address and its adapter card to be able to forward packets destined for it up the TCP/IP stack. The host must be able to cache the keys sent to it by the MHGW during authentication so that it could be used for future multicast traffic.

When a HH wishes to be a member of a group, it sends a group membership report on its terrestrial link to the MHGW. The MHGW will then construct a delivery tree if needed, add the HH in the MDB, and then unicast an authentication key to the HH. The HH then listens on the satellite interface for packets destined for that group. When it receives packets, it sends acknowledgments to the MHGW via its terrestrial link.

It is important to note that as HHs join or leave groups, new keys may be broadcast by the MHGW. Therefore, a process running on the HH would need to renew keys for the HH. This process may need to periodically compare the checksum of its current key to that broadcast on satellite. If they are different, then it should trigger a request for new keys from the MHGW to be sent via unicast to prevent other non-member HHs from receiving key.

The multicast packets are broadcast to the HHs similar to the way multicast packets are sent to hosts attached to an Ethernet LAN. Hence, one way of receiving the multicast packets would be to make the HH physical interface (adapter) act like a single Ethernet link for the sake of carrying a multicast address. To achieve this, a socket has to be opened through which the relay application running on top of TCP or UDP can receive multicast packets. In the only hybrid Internet access product, DirecPC, developed by Hughes Network Systems for Windows using the architecture described in [13], there is a "special"

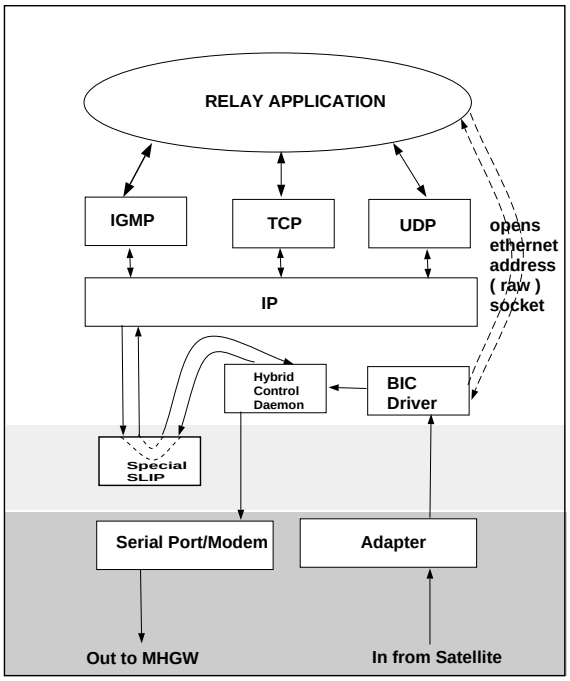

Figure 3.4: Flow control in the Hybrid Host

SLIP driver in the HH that communicates with the two physical networks to make the TCP/IP package believe that is connected to an Ethernet card when it is actually connected to a satellite dish and modem. A hybrid control daemon manages the flow of data between this special driver and a BIC driver. The BIC driver does all the call handling by scanning all packets transmitted over the satellite channel for one with a header corresponding to the IP address of the satellite interface. In addition, the BIC driver performs some error detection and correction on the packet and buffers the received packet before passing it to the special driver. Similarly for our system, the BIC driver call handler can be modified to support raw sockets and filter out UDP or TCP packets destined for multicast groups that relay applications have joined. Figure 3.4 shows the flow control in the HH subsystem.

When a HH is the source of multicast traffic, packets are encapsulated through the terrestrial tunnel to the MHGW. At the MHGW, they are decapsulated revealing their true multicast address destination and routed to other group members according to the distribution tree of the group. If group membership includes other HHs, the packets are also broadcast on satellite. Hence, additional filtering has to be done by the BIC driver to discard packets with HH address as the source.

3.4 Multiple MHGWs

With all multicast traffic to and from HHs routed through the MHGW, it is inevitable that there would be traffic congestion problems as the number of HH group members grow. Fortunately, provision has already been made in the current hybrid Internet access architecture to support multiple hybrid gateways (HGW) where each subnet of HHs are represented by different hybrid gateways. Packets to and from HH are routed first to the hybrid LAN gateway which broadcast it on the Ethernet LAN connecting all HGWs. The HGW takes up all routing tasks for all packets to and from HHs on its subnetwork.

To project this scenario to the multicasting case, the MHGW could be implemented at the HGW with one of the HGWs designated as the IGMP querier (DR-MHGW) responsible for joining multicast trees. When a HGW receives an IGMP group report from a HH, it will include the HH in its MDB module and broadcast a copy of the report on the LAN. The IGMP Querier will pick up the report and consult its FIB to determine if it has already joined the corresponding delivery tree. If not, it will trigger its HCBT module to send a join

message towards the core of the tree as specified in section 3.3.2. Similarly, when a multicast packet arrives, the DR-MHGW will broadcast it on the LAN and forward a copy over all interfaces (including satellite interfaces) as dictated by its FIB. All MHGWs with group members will buffer a copy until acknowledgment are received from all group members on its subnet. HHs with errored or missed packets will request their MHGW for retransmission of packets. This will significantly reduce the buffering management complexity at the MHGW. Also, since the DR-MHGW will be the only one attached to the delivery tree, all other MHGWs need not run a HCBT module.

3.5 Core Selection and Migration

A major problem of CBTs is that shared trees built incur a high traffic concentration on the shared path. Furthermore, the tree built is not always the shortest path tree.. It is believed that strategic core placement would help eliminate these problems completely. This would require developing core migration techniques that allow the dynamic transition from an initial CBT tree constructed around a pre-configured set of cores to another tree with different set of cores. The authors of CBT have not completely solved the core placement or core advertisement problem, but have defined a dynamic source migration mechanism in the appendix of [8]. This strategy allows a CBT tree to dynamically reconfigure itself around the source's local CBT router to emulate a shortest path tree.

