1 Interdomain Multicast Fundamentals

This chapter introduces and describes the fundamental concepts of multicast. Subsequent chapters build upon these concepts, illustrating how they are specifically used in the protocols and technologies that enable the operation of interdomain multicast. This chapter also defines terms and conventions that will be used throughout the book.

1.1 What Is Multicast?

The three main methods of data delivery are **unicast**, **broadcast**, and **multicast**. These methods are summarized as follows:

- **Unicast:** Data is delivered to one specific recipient, providing one-to-one delivery.
- **Broadcast:** Data is delivered to all hosts, providing one-to-all delivery.
- **Multicast:** Data is delivered to all hosts that have expressed interest. This method provides one-to-many delivery.

The Internet was built primarily on the unicast model for data delivery (see Figure 1-1). However, unicast does not efficiently support certain types of traffic.

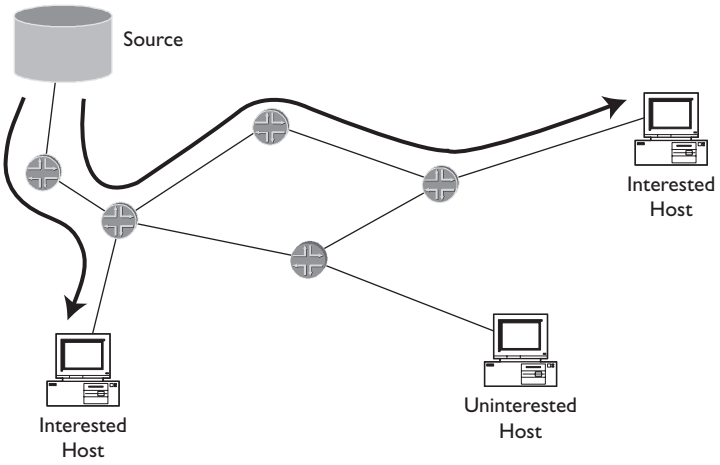


Figure 1-1 Unicast delivery

Multicast, originally defined in RFC 1112 by Steve Deering, provides an efficient method for delivering traffic that can be characterized as "one-to-many" or "many-to-many."

Radio and television are examples of traffic that fit the one-to-many model. With unicast, a radio station would have to set up a separate session with each interested listener. A duplicate stream of packets would be contained in each session. The processing load and the amount of bandwidth consumed by the transmitting server increase linearly as more people tune in to the station. This might work fine with a handful of listeners; however, with hundreds or thousands of listeners, this method would be extremely inefficient. With unicast, the source bears the burden of duplication.

Using broadcast (see Figure 1-2), the radio station would transmit only a single stream of packets, whether destined for one listener or for one million listeners. The network would replicate this stream and deliver it to every listener. Unfortunately, people who had not even tuned in to the station would be delivered this traffic. This method becomes very inefficient when many uninterested listeners exist. Links that connect to uninterested end hosts must carry unwanted traffic,

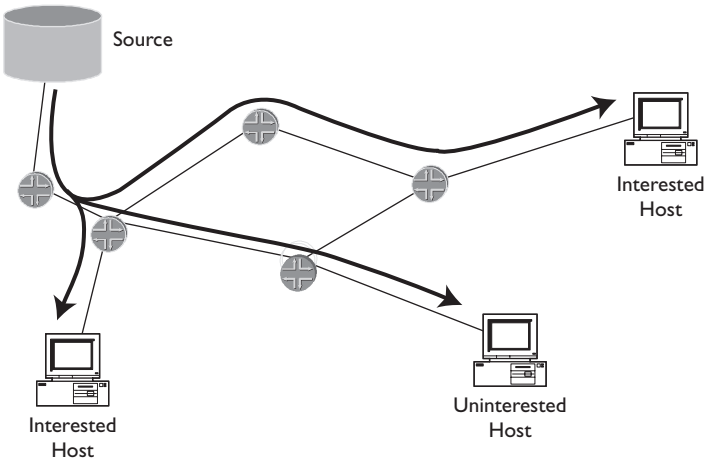


Figure 1-2 Broadcast delivery

wasting valuable network resources. With broadcast, the network carries the burden of delivering the traffic to every end host.

Multicast, on the other hand, provides the best of both worlds without introducing the disadvantages of each (see Figure 1-3). Multicast enables the radio station to transmit a single stream that finds its way to every *interested* listener. As in the case of broadcast, the processing load and the amount of bandwidth consumed by the transmitting host remain constant, regardless of audience size. The network is responsible for replicating the data and delivering it only to listeners who have tuned in to the station. Links that connect to uninterested listeners do not carry the traffic. This method provides the most efficient use of resources because traffic flows only through links that connect to end hosts that want to receive the data.

To deliver data only to interested parties, **routers** in the network build a **distribution tree**. Each subnetwork that contains at least one interested listener is a *leaf* on the tree. When a new listener tunes in, a new branch is built, *joining* the leaf to the tree. When a listener tunes out, its branch is *pruned* off the tree. Where the tree branches, routers replicate the data and send a single flow down each branch. Thus no link ever carries a duplicate flow of packets.

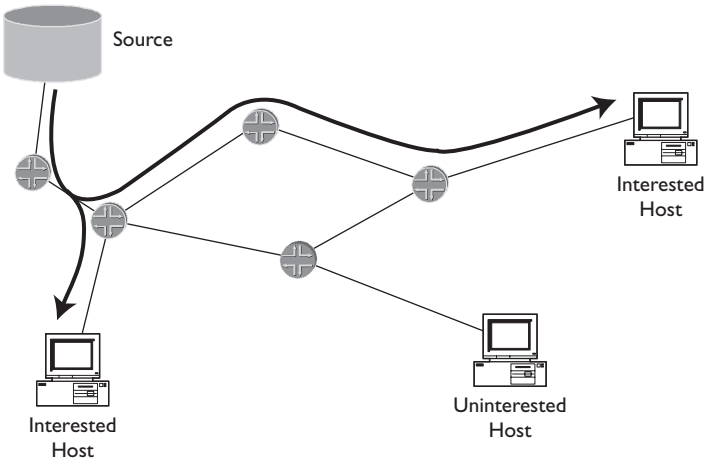


Figure 1-3 Multicast delivery

With multicast, the source is not burdened because it must transmit only a single stream of data, and the network is not burdened because it must deliver traffic only to end hosts that have requested it. However, in the zero-sum world of networking, where nothing is free, the burden of multicast falls on network engineers who must design and manage the mechanisms that make it work!

1.2 INTERNETWORKING BASICS

To facilitate the reader's understanding, this section covers some of the notation and conventions used in the book and thus indicates the level of the typical reader's internetworking knowledge anticipated by the authors.

Throughout the book we use the slash notation for bit mask when describing **IP** address ranges. The slash notation indicates how many bits of the address remain constant throughout the range of addresses. For example, 10.0.0.0/8 indicates a range of IP addresses all with the first 8 bits equal to 10. The range is from address 10.0.0.0 to 10.255.255.255.

We also make reference to **classful** networks. The class A, B, and C networks constitute all unicast IP addresses as follows:

- **Class A networks:** Describe the range of networks from 1.0.0.0/8 through 126.0.0.0/8.
- **Class B networks:** Describe the range of networks from 128.0.0.0/16 through 191.255.0.0/16.
- **Class C networks:** Describe the range of networks from 192.0.0.0/24 through 223.255.255.0/24.

Originally, networks were assigned to organizations along classful boundaries. That meant class A networks were assigned in /8 blocks, class B in /16 blocks, and class C in /24 blocks. Classful allocation was inefficient because organizations that required slightly more than 254 addresses could be assigned an entire class B. **Classless interdomain routing** (**CIDR**) enabled the assignment and routing of addresses outside of classful boundaries. An organization that needed enough addresses for 500 hosts could be assigned one /23, instead of an entire class B network.

All **multicast addresses** fall in the class D range of the **IPv4** address space. The class D range is 224.0.0.0 through 239.255.255.255. Multicast addresses do not have a **mask** length associated with them for **forwarding** purposes. Each address is treated independently so the mask used for forwarding is always assumed to be /32. We use shorter mask lengths on multicast addresses in some parts of the book for reasons other than forwarding. These masks generally are used to describe ranges of multicast addresses. For example, the address range reserved for **Source-Specific Multicast** (**SSM**) is 232.0.0.0/8.

