PIM Protocol Extensions

Module 5

Agenda

- Source Specific Multicast
- Bidirectional (Bidir) PIM

Source-Specific Multicast

Barriers to Multicast Deployment

Cisco.com

- **Global Multicast Address Allocation**
  - **Dynamic Address Allocation**
    - **No adequate dynamic address allocation methods exist**
    - **SDR – Doesn’t scale**
    - **MASC – Long ways off!**
  - **Static Address Allocation (GLOP)**
    - **Based on AS number.**
    - **Insufficient address space for large Content Providers.**
- **Multicast Content “Jammers”**
  - **Undesirable sources on a multicast group.**
    - **“Capt. Midnight” sources bogus data/noise to group.**
    - **Can cause DoS attack by congesting low speed links.**

- **Barriers to Multicast Deployment**
  - Global Multicast Address Allocation

    This has been a barrier to deployment of multicast for some time. Content Providers need some way to allocate multicast address in order to multicast their content.

    - Historically, SDR has been the standard of dynamic multicast address allocation. Unfortunately, SDR makes use of a flat address space and uses periodic announcements of individually allocated addresses. As more addresses are allocated by individual servers, the period between these announcements drops. At some point, the period is too small to reliably keep other SDR hosts updated. This means that SDR doesn’t scale well.
    - MASC (Multicast Address Set-Claim) will scale to handle the entire Internet but the problem is that this mechanism is very complex to implement and has not seen sufficient deployment to be of general use.
  - Static Address Allocation
    - As a interim to MASC, GLOP Addressing (RFC 2770) describes a method of address allocation that statically assigns 256 global multicast addresses to each AS. Still, large content providers have already exhausted their range of addresses and need more.
  - Multicast Content “Jammers”
    - Since hosts join a group to receive content, any source can send to that group and the hosts will receive it. In some cases, this is undesirable such as when the CEO uses an IP/TV like application to broadcast a message to his employees. It is very easy for anyone to “jam” the broadcast by sending their own content on the same group. While some applications will ignore this “bogus” content, it can still result in a Denial-of-Service attack if the bogus traffic is sent at high rates that exceed link speeds.

Source Specific Multicast (SSM)

Cisco.com

- **Uses Source Trees only.**
- **Assumes One-to-Many model.**
  - **Most Internet multicast fits this model.**
  - **IP/TV also fits this model.**
- **Hosts responsible for source discovery.**
  - **Typically via some out-of-band mechanism.**
    - **Web page, Content Server, etc.**
  - **Eliminates need for RP and Shared Trees.**
  - **Eliminates need for MSDP.**

- **Source Specific Multicast**
  - Another variant of a PIM Sparse mode supports Source Specific Multicast (SSM) applications. The PIM SS (Source Specific) utilizes all the benefits of sparse mode protocols but eliminates shared trees at all and only builds source specific shortest path trees. These trees are built directly on receiving group membership reports that request a given source. The PIM SS is a draft proposal (draft-bhaskar-pim-ss-00.txt).
  - The SSM is suitable for well known sources within a domain or in another domain. The Multicast Source Discovery Protocol (MSDP) which is needed for interdomain multicast routing when regular PIM Sparse Mode is used within a domain is no longer needed for SSM.
  - A dedicated multicast group address range 232/8 is used exclusively for shortest-path trees for SSM. Routers are prevented to build a shared tree for any of the groups from this address range. The address range 232/8 is assigned for global well-known sources.
  - Source specific multicast (SSM) is a datagram delivery model that best supports one-to-many applications, also known as broadcast applications.

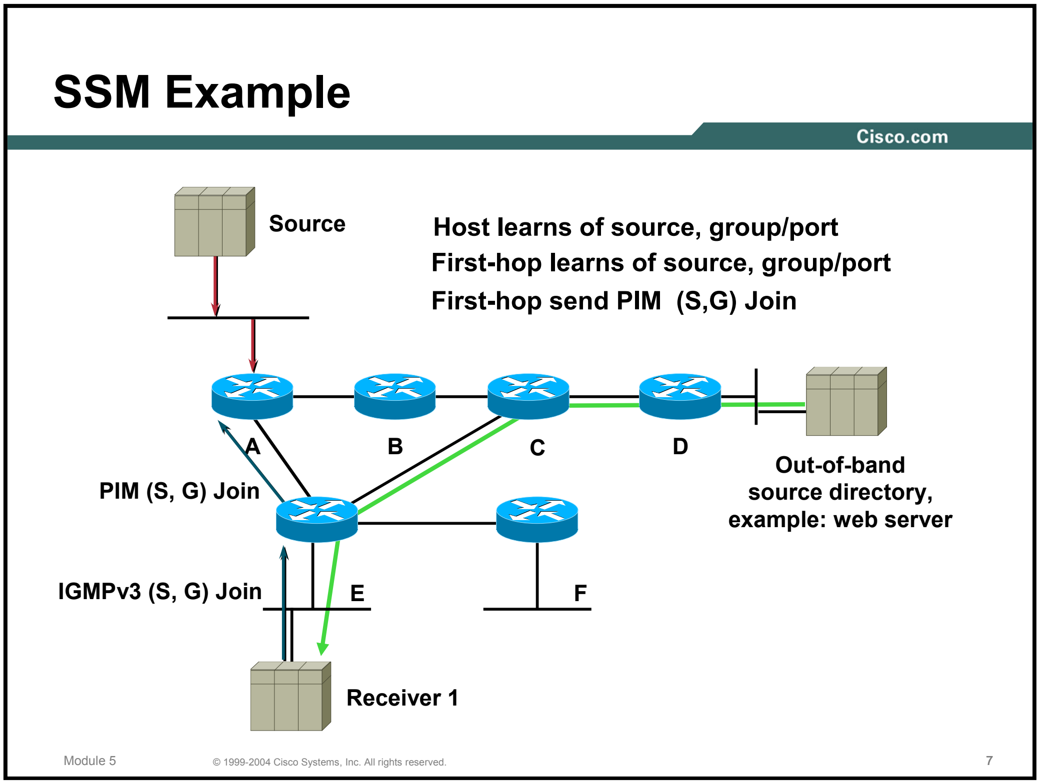
SSM Overview

Cisco.com

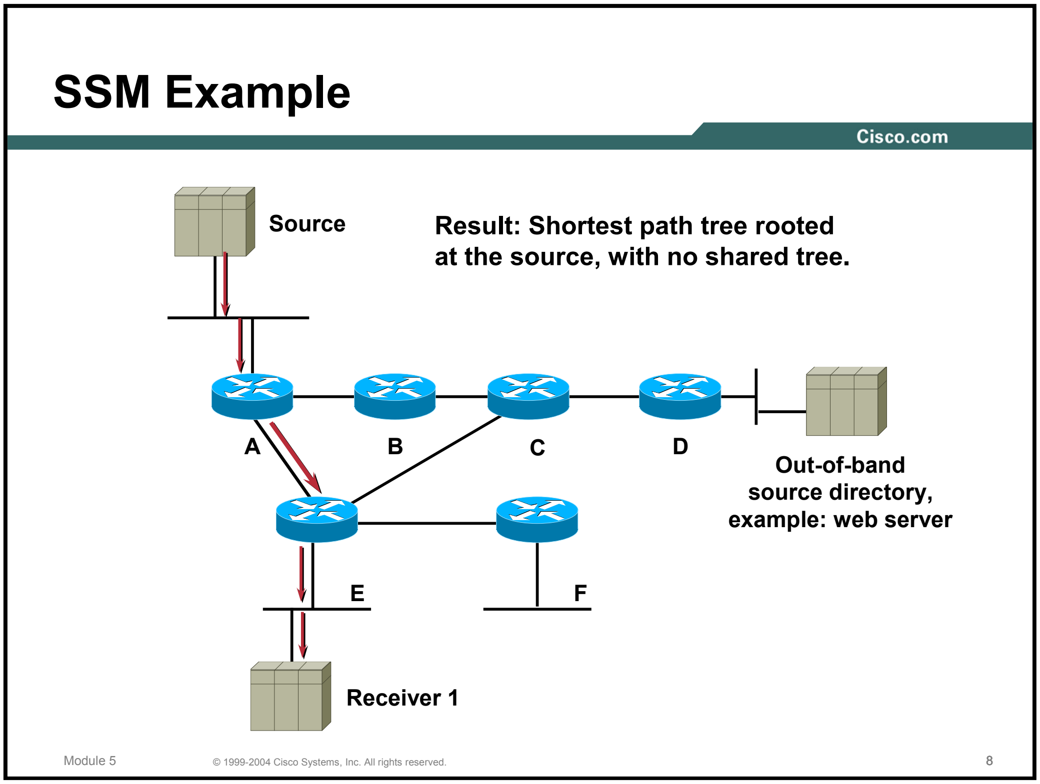
- **Hosts join a specific source within a group.**
  - **Content identified by specific (S,G) instead of (*,G).**
  - **Hosts responsible for learning (S,G) information.**
- **Last-hop router sends (S,G) join toward source**
  - **Shared Tree is never Joined or used.**
  - **Eliminates possibility of content Jammers.**
  - **Only specified (S,G) flow is delivered to host.**
- **Simplifies address allocation.**
  - **Dissimilar content sources can use same group without fear of interfering with each other.**

Module 5 6

- **SSM: For Well-know Sources**
  - The Source Specific Multicast allows last-hop router to immediately send (S,G) Join towards the source. Thus the PIM Sparse Mode (*,G) Join towards the RP is eliminated at all and first-hop routers start forwarding the multicast traffic donw the shortest-path tree (SPT) from the very beginning - as soon as the SPT is built by receiving first (S,G) Join.
  - The assigned address range 232/8 also simplifies the address allocation problems since the range is a global range for sorces that have to be well-known. Implementations in routers must not build any shared tree for those groups.
  - Source specific groups can coexist with other groups in PIM Sparse mode domains.



- **SSM – Example**
  - The prerequisite for SSM deployment is a mechanism that allows hosts not only to report the group they want to join but also the source for the group. This mechanism is built into emerging IGMP version3 standard. With IGMP v3 last-hop routers learn from the report for the multicast source and the group. It then simply creates (S,G) Join and forwards it directyl to the source.
  - The ways how hosts learn about existence of sources can be different – normally via some directory services (session announcements directly from sources or some out-of-band mechanisms, e.g. web pages). Most of those mechanisms distribute the information via multicast.