The network architecture assumed for this solution routes all multicast traffic to and from hybrid hosts through the MHGW. As suggested in Section 3.2, this compliments the CBT design that routes multicast traffic along a shared

tree towards core. A lot of research has been done on determining core selection methods for multicast routing and how it affects performance in [16], [17], and [18]. Specifically, three performance criteria, bandwidth, delay, and traffic concentration, are considered to investigate the effect core choice has on them. In their evaluation, the authors of [16] considered instances of three different types of scenarios, reflecting distributions and numbers of sources and receivers. An "All Receivers Sources" scenario modeled applications such as video conferencing where receivers are distributed randomly throughout the network and a user is both a source and a receiver. "Single Source, Distributed Receivers" covered applications such as a video broadcast of a lecture or meeting where most members are receivers. Finally, "Localized Receivers" modeled distributed resource discovery applications where sources (clients) are randomly distributed and request information from receivers (servers) via multicast. In addition, core selection methods were classified into one of the following categories in increasing order of information required about the network: arbitrary, random, topology-based, or group-based, where arbitrary requires no information about the network and group-based requires information on both network topology and location of nodes. From the studies in [16] and [17] it was established that the best performance - maximum bandwidth improvement and minimum delay degradation is obtained from a core chosen based on both the network topology and location of nodes (receivers), although the improvement was not significant for certain distribution scenarios. Furthermore, it was established that the core should be the center of the portion of the shortest path trees that spans all group members and sources.

Traffic distribution in hybrid networks can be best modeled by a "single

source, distributed receivers" since it was developed based on the assumption that traffic is asymmetric with most users receiving much more than they are sending. Hence, multicast applications in hybrid networks would mostly be of video broadcasting nature. Since in the HCBT architecture described, the MHGW is responsible for routing of all multicast packets, it acts as a source to the HHs and the rest of the multicast network is hidden from the hosts. On the other hand, CBT mode allows users to unicast all multicast packet towards the core of the group using encapsulation. Once the packet reaches the core, it is decapsulated and forwarded out of all outgoing interfaces. Therefore, to emulate the shortest path tree and minimize delay for HHs, it makes sense to select the MHGW as a core for all groups joined. The MHGW could be configured to be a core for all groups joined. Alternatively, since dynamic core migration has been specified in [8], the MHGW could be configured to trigger a core migration to itself after it joins a tree for a group. The disadvantage of making the MHGW a core is that additional processing power may be required to process CBT protocol messaging. Introducing multiple cores would keep this to a minimum and would also reduce the traffic concentration problems inherent in shared links.

Chapter 4

Analysis and Evaluation

4.1 Performance Metrics of Multicast Protocols

To evaluate performance of a multicast protocol, several indicators are used to see how well the protocol performs under different scenarios. For dynamic multicast routing, it is important to determine the latency involved in joining the multicast group, from the time the request is sent by a host to the time the first multicast packet is received. It is desirable that this latency be kept to a minimum.

However, the main performance metric used is the time it takes for each member of the group to receive packets sent, i.e. transfer time. The transfer time depends on the throughput of the multicast session which in turn depends on both the available bandwidth and the probability of packet loss. Thus reducing the transfer time involves using a congestion control scheme to ensure that available bandwidth is not exceeded, and at the same time controlling packet loss in the delivery path. In order to control packet loss, it is essential to first

understand the underlying process and identify the source of losses so that the appropriate error control measures can be taken. Packet loss can be due to transmission or switching errors and buffer overflows at routers and hosts. A lot of studies have been done to determine packet loss correlation in multicast networks, and in [19], it was shown that losses on the MBONE are in fact "temporally" correlated, i.e., most losses occur at receivers and routers and not on links.

Topology of the multicast distribution tree also affects the packet loss characteristics and consequently, the transfer time. Mishra et. al. in [20], a study done to evaluate the effects of topology on reliable multicast routing, conclude that as a general rule, a topology which increases "fanout" of the distribution tree performs better in the asymptotic case.

Traffic concentration on the links in the distribution tree is also used as a performance indicator. Multicast routing protocols that construct shared trees experience a higher concentration when compared to source-based trees [17]. Path cost in terms of the the number of links transversed when delivering a packets to all group members also gives a good estimate of the bandwidth used. Other metrics used include overhead traffic of protocol, scalability, and protocol algorithmic complexity.

Because of the high-delay satellite link involved, the most important, metric for the protocol proposed in this thesis is the transfer delay in delivering multicast packets to all hybrid hosts since the remainder of the delivery path is terrestrial. We will assume when estimating the delay in subsequent sections that packet losses on the satellite link are insignificant and that most losses occurs at the hybrid host receivers due to overflow of buffers. This is actually quite close

to reality because most current and future satellite systems incorporate strong Forward Error Correction (FEC) protection so that up to a certain signal-to-noise ratio (SNR), the satellite channel can be modeled as an on-off switch.

4.2 Traffic Model of MHGW

IP multicast applications on the MBONE are implemented above the UDP layer which does not guarantee delivery of packets to all group members. Reliability is generally introduced in the application layer which takes up the responsibility of requesting for lost or errored packets.

For our hybrid multicast system, reliable delivery is an essential feature since most of the group recipients are HHs and packets are transmitted over a satellite link with high delay. Therefore, it becomes necessary for the MHGW to take up the responsibility of ensuring reliable delivery of packets to the HHs since the round-trip-time would be significant if the HH has to acknowledge each packet back to source. This approach has been taken by reliable multicast protocols such as Reliable Multicast Transport Protocol (RMTP) [21], which aggregates group members and assigns select nodes in delivery tree the responsibility of assuring reliable delivery of packets to downstream nodes.

As mentioned in Section 3.3, the MHGW will buffer a copy of all multicast packets destined for group until they are delivered to all group members. Therefore for each multicast session, two queues are maintained at the MHGW, one for forwarded packets and another for buffered copies. Figure 4.1 shows the traffic model for delivery of packets to the HHs. The notation used is given in Table 4.1.