We refer throughout the book to unicast and **multicast routing protocols**. **Unicast routing protocols** are used by routers to exchange routing information and build routing tables. Unicast IP routing protocols are further categorized into **interior gateway protocols** (**IGPs**) and **exterior gateway protocols** (**EGPs**).

IGPs provide routing within an administrative **domain** known as an **autonomous system** (**AS**). EGPs provide routing between ASs. **Routing Information**

Protocol (RIP), **Open Shortest Path First (OSPF)**, and **Intermediate System to Intermediate System (IS-IS)** are examples of IGPs, while **Border Gateway Protocol (BGP)** is an example of an EGP. Multicast routing protocols are used by routers to set up multicast forwarding **state** and to exchange this information with other multicast routers. Examples of multicast IP routing protocols are **Distance Vector Multicast Routing Protocol (DVMRP)**, **Protocol Independent Multicast–Dense Mode (PIM-DM)**, and **Protocol Independent Multicast–Sparse Mode (PIM-SM)**.

The terms **control packets** and **data packets** are used to differentiate the types of packets being routed through the network. Control packets include any packets sent for the purpose of exchanging information between routers about how to deliver data packets through the network. Control packets are typically protocol traffic that network devices use to communicate with one another to make such things as routing possible.

Data packets use the network to communicate data between hosts; they do not influence the way the network forwards traffic. Letters delivered via postal mail are analogous to data packets. Information exchanged between post offices to describe what ZIP codes mean is analogous to control packets. In the IP world, all packets sent for an **FTP** session between hosts are considered data packets, while a BGP **Update message** is an example of a control packet.

1.3 MULTICAST BASICS

A multicast address is also called a multicast **group address**. A group member is a host that expresses interest in receiving packets sent to a specific group address. A group member is also sometimes called a **receiver** or **listener**. A multicast *source* is a host that sends packets with the destination IP address set to a multicast group. A multicast source does not have to be a member of the group; sourcing and listening are mutually exclusive.

Because there can be multiple receivers, the path that multicast packets take may have several branches. A multicast data path is known as a *distribution tree*. Data

flow through the multicast distribution trees is sometimes referenced in terms of **upstream** and **downstream**. Downstream is in the direction toward the receivers. Upstream is in the direction toward the source. A downstream interface is also known as an *outgoing* or *outbound* interface; likewise, an upstream interface is also known as an *incoming* or *inbound* interface.

Routers keep track of the incoming and outgoing interfaces for each group, which is known as *multicast forwarding state*. The **incoming interface** for a group is sometimes referred to as the **IIF**. The **outgoing interface list** for a group is sometimes referred to as the **OIL** or **olist**. The OIL can contain 0 to N interfaces, where N is the total number of logical interfaces on the router.

Multicast forwarding state in a router is typically kept in terms of "**(S,G)**" and "**(*,G)**" state, which usually are pronounced "ess comma gee" and "star comma gee," respectively. In (S,G), the "S" refers to the unicast IP address of the source. The IP header of the multicast data packet contains S as the packet's source address. The "G" represents the specific multicast group IP address of concern. The IP header of the multicast data packet contains G as the packet's destination address. So for a host whose IP address is 10.1.1.1 acting as a source for the multicast group 224.1.1.1, (S,G) state would read (10.1.1.1,224.1.1.1).

In (*,G) notation, the asterisk (*) is a wild card used to denote the state that applies to any source sending to group G. A multicast group can have more than one source. If two hosts are both acting as sources for the group 224.1.1.2, (*,224.1.1.2) could be used to represent the state a router could contain to forward traffic from both sources to the group. The significance of (S,G) and (*,G) state will become more apparent when we discuss shortest path and **shared trees** in Chapters 2 and 3.

1.3.1 REVERSE PATH FORWARDING

Multicast routing involves a significant paradigm change from standard unicast routing. In general, routers make unicast routing decisions based on the destination address of the packet. When a unicast packet arrives, the router looks up the

destination address of the packet in its **routing table**. The routing table tells the router out from which interface to forward packets for each destination network. Unicast packets are then routed from source to destination.

In multicast, routers set up forwarding state in the opposite direction of unicast, from receiver to the root of the distribution tree. Routers perform a **reverse path forwarding (RPF)** check to determine the interface that is topologically closest to the root of the tree (see Figure 1-4). RPF is a central concept in multicast routing. In an RPF check, the router looks in a routing table to determine its *RPF interface*, which is the interface topologically closest to the root. The RPF interface is the incoming interface for the group.

In a **shortest path tree (SPT)**, the root of the distribution tree is the source. If a router learns that an interested listener for a group is on one of its directly connected interfaces, it tries to join the tree for that group. In Figure 1-5, this router somehow knows the IP address of the source of this group. To build an SPT, it executes an RPF check by scanning its routing table for the source address. The RPF check tells the router which interface is closest to the source. The router now knows that multicast packets from this source to this group should flow into the router through this RPF interface.

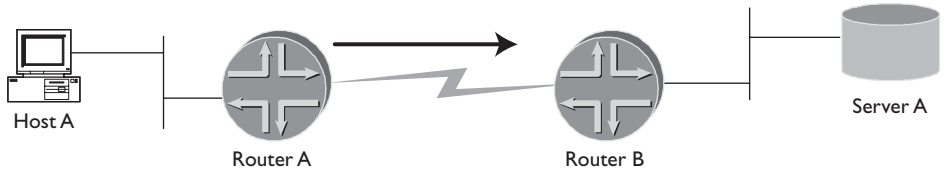
The router sends a **Join message** out the RPF interface to inform the next router upstream it wants to receive packets for this group from this source. This message is an (S,G) Join message. The router receiving the (S,G) Join message adds the interface on which it was received to the OIL for the group and performs an RPF check on the source. This upstream router sends an (S,G) Join message out its RPF interface for the source informing its upstream router that it wants to join the group.

Each upstream router repeats this process of propagating Joins out the RPF interface until this new branch of the tree either a) reaches the router directly connected to the source or b) reaches a router that already has multicast forwarding state for this source-group pair. In this way, a new branch of the tree is created from receiver to source. Once this branch is created and each of the routers has

1. Server A sends data packets to a specific multicast group, but at this point, router B does not know of any hosts interested in receiving them, so router B discards them.

2. Host A announces to router A its interest in receiving from server A multicast data packets that are destined for the specific multicast group.

3. Router A does an RPF lookup for server A's address revealing that router B is the RPF neighbor for server A's address. Router A requests that router B forward the data packets for the multicast group.



4. Now both routers know the correct interfaces out of which to forward the data packets. The data packets are delivered successfully from server A to host A.

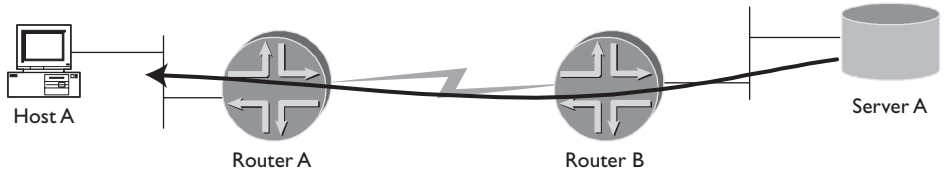


Figure 1-4 Reverse path forwarding (RPF)