- **SSM – Example**
  - The result of building source-rooted tree (shortest-path tree) right from beginning is that RP mechanisms for source-specific groups are completely eliminated. The RPs for those groups are not needed any more and routers must not build shared trees for groups in the range 232/8.
  - The benefits of building shortest-path trees directly (and not via PIM Sparse mode switchover mechanism) are evident – the latency of multicast traffic is decreased and less multicast state is kept in multicast forwarding tables.
  - Another major benefit of SSM in in address management. Traditionally multicast applications had to acquire a unique IP multicast group address because traffic distribution was based only on the group address used. When two applications with different sources and receivers used the same IP multicast group address, the receivers received the traffic from both sources.
  - In SSM, traffic from each source is forwarded between routers in the network independent of traffic from other sources. Thus different sources can reuse multicast group addresses in the SSM range

SSM Configuration

Cisco.com

- **Global command**

```
ip pim ssm {default | <acl>}
```

  - **Defines SSM address range**
    - **Default range = 232.0.0.0/8**
    - **Use ACL for other ranges**
  - **Prevents Shared Tree Creation**
    - **(*, G) Joins never sent or processed**
    - **PIM Registers never sent or processed**
  - **Available in IOS versions**
    - **12.1(5)T, 12.2, 12.0(15)S, 12.1(8)E**

Module 5 9

- **SSM Global Configuration Command**

  ip pim ssm {default | <acl>}

  – Enables SSM operation on the router and defines the SSM address range. If the keyword **default** is used, the standard 232/8 address range is assigned as the SSM range. This may be overridden and another address range specified via the of the **<acl>** option.

  – When SSM is enabled for a range on a router, the router will not create Shared Trees for this range. This means that no (*,G) Joins or PIM Registers are sent or processed by the router.

SSM Configuration of Legacy Routers

Cisco.com

- Only Last-Hop routers ***must*** be upgraded.
  - Core may be upgraded later.
- Must insure no Shared Trees in SSM range.
  - Use '`ip pim accept-register`' at RP.
    - Prevents sources from registering in 232/8.
  - Use '`ip pim accept-rp`' on all routers.
    - Prevents (*,G) Joins from being processed for 232/8.
  - Use '`ip msdp sa-redistribute`' at RP.
    - Stops SA message origination in 232/8.
  - Use '`ip msdp sa-filter`' on MSDP peers.
    - Prevents forwarding of SA messages in 232/8.

Module 5 10

- **SSM Configuration with Legacy Routers.**

  It is not necessary to completely upgrade every router in the network before SSM is deployed. The last-hop routers are the only routers that *must* be upgraded to support SSM. However, care must be taken that the other routers in the core do not accidentally create Shared Trees for the SSM range. This can be accomplished using the following commands:

  ip pim accept-register

  - Configure this command with the appropriate ACL on a dummy RP to prevent inadvertent (*,G) Joins and Registers from creating a Shared Tree.

  ip pim accept-rp

  - Use this command on all routers to insure that the router will only use the dummy RP for the SSM group range. Care must be taken here as misconfiguration of this command can result in the router thinking there is ***no*** RP for this group range and will force the groups in this range into Dense Mode.

  ip msdp sa-redistribute

  - Use this command on all MSDP enabled RP's in the network to prevent accidental SA advertisement of SSM sources inside of the network to other MSDP routers.

  ip msdp sa-filter

  - Use this command on all MSDP enabled routers in the network to insure that no SA messages are sent or accepted for sources in the SSM range.

SSM – Summary

Cisco.com

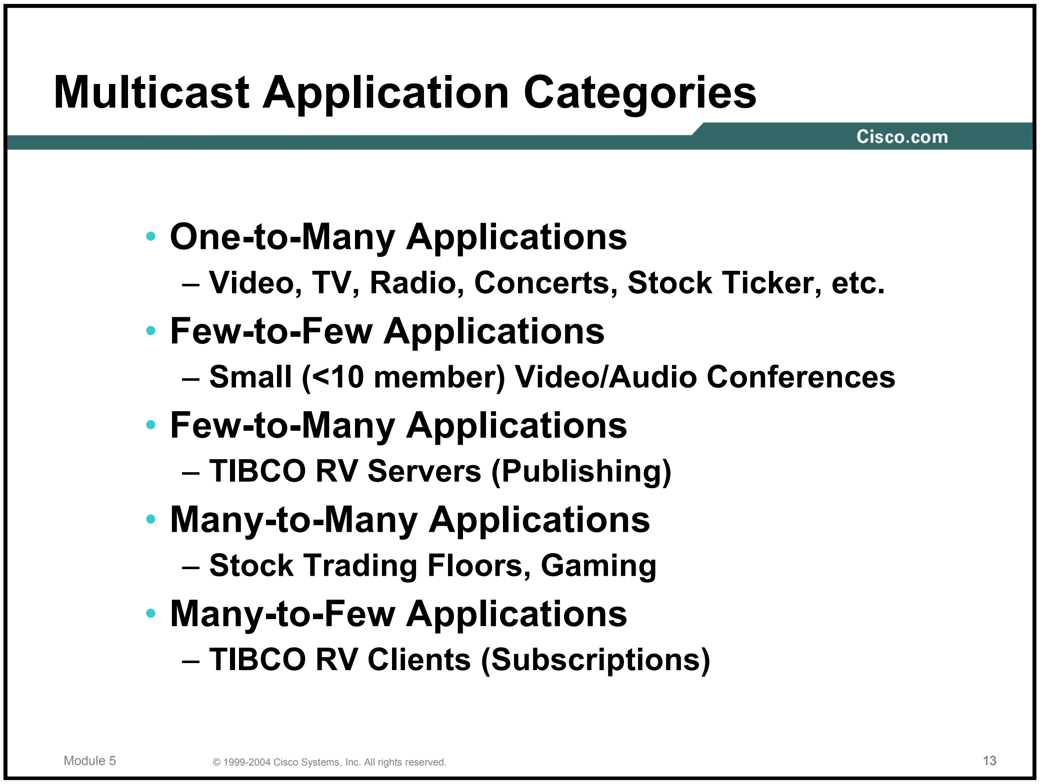
- **Uses Source Trees only.**
  - **Hosts are responsible for source & group discovery.**
  - **Hosts must signal router which (S,G) to join.**
- **Solves multicast address allocation problems.**
  - **Flows differentiated by *both* source and group.**
  - **Content providers can use same group ranges.**
    - **Since each (S,G) flow is unique.**
- **Helps prevent certain DoS attacks**
  - **"Bogus" source traffic:**
    - **Can't consume network bandwidth.**
    - **Not received by host application.**

Module 5 11

- **SSM Summary**
  - SSM assumes that the hosts are now responsible for determining the IP address of the active source(s) that they wish to receive and that they will signal the router using IGMPv3 (or other similar means) as to exactly which (S,G) flow that they wish to receive. This eliminates the need for the Shared Tree which is primarily used to make last-hop routers aware of active source for the group so that they can join the SPT and bypass the RP and the Shared Tree.
  - In SSM all flows are now uniquely defined by both "S" and "G" (as opposed to only "G"), and only Shortest Path Trees are used to deliver the specific (S,G) traffic to the end host. This means that sources can share the same group address without fear that their traffic will get merged on a Shared Tree somewhere in the network. This means content providers are free to use any and all addresses in the global 232/8 SSM address range and therefore offer a nearly unlimited number of address for them to source content to the Internet.
  - Since there is no Shared Tree in SSM, it is (theoretically) not possible for a host to receive traffic other than (S,G) flow that it requested.

    I say theoretically here as it is still possible – albeit unlikely – that traffic can get merged at the last hop subnet if the switches do not fully support IGMPv3. Consider the case where one host on the switch issues an Include(S1, G) IGMPv3 request while another issues an Include(S2, G) IGMP request. If the switch does not maintain state on an (S,G) basis at Layer 3, and instead only maintains "G" state, then it is possible for the hosts to receive both S1 and S2 flows.

Bidirectional (Bidir) PIM



- **Multicast Application Categories**
  - One-to-Many Applications
    - This is by far the most common form of IP multicast traffic and consists of Video, TV, Radio, Concerts, Stock Ticker, etc. sent from a single source to a set of receivers.
  - Few-to-Few Applications
    - The original Mbone Multi-media conference tools are a good example of this sort of application where a small number (<10 member) of hosts join the Video/Audio Conference.
  - Few-to-Many Applications
    - TIBCO RV Servers are a good example of this sort of application. In this case a small number of RV Servers “Publish” data to a large number of client workstations via a multicast group.
  - Many-to-Many Applications
    - Stock Trading Floors and multicast based Gaming are good examples of this sort of application. In this environment, every participant in the multicast group is a source of multicast traffic.
  - Many-to-Few Applications
    - TIBCO RV Clients are good examples of this sort of application. Here the RV Clients “Subscribe” to certain data categories by multicasting “subscriptions” to the RV Servers via a multicast group.

Multicast Application Categories

PIM-SM (S, G) State

Cisco.com

- **One-to-Many Applications**
  - **Single (S,G) entry**
- **Few-to-Few Applications**
  - **Few (<10 typical) (S,G) entries**
- **Few-to-Many Applications**
  - **Few (<10 typical) (S,G) entries**
- **Many-to-Many Applications**
  - ***Unlimited (S,G) entries***
- **Many-to-Few Applications**
  - ***Unlimited (S,G) entries***

Module 5 14

- **PIM-SM (S,G) State Impacts**
  - One-to-Many Applications
    - Requires a single (S,G) entry.
  - Few-to-Few Applications
    - Requires a few (<10 typical) (S,G) entries
  - Few-to-Many Applications
    - Requires a few (<10 typical) (S,G) entries
  - Many-to-Many Applications
    - *Requires an Unlimited amount of (S,G) entries*
  - Many-to-Few Applications
    - *Requires an Unlimited amount of (S,G) entries*

Many-to-Any State Problem

Cisco.com

- **Creates huge amounts of (S,G) state**
  - **State maintenance workloads skyrocket**
    - **High OIL fanouts make the problem worse**
  - **Router performance begins to suffer**
- **Using Shared-Trees only.**
  - **Provides some (S,G) state reduction**
    - **Results in (S,G) state only along SPT to RP**
    - **Frequently still too much (S,G) state**
    - **Need a solution that only uses (*,G) state**

Module 5 15

- **Many-to-any State Problem**
  - Because these types of applications result in an (S,G) entry for the many sources, the number of (S,G) entries in the mroute table in the routers can skyrocket. As the number of entries increase, the routers work harder and harder trying to maintain this state.
    - One of the the biggest impacts is the recalculation of RPF information for every entry in the mroute table once every 5 seconds.
  - Using Shared-Trees only.
    - While disabling the switch to the Shortest-Path Tree *can* result in some state reduction, it is generally insufficient to reduce the amount of overall state in the network to the desired levels. This is because in PIM-SM, (S,G) state is *always* created along the path from the source to the RP. What is needed is some solution that does not make use of (S,G) state to forward traffic to the RP.

