Deploying IP Multicast

Module 6

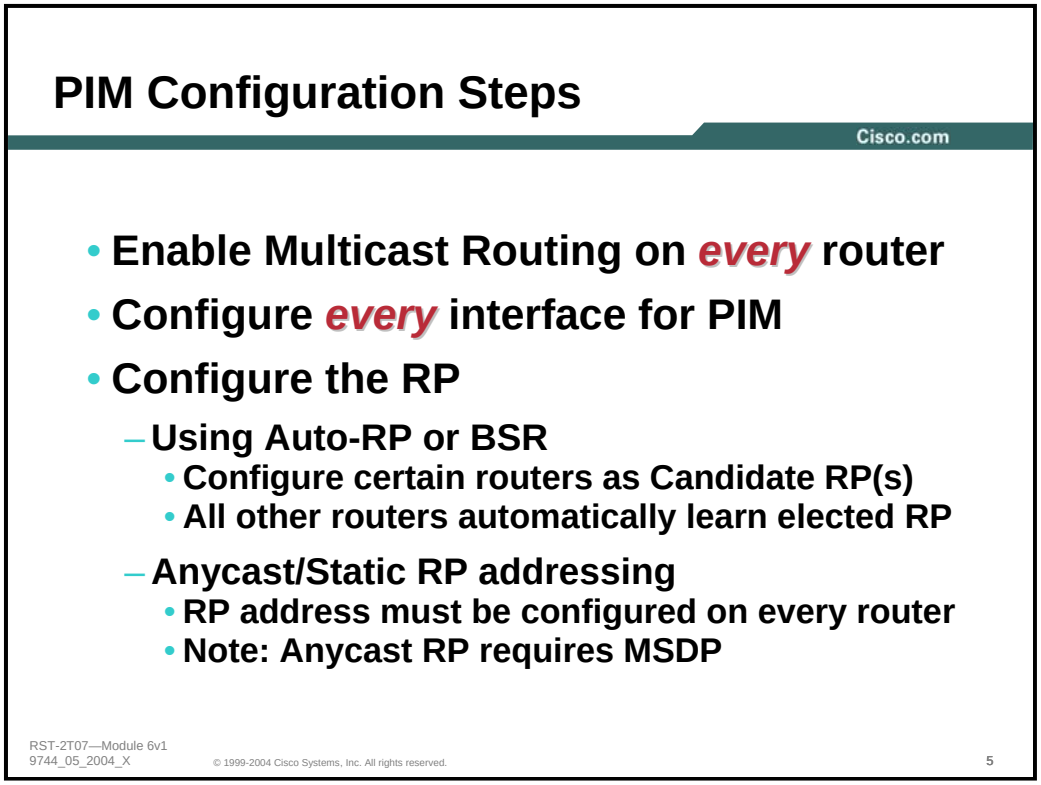
Agenda

Cisco.com

- **Basic Multicast Engineering**
- **Advanced Multicast Engineering**
- **Case Study – ACME Financials**

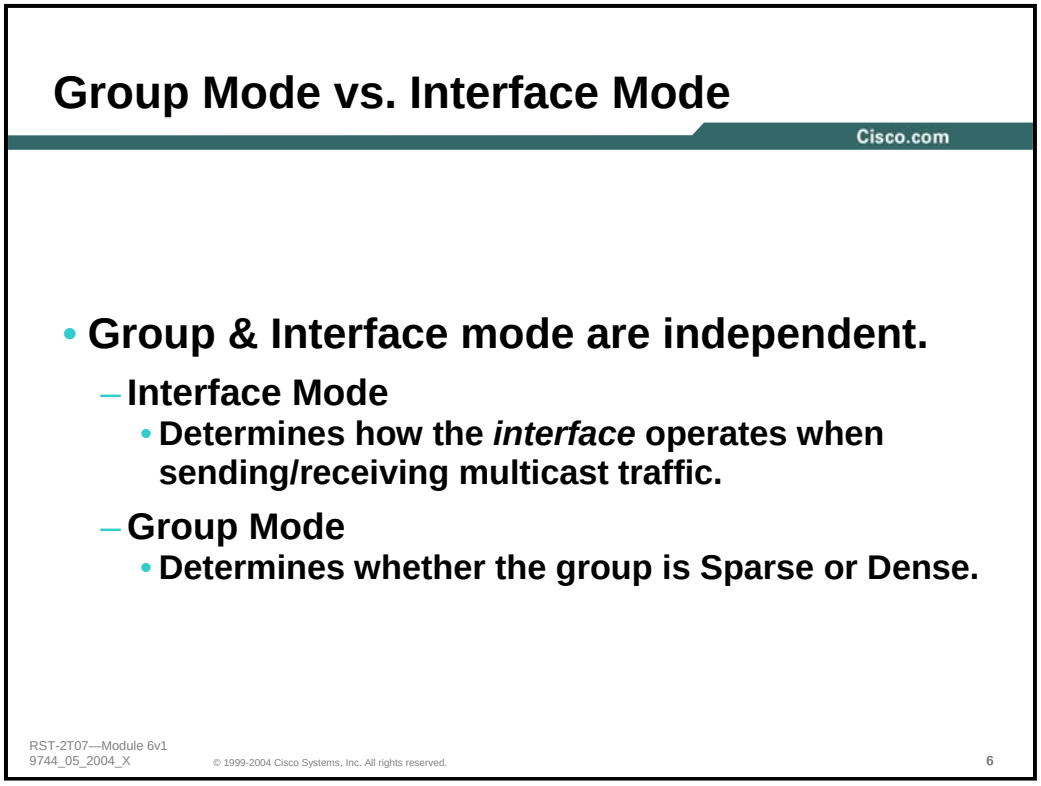
Basic Multicast Engineering

Basic Multicast Configuration – PIM Configuration Steps

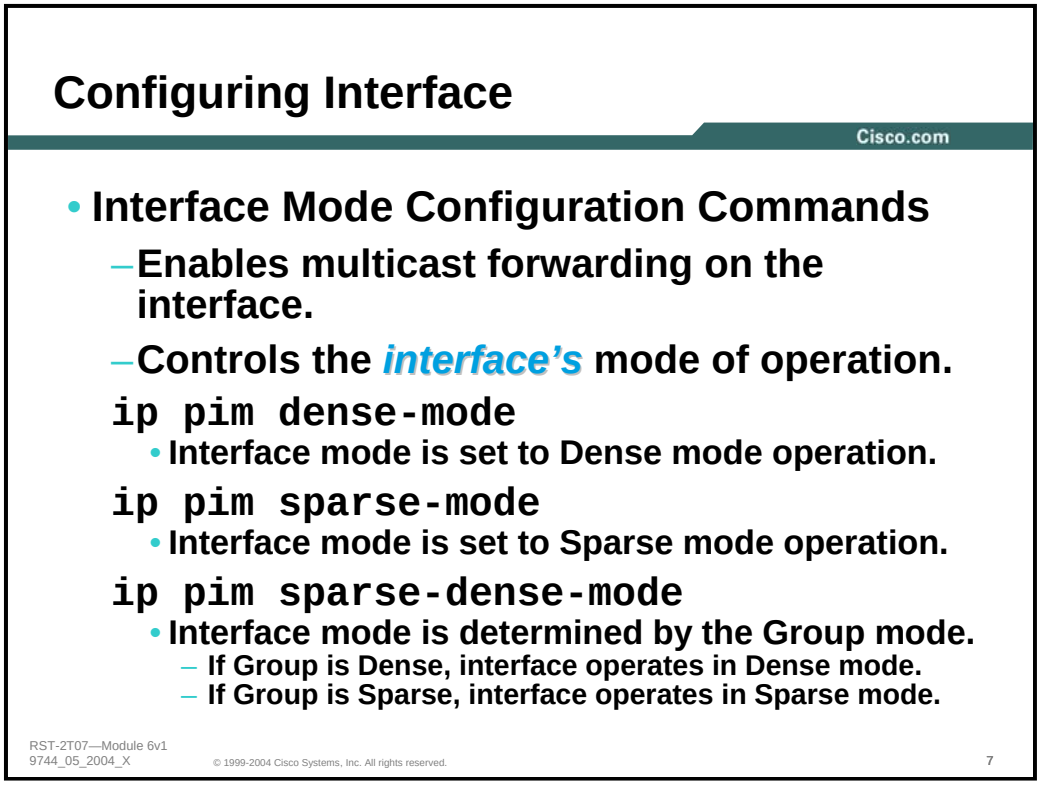


- **Enable Multicast on every router**
  - This is necessary because the PIM RPF (Reverse-Path Forwarding) check depends on the unicast route table to calculate the proper incoming interface for each multicast forwarding entry. Unless extraordinary measures have been taken in the network design (using static mroutes, DVMRP routes, MBGP or other means) to insure that the RPF calculation always resolves to an upstream router that is multicast enabled, it is best to enable multicast routing on every router in the network. This avoids the problem of RPFing to an upstream router that is not multicast enabled which will result in multicast traffic being “black-holed”.

    Note: The RPF check is used in IP Multicast to prevent multicast route loops from forming. It does this by accepting multicast traffic only on the interface that is on the best path back to the source. In other words, the RPF check insures that multicast traffic flows down the distribution tree and never loops around and goes back up the tree at any point. This is similar to the Spanning Tree mechanism which prevents “bridge-loops” from forming in bridgednetworks.
- **Enable PIM on every interface on each router**
  - This is necessary for the same reason above. If the RPF calculation resolves to an interface (using the unicast routing table) that is NOT PIM enabled, multicast traffic will not flow.
- **Configure an RP.**
  - Assuming that the network is to run in Sparse mode (and virtually all production networks should run in Sparse mode), it will be necessary to configure one or more RP’s.
  - RP’s may be configured using several different techniques such as Auto-RP, BSR, Static configuration and Anycast RPs. These techniques are explored in later sections.



- **Group mode vs. Interface mode.**
  - These are totally independent concepts and often a source of confusion for newcomers to IP Multicast.
  - Interface mode
    - As previously discussed, Interface mode controls the behavior of the interface itself and determines how the interface sends and receives multicast control and data packets. It is completely independent from a multicast group's operating mode.
  - Group mode
    - This is the actual mode in which a particular multicast group is operating (Sparse or Dense) and is independent of how any interface is configured on the router.



- **Interface Configuration Commands**
  - There are three different commands that can be used to enable an interface for PIM. Each of these three commands controls the interface's multicast behavior.

    ip pim dense-mode

    This command hardwires the interface to behave as a Dense mode interface. This means that it will apply Dense mode rules to the sending and receiving of multicast data and control traffic.

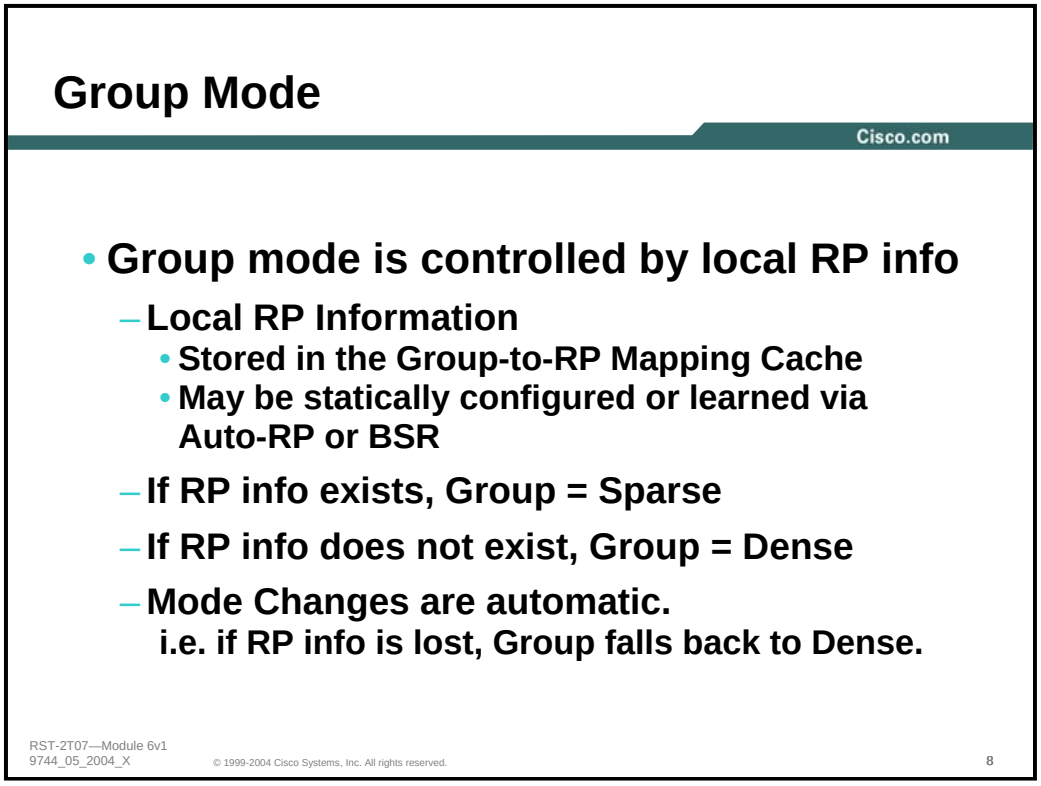
    ip pim sparse-mode

    This command hardwires the interface to behave as a Sparse mode interface. This means that it will apply Sparse mode rules to the sending and receiving of multicast data and control traffic.

    ip pim sparse-dense-mode

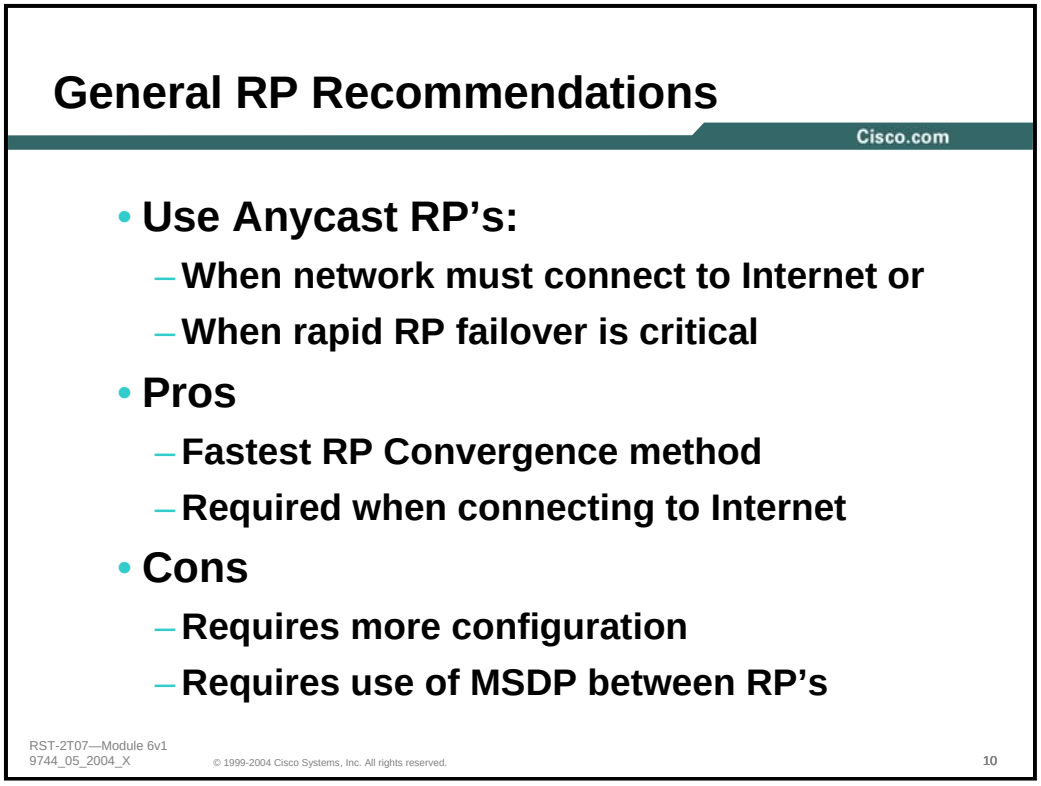
    Instead of hardwiring this interface to behave in one mode or the other, this command causes the interface to behave as a Sparse mode interface when forwarding traffic for Sparse mode groups and to behave as a Dense mode interface when forwarding traffic for Dense mode groups. This allows the interface to switch modes "dynamically" between Sparse and Dense on a packet by packet basis thereby providing support for both Sparse and Dense mode forwarding models in the network. This mode can be thought of as 'ip pim dynamic-mode' since it switches between Sparse and Dense modes dynamically.

    Note: Support for both Dense mode and Sparse mode groups is important in networks that configure RP's using the Auto-RP mechanism. This is because Auto-RP uses two multicast groups which are normally operated in Dense mode. (We will discuss Auto-RP in more detail in a later section.)

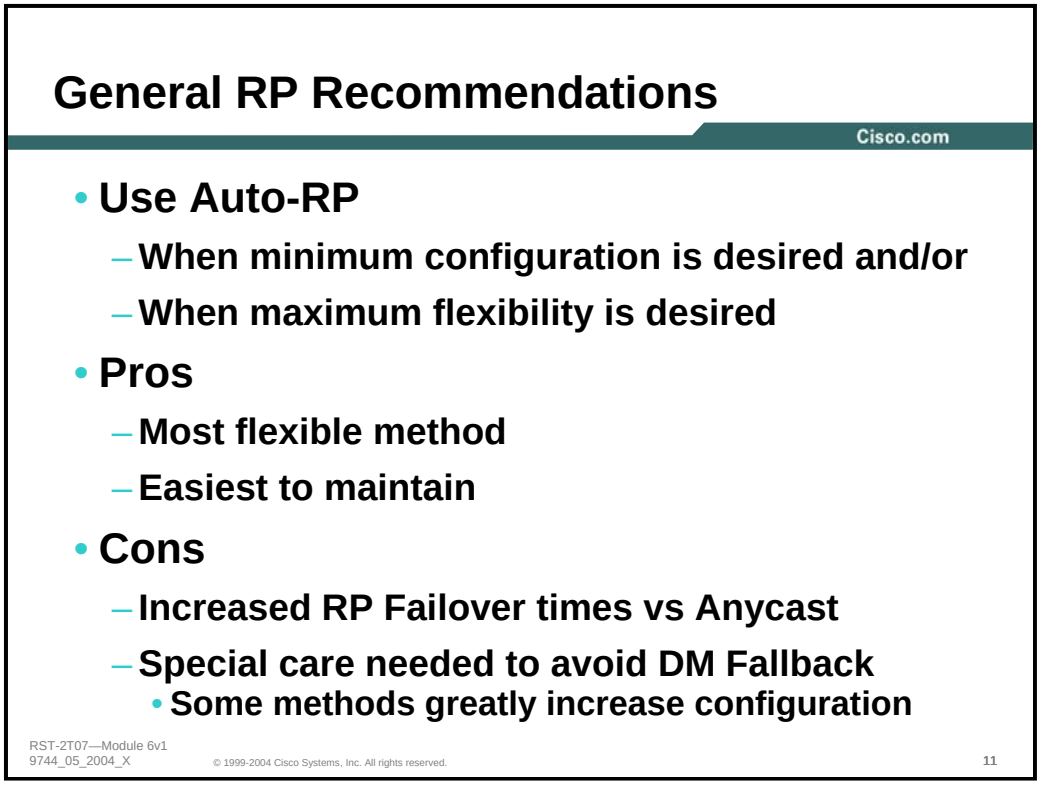


- **Group Mode**
  - Group mode in a router is strictly controlled by its knowledge of a RP for a particular group. This knowledge is stored in the “Group-to-RP” Mapping Cache which can be displayed via the show ip pim rp-mapping IOS command. RP information in this cache can be static configured or learned from the network via the Auto-RP or BSR mechanisms.
  - The rule is simple, if there is an RP in the Group-to-RP mapping cache that covers the group in question, then the router will create a Sparse mode forwarding entry for this group and the group mode will be Sparse. If there is no matching entry in the Group-to-RP mapping cache for a particular group, the router will create a Dense mode forwarding entry for this group and the group will be a Dense mode group.
  - This assignment of mode to a group is dynamic. If RP information is configured or learned on a router where previously there was none, any existing multicast forwarding state for that group will be converted from Dense to Sparse mode. If for some reason the router looses all RP information about a group, then that group will change from Sparse to Dense automatically and the multicast forwarding entry will be marked appropriately in the multicast forwarding table. On the other hand, This behavior