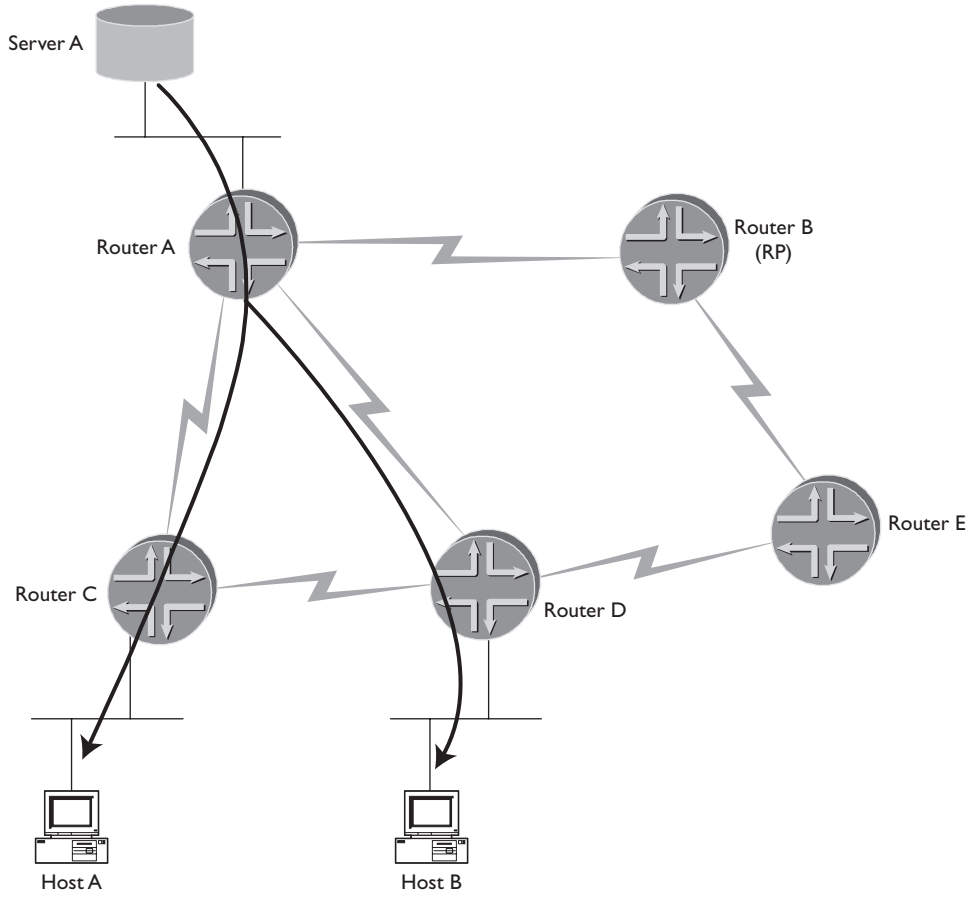


Figure 1-5 Shortest path tree (SPT)

forwarding state for the source-group pair, multicast packets can flow down the tree from source to receiver.

In a shared tree, the root of the distribution tree is a router somewhere in the core of a network. In PIM-SM, this core router is called a **rendezvous point** (**RP**). If a router learns that an interested listener for a group is on one of its directly connected interfaces, it tries to join the tree for that group. In Figure 1-6, this router

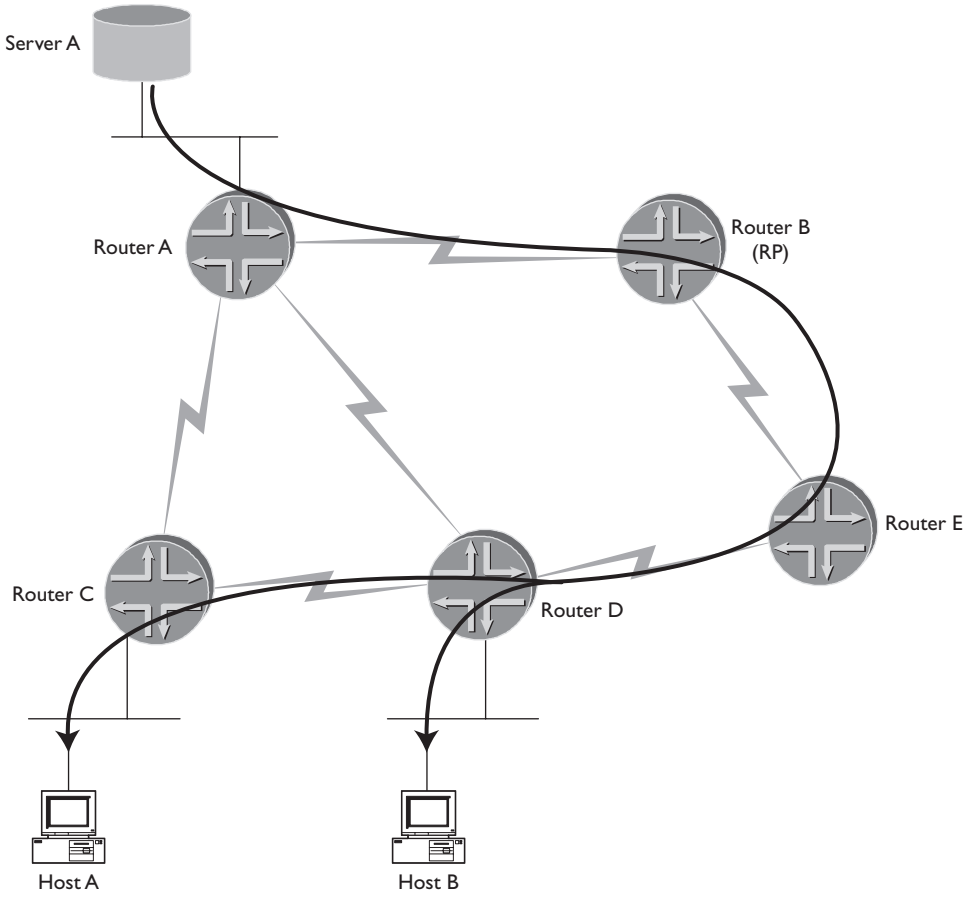


Figure 1-6 Rendezvous point tree (RPT)

does not know the address of the source of this group. However, it does know that another router in the network is aware of the source. The router that somehow knows the sources for all multicast groups is the RP (we will find out just how it knows this in Chapters 2 and 3).

The router with the directly connected listener, or the last-hop router, performs an RPF check for the IP address of the RP. This RPF check yields the RPF

interface for the RP, and a (*,G) Join is sent out from this interface toward the RP. Notice that this is a (*,G) Join instead of an (S,G) Join because the last-hop router does not know the source. It needs to know only a) that the RP should know the source and b) how to get to the RP. By sending the (*,G) Join toward the RP, the last-hop router is telling the upstream router it is interested in receiving multicast packets for the group via the shared tree, or **rendezvous point tree (RPT)** as it called in PIM-SM.

The router receiving the (*,G) Join message adds the interface on which it was received to the OIL for the group and performs an RPF check on the RP. This upstream router sends a (*,G) Join message out its RPF interface for the RP, informing its upstream router that it wants to join the group. Each upstream router repeats this process of propagating Joins out of the RPF interface until this new branch of the tree either a) reaches the RP or b) reaches a router that already has multicast forwarding state for the group along the RPT. In this way, a new branch of the tree is created from receiver to RP.

To forward multicast packets down the RPT, the RP itself must be receiving the multicast packets. To receive this traffic, the RP can execute an RPF for the source and send an (S,G) Join toward the source. By joining the SPT, the RP is able to transmit packets down the RPT. Multicast packets now flow from the source to the RP via the SPT and then from RP to the receiver down the RPT.

Further details of SPT, RPT, and PIM-SM operation are examined in greater depth in subsequent chapters. For now, it is most important to understand the concept of reverse path forwarding.

1.3.2 POPULATING THE RPF TABLE

The routing table used for RPF checks can be the same routing table used to forward unicast data packets, or it can be a separate routing table dedicated to multicast RPF. In either case, this RPF table contains only unicast routes. It does not contain multicast group addresses because RPF checks are performed only on unicast addresses (either the source or the RP).

If the same routing table used to forward unicast data packets is used for RPF, it is populated by the traditional unicast routing protocols (RIP, OSPF, IS-IS, BGP, and so on). If a dedicated multicast RPF table is used, it must be populated by some other means. Some multicast routing protocols (for example, DVMRP) include mechanisms for populating a dedicated RPF table. Others (for example, PIM-SM and PIM-DM) rely on other protocols to set up this table.

Some traditional routing protocols (such as BGP and IS-IS) now have extensions that can be used to differentiate between different sets of routing information (for example, **Multiprotocol Extensions to Border Gateway Protocol—MBGP—**and **Multitopology Routing in IS-IS—M-ISIS**). Routes can be tagged as multicast RPF routes and thus distinguished from unicast routes. The advantage of having a dedicated RPF table is that a network administrator can set up separate paths and policies for unicast and multicast traffic. Chapters 7 and 8 examine in detail MBGP and M-ISIS, respectively.

1.4 INTERDOMAIN MULTICAST ROUTING

For years multicast has enjoyed niche success in many financial and enterprise networks. Financial institutions have applications, such as stock tickers, that require sharing the same data across the network. Using unicast for these applications is inefficient and not cost effective. Likewise, some enterprise networks serve companies with applications ideally suited to multicast delivery—for example, a central headquarters that must feed hundreds of branch sites with price lists and product information. Transferring these identical files to all sites individually with unicast simply is not efficient.

In the past, enterprise networks have frequently looked much different than the networks managed by **Internet service providers (ISPs)**. This difference existed because these networks had to meet a set of radically different requirements. Enterprise networks connect the offices of a single company, which often involves transporting primarily a single type of data (for example, file transfer). Transporting only a single type of data enables the network to be built in a way that optimizes delivery of that type of traffic. Also, few, if any, of the routers in an enterprise network connect to routers controlled by another entity.