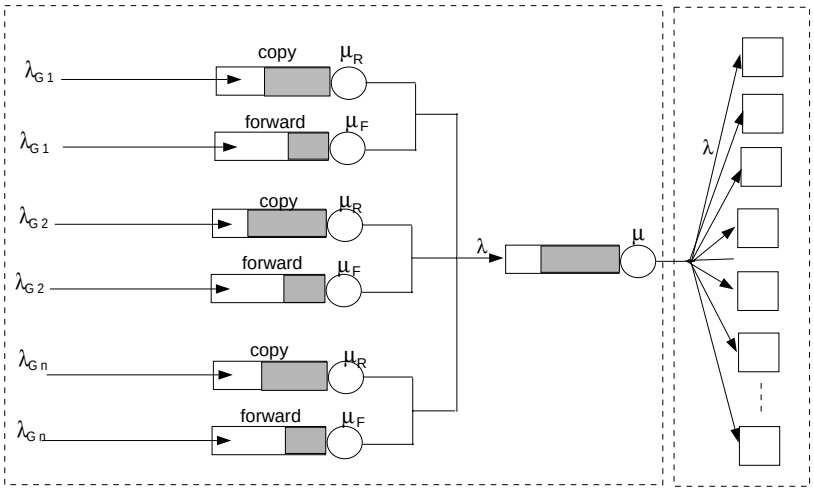

Figure 4.1: Traffic Model of MHGW.

λ_i	= arrival rate for group Gi
λ	= arrival rate at the satellite gateway
μ_F	= service rate for packets destined for HH
μ_R	= service rate for retransmission packets
μ	= service rate of packets at satellite gateway

Table 4.1: Notation for MHGW Traffic Model.

When a HH group member receives a packet it sends an ACK message back to the MHGW. There are two generic acknowledgment schemes currently used to inform sources about the status of lost packets. ACK_based schemes send information about packets correctly received while NACK_based schemes messages are sent only when packets have been lost. Even though NACK_based schemes generate a lesser amount of traffic, they do not guarantee reliability since in some situations, the sender may not be aware of lost packets. For complete reliability most systems employ block_based_ACK schemes where blocks of packets rather

P_i	= percentage of packets missed by at least one group member Gi
P_{Mi}	= percentage of packets missed by MHGW from source for group i
μ_i	= service rate for all packets destined for HH

Table 4.2: Notation for Alternative MHGW Traffic Model.

than individual packets are acknowledged by the receiver.

Once an ACK packet is received from all HH that are group members, the copy of the packet in the retransmission queue is discarded. If a packet is missed, it is retransmitted back to the group. Hence, the service rate of the retransmission queue, μ_R, depends on the packets requiring retransmission which in this case depends on the probability of buffer overflow. Alternatively, for each group, a single queue could be maintained but with different arrival rates for new packets and missed packets. It is assumed that any missed packet is transmitted back to the entire group and not to individual hosts . The arrival rate of missed packets would then depend on the percentage of missed packets by at least one member of the group. Figure 4.2 shows the new traffic model with the additional notation given in Table 4.2

It is obvious that there is high correlation between the various processes in the MHGW. Therefore, it may be difficult to do a precise queuing analysis. Instead, a steady state analysis could be done. Suppose we assume an M/M/1 queuing model, even though we do not have Poisson arrivals for all nodes (since there is feedback), it is still possible to obtain a **product form solution** for this model network.

The flow balance equations obtained are:

$$\lambda_i \quad = \quad \lambda_{Fi} + P_i\lambda$$

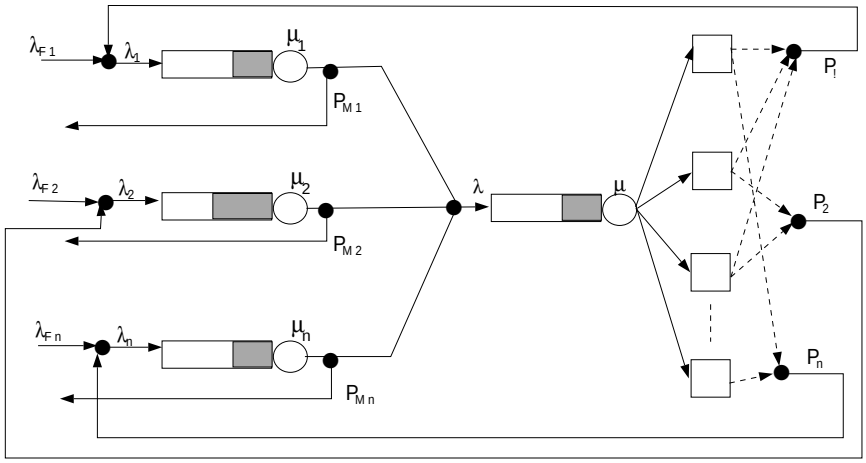

Figure 4.2: Alternative Traffic Model of MHGW.

$$\lambda = \sum_{i=1}^{n}(\lambda_i(1 - P_{Mi}))$$

Therefore

$$\lambda = \frac{\sum_{i=1}^{n}(\lambda_{Fi}(1 - P_{Mi}))}{1 - \sum_{i=1}^{n}(P_i(1 - P_{Mi}))}$$

$$\lambda_i = \lambda_{Fi} + P_i \frac{\sum_{i=1}^{n}(\lambda_{Fi}(1 - P_{Mi}))}{1 - \sum_{i=1}^{n}(P_i(1 - P_{Mi}))}$$