Bidirectional (Bidir) PIM

Cisco.com

- **Idea:**
  - **Use the same tree for traffic from sources towards RP and from RP to receivers**
- **Benefits:**
  - **Less state in routers**
    - **Only (*, G) state is used**
    - **Source traffic follows the Shared Tree**
      - **Flows up the Shared Tree to reach the RP.**
      - **Flows down the Shared Tree to reach all other receivers.**

Module 5 16

- **Bidir PIM**
  - PIM Sparse Mode in its native form is unidirectional – the traffic from sources to the RP initially flows encapsulated in Register messages wich presents a significant burden due to encapsulation / decapsulation mechanisms. Additionally, shortest path tree is built between the RP and the source (initiated by the RP) which results in (*,G) and (S,G) entries at least on the way between the RP and the source.
  - Several multicast applications use many-to-many model where each participant is receiver and sender as well. In such an environment (*,G) and (S,G) entries appear everywhere along the path from participants and the associated RP in a PIM Sparse Mode domain resulting in increased memory and protocol overhead. It is also possible that the path from the source to the RP and the opposite path (from the RP to the source which is a receiver as well) are incongruent.
  - Bi-directional PIM dispenses with both encapsulation and source state by allowing packets to be natively forwarded from a source to the RP using shared tree state only. This ensures that only (*,G) entries will appear in multicast forwarding tables and that the path taken by packets flowing from the participant (source and/or receiver) to the RP and vice versa will be the same.

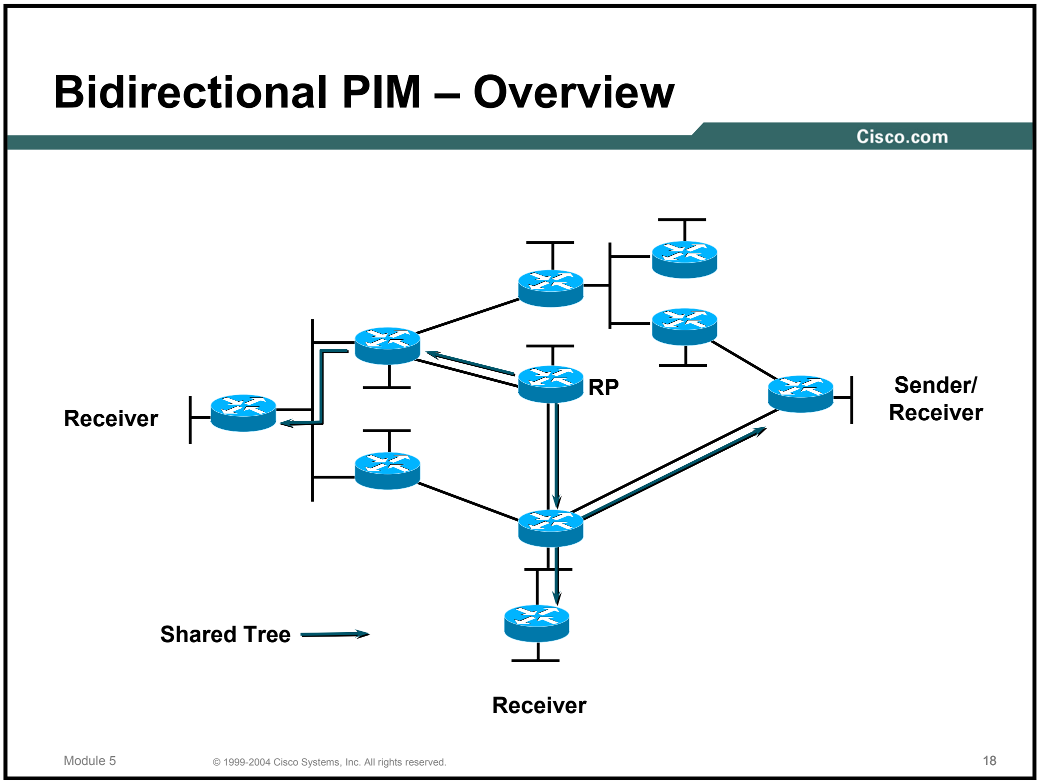
Bidirectional (Bidir) PIM

Cisco.com

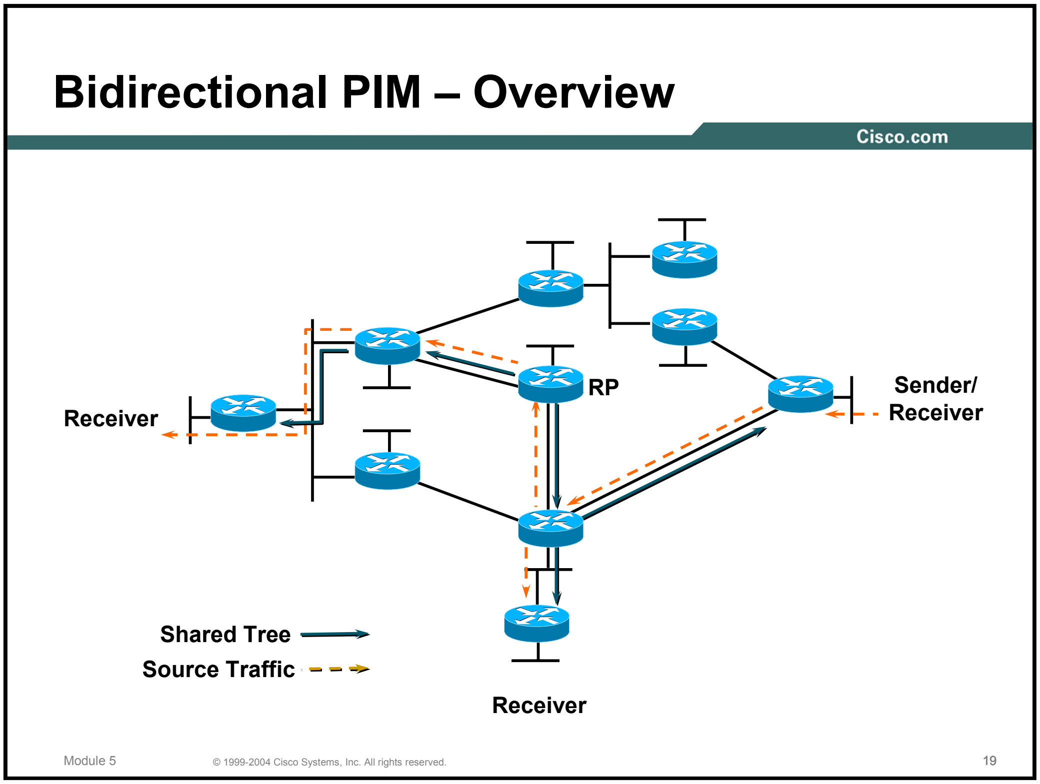
- **Bidirectional Shared-Trees**
  - **Violates current (*,G) RPF rules**
    - **Traffic often accepted on *outgoing* interfaces.**
    - **Care must be taken to avoid multicast loops**
  - **Requires a Designated Forwarder (DF)**
    - **Responsible for forwarding traffic up Shared Tree**
      - **DF's will accept data on the interfaces in their OIL.**
      - **Then send it out all other interfaces. (Including the IIF.)**

Module 5 17

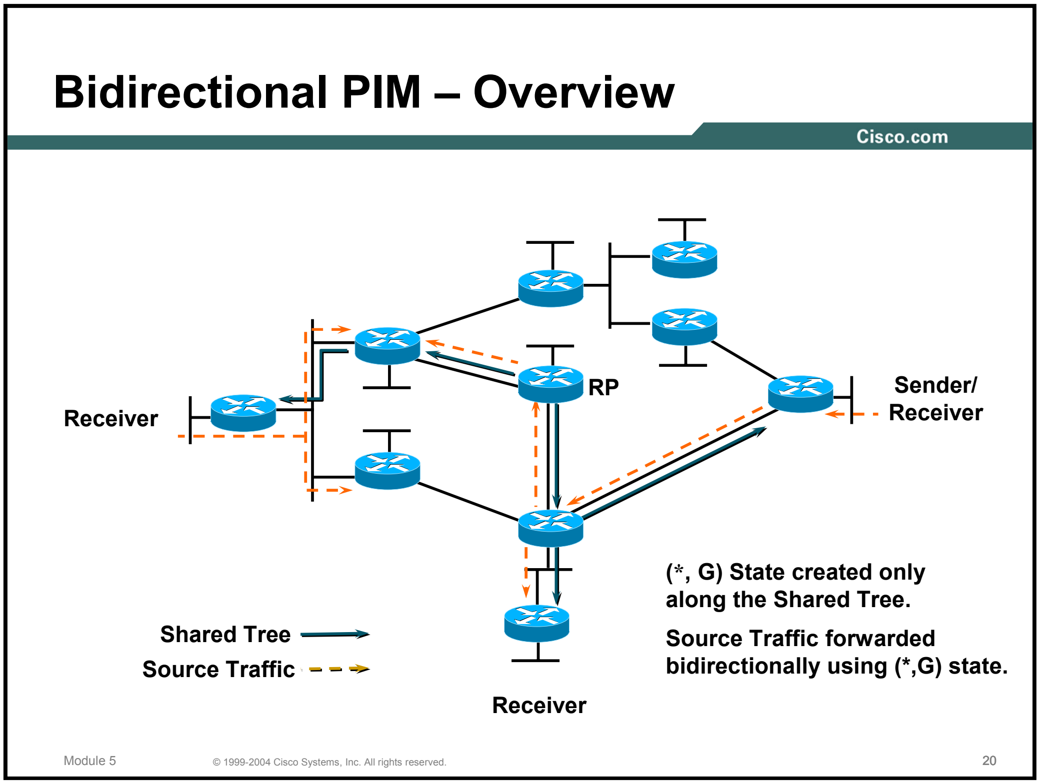
- **Bidirectional Shared-Trees**
  - Unfortunately, the use of bidirectional Shared Trees violates normal (*,G) RPF rules as traffic is allowed to flow *up* the Shared Tree. This means that in some cases, traffic will be accepted on an *outgoing* interface. This can lead to multicast route loops which can meltdown the network.
  - In order to solve this problem, Bidir PIM makes use of the Designated Forwarder (DF) concept. The DF is the router on each subnet that is responsible for forwarding traffic *up* the Shared Tree and hence are allowed to suspend the normal RPF rules and accept traffic on an outgoing interface.
  - When the DF receives traffic on an interface (be it incoming *or outgoing*), it forwards the traffic out all *other* interfaces (besides the one on which the traffic arrived) in the OIL as well as the Incoming interface.