ISP networks couldn't be more different. ISPs can have up to thousands of different customers, each a separate administrative entity. The data can include an unclassifiable mix of voice, video, e-mail, Web, and so on. Providing ubiquitous support for these various traffic types across the interdomain world of the Internet has always set ISPs apart from enterprises in the way they are designed and operated.

Unicast and multicast routing on enterprise and financial networks has often involved deploying protocols and architectures that best meet the needs of the companies they connect. These protocols and architectures often do not address the scalability and interdomain requirements of ISPs. However, recent trends have shown that the networking needs of enterprises have evolved to more closely resemble those of ISPs. Accordingly, many enterprise networks today are beginning to use the same principles and philosophies found in the engineering of ISPs' networks, albeit on a smaller scale.

The focus of this book is to describe the technologies and challenges faced by ISPs when deploying and operating multicast across the Internet. The first reason for this focus is neglect. Most networking books concentrate on enterprise networks rather than the unique demands of service provider networks. Second, ISP networks generally possess the superset of requirements that are found on other types of networks. For example, financial networks typically need to support many-to-many applications. Other enterprise networks may need to support only one-to-many applications. Because ISPs may be delivering service to both types of networks, they must be equipped to handle *both* types of applications. Additionally, ISP networks have scalability demands that are rarely found on any other types of networks.

While ISPs continue to have unique requirements for scalability and interdomain stability, most of the same multicast technologies found in ISP networks can be applied for use on other networks. By adopting these ISP philosophies, financial and enterprise networks are capable of ubiquitously supporting all types of multicast traffic. This flexibility enables a network to be prepared if traffic types change in the future.

The scope of this book is confined to the protocols and technologies currently used in the production networks of service providers. In order to provide a prag-

matic examination of the challenges faced by ISPs today, little to no mention is made of protocols that have not been implemented by routing vendors or deployed by service providers at the time of writing. Accordingly, IPv6 is outside the scope of this book.

1.5 WHERE IS MULTICAST?

The **Multicast Backbone**, or **MBone**, refers to the networks on the Internet that are enabled for multicast. The original MBone was built in the early 1990s as a network of multicast-enabled routers that were connected by **tunnels**. These routers were frequently UNIX servers running multicast routing software developed before router vendors had stable implementations of multicast software.

Tunnels allowed these early multicast-enabled "islands" to appear to be virtually connected to one another. Multicast packets were encapsulated within unicast packets and sent in the tunnel. Routers that were not multicast-enabled simply saw the unicast IP packet and routed it toward the tunnel destination. When the unicast packet reached the tunnel destination, the router decapsulated the unicast header to find the multicast packet within. If that packet had to be forwarded to another tunneled router, it was once again encapsulated and sent out another tunnel.

As router vendors implemented more stable multicast routing code, ISPs began to replace tunnels with *native* multicast routing in the late 1990s. Native multicast routing means routers forward raw multicast packets without encapsulating the multicast data within unicast packets. Most of the world's largest ISPs are multicast-enabled in at least some portion of their production networks today.

Multicast Internet Exchanges (**MIXs**) were built to connect multicast-enabled ISPs. MIXs are usually found in **network access points** (**NAPs**) where ISPs publicly peer with one another. A MIX enables ISPs to exchange multicast traffic on separate equipment from what is used for unicast peering. SprintNAP, in Pennsauken, New Jersey, and the NASA Ames Research Center **Federal Internet Exchange** (**FIX**-West), in Mountain View, California, contain two of the most popular MIXs used for public multicast peering.

Most people think of the old tunneled network of UNIX boxes when they hear the word "MBone," but it technically refers to any network that is multicast-enabled. Unanimous agreement has not been reached on a catchy word or phrase to colloquially refer to the native multicast-enabled portion of the Internet.

1.6 MULTICAST ON THE LAN

Throughout this book we focus primarily on the protocols that enable multicast packets to be forwarded within and between different domains. However, to provide a complete picture, we should examine what occurs on the link, or **local area network** (**LAN**), on which group members reside.

1.6.1 IGMP

When a host wants to become a multicast receiver, it must inform the routers on its LAN. The **Internet Group Management Protocol** (**IGMP**) is used to communicate group membership information between hosts and routers on a LAN.

To join a multicast group that is not already being forwarded on its LAN, a host sends an IGMP Report to a well-known multicast group. All IGMP-enabled routers on that LAN are listening to this group. Upon hearing a host's IGMP Report for a multicast group, G, one of the routers on the LAN uses a multicast routing protocol to join that group. In the case of PIM-SM, this router sends a (*,G) Join toward the RP for the specified group.

IGMP versions 1 and 2 allow a host to specify only the group address that it is interested in receiving. IGMP version 3 allows a host to express interest in only specified sources of a group, triggering an (S,G) Join by a PIM-SM router on the LAN. This is a key component of Source-Specific Multicast, which we examine in section 1.7.

A host must support IGMP in order to receive multicast packets. The version of IGMP supported is a function of the host's operating system. For example, unless otherwise modified, PCs running Windows 95 support IGMPv1. Likewise, PCs running Windows 98 or 2000 support IGMPv2, while IGMPv3 is available in Windows XP.

1.6.2 IGMP PROXYING

When a host reports interest in a multicast group from a source outside its LAN, it is the responsibility of a router on the LAN to join that group using a multicast routing protocol like PIM-SM. However, some routers do not support any multicast routing protocols. Low-end routers and legacy equipment such as dialup **remote access servers (RAS)** are examples of routing devices that sometimes do not support any multicast routing protocols.

Nearly all routing devices support IGMP. A common technique used in routers that do not support any multicast routing protocols is **IGMP proxying**. A router that hears an IGMP Report from a host simply relays that IGMP message to an upstream router that does support a multicast routing protocol. IGMP messages simply "hop over" a local router and reach a router that is capable of joining the group via a protocol like PIM-SM. IGMP proxying lowers the bar that low-end routing devices need to meet in order to deliver multicast.

1.6.3 LAYER 3 TO LAYER 2 MAPPING

The layers of the **OSI** reference model that we are most concerned with in this book are the data link, or layer 2, and the network, or layer 3. Here we focus on Ethernet, by far the most common layer 2 LAN technology. All layer 3 packets, in this case IP, are encapsulated with an Ethernet header and trailer and transmitted onto a LAN as an Ethernet frame.

All devices on the Ethernet have a unique 48-bit **Media Access Control (MAC) address**. To speak to one another, devices on the LAN keep a table that maps unicast IP addresses to MAC addresses. When packets are encapsulated in frames, the destination MAC address in the frame header is set to the MAC address corresponding to the IP address in the header of the IP packet.

IP multicast packets are destined to class D group addresses, which do not correspond with a single end host. Likewise, the MAC address used for multicast packets cannot be the address of a single station on the LAN. A special range of MAC addresses must be used for multicast.

The high order four bits of the first octets of class D addresses are always the same. Thus 28 bits may be varied in a multicast IP address. To provide a 1:1 mapping between MAC addresses and multicast IP addresses, the MAC address range must allow up to 28 bits to be varied.

The MAC address range that is assigned for multicast is only 24 bits long. One other bit in the address range is reserved, so that leaves only 23 bits of a MAC address to map to 28 bits of an IP address. As legend has it, Steve Deering, the father of multicast and a graduate student at the time, had only enough funding to purchase a 24-bit block of MAC addresses from IEEE. Because of this, every multicast MAC address corresponds to 32 different IP addresses.

As we see in Figure 1-7, the first 24 bits of a MAC address corresponding to an IPv4 multicast address is always 01-00-5E (in hexadecimal). The remaining 24 bits can vary from 00-00-00 to 7F-FF-FF (the first bit is always 0). The low-order 23 bits of the multicast IP address map to the MAC address. So an IP address of 224.1.1.1 maps to 01-00-5E-01-01-01. With 32:1 oversubscription, 224.129.1.1, 225.1.1.1, and 29 other IP addresses also correspond to this MAC address.