The state of the system is given by

$$S(n_1, ...n_i) = (1 - \rho)\rho^n \prod_{i=1}^{n}(1 - \rho_i)\rho_i^{n_i}$$

where $\rho_i = \frac{\lambda_i}{\mu_i}$

The expected length and waiting time for each queue are given respectively

by

$$E[X_i] = \frac{\rho_i}{1-\rho_i}$$
$$E[W_i] = \frac{\rho_i}{\mu_i(1-\rho_i)}$$

Thus the expected queue length and waiting time for the system are:

$$E[X] = \sum_{all\ i} \frac{\rho_i}{1-\rho_i}$$
$$E[W] = \sum_{all\ i} \frac{\rho_i}{\mu_i(1-\rho_i)}$$

thereby giving a total system time of

$$E[T] = \sum_{all\ i} \frac{\frac{1}{\mu_i}}{\mu_i(1-\rho_i)}$$

Suppose that there are two traffic types arriving into the system for each group, γ_h and γ_l, representing high_data_rate and low_data_rate traffic respectively, then

$$\lambda_{Fi} = \gamma_h + \gamma_l$$

If the service times for the two classes are independent and exponentially distributed with rates μ_h and μ_l respectively, it has been shown in [22] that the number of departures from the queue in an interval [0, t], for $t \geq 0$, is not a Poisson process, under any initial distribution π_0 of the state of the system. Thus we can no longer assume an M/M/1 model, but instead, an M/G/1 analysis needs to be done. The **Pollaczek-Khinchin formula** [23] gives the result of the mean value analysis of an M/G/1 queuing system with average queue length

$$E[X_G] = \frac{\rho}{1-\rho} - \frac{\rho^2(1-\mu_{trans}^2\sigma^2)}{2(1-\rho)}$$

where σ^2 = variance of service time distribution

and $\frac{1}{\mu_{trans}}$ = average service time

The traffic intensity at queues for each group would be given by

$$\rho = \frac{1}{1 - P_i}(\frac{\gamma_h}{\mu_h} + \frac{\gamma_l}{\mu_l})$$

The mean waiting times at both the source and receivers can be obtained, and consequently, the overall delay involved in delivering multicast packets to all group members can be obtained. Kurose et.al. show in [24] that performance depends on the acknowledgment scheme used and provided numerical results to prove that a NACK-based protocol that limits NACK generation by intentionally and randomly delaying NACK packets can achieve substantially higher throughput than other point-to-point NACK or ACK-based schemes.

4.3 Analytic Delay Model

A thorough analytic model has been derived in [20] to compute the average throughput seen by a multicast session from a source and this basic model has been adopted to determine the throughput at the MHGW.

Let $R(G)$ be the number of transmissions of a packet from a group G until all HH group members receive the packet and let p^i be the probability that all i attempts to deliver packets to all HH receivers will fail during a satellite broadcast. Then the probability distribution function $F_G(i)$ is given by

$$F_G(i) = P[R(G) \leq i] = 1 - p^i$$

Therefore, the average number of times a packet has to be retransmitted by the MHGW before it is received by all HH receivers in a multicast group is

$$E[R(G)] = \sum_{i=0}^{\infty} (1 - F_G(i))$$

Thus the effective packet loss as seen by the source would be given by

$$P_G^{eff} = 1 - \frac{1}{E[R(G)]}$$

Since packets are transmitted in block sizes of β_G before waiting for acknowledgments from HHs, if the interarrival time between packets is t_r, the time taken to transmit β_G packets is $\beta_G * t_r$. Out of the β_G packets transmitted in each round, a few may be lost at the receivers due to buffer overflow. If $rounds(\beta_G)$ denotes the number of transmission of a block containing β_G packets until all packets are received by HH group receivers, then

$$P[rounds(\beta_G) \leq k] = F_G(k)^{\beta_G}$$

It follows that the expected number of rounds will be

$$\begin{aligned} E[rounds(\beta_G)] &= \sum_{i=0}^{\infty} P[rounds(\beta_G) > k] \\ &= \sum_{i=0}^{\infty} (1 - F_G(k)^{\beta_G}) \end{aligned}$$

Each packet requires an average $E[R(G)]$ transmissions for successful delivery. Therefore the average transmission delay for a block of size β_G is

$$E[D_{trans}] = \beta_G t_r E[R(G)]$$

and the waiting delay is given by

$$E[D_{wait}] = RTT_{max} E[rounds(\beta_G)]$$

where RTT_{max} is the round_trip_time from the source to the farthest receiver. Therefore the average delay to multicast N packets in blocks of β_G is

$$D_{\beta_G}(N) = (E[D_{trans}] + E[D_{wait}]) \frac{N}{\beta_G} \tag{4.1}$$

$$= N t_r E[R(G)] + RTT_{max} E[rounds(\beta_G)] \frac{N}{\beta_G} \tag{4.2}$$

The block size β_G is limited by the size of the buffers at the receiver HH. Biersack and Nonenmacher [25] derived formulas for computing both the probability mass function (pmf) of the number of receivers that successfully receive a multicast packet and the mean number of retransmissions until all receivers successfully received a packet. The main result they arrived at was

$$E[R(G)] \approx pL$$

for $pL \leq 1$, where L is the number of links in the multicast tree and p is the link loss probability due to loss in receiver buffers. This approximation was used to compute the transfer delay (see Equation 4.2) for multicasting in a terrestrial and hybrid network and the results are graphed in Figure 4.3. The interarrival time between packets, t_r of 0.001 secs with a fixed packet size of 9000 bits, corresponding to a maximum sending rate of 9 Mb/s. A loss probability of 0.001 and 0.03 was assumed for hybrid and terrestrial routing respectively, representative of losses on the MBONE [19]. The satellite delay was assumed to be 250ms and 3 hops was assumed between source of multicast packets and group recipients in both cases with the same topology.

Case 1: Terrestrial Routing

In the case where multicasting is done terrestrially, the effective packet loss probability seen by the source increases as the number of receivers increases, especially if the receivers are widely distributed as expected for Single-Source Many-Receivers applications under consideration. This is because the number of links needed to reach all receivers is high and thus $E[R(G)]$ becomes high. Since $E[rounds(\beta)]$ can be computed from the probability distribution function, it follows that both transmission and waiting components have an effect on