- **Bidirectional PIM – Overview**
  - In this example, a Bidir Shared Tree has been built that connects each receiver that has joined the group.



- **Bidirectional PIM – Overview**
  - When traffic is sent by a source on the Bidir Shared Tree, it flows both up and down all branches of the Shared Tree in order to reach every receiver in the network.



- **Bidirectional PIM – Overview**
  - Bidir PIM Shared Trees only use (*,G) forwarding state to forward the traffic up and down the Shared Tree. This result in only a single mroute entry per group regardless of the number of senders in the network.

PIM Modifications for Bidir Operation

Cisco.com

- **Designated Forwarders (DF)**
  - **One DF per link.**
  - **Router with best path to the RP is elected DF.**
    - **Note: Designated Routers (DR) are not used for bidir groups.**
  - **In addition to normal (*,G) forwarding rules:**
    - **Accepts traffic on outgoing interfaces.**
    - **Forwards traffic out all other interfaces**

Module 5 21

- **PIM Modifications for Bidir Operation**
  - The major modification of PIM Sparse Mode to support bidirectional mode is an addition of a Designated Forwarder, which takes over the role of a Designated Router (DR) and has the following responsibilities:
    - It is the only router that forwards packets travelling downstream (towards receiver segments) onto the link
    - It is the only router that picks-up upstream traveling packets (away from the source) off the link and forwards them towards the RP
  - There is one DF per RP for bidirectional group(s) on each link. One and only one election is performed at RP discovery time. There is no constant control traffic and control messages appear only on changes. The election is robust and enforces consistent view on all routers on link. The router with the best unicast route to the RP is elected as a DF.
  - There is no effect of this election on local sources – their traffic reaches locally attached receivers directly and special treatment is no longer required when the sources are directly connected to a router. Data from those sources will automatically be picked up by the DF and forwarded towards the RP.

Designated Forwarder Election

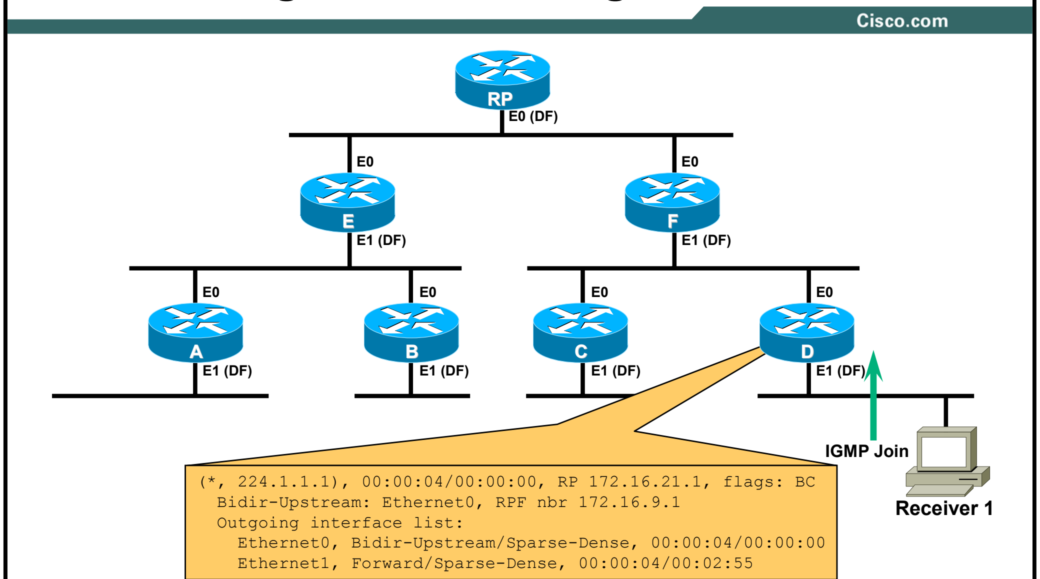
Cisco.com

- **Automatically performed on every link.**
  - **When Bidir Group-range/RP is learned or configured.**
  - **Router with the best path to the RP elected DF.**
  - **Uses assert-like metric comparison to pick best path.**
- **Purpose:**
  - **Ensures all routers on link agree on who is DF.**
    - **Prevents route loops from forming.**

Module 5 22

- **DF Election**
  - The election of a Designated Forwarder on each link follows similar principles known from the Assert process in PIM Dense Mode. The mechanism ensures that all the routers on the link have consistent view of the same RP. To perform the election of the DF for a particular RP, routers on a link need to exchange their unicast routing metric information for reaching the RP.
  - Note: The election of a DF is per RP and not per individual group.
  - The election process happens once only - when information on a new RP becomes available. There are however some conditions where an update to the election is needed:
    - A change in unicast metric to reach the RP for any of the routers on the link
    - The interface on which the RP is reachable changes to an interface for which the router was previously the DF
    - A new PIM neighbor on a link
    - The elected DF dies

Forwarding / Tree Building



Receiver 1 joins group causing router "D" to create (*, G) state.

Module 5 23

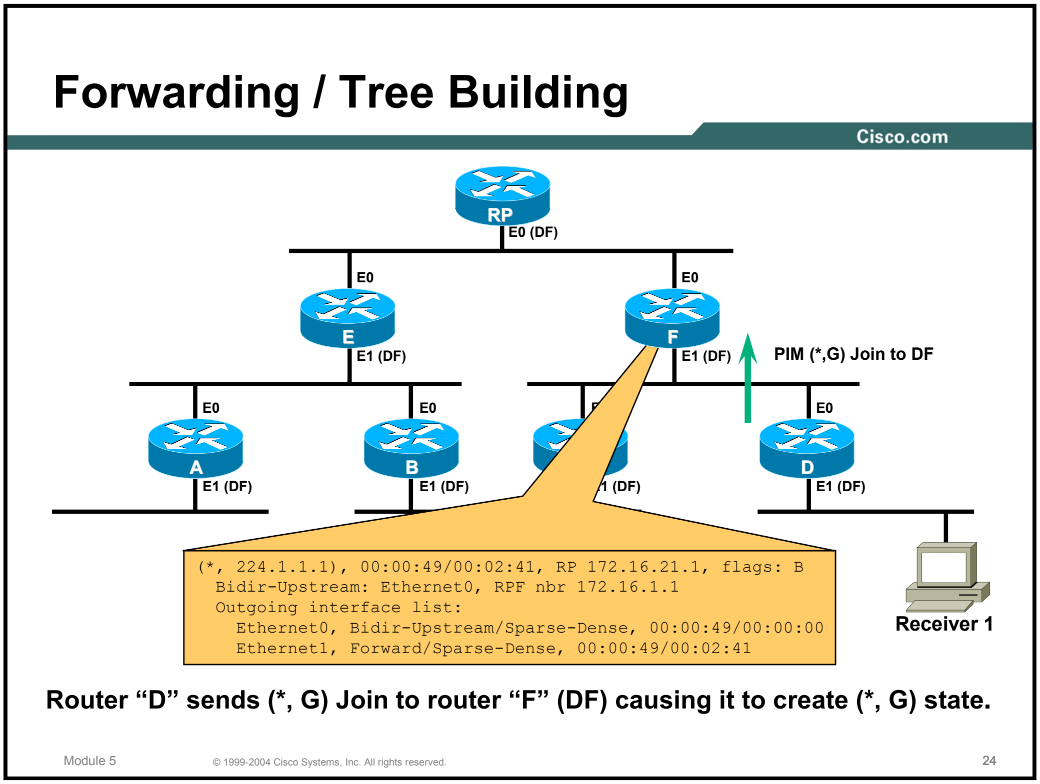
- **Forwarding / Tree Building**

  On this and the following pages we will observe an example of Bidir PIM in action. We will see how the Shared Tree is built and examine the associated Bidir (*,G) mroute state that is created as both receivers and sources become active in the network.

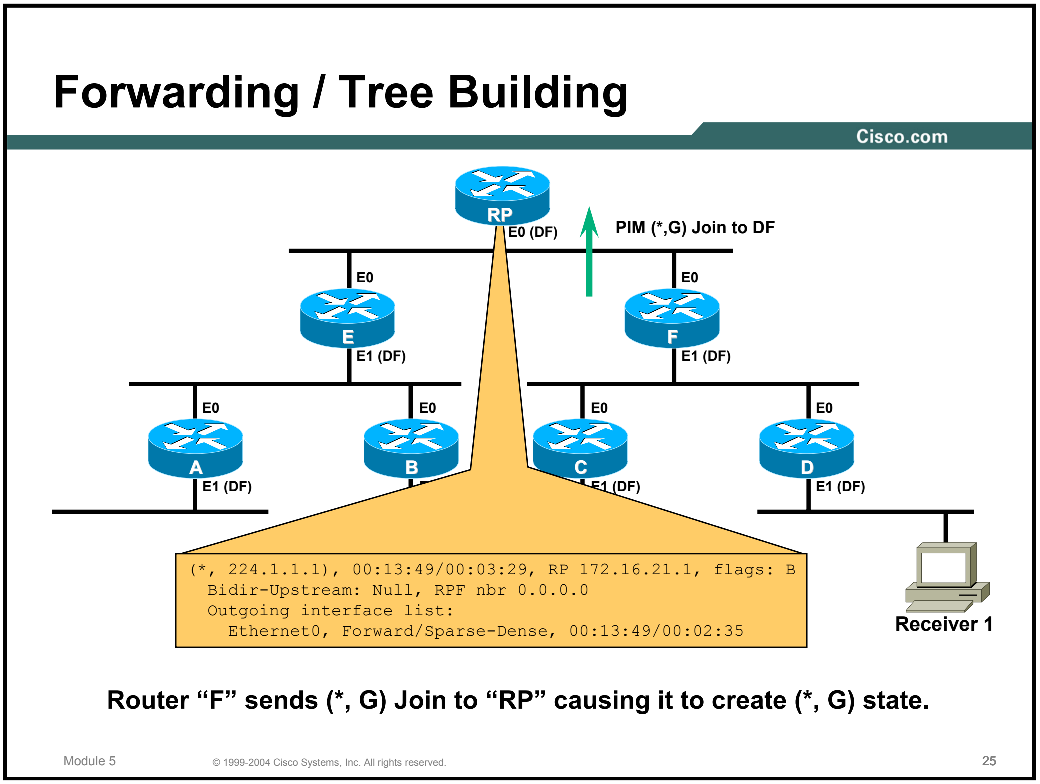
  – Initially, there is no state in the network when a receiver joins Bidir group 224.1.1.1.