Collisions caused by oversubscription are handled by the IP stack of the receiving host. That is, a host interested in receiving 224.1.1.1 also receives Ethernet frames containing packets for 225.1.1.1 if they are on the LAN. After decapsulating the

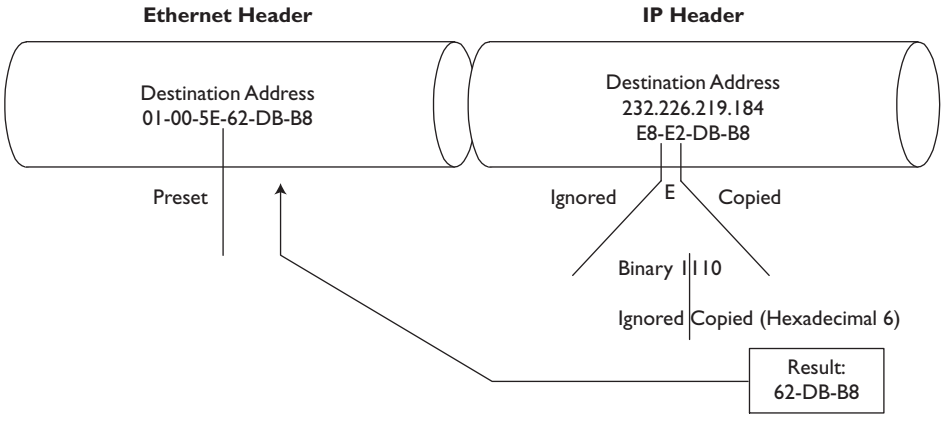


Figure 1-7 Layer 3 to layer 2 address mapping

Ethernet frame and discovering the IP address of the unwanted packet, the host discards the packet.

1.6.4 LAYER 2 SWITCHES

Ethernet switches have replaced hubs as the most popular way to connect hosts on a LAN because they inexpensively create a separate collision domain on each port. That is, while hubs transmit all traffic out all ports, switches transmit traffic only out a port that is destined for the station connected to that port. Frames destined for MAC addresses that a switch does not know the owner of are broadcast out all switch ports. Multicast packets fit in this category and, accordingly, are broadcast out all ports by a switch.

Flooding multicast packets out all switch ports wastes valuable network resources. Also, hosts that receive this unwanted traffic must use processing cycles to examine packets that they will eventually discard. **IGMP snooping** is one way to eliminate this inefficiency. Switches that support IGMP snooping can usurp responsibilities not usually associated with switches. An IGMP snooping switch looks at IGMP messages to determine which hosts are actually interested in receiving multicast traffic. Multicast packets are forwarded only out ports that connect to a host that is an interested listener of a specified group.

Cisco Group Management Protocol (CGMP) is a proprietary mechanism that provides the same functionality as IGMP snooping. CGMP enables Cisco System routers and switches to communicate with one another to determine which hosts are interested in each multicast group. CGMP works only in an environment of Cisco routers and Cisco switches. No other router or switch vendors are known to provide implementations of CGMP.

In switched environments that connect hosts to routers, IGMP snooping and CGMP generally provide a sufficient solution for eliminating broadcast traffic. However, not all switched environments involve IGMP. Switches are commonly used to connect routers together across multiaccess interfaces, forming transit LANs.

When only routers are connected together on a transit LAN, multicast routing protocols, not IGMP, are involved in controlling transit multicast traffic. The switch does not look at the multicast routing protocol packets, so there is no way

to see which port connects to a router that has joined a group. There has been some discussion of "PIM snooping," but it would be very difficult to implement because PIM is far more complex than IGMP.

The only way to prevent multicast traffic from being flooded out all switched ports in this environment is to change the logical topology with **virtual LANs (VLANs)**. VLANs can be used to create a point-to-point logical interface, or *sub-interface*, between every router connected to the switch. By changing the logical topology from broadcast to point-to-point, multicast traffic is sent only to routers interested in receiving it. However, using VLANs to create this kind of logical topology can force routers to perform more replication out these logical interfaces and consume more bandwidth over the physical interface than the logical interfaces use. Additionally, creating all of these VLANs can be a significant administrative and operational burden.

1.7 ASM VERSUS SSM

The original vision for multicast in RFC 1112 supported both one-to-many and many-to-many communication models and has come to be known as **Any-Source Multicast (ASM)**. Radio and television, as we have already discussed, are obvious examples of the one-to-many model. Applications such as online gaming and videoconferencing, in which some or all of the participants become sources, are examples of the many-to-many model. To support the many-to-many model, the network is responsible for source discovery. When a host expresses interest in a group, the network must determine all of the sources of that group and deliver them to the receiving host.

The mechanisms that provide this control plane of source discovery contribute the majority of the complexity surrounding interdomain multicast. However, applications that are believed to possess the greatest potential for commercial viability on the Internet use the one-to-many model. Since the bulk of the complexity is providing the least important functionality, the "ratio of annoyance" is disproportionately high in ASM.

It recently has been suggested that by abandoning the many-to-many model, multicast could deliver more "bang for the buck" on the Internet. By focusing on

the one-to-many model, the most appealing of multicast applications could be supported while vastly reducing the amount of complexity required. Source-Specific Multicast (SSM) is a **service model** that supports multicast delivery from only one specified source to its receivers.

By sacrificing functionality that many may consider less important on the Internet, the network no longer needs to provide the control plane for source discovery. This control plane is now the responsibility of receivers. Typically, the application layer (via a mouse click, for example) informs the receiver who the source is. When the receiver informs its directly connected router that it is interested in joining a group, it specifies the source as well as the group. This last-hop router is then able to join the SPT directly, instead of having to join the RPT.

SSM eliminates the need for RPTs, RPs, and **Multicast Source Discovery Protocol (MSDP)**, radically simplifying the mechanisms needed to deliver multicast. Best of all, this service model is realized through a subset of functionality already present in existing protocols. Very little needs to be added.

It is important to note that ASM and SSM are service models, not protocols. Different protocols are implemented and configured to deliver the service model. For example, SSM is a service model that is realized through a subset of functionality of PIM-SM and IGMPv3. The first five chapters of this book examine interdomain multicast generally from an ASM point of view because ASM is much more interesting from a protocol perspective. With a clear understanding of ASM, the operation and benefits of SSM become apparent. Chapter 6 describes SSM in detail.

1.8 ADDRESSING ISSUES

The addresses available for multicast usage range from 224.0.0.0 to 239.255.255.255. This plentiful, but finite, range is controlled by the **Internet Assigned Numbers Authority (IANA)**. Certain subranges within the class D range of addresses are reserved for specific uses:

- **224.0.0.0/24:** The link-local multicast range
- **224.2.0.0/16:** The **Session Announcement Protocol (SAP)/Session Description Protocol (SDP)** range

- **232.0.0.0/8:** The SSM range
- **233.0.0.0/8:** The AS-encoded, statically assigned **GLOP** range (RFC 3180)
- **239.0.0.0/8:** The administratively scoped multicast range (RFC 2365)

For a complete list of IANA assigned multicast addresses, refer to the *http://www.iana.org/assignments/multicast-addresses* Web site.

If class D addresses had been assigned in the same manner unicast addresses were allocated, this address space would have been exhausted long ago. In general, IANA allocates static multicast addresses only used for protocol control. Examples of this type of address include

- **224.0.0.1:** All systems on this subnet
- **224.0.0.2:** All routers on this subnet
- **224.0.0.5:** OSPF routers
- **224.0.0.6:** OSPF **designated routers** (**DRs**)
- **224.0.0.12:** DHCP server/relay agent

To protect against address exhaustion, a simple dynamic address allocation mechanism is used in the SAP/SDP block. Applications such as **Session Directory Tool** (**SDR**) that use this mechanism randomly select an unused address in this range. This dynamic allocation mechanism for global multicast addresses is somewhat analogous functionally to DHCP, which dynamically assigns unicast addresses on a LAN.

Unfortunately, some applications require the use of static multicast addresses. GLOP, described in RFC 3180, provides static multicast ranges for organizations that already have reserved an AS number. In GLOP, an AS number is used to derive a /24 block within the 233/8 range. The static multicast range is created in the following form:

```
233.[first byte of AS number].[second byte of AS number].0/24
```

For example, AS 12345 is automatically allocated 233.48.57.0/24. Here is an easy way to compute this:

1. Convert the AS number to hexadecimal: 12345 = 0x3039.
2. Convert the first byte back to decimal: 0x30 = 48.
3. Convert the second byte back to decimal: 0x39 = 57.

Thus any organization with an AS number is automatically assigned a /24 of multicast addresses. GLOP is not an acronym or abbreviation; for some odd reason it was selected as the name for this clever mechanism.