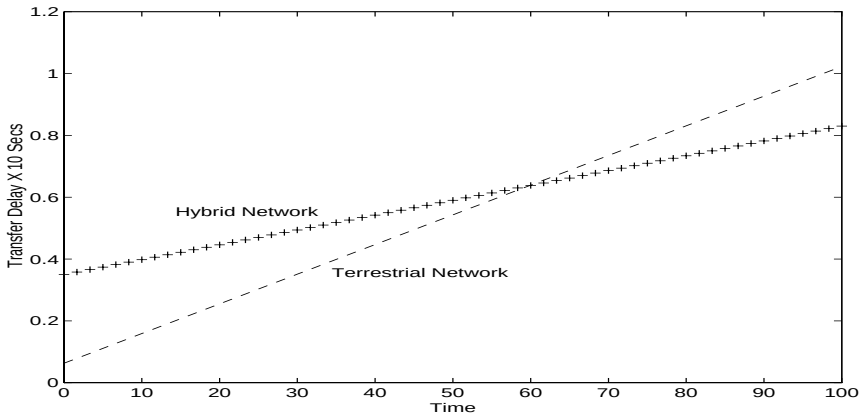

Figure 4.3: Analytic Transfer Delay Computation

the average time taken to multicast packets. When the number of packets N is small corresponding to "short-length" or low_data_rate packets, the average time remains small. However, as the the length of the session increases, the delay increases considerably since the intermediate nodes have more processing burden, thus increasing both $E[R(G)]$ and $E[rounds(\beta)]$.

Case 2: Hybrid Routing

In the hybrid case where multicast packets are routed via a high delay satellite link to HHs, the dominating term in Equation 4.2 becomes RTT_{max}. Because multicasting packets to most of the group receivers utilize the same link, L is small compared to the terrestrial case, and consequently $E[R(G)]$ and the effective packet loss seen by the source is considerably smaller. Thus, even as N increases, the average delay remains fairly constant, thus making hybrid networks more favorable for high_data_rate or lengthy sessions.

Another performance bottle neck is the buffer capacity of the MHGW and HHs. Because the MHGW acknowledges packets on behalf of the HHs, if the

buffer size is high and the block size is small, performance can be considerably improved by buffer space utilization techniques such as overlapping transmission.

Performance could also be improved by reducing the RTT_{max} using "good" core selection method. As previously discussed in Section 3.5, core selection can be used to alleviate both delay and traffic concentration problems. Dynamic core migration techniques have already been described for migrating core to the source to emulate a shortest path tree. However, in our case, it would be advantageous to designate the MHGW as the primary core since the shared tree built will would delivery packets from most members in the the shortest possible time. In addition this would reduce the computational load on the MHGW since the primary core is not required to send control JOIN messages. Thus, when an IGMP report is received from the HH, the MHGW would not have to do anything to join the tree. All it has to do is to include the satellite interface in the FIB table. Core migration to the MHGW should only be triggered by existence of HH members and should be disabled when there are no longer HH members for that group.

4.4 Performance Evaluation

The systems engineering approach to design and development of systems guides us to revisit out design once it is complete, build a prototype, and validate it against our initial design requirements and specifications. Optimization tools can then be used to enhance system performance. Simulation techniques were used to verify and validate our design instead of an actual prototype because it gives us a quicker methodology for evaluating performance and more flexibility

with modifying design parameters once the model is built.

One of the motivating factors of supporting multicasting in hybrid satellite terrestrial networks is to allow companies with limited gateway bandwidth to engage in high-bandwidth multicast applications. Therefore, a simulation was done to evaluate the bandwidth savings in multicasting over a hybrid network over traditional terrestrial wireline multicasting. Since some of these applications may have a time constraint on the transmission time, further studies were done to find the effects of high-delay satellite link on a multicast session. Finally, we investigate the use of traffic type and size in deciding whether it is advantageous to route multicast packets through satellite or not.

Simulation Model

All simulations were done using OPtimized Network Engineering Tools (OPNET), a comprehensive engineering system capable of simulating communications networks with detailed modeling and performance analysis. OPNET features include: graphical specification of models; a dynamic, event-scheduled Simulation Kernel; integrated data analysis tools; and hierarchical, object-based modeling. OPNET's hierarchical modeling structure accommodates special problems such as distributed algorithm development.

Two OPNET network models were built to simulate two environments: one in which multicasting is done terrestrially, and another in which all multicast packet are routed through a hybrid network over satellite to HH group members. The scenario under consideration is "single-source distributed receivers" typical of applications such as video lecture broadcast, with listeners (HHs) allowed to send only low_data_rate traffic to group since they may have limited uplink bandwidth.

Parameter	Description
Traffic Type	Rate of packet generation for the multicast source (Server) and hybrid host (Source) to model high-data-rate and low-data-rate type traffic.
Traffic Size	Size of transaction requested to model "short-length" and "bulky" traffic.
Service Rate	Service rate of packets destined for HHs at MHGW and Corporate Gateway buffers

Table 4.3: Simulation Parameters.

It is assumed that there are additional group members in the terrestrial network in the viscinity of the source. The same number of hops are used between the HHs and source in both environments. Standard OPNET TCP/IP processes were modified when appropriate to build the simulation model. For the hybrid network model, some of the processes used in [26], a simulation of a hybrid network, were also modified to support routing of multicast packets. The same network topology assumed in the analysis section was used in the simulation. The simulation parameters used are given in Table 4.3 and their values are given in Appendix B.

Simulation Results

Figure 4.4 compares the corporate link utilization for each of the two environments considered: multicasting in a terrestrial network (wireline) versus a hybrid network with a satellite downlink. As expected, introducing hybrid terminals in a corporate LAN preserves corporate gateway bandwidth for other

traffic. Since all incoming packets for hybrid case are routed through satellite, available bandwidth for other traffic is more than twice the bandwidth available when incoming multicast packets are routed terrestrially.

Figure 4.5 shows the round-trip-time (RTT) of packets for both hybrid host and terrestrial host group members. As the session length increases, the performance of the terrestrial network considerably declines while that of the of hybrid network remains stable, following the same trend obtained in the analytic studies (see Figure 4.3). In the terrestrial network, the RTT of packets initially slows down but increases quickly because the corporate gateway is slowed down by the additional packets to be processed. Thus, more packets are transfered as indicated by the increase in throughput at the HHs observed in Figure 4.6 for the hybrid network case.