  – The IGMP Membership Report (aka IGMP Join) is received by the local Designated Forwarder on the subnet (Router D) which in turn, creates a new (*,G) entry in the mroute table. Notice that the associated (*,G) entry is nearly identical to a standard PIM-SM (*,G) entry with the following exceptions:

    - The **B** flag is set on the (*,G) entry to denote that this is a Bidir group.
    - Instead of an "Incoming" Interface, Bidir PIM (*,G) entries have a "Bidir-Upstream" interface. This is the interface that is in the direction of the Bidir RP for this group. In this case, **Ethernet0** is the "Bidir-Upstream"
    - The "Bidir-Upstream" interface will always appear in the Outgoing Interface List (OIL) of a Bidir (*,G) entry. Notice that in this case this meant that not only was **Ethernet1** placed in the OIL (the interface where the IGMP Join was received), but the "Bidir-Upstream" interface, **Ethernet0**, was placed in the OIL as well.



- **Forwarding / Tree Building**
  - The most of the standard PIM-SM (*,G) state maintenance mechanisms apply to Bidir PIM. Such is the case for when the OIL of a (*,G) mroute entry goes from Null (empty), to non-Null. (This effectively occurred when the (*,G) entry was created.) This causes Router D to trigger an immediate (*,G) Join toward the RP. This (*,G) Join is sent with the “Neighbor Address” of the PIM Join/Prune packet set to the address of the DF on this upstream subnet. In this case, it is the address of **Ethernet1** of Router F.
  - When Router F receives this PIM (*,G) Join, it too creates a Bidir (*,G) mroute entry. Once again, because this is a Bidir PIM (*,G) entry, the “Bidir-Upstream” interface, **Ethernet0**, is included in the OIL along with **Ethernet1**, the interface where the (*,G) Join was received.

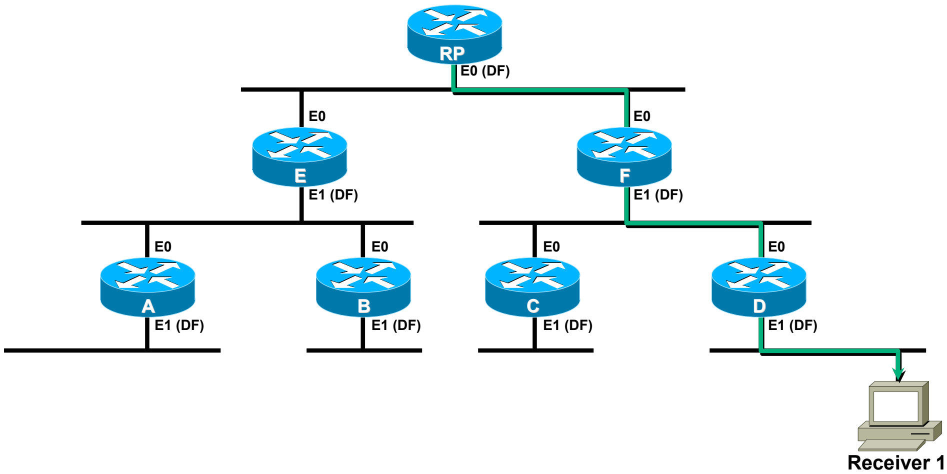


- **Forwarding / Tree Building**
  - Since the OIL of this (*,G) entry went from Null to Non-Null, Router F triggers a (*,G) Join in the direction of the RP.
  - Since the OIL of the (*,G) mroute entry on Router F went from Null to non-Null (this effectively occurred when the (*,G) entry was created), Router F triggers an immediate (*,G) Join toward the DF on this subnet. (This (*,G) Join is sent with the "Neighbor Address" of the PIM Join/Prune packet set to the address of the DF on this upstream subnet. In this case, the DF is the RP itself and the "Neighbor Address" field is set to the address of **Ethernet1** of the RP.

  - When the RP receives this PIM (*,G) Join, it too creates a Bidir (*,G) mroute entry. Because the RP is the root of the Bidir Shared Tree, the "Bidir Upstream" interface is **Null**. This results in only **Ethernet1**, the interface where the (*,G) Join was received, appearing in the OIL of the (*,G) mroute entry on the RP.

Forwarding / Tree Building

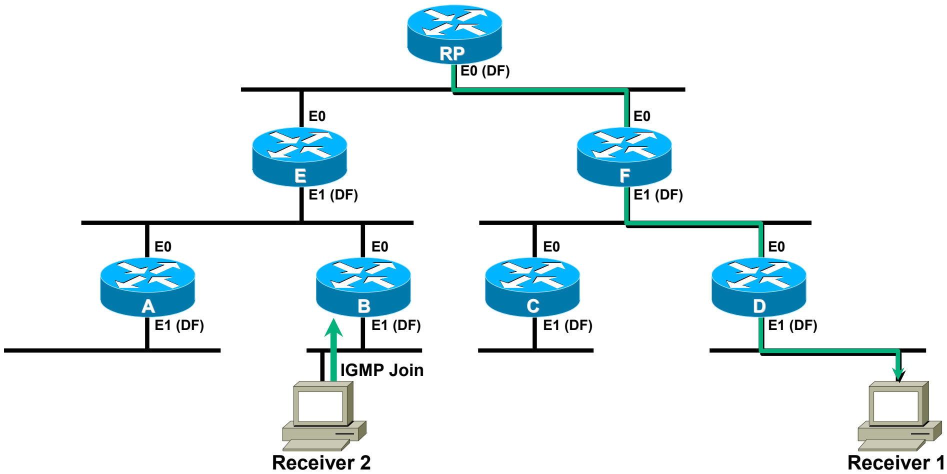
Cisco.com



Branch of Shared Tree is now built down to Receiver 1.

- **Forwarding / Tree Building**
  - At this point, a single branch of the Bidir Shared Tree for group 224.1.1.1 has been built from the RP down to Receiver 1.

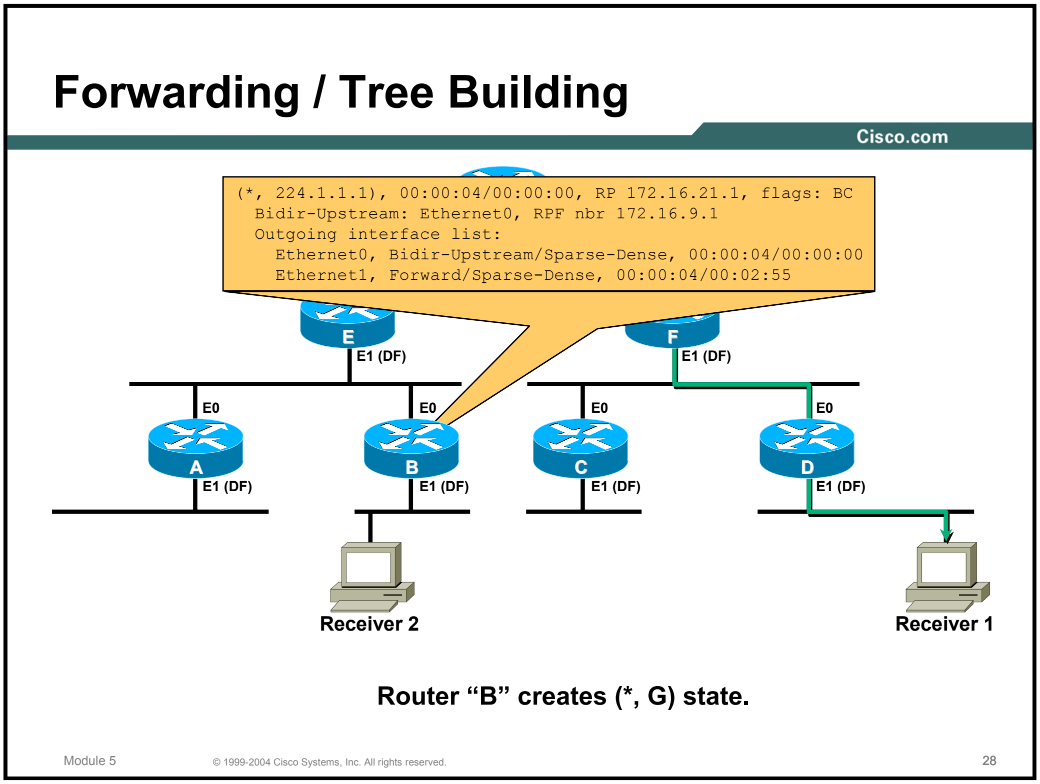
Forwarding / Tree Building



Receiver 2 also joins group.

- **Forwarding / Tree Building**
  - At this point in our example, Receiver 2 goes active and joins group 224.1.1.1 by sending an IGMP Membership Report (aka IGMP Join).

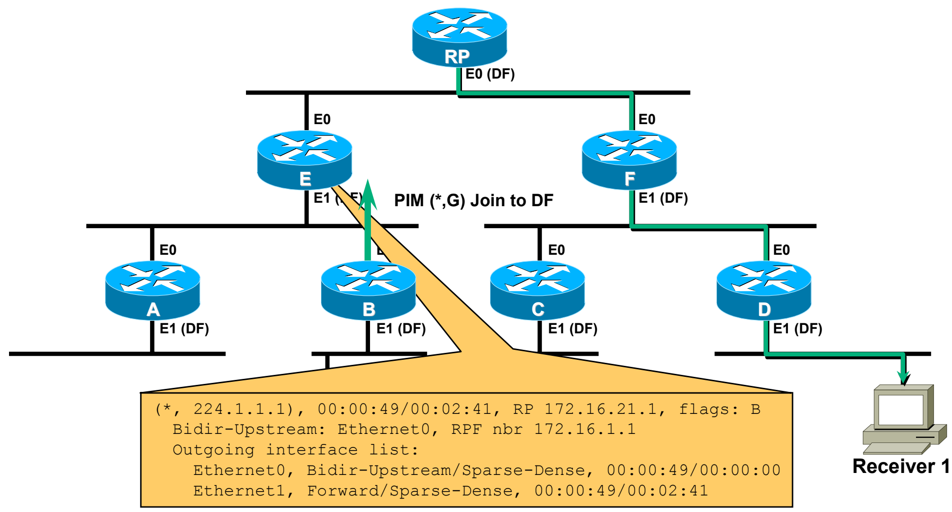
Forwarding / Tree Building



- **Forwarding / Tree Building**
  - The IGMP Membership Report (aka IGMP Join) is received by the local Designated Forwarder on the subnet (Router B) which in turn, creates a new (*,G) entry in the mroute table in the same fashion as was done earlier for Receiver 1.

Forwarding / Tree Building

Cisco.com



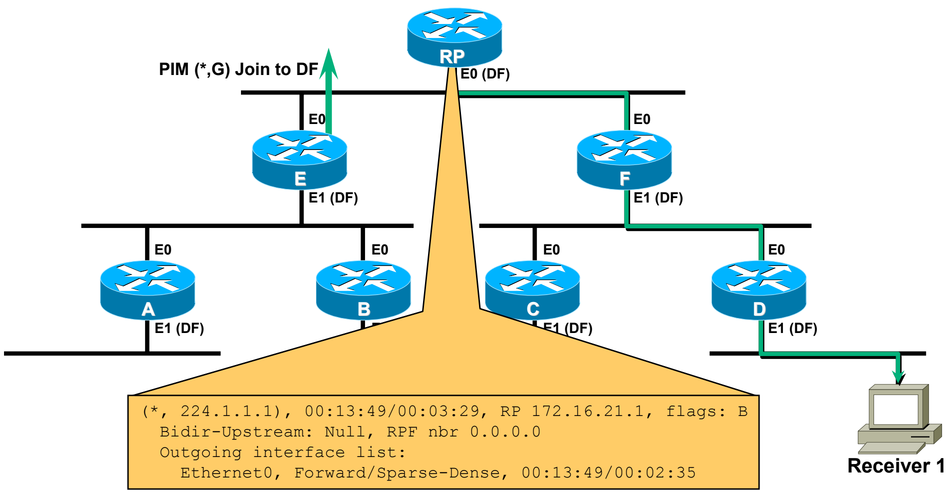
Router "B" sends (*, G) Join to "E" (DF) causing it to create (*, G) state.

Module 5 29

- **Forwarding / Tree Building**
  - Since the OIL of the (*,G) mroute entry went from Null to non-Null (this effectively occurred when the (*,G) entry was created), Router B triggers an immediate (*,G) Join toward the RP. This (*,G) Join is sent with the "Neighbor Address" of the PIM Join/Prune packet set to the address of the DF on this upstream subnet. In this case, it is the address of **Ethernet1** of Router E.
  - When Router E receives this PIM (*,G) Join, it too creates a Bidir (*,G) mroute entry. Once again, because this is a Bidir PIM (*,G) entry, the "Bidir-Upstream" interface, **Ethernet0**, is included in the OIL along with **Ethernet1**, the interface where the (*,G) Join was received.

Forwarding / Tree Building

Cisco.com

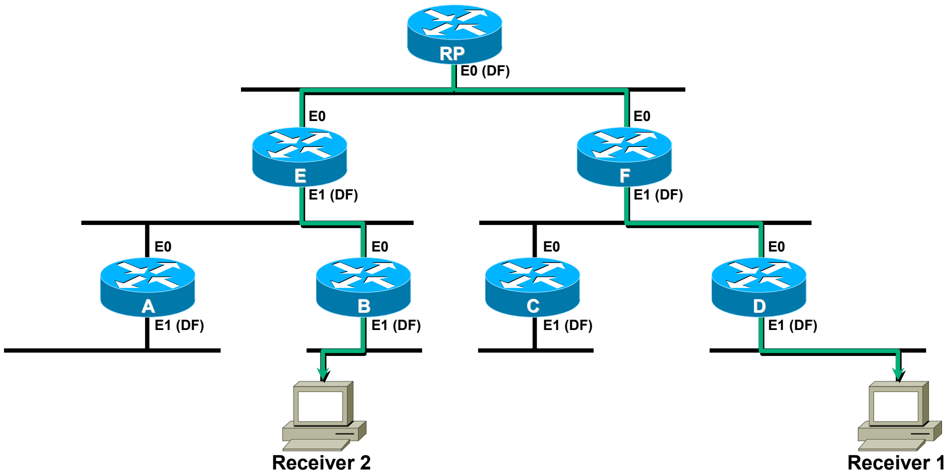


Router "E" sends (*, G) Join to "RP". (State on RP remains unchanged.)

- **Forwarding / Tree Building**
  - Since the OIL of the (*,G) mroute entry on Router E went from Null to non-Null (this effectively occurred when the (*,G) entry was created), Router E triggers an immediate (*,G) Join toward the RP. (This (*,G) Join is sent with the "Neighbor Address" of the PIM Join/Prune packet set to the address of the DF on this upstream subnet. In this case, the DF is the RP itself and the "Neighbor Address" field is set to the address of **Ethernet1** of the RP.
  - When this (*,G) Join is received by the RP, the state remains basically unchanged since **Ethernet0** was already in the OIL of the (*,G). (Actually, the only thing that happens is that the interface timer on **Ethernet0** is reset.)

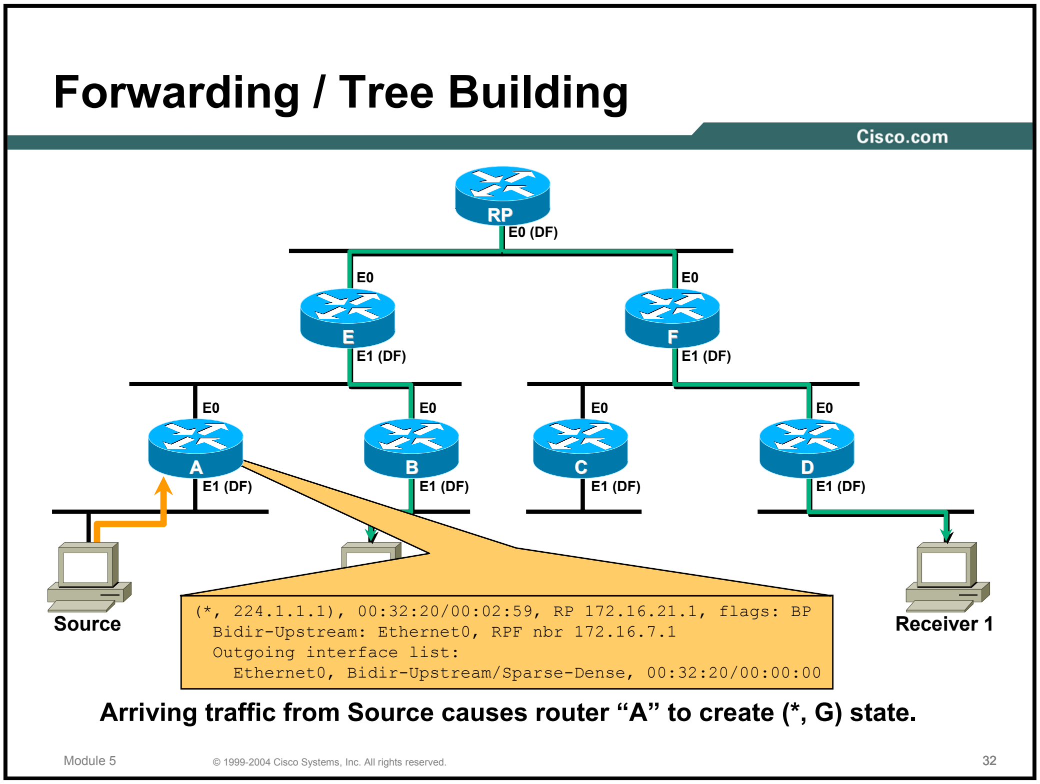
Forwarding / Tree Building

Cisco.com

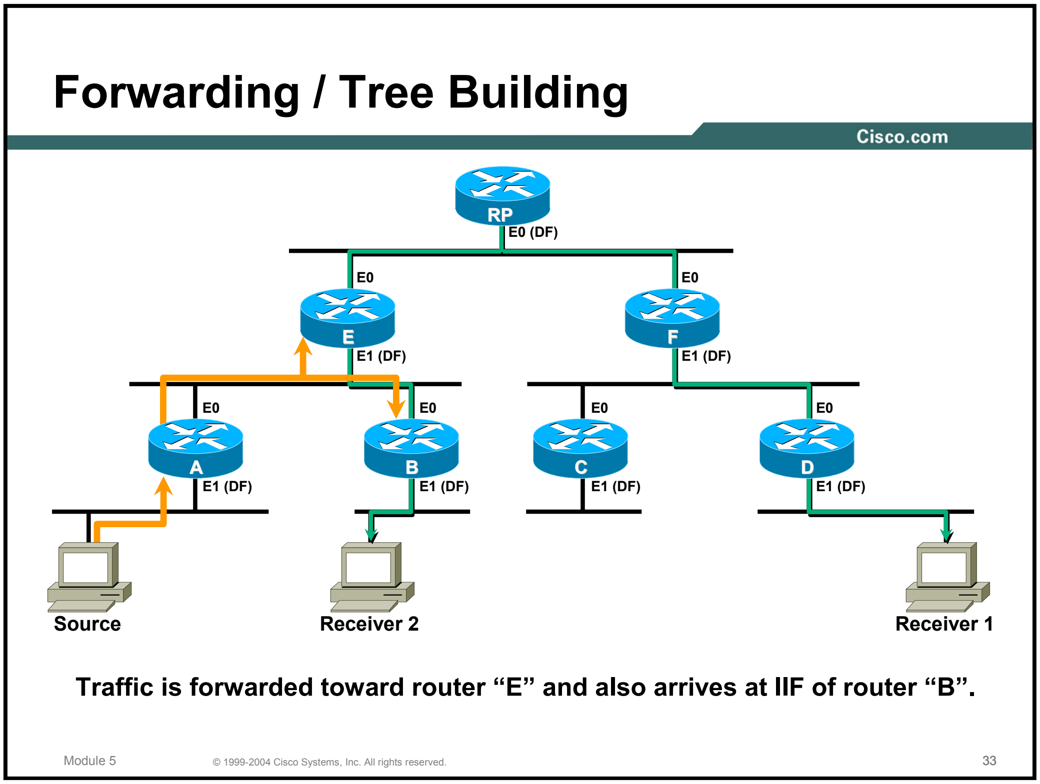


New branch of Shared Tree is built to Receiver 2.

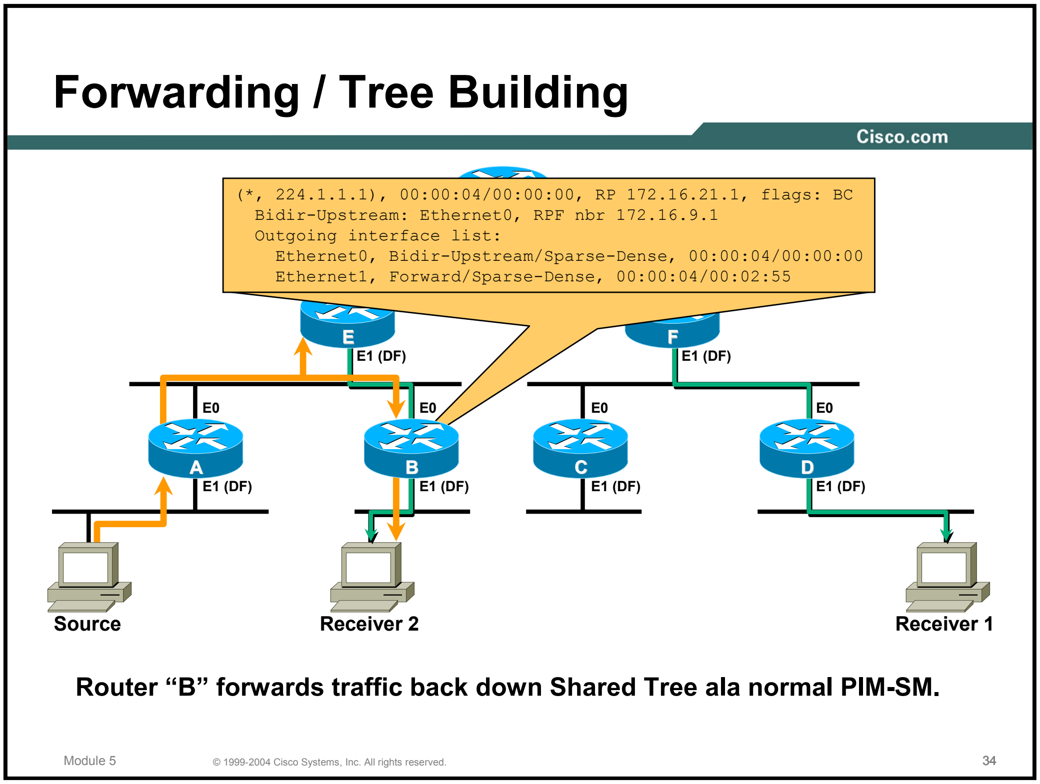
- **Forwarding / Tree Building**
  - At this point, a second branch of the Bidir Shared Tree has been built down to Receiver 2.



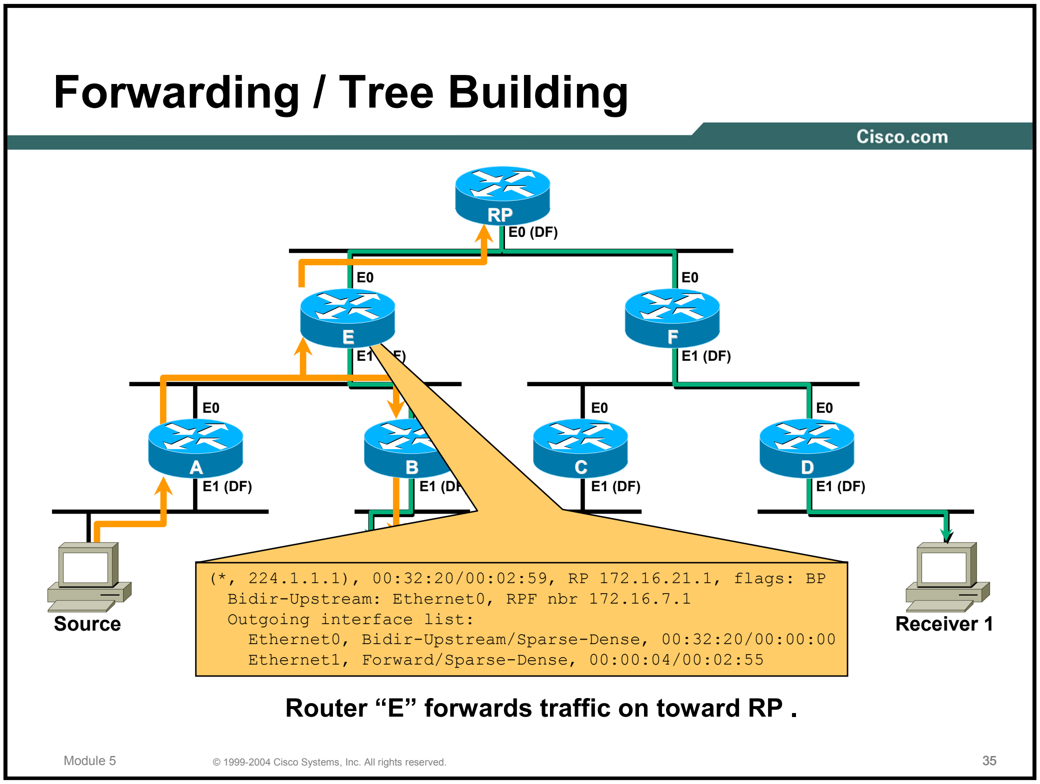
- **Forwarding / Tree Building**
  - When our source goes active, the arriving multicast traffic causes Router A to create Bidir (*,G) state ***because it is the elected Designated Forwarder on this subnet.***
  - Notice that since the source has not joined Bidir group 224.1.1.1 (this is a send-only host), only the "Bidir-Upstream" interface appears in the OIL of the (*,G) entry.



- **Forwarding / Tree Building**
  - The forwarding rules for Bidir PIM is to forward the multicast traffic out all interfaces that appear in the (*,G) OIL *except* the interface on which the traffic was received. This results in the source traffic being forwarded on the “Bidir-Upstream” interface, **Ethernet0**, in the direction of the RP. (See the previous page for the details on the state in Router A.)
  - Not only does the traffic arrive at Router E (the upstream DF for this subnet) but it also arrives at Router B.

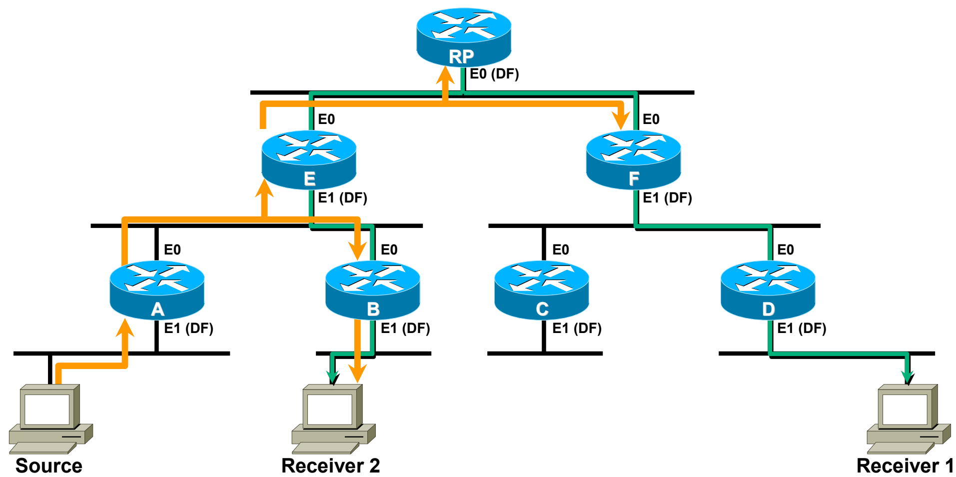


- **Forwarding / Tree Building**
  - Although Router B is not the DF for this subnet, it does have a matching (*,G) entry and the multicast traffic ***is*** arriving on the “Bidir-Upstream” interface. As a result, Router B accepts the traffic and forwards it out all interfaces in the OIL *except the interface it arrived on,* which is the “Bidir-Upstream” interface. Hence, Router B forwards the traffic down the Shared Tree towards Receiver 2.



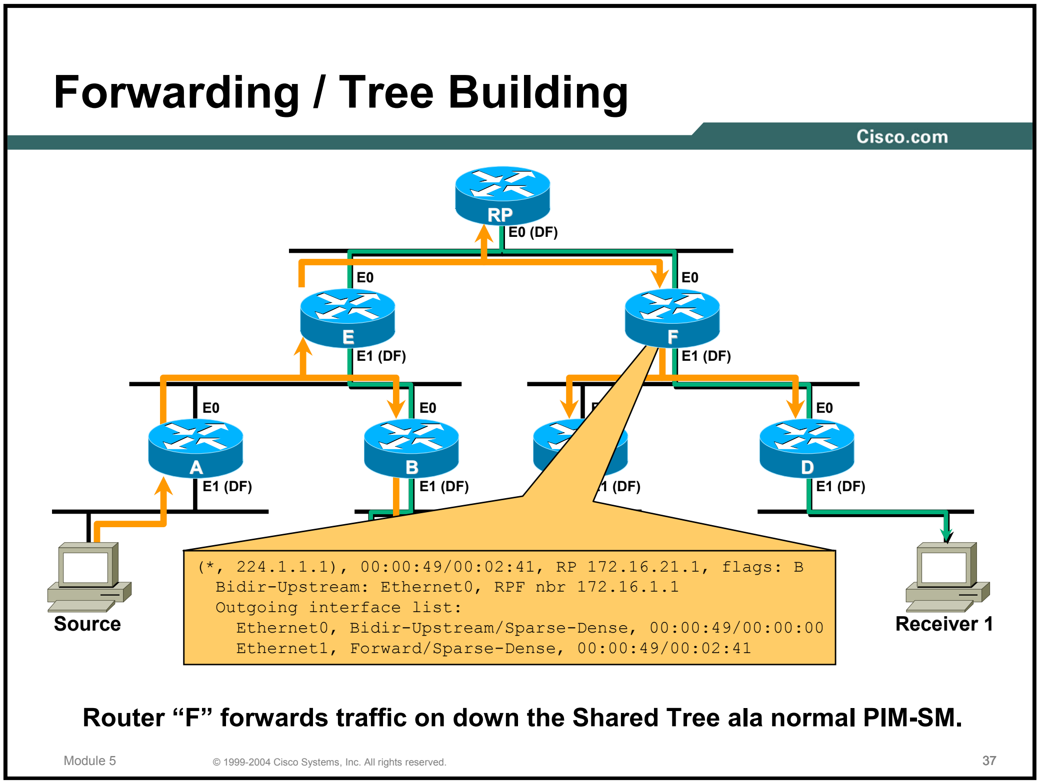
- **Forwarding / Tree Building**
  - The arrival of multicast traffic at Router E causes it to create Bidir (*,G) state because it is the elected Designated Forwarder on this subnet.
  - Again, the forwarding rules for Bidir PIM is to forward the multicast traffic out all interfaces that appear in the (*,G) OIL *except* the interface on which the traffic was received. This results in the source traffic being forwarded on the "Bidir-Upstream" interface, **Ethernet0**, in the direction of the RP. (See the previous page for the details on the state in Router E.)

Forwarding / Tree Building

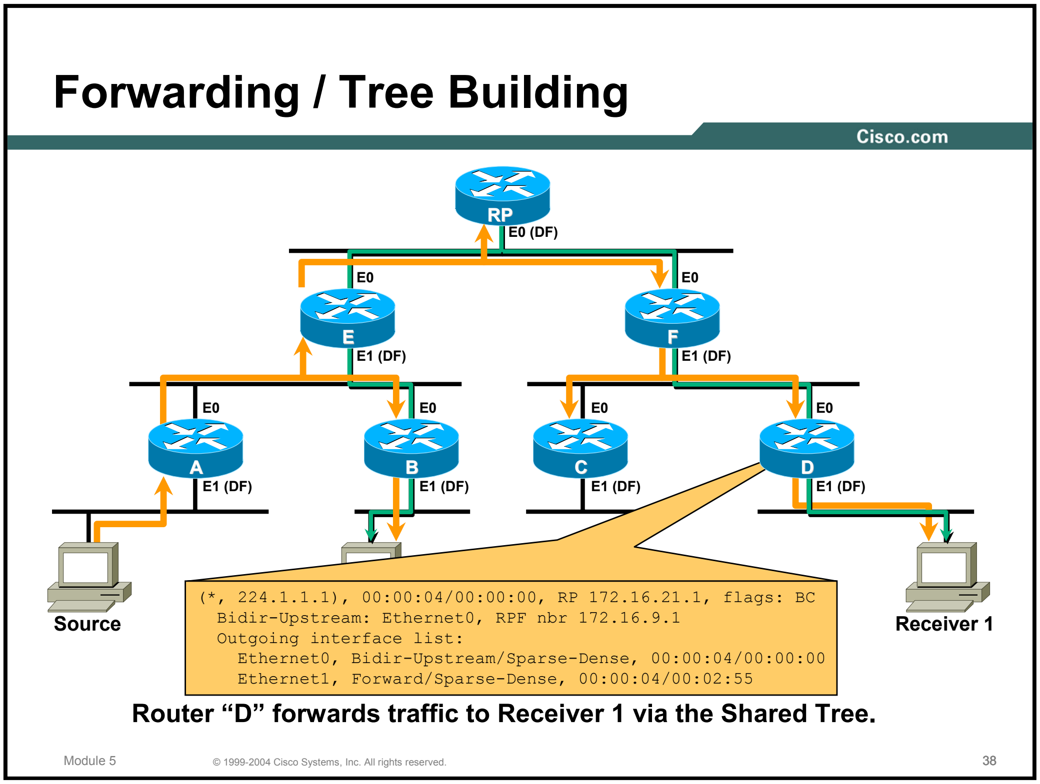


Traffic forwarded toward RP also arrives at the IIF of router "F".

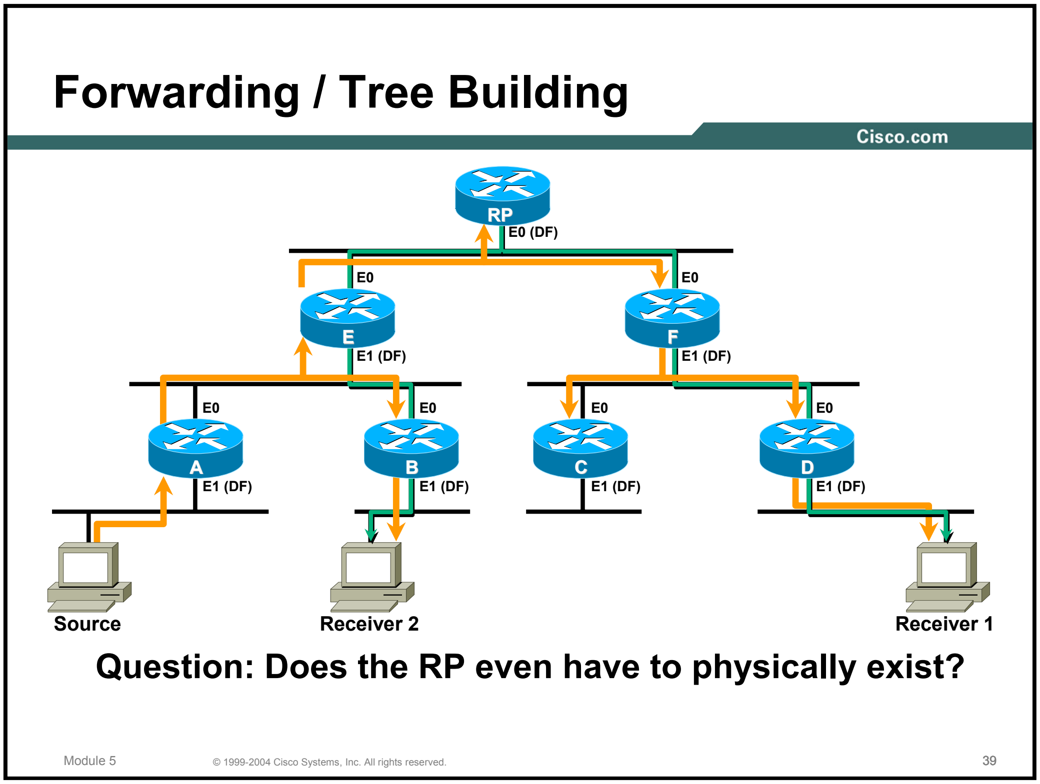
- **Forwarding / Tree Building**
  – The multicast traffic is now flowing to the RP. In addition, the traffic is arriving at Router F via the common LAN segment.



- **Forwarding / Tree Building**
  - The traffic arriving at Router F is arriving on the "Bidir-Upstream" interface towards the RP. As a result, Router F accepts the traffic and forwards it out all interfaces in the OIL *except the interface it arrived on,* the "Bidir-Upstream" interface. This results in Router F forwarding the traffic out **Ethernet1** and down the Shared Tree towards Router D.



- **Forwarding / Tree Building**
  - The traffic arriving at Router D is also arriving on its "Bidir-Upstream" interface towards the RP. As a result, Router D accepts the traffic and forwards it out all interfaces in the OIL *except the interface it arrived on,* the "Bidir-Upstream" interface. This results in Router F forwarding the traffic out **Ethernet1** down the Shared Tree to Receiver 1.



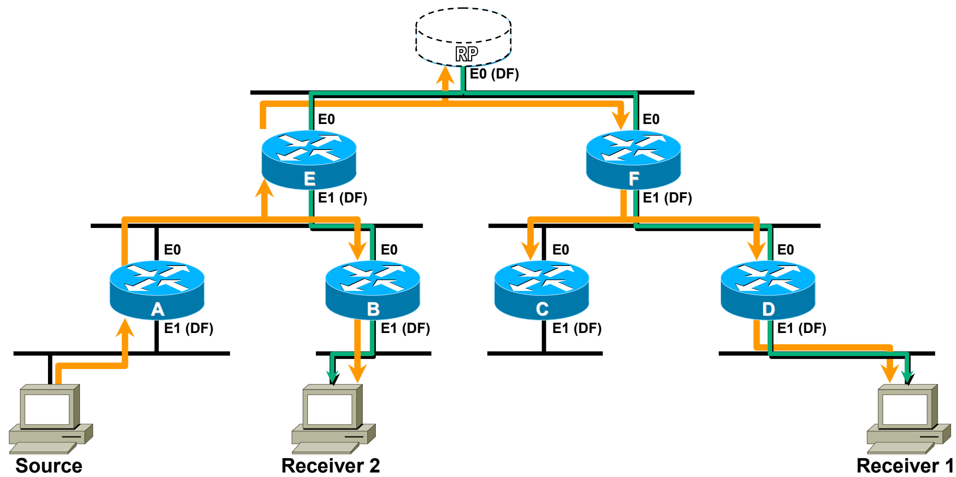
- **Forwarding / Tree Building**
  - Test Time!!

    If the RP failed or was otherwise non-existent, how would this effect the Forwarding and Tree Building in the previous example?

    - Hint: Consider what functions the RP actually performed in the previous example.

Forwarding / Tree Building

Cisco.com



Question: Does the RP even have to physically exist?
Answer: No. It can just be a phantom address.

- **Forwarding / Tree Building**
  - The Answer is, “No”.
  - Since there is neither a Source Registration process nor a Shortest-Path Tree from the RP to the active sources in Bidir PIM, all the RP did was to lend its address as the “core” of the Bidir Shared Tree.
    - This address was used in the DF Election process to determine which router on each subnet had the best metric to the core of the Bidir Shared Tree, i.e. the RP.
    - Once the DF’s were elected for every subnet in the network, the Bidir Shared tree could be built to the core of the network.

Bidir PIM – Summary

- **Uses Shared Trees only.**
  - **Single (*, G) forwarding entry per group.**
  - **Source traffic flows up and down Shared Tree.**
- **Drastically reduces network mroute state.**
  - **Eliminates ALL (S,G) state in the network.**
    - **By eliminating SPT between source and RP.**
  - **Allows Many-to-Any applications to scale.**
    - **Permits virtually an unlimited number of sources.**

- **Bidir PIM – Summary**
  - Bidir PIM forwards traffic both up and down the Bidirectional Shared Trees to reach each receiver in the network. As a result, only (*,G) mroute entries (one per Bidir group) is used and no (S,G) state is created anywhere in the network.
  - Bidir PIM drastically reduces the amount of multicast forwarding state in the network by eliminating all (S,G) entries which eliminates all Shortest-Path Trees in the network.
  - By using only a single (*,G) entry to forward all source traffic, Many-to-Any applications are able to scale to large numbers of hosts.