Addresses in the 239/8 range are defined as administratively scoped. Packets destined for these addresses should not be forwarded outside an administratively defined boundary (typically a domain border), which is somewhat analogous to unicast private address space, such as 10/8. Scoping is discussed in further detail in Chapter 4.

Addresses in the 232/8 range are reserved for SSM. A wonderful byproduct of SSM is that the group address no longer needs to be globally unique. The source-group tuple, or *channel,* provides all the required uniqueness because the receiver is specifying interest in only one source for the group.

Multicast addressing allocation had long been a headache for multicast engineers. The recent addition of of SSM finally provided the long-sought *coup de grace* in this struggle. It is now generally agreed that between SSM, GLOP, administrative scoping, and SAP/SDP, current multicast address allocation schemes are sufficient for the Internet until IPv6 becomes prevalent. In IPv6, the number of multicast and unicast addresses available is practically infinite.

1.9 APPLICATIONS

The most widely used application on the old MBone was SDR. By launching SDR, a host listens to the well-known SAP group, 224.2.127.254. Any source host that wants to advertise a session (usually audio and/or video) describes its

session in SDP messages. These SDP messages contain the address of the source, type of session, contact information for the source, and so on and are transmitted on the SAP multicast group.

Thus every host running SDR learns about every session on the Internet by receiving these SDP messages on the SAP group. By clicking one of the sessions listed in the SDR window, applications such as VIC (video conferencing tool) or VAT (visual audio tool) are launched to display the video or audio. An interesting feature of many of the applications launched by SDR is that by joining a session, you also become a source for that session. Because every receiver is also a source, each participant can see the others, which makes these applications ideal for collaboration and videoconferencing (and unscalable for sessions with lots of participants!).

Most agree that SDR is a "neat little toy" but not really a commercially viable application. Most SDR sessions are "cube-cams" or video camera shots of ISP parking lots. Because SDR acts as a global directory service for all multicast content on the Internet, it is not expected to scale to support large numbers of sessions.

Windows Media Player (WMP) is currently a popular application for accessing multicast audio and video content. WMP has excellent scaling potential for the Internet because, unlike many SDR-launched applications, receivers do not become sources to the group they join. Also, WMP has the capability to attempt to join a multicast session first, failing over to a unicast session if unsuccessful, which is ideal for content providers seeking the efficiency of multicast and the availability of unicast. Cisco System's **IP/TV** is another promising application for delivering multicast multimedia content. IP/TV supports multicast content only.

Juniper Networks and Cisco System routers can be configured to listen to the SAP group and keep a cache of SDR sessions. Joining the SAP group is useful in troubleshooting. It is a quick and easy way to determine whether the router has multicast connectivity with the rest of the Internet.

1.10 MULTICAST PERFORMANCE IN ROUTERS

When deploying multicast, it is important to consider whether the routers in a network are well suited to support multicast. Just as some cars provide speed at a

cost of safety, some routers provide unicast performance at a cost of multicast. As high-end routers are built to scale to terabits and beyond, router designers sometimes compromise multicast performance to optimize unicast forwarding. The two most important considerations when evaluating a router for multicast are state and forwarding performance.

A router must keep forwarding state for every multicast group that flows through it. Pragmatically, this means (S,G) and (*,G) state for PIM-SM. It is important to know how many state entries a router can support without running out of memory. MSDP-speaking routers typically keep a cache of Source-Active messages. Likewise, knowing the maximum number of Source-Active entries a router can hold in memory is crucial.

The obvious next question is "how many entries should a router support?" Like many questions in life, there is no good answer. Past traffic trends for multicast are not necessarily a reliable forecast for the future. Traffic trends for the Internet in general are rarely linear. Growth graphs of Internet traffic frequently resemble step functions, where stable, flat lines suddenly yield to drastic upward surges that level off and repeat the cycle.

The best policy is to select a router that can hold far more state than even the most optimistic projections require and monitor memory consumption. When state in a router begins to approach maximum supportable levels, take appropriate action (upgrade software or hardware, redesign, apply rate limits or filters, update your resume, and so on). With the exception of the Ramen **worm** attacks (see Chapter 5), state has not been much of a problem yet. Of course, as with mutual funds, past performance does not ensure future success.

Forwarding performance is characterized by **throughput** and **fanout**. Throughput describes the maximum amount of multicast traffic a router can forward (in packets per second or bits per second). Fanout describes the maximum number of outgoing interface for which a router can replicate traffic for a single group. As port densities in routers increase, maximum supported fanout becomes a critical factor. Also, it should be understood how increasing fanout levels affects throughput. As is the case with state, it is important to be aware of the performance limits, even if the exact amount of multicast traffic on the network is not known.

Forwarding performance is primarily a function of hardware. The switching architecture a router uses to forward packets is usually the most important factor in determining the forwarding performance of a hardware platform. *Shared memory* switching architectures typically provide the best forwarding performance for multicast. A shared memory router stores all packets in a single shared bank of memory.

Juniper Networks' M-series routers employ a shared memory architecture that is very efficient for multicast. In this implementation, multicast packets are written into memory once and read out of the same memory location for each outgoing interface. Because multicast packets are not written across multiple memory locations, high throughput levels can be realized regardless of fanout.

Some routers are based on a *crossbar* switching architecture. The "crossbar" is a grid connecting all ports on the router. Each port shows up on both the X and Y axes of the grid, where the X axis is the inbound port and the Y axis is the outbound port. With the crossbar architecture, packets wait at the inbound port until a clear path is on the crossbar grid to the outbound port. Inbound traffic that is destined for multiple egress ports must be replicated multiple times and placed in multiple memory locations. Because of this, routers with crossbar architectures usually exhibit multicast forwarding limitations.

Router designers sometimes work around this inherent challenge by creating a separate virtual output queue dedicated to multicast and giving the queue higher priority than the unicast queues. Unfortunately, this technique can cause multicast traffic to suffer head-of-line blocking, which occurs when packets at the head of the queue are unable to be serviced, preventing the rest of the packets in the queue from being serviced as well. Such a design assumes multicasts are a small percentage of total traffic because a router incorporating this design would be inefficient under a high multicast load.

1.10.1 RP LOAD

A cursory look at PIM-SM suggests that RPs should experience high load because they provide the root of all the shared trees in their domain. However, last-hop routers usually switch to the SPT immediately (SPT switchover is described in

Chapters 2 and 3), so the shared tree is typically short-lived. One mechanism that can cause RPs to experience high load, though, is the PIM-SM register process.

As we will see in forthcoming discussions of PIM-SM (see Chapter 4), routers that learn of a new source inform the RP in their domain by encapsulating the multicast packets into unicast packets and sending them to the RP. The RP must decapsulate and process these packets. If a router sends these encapsulated packets at a very high rate, the RP can be overrun while trying to process them. To prevent this from occurring, Juniper Networks routers configured as RPs require a special interface that is used to decapsulate these packets in hardware.

1.11 DISCLAIMERS AND FINE PRINT

Throughout this book, reference is made to **RFCs** (Request for Comments) and **Internet Engineering Task Force (IETF) Internet-Drafts**. Internet-Drafts are submitted to the IETF as working documents for its working groups. If a working group decides to advance an Internet-Draft for standardization, it is submitted to the Internet Engineering Steering Group (IESG) to become an RFC. RFCs are the closest things to the official laws of the Internet. For a good description of Internet-Drafts and the various types of RFCs, visit *http://www.ietf.org/ID.html*.

It is not uncommon for protocol-defining Internet-Drafts never to reach RFC status. Likewise, vendors do not always implement protocols exactly as they are defined in the specification. Internet-Drafts that are not modified after six months are considered expired and are deleted from the IETF Web site. All RFCs and current Internet-Drafts can be found at the IETF's Web site. A good way to find an expired Internet-Draft is by searching for it by name at *http://*www.google.com. A search there will usually find it on a Web site that mirrors the IETF Internet-Drafts directory without deleting old drafts. Unless otherwise stated, all Internet-Drafts and RFCs mentioned in this book are current at the time of writing. These documents are constantly revised and tend to become obsolete very quickly.

Similarly, the implementations of Juniper Networks and Cisco System routers, the routers most commonly found in ISP networks, are described throughout this book. The descriptions and configurations are meant to assist engineers in

understanding the predominant implementations found in production networks and provide a starting point for configuration. They are not the official recommendations of these vendors. It is also important to note that these vendors are constantly updating and supplementing their implementations. For officially supported configurations, it is best to contact these vendors directly.

1.12 WHY MULTICAST?

In less than a decade, the Internet has gone from a little known research tool to a dominant influence in the lives of people around the globe. It has created an age in which information can be disseminated freely and equally to everyone. The Internet has changed the way people communicate, interact, work, shop, and even think. It has forced us to reconsider many of our ideas and laws that had been taken for granted for decades.

Any person on earth with a thought to share can do so with a simple Web page, viewable to anyone with a connection onto the network. When considering the revolutionary impact their achievements have had on the way people interact, it is not ludicrous to mention names like Cerf, Berners-Lee, and Andreessen in the same breath as Gutenberg and Bell.

Nearly every aspect of communication in our lives is tied in one way or another to the Internet. Noticeably absent, however, in the amalgamation of content that is delivered prominently across the Internet is video. Video is an ideal fit for the Internet. While text and pictures do well to convey ideas, video provides the most natural, comfortable, and convenient method of human communication.

Even the least dynamic examples of video reveal infinitely more than the audio-only versions. For example, accounts of the 1960 Nixon-Kennedy debates varied widely between those who had watched on TV and those who had listened on the radio. So why then is video restricted primarily to the occasional brief clip accessible on the corner of a Web page and not a dominant provider of content for the Internet?

The answer is simple: The unicast delivery paradigm predominant in today's Internet does not scale to support the widespread use of video. Earlier attempts,

such as the webcasts of the Starr Hearings and the Victoria's Secret fashion show, have failed to demonstrate otherwise.

The easiest target for video's lack of pervasiveness on the Internet has always been the limited bandwidth of the "last mile." It has often been argued that potential viewers simply do not have pipes large enough to view the content. However, with the proliferation of technologies like digital subscriber line (DSL) and cable modems, widespread residential access to video of reasonably adequate quality exists. Furthermore, for years, the number of people employed in offices with broadband Internet connectivity has been substantial. Finally, with nearly every college dorm room in the United States (and increasingly throughout the world) equipped with an Ethernet connection, client-side capacity is quickly becoming a nonissue.

The server side, on the other hand, has principally relied on unicast to deliver this content. The cost required to build an infrastructure of servers and networks capable of reaching millions of viewers is simply too great, if even possible. Compare that to the cost of delivery with multicast, where a content provider with only a server powerful enough and bandwidth sufficient to support a single stream is potentially able to reach every single user on the Internet.

Interestingly, while it has always been viewed as a bandwidth saver, the previously mentioned efficiency underscores multicast's capability as a bandwidth multiplier. With a multicast-enabled Internet, every home can be its own radio or television station with the ability to reach an arbitrarily large audience. If Napster created interesting debates on copyright laws, imagine the day when everyone on earth will be able to watch a cable television channel multicast from your very own PC.

It is worth noting that multicast need not be used solely for video. Multicast provides efficient delivery for any content that uses one-to-many or many-to-many transmission. File transfer, network management, online gaming, and stock tickers are some examples of applications ideally suited to multicast. However, multimedia, and more specifically video, is widely agreed to be the most interesting and compelling application for this delivery mechanism.

The brief history of the Internet suggests the inevitability that it someday will be a prevalent vehicle for television and radio, as all data networks converge onto a single common IP infrastructure. Accepting this, multicast provides the only scalable way to realize this vision. With such great potential for providing new services, it is logical to wonder why multicast has not been deployed ubiquitously across the Internet. In fact, to this point, the deployment has actually been somewhat slow.

The current number of multicast-enabled Internet subnets is miniscule compared to the overall Net. There is no single, simple answer why this is the case. The reasons include a collection of realities, concerns, and myths. Any discussion of multicast's benefits should also address these issues. In most cases, recent developments have been made that allay these concerns.

1.12.1 MULTICAST LACKS THE "KILLER APP"

It took Mosaic, the first modern browser, to truly harness the power of the World Wide Web, resulting in unparalleled permeation. Many have argued that multicast needs the same "killer app" to fuel an explosion of growth. However, a closer look reveals that many of today's multicast applications are more than sufficient; they just happen to work without multicast.

A common technique used by some of the most popular multimedia applications is to attempt to access the content first via multicast, then failing over to unicast, if unsuccessful. To the end user, the result is the same. The selected show looks the same, and the favorite song sounds the same, whether delivered through unicast or multicast. The true difference exists in the amount of content available. Because of unicast's inability to scale, there are fewer shows to view and fewer songs to hear.

But the applications are plenty "killer."

1.12.2 THE CONTENT VERSUS AUDIENCE CHICKEN-AND-EGG SCENARIO

An intriguing phenomenon has emerged that has been a significant hindrance to deployment. Many multimedia content providers have been slow to provide

multicast content because of the limited number of capable viewers. Conversely, because of this limited amount of enticing content, there has been a perceived lack of demand from end users for multicast availability, thus resulting in a small audience.

This deadlock can be broken by multicast-enabled ISPs, partnering with content providers, to market this content to end users. This type of content provides a differentiator for these ISPs to attract more customers. To compete for these customers, more ISPs deploy multicast. Soon, multicast becomes a standard part of Internet service, expected by all end users. Eventually, ISPs that are not multicast-enabled are at a distinct, competitive disadvantage. In the meantime, content continues to increase, fueling the demand cycle.

Content providers can use the example of HDTV as inspiration. Soon after the introduction of HDTV, some TV stations began to broadcast their programming in the new format, even though very few people had the hardware that could take advantage of this technology. Despite having a miniscule audience to enjoy HDTV, these pioneering broadcasters made content available, which began to give consumers the incentive to purchase the new TV sets. Likewise, by providing an abundance of multicast content on the Internet, content providers give end users the incentive to demand access to this content from their ISPs.

1.12.3 THE "HOW DO WE CHARGE FOR IT?" SYNDROME

The first question most ISP product managers ask when considering deployment of multicast is nearly always, "How do we charge for it?" The question that should be asked, however, is "How do we make money from it?" For years ISPs have struggled with the business case for multicast. The early model was somehow to charge the users of the service. ISPs adopting this model have generally met disappointing results. While they may have found a market of enterprise and virtual private network (VPN) customers willing to pay for the service, Internet users found this model to be less than enticing.

This lack of success is predictable because it neglects to consider one of the paramount philosophies making the Internet so popular: Delivering a raw IP

connection to end users, through which many services can be derived, will be far more profitable than trying to charge users for each of the services they consume.

Imagine if, in the first few years after the Web was invented, ISPs had decided to charge their customers extra fees for the **HTTP** packets that traversed their connection. It might have changed the way people used the Web. Users may not have surfed so freely from site to site. Instead, ISPs quickly discovered that if they provided a simple connection, with no stifling rules or extra charges, people used the network more. In sacrificing revenue from "toll-taking," they enjoyed explosive growth as more customers used the network for more services. Unfortunately, many ISPs view multicast along this toll-taking model.

By deploying multicast, ISPs are enabling new services to be provided. It brings traffic onto the network that wasn't previously deliverable. ISPs that have provided multicast as a free part of their basic IP service have realized little revenue directly from multicast. *But they have gained customers they would not have otherwise attracted.* Moreover, providing multicast has lured the most valuable of customers—content providers. ISPs have long known that content begets customers. Internet users recognize the value and performance benefits of being able to access sites directly connected to their ISP's network.

ISPs that have offered multicast as just another basic, value-added service, like **DNS**, have been viewed by many as leaders, but that does not mean direct revenues from multicast cannot be realized. As in the case of unicast, the higher layers should provide advanced billable multicast services, while the network layer should be responsible for simply routing packets. Following the example of the Web, providers of higher-layer services, such as content hosting and **application service providers** (**ASPs**), will likely find a significant market for multicast content hosting.

1.12.4 MULTICAST PROTOCOLS ARE COMPLEX AND MAY BREAK THE UNICAST NETWORK

The protocols used to deploy multicast in a scalable way on the Internet today can certainly be considered nontrivial (enough to warrant the necessity for this

book!). RPF, a central concept in multicast, represents a significant change of paradigm from the traditional destination-based unicast routing.

Designers and operators of networks agree that a cost that cannot be ignored is included in deploying and maintaining multicast routing protocols, even if it involves no new hardware and simply "turning on" features already available in software. They also agree that the addition of any new protocol into a network offers the potential to introduce new bugs that can impact the stability of the network. This dilemma is faced when introducing any new technology into a network. Ultimately, the benefits provided by the new features must be weighed against the risk and cost of deployment.

Much of the complexity of multicast routing protocols has stemmed from the traditional view that multicast should provide many-to-many delivery in addition to one-to-many. To support this ASM model, the network must provide the control plane of source discovery. Recently, it has been widely agreed that the most "interesting" and commercially viable applications for multicast require only one-to-many delivery. By sacrificing functionality that may be considered somewhat less important on the Internet, much of the complexity of these protocols can be eliminated.

SSM is a service model that guarantees one-to-many delivery and can be realized with a subset of functionality from today's multicast protocols. By moving the control plane of source discovery to higher-layer protocols (like a click in a browser), the required multicast routing protocols become radically simpler. *This enables a reduction of operating and maintenance costs that cannot be overstated.*

1.12.5 CANNIBALIZATION OF UNICAST BANDWIDTH REVENUES

Throughout history, new technologies have evolved that have forced businesses to consider cannibalizing profitable incumbent technologies for new products. Generally, those who fail to embrace change get surpassed by those who do. When the automobile was first invented, imagine the dilemma faced by horse-drawn carriage makers as they pondered whether they should start building cars. Because multicast provides such efficient use of resources, some ISPs have been

concerned that they will lose revenue as their customers consume less bandwidth. This view is no less shortsighted than that held by our unwise carriage-building friends.

While multicast reduces the resources required for a single session of content, it brings new content on the network. It brings more customers who will eventually demand more bandwidth for higher-quality streams. And, as mentioned earlier, multicast can be used as a traffic multiplier, consuming more bandwidth through the network as more receivers join. The lessons learned on the Internet are no different than those of previous revolutionary technical breakthroughs. History does not look favorably upon the unwillingness to sacrifice limited short-term revenues in favor of products with limitless growth potential.

1.12.6 End-to-End Connectivity Required

For multicast to work properly, every layer 3 device on the path from source to receiver must be configured to be multicast-enabled. Pragmatically, this means every link on the Internet must be configured for PIM-SM, the de facto standard multicast routing protocol. If even one link in this path is not configured properly, multicast traffic cannot be received. This barrier can be a significant one as this path may transit many networks, each run by a different entity.

Because of this restriction, many consider multicast to be relegated to a hobbyist toy until the entire Internet is enabled. However, end-to-end multicast connectivity may not always be a requirement for applications to enjoy the benefits of multicast.

A hybrid unicast-multicast content delivery infrastructure can be built that provides the best of both worlds. A deployment of unicast-multicast "gateways" can be used to support the ubiquity of unicast with the scalability of multicast. Content can be multicast across an enabled core network to devices that can relay it to unicast-only hosts. This distributes the load that unicast must handle, relying on multicast to simply provide a back-end feeder network for the content gateways.

1.12.7 Lack of Successful Models

Some multicast critics have suggested that no profitable services have ever been based on multicast. This observation fails to notice two communications media

that have enjoyed commercial success for decades. Radio and broadcast television are based on a delivery mechanism that can be considered a special case multicast. Radio and television stations transmit data (their audio and/or video signal) across a one-hop, multiaccess network (the sky). Receivers join the group by tuning in their radio or TV to the group address (channel) of the station.

While radio and broadcast television do not use a packetized IP infrastructure (yet), the delivery mechanism used to provide content to receivers is decidedly multicast.

1.12.8 NOT READY FOR PRIME-TIME TELEVISION

After watching a 300Kbps Internet video stream on a 6-square inch section of a PC monitor, one's first inclination is definitely not to get rid of the family's 25-inch TV. While this can be considered reasonably good quality to expect on the Internet, it doesn't begin to compare to the quality and dependability that are expected from broadcast television. The bandwidth needed to approach this level of quality is orders of magnitude greater than that commonly found in most homes.

The quality and reliability of voice on the century-old **public switched telephone network (PSTN)** well exceeds that found in mobile phones. However, the functionality and limitless potential for features have enabled people to tolerate a lower voice quality in return for greater flexibility.

Likewise, the Internet has many inherent benefits that are difficult to match with broadcast communications. Despite having limited reach and no way to charge or exactly measure its audience, radio has been a viable business for the better part of a century. The Internet, with its bidirectional communication, provides the capability to log the exact behavior of every single viewer. After gazing upon an enticing advertisement, the viewer can instantly order the promoted product with the click of a mouse.

Additionally, the content that is available on television and radio is provided only by those with expensive studios and stations. On the Internet, anyone with a server and a connection can provide content accessible across the globe. Finally, multicast video on demand, generally believed to be impossible, is becoming a

reality thanks to clever techniques that are being pioneered by innovative content delivery companies.

Initially, it is likely multicast video will be primarily niche content not commonly found on television, such as foreign TV channels or high school sporting events. As new technologies evolve, such as set-top boxes and hand-held devices, and as bandwidth to the home increases, the Internet will become an extremely attractive vehicle for television and radio. *Multicast provides the scalability to make this a reality.*

1.12.9 Susceptibility to DoS

In the ASM service model, receivers join all sources of a group. While this functionality is ideal for applications such as online gaming, it leaves receivers open to **denial-of-service (DoS)** attacks. Any malicious user can send traffic to a multicast group, flooding all the receivers of that group, which greatly concerns content providers.

It is first worth noting that all IP traffic is susceptible to DoS, a reality in a network providing any-to-any connectivity. In fact, DoS is not even unique to the digital world. Throwing a brick through a storefront window, putting eggs in a mailbox, or parking a car in the middle of the street are only a few of an infinite number of analogs in the brick-and-mortar world. It just so happens that ASM DoS attacks are a bit easier to execute and have the potential to affect more users than their unicast counterparts. SSM, however, guarantees that the receivers will join only a single source. While DoS is not impossible with SSM, it is far more difficult to attack SSM receivers.

1.12.10 Unfriendly Last Mile Technologies, Less Friendly Firewalls

Multicast provides its benefits at the network layer. It is generally transparent to layer 2 technologies such as frame relay, ATM, and Ethernet, which means it is sometimes broadcast out all ports. Many of the high-speed last mile deployments of DSL and cable modems utilize primarily layer 2 infrastructures. Many of these architectures will be unable to realize the efficiencies supplied by multicast. For-

tunately, in the world of data communications, the only constant is change. Service providers realize they must be agile enough to modify their offerings when needed to contend in this fiercely competitive landscape. As multicast becomes a standard part of the Internet, these providers will be motivated to make the necessary software or hardware upgrades to support it.

Multicast is predominantly delivered via **User Datagram Protocol** (**UDP**). Those concerned with security find UDP traffic inherently scarier than its connection-oriented counterpart, **Transmission Control Protocol** (**TCP**). Many firewalls and other security devices do not even support multicast. Once again, as multicast becomes ubiquitous across the Internet, makers of these devices will add support for the services their customers demand. Similarly, common practices will be developed to allay the security vulnerabilities that exist today with multicast traffic.

1.12.11 THE NEED FOR MULTICAST

In global emergency situations, multicast can play a crucial role in delivering vital communication to millions of Internet users, providing extra communications capacity at a time when heretofore conventional methods are strained to the breaking point. Indeed, nowhere has this been more precisely demonstrated than in the tragic events of September 11, 2001. In the early hours following the terrorist attacks in New York and Washington, most news Web sites were inaccessible as extraordinarily large numbers of users attempted to simultaneously access these sites.

At Northwestern University, CNN was rebroadcast as a multicast feed on the Internet and quickly gathered an audience of over 2,000 viewers. At the time, this multicast audience was believed to be the largest for a single feed in history. However, the size of this audience was infinitesimal compared to the number of users that wanted desperately to view this coverage and learn what was happening. As millions tried in vain to view pictures, video, text, anything that could have described the horrific events unfolding that day, users on multicast-enabled networks were able to watch real-time video accounts throughout the entire day.

Users on networks not enabled for multicast were forced to scramble to find radios and televisions. On September 11, 2001, multicast enabled Internet users to

stay informed; in the future, multicast can be used to deliver critical information regarding public safety and security.

1.12.12 FINAL OUTLOOK

The free and open dissemination and collaboration of information provided by the Internet is among humankind's most powerful achievements. While the Internet has enjoyed unparalleled growth and has saturated nearly every element of our culture, it is poorly equipped to support multidestination traffic without multicast.

On enterprise and financial networks, multicast has enjoyed modest success for years; on the Internet, it has the capability to support content with the potential to be no less revolutionary than the World Wide Web. The reasons for its slow deployment across the Internet vary widely from validity to misunderstanding. In all cases, these obstacles are surmountable, especially given recent enhancements such as SSM. Finally, history has suggested the eventual convergence of all data networks onto a single IP infrastructure; multicast makes this forecast attainable.