Figure 4.7 shows the effect of traffic type and size on the transfer time of multicast packets. From the figure, it can be seen that the delay is less in the terrestrial network for "short-length" or low_data_rate sessions (see Table B.3). Thus under such a scenario, it is not advantageous to route multicast packets through satellite. This clearly demonstrates the need for an intelligent routing scheme at the MHGW as suggested in Section 3.1 that would allow only high_data_rate or "bulky" traffic to be routed via satellite.

Our analytic studies suggest that a major performance bottleneck through the hybrid network is the buffer capacity of the MHGW and HHs. Figure 4.8 shows a comparison of the throughput for different MHGW buffer sizes (see Table B.4). From the figure, it can be seen that the achievable throughput is higher when a large buffer size is used since this allows the source to send larger amounts of data by advertising a larger window size.

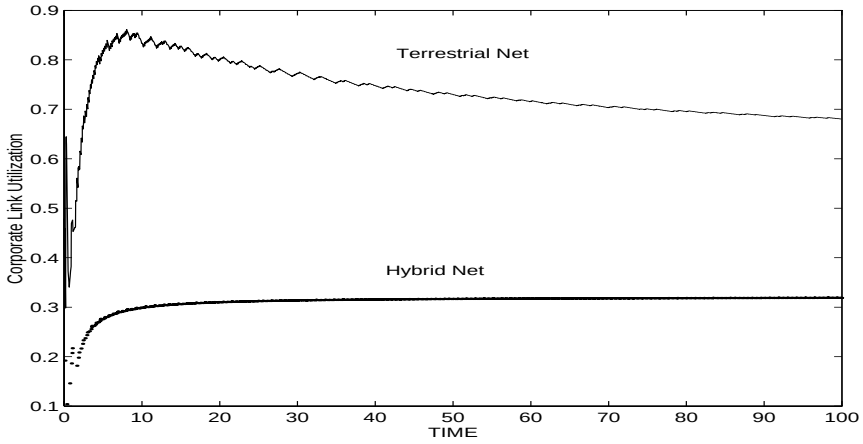

Figure 4.4: Corporate Link Utilization Comparison

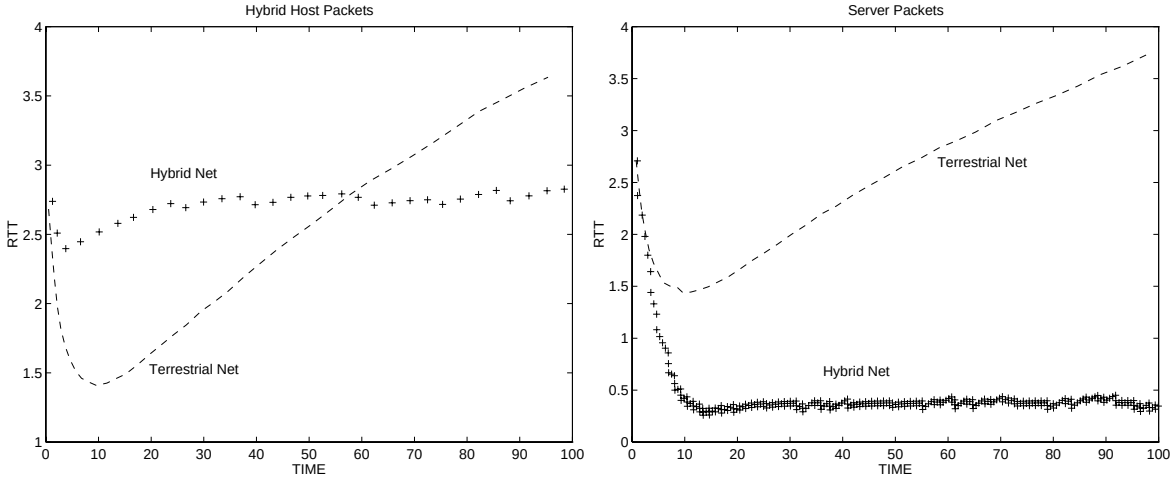

Figure 4.5: Round-Trip-Time Comparison

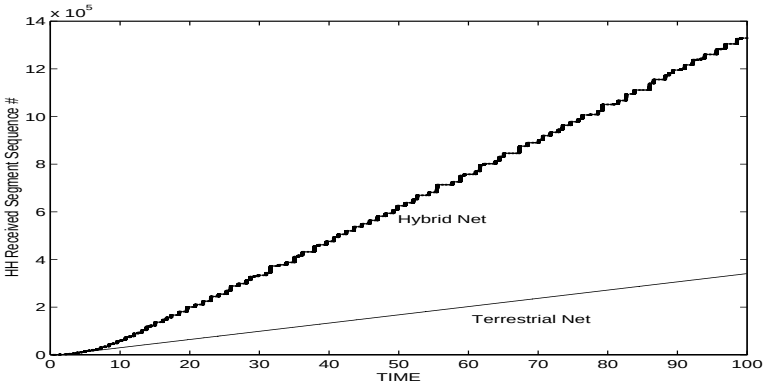

Figure 4.6: Received Segment Sequence Number of HH Packets

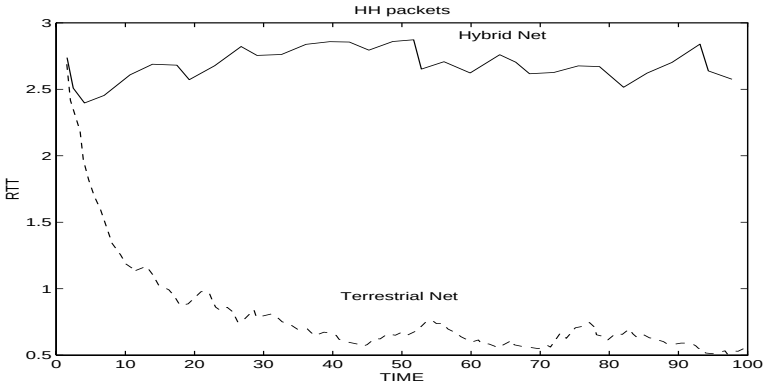

Figure 4.7: Effect of traffic type on RTT

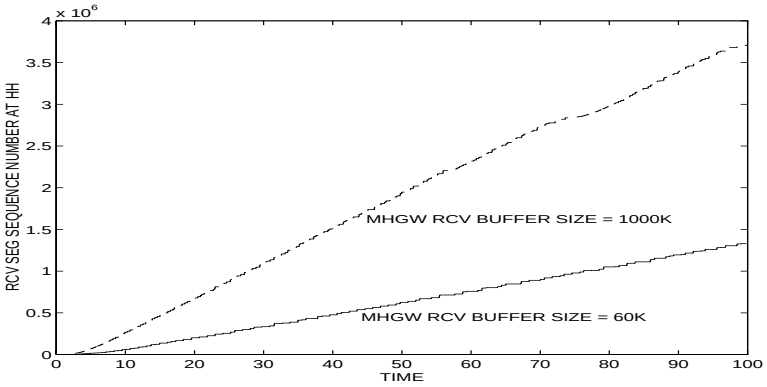

Figure 4.8: Effect of buffer size on Throughput

Chapter 5

Conclusions & Further Research

The asymmetric nature of traffic in most networks, as evident in the Internet, is shifting current networking technology trends more towards the development of hybrid networks. Emerging group communication applications such as video broadcasting and teleconferencing that demand high bandwidth have driven the development of multicast protocols on the MBONE. Thus hybrid terminals can be deployed for receiving IP multicast packets as a means of preventing congestion on the Internet backbone and preserving Corporate gateway bandwidth.

The goal of this thesis was to develop a system-level design of a "demand-assigned" multicast routing protocol that would allow hybrid hosts in satellite-terrestrial hybrid networks to dynamically join and leave multicast groups. In coming up with the design, the systems engineering approach which emphasizes translation of a systems needs to a set of formally written requirements was taken. The design presented minimized changes to the current hybrid Internet access architecture [12] and eliminated changes to other multicast routers on the Internet.

The design entailed implementing a Multicast Hybrid Gateway that would

be responsible for keeping track of group membership of hybrid hosts, join delivery trees of multicast sessions, and route multicast packets to hybrid host group members. In addition, this subsystem could also handle authentication or intelligent routing schemes. The IGMP protocol was modified to emulate a virtual link between hybrid hosts and the gateway. Also, the multicast routing protocol employed in the terrestrial part of the network was assumed to be CBT and thus appropriate changes were done in the CBT module of the hybrid gateway.

We studied the steady state behavior of the Multicast Hybrid Gateway and did an analytic evaluation of our system. Our analysis showed that as the number of packets to be multicast increases, despite the high delay in the satellite link, the average delay in multicasting packets to all group members is lower in a hybrid network than in a pure terrestrial network. Our results also suggests that the bottleneck on delivery of multicast packets to hybrid hosts is the buffer capacity at both the gateway and hosts. Thus, delay could be improved by overlapping block transmission cycles or using good core selection techniques that would place gateway virtually next to the source of multicast packets.

Simulation techniques were used to demonstrate the bandwidth savings in multicasting in hybrid networks over terrestrial counterpart. Our simulation results agreed with our mathematical analysis that it is only advantageous to route packets over satellite if there is high_data_rate or "bulky" traffic, thus indicating a need to implement intelligent routing at the gateway, Finally, our results also agree with analytic studies that show that an increase in buffer size also improves performance.

Further studies should include a detailed protocol design of the group membership protocol including the protocol messaging format. Also, a thorough

study of the buffer management technique used to ensure reliability as well as the authentication scheme used to prevent intruders from receiving broadcast packets should be done. More analytic studies on other performance metrics should be done and simulation carried out to support study. Specifically, the scalability of the design as the group membership of hybrid hosts and number of groups increases should be carefully studied as well as its effect on traffic concentration on links leading to the gateway. Furthermore, the protocol overhead and latency in joining and leaving groups would give a good assessment of the protocol.

Finally, with the ever growing popularity of hybrid networks, more studies should be done to determine how to extend multicasting to other hybrid networks, such as fiber-coaxial networks, using the same concept proposed in this thesis.

Appendix A

Acronyms

ACK - Acknowledgement Message

CBT - Core-Based Trees

DR - Designated Router

DVMRP - Distance Vector Multicast Routing Protocol

FIB - Forwarding Information Base

GDKC - Group Key Distribution Center

HCBT - Hybrid Core-Based Trees

HGW - Hybrid Gateway

HH - Hybrid Host

IETF - Internet Engineering Task Force

IGMP - Internet Group Management Protocol

IP - Internet Protocol

MBONE - Multicast Backbone

MDB - Multicast Database

MHGW - Multicast Hybrid Gateway

MOSPF - Multicast Open Shortest Path First

PIM-SM - Protocol Independent Multicast Sparse Mode

PIM-DM - Protocol Independent Multicast Dense Mode

OPNET - Optimized Network Engineering Tools

OSPF - Open Shortest Path First

RIP - Routing Independent Protocol

RTT - Round Trip Time

RP - Rendezvous Point

SPT - Shortest Path Tree

TCP - Transport Control Protocol

TTL - Time to Live

UDLR - Uni-directional Link Routing

UDP - User Data Protocol

Appendix B

Simulation Parameter Values

Table B.1 and B.2 shown below present the important parameter values used in the simulation.

Parameter	Value
Source Application Interarrival Rate	0.001secs/pk
Packet Size	9 Kbits
Hybrid Host Application Interarrival Rate	0.5 secs/pk
HH RCV Buffer Size	45 Kbytes
MHGW RCV Buffer Size	64 Kbytes
Source RCV Buffer Size	4 Kbytes
Modem Speed	28.8 Kbits/sec

Table B.1: Parameter Values for Hybrid Network.

Parameter	Value
Source Application Interarrival Rate	0.001secs/pk
Packet Size	9 Kbits
Hybrid Host Application Interarrival Rate	0.5 secs/pk
HH RCV Buffer Size	45 Kbytes
Source RCV Buffer Size	4 Kbytes

Table B.2: Parameter Values for Terrestrial Network.

Parameter	Value
Source Application Interarrival Rate	0.5secs/pk
Hybrid Host Application Interarrival Rate	0.5 secs/pk

Table B.3: Parameter Values for Low-Data-Rate Traffic.

Parameter	Value
HH RCV Buffer Size	45 Kbytes
MHGW RCV Buffer Size	1000 Kbytes
Source RCV Buffer Size	4 Kbytes
Source Application Interarrival Rate	0.001secs/pk
Packet Size	9 Kbits
Hybrid Host Application Interarrival Rate	0.5 secs/pk

Table B.4: Parameter Values To Show Buffer Size Effect.

Bibliography

[1] S. Deering. *"Host Extensions to IP multicasting," RFC 1112.* Stanford University, August 1989.

[2] W. Fenner. *"Internet Group Membership Protocol Version 2," INTERNET DRAFT.* Xerox PARC, May 1996.

[3] A. Thyagarajan B. Cain and S. Deering. *"Internet Group Membership Protocol Version 3,"INTERNET DRAFT.* Stanford University, September 1996.

[4] T. Pusateri. *"Distance Vector Multicast Routing Protocol (DVMRP Version 3)," INTERNET DRAFT: To replace RFC 1075.* Juniper Networks, February 1997.

[5] J. Moy. *"Multicast Extensions to OSPF," RFC 1584.* Proteon, Inc, March 1994.

[6] J. Moy. *"OSPF Version 2, " RFC 1583.* Proteon, Inc, March 1994.

[7] A. Ballardie. *"Core Based Trees (CBT) Multicast Architecture," INTERNET DRAFR.* University College London, July 1996.

[8] A. Ballardie, S. Reeve, and N.Jain. *"Core Based Trees (CBT) Multicast Protocol Specification," INTERNET DRAFT.* University College London, Bay Networks, July 1996.

[9] S. Deering et.al. *"Protocol Independent Multicast-Sparse Mode (PIM-SM): Motivation and Architecture, " INTERNET DRAFT.* Xerox PARC, et. al., October 1996.

[10] S. Deering et.al. *"Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification," INTERNET DRAFT.* Xerox PARC, et. al., March 1997.

[11] S. Deering et.al. *"Protocol Independent Multicast-Sparse Mode (PIM-DM): Protocol Specification," INTERNET DRAFT.* Xerox PARC, et. al., September 1996.

[12] N. Suphasindhu D. Dillon A.D. Falk V. Arora and J.S. Baras. "Hybrid Internet Access". *Proc Conference on NASA Centers for Commercial Development of Space, AIP No.325*, pages 69–74, July 1995.

[13] J.S. Baras V. Arora, N. Suphasindhu and D. Dillon. "Asymmetric Internet Access Over Satellite-Terrestrial Networks". *American Institute of Aeronautics and Astronautics*, June 1995.

[14] A. Ballardie. *"Core Based Trees (CBT) Multicast Interoperability," INTERNET DRAFT.* University College London, April 1996.

[15] A. Ballardie. *"Scalable Multicast Key Distribution," RFC 1949.* University College London, May 1996.

[16] E. Zegura K. Calvert and M. Donahoo. *"Core Selection Methods for Multicast Routing," Technical Report GIT-CC-95/15.* Georgia Institute of Technology, June 1995.

[17] E. Zegura K. Calvert and M. Donahoo. *"A Comparison of Two Practical Multicast Routing Schemes," Technical Report GIT-CC-94/25.* Georgia Institute of Technology, February 1994.

[18] L. Wei and D. Estrin. *"The Trade-off of Multicast Trees and Algorithms, " Technical Report USC-CS-93-560.* University of Southern California, June 1993.

[19] J. Kurose M. Yajnik and D. Towsley. "Packet Loss Correlation in the MBONE Multicast Network". *IEEE Global Internet Conference*, December 1996.

[20] P. Mishra P. Bhagwat and S. Tripathi. "Effect of Topology on Performance of Reliable Multicast Communication". *Proceedings of INFOCOM 94*, pages 602–609, June 1994.

[21] J. Lin and S. Paul. "RMTP: A Reliable Multicast Transport Protocol". *IEEE Proceedings of INFOCOM 96*, June 1996.

[22] Jean Warland. *An Introduction to Queuing Networks.* Prentice Hall, 1st. edition, 1988.

[23] Christos G. Cassandras. *Discrete Event Systems: Modeling and Performance Analysis.* Aksen and IRWIN Associates, 1st. edition, 1993.

[24] J. Kurose M. Yamamoto, D. Towsley and H. Ikeda. "A Delay Analysis of Sender-Initiated and Receiver-Initiated Reliable Multicast Protocols". *IEEE Proceedings of INFOCOM 97*, April 1997.

[25] J. Nonnenmacher and E. Biersack. "Performance Modeling of Reliable Multicast Transmission". *IEEE Journal on Selected Areas in Communication*, pages 548–558, May 1997.

[26] A. Gaid. *"Simulation of a Hybrid Network in Order to Enhance Performance of Hybrid Internet Service," Technical Report.* University of Maryland CSHCN, December 1996.

[27] Christian Huitema. *Routing In The Internet.* Prentice Hall, 1st. edition, 1995.

[28] Douglas Comer. *Internetworking With TCP/IP VOL I.* Prentice Hall, 3rd. edition, 1995.

[29] D. Stevens D. Comer. *Internetworking With TCP/IP VOL II.* Prentice Hall, 2nd. edition, 1994.

[30] J. Kurose D. Towsley and S. Pingali. "A Comparison of Sender-Initiated and Receiver-Initiated Reliable Multicast Protocols". *IEEE Journal on Selected Areas in Communication*, pages 398–406, April 1997.

[31] Y. Zhang and S. Dao. *"Integrating Direct Broadcast Satellite with Wireless Local Access, " First Intl Workshop on Satellite Based Information Services.* Hughes Research Laboratories, November 1996.

[32] J.S. Baras V. Arora, N. Suphasindhu and D. Dillon. *"Effective Extensions of Internet in Satellite-Terrestrial Networks," Technical Report.* University of Maryland CSHCN, Hughes Network System, June 1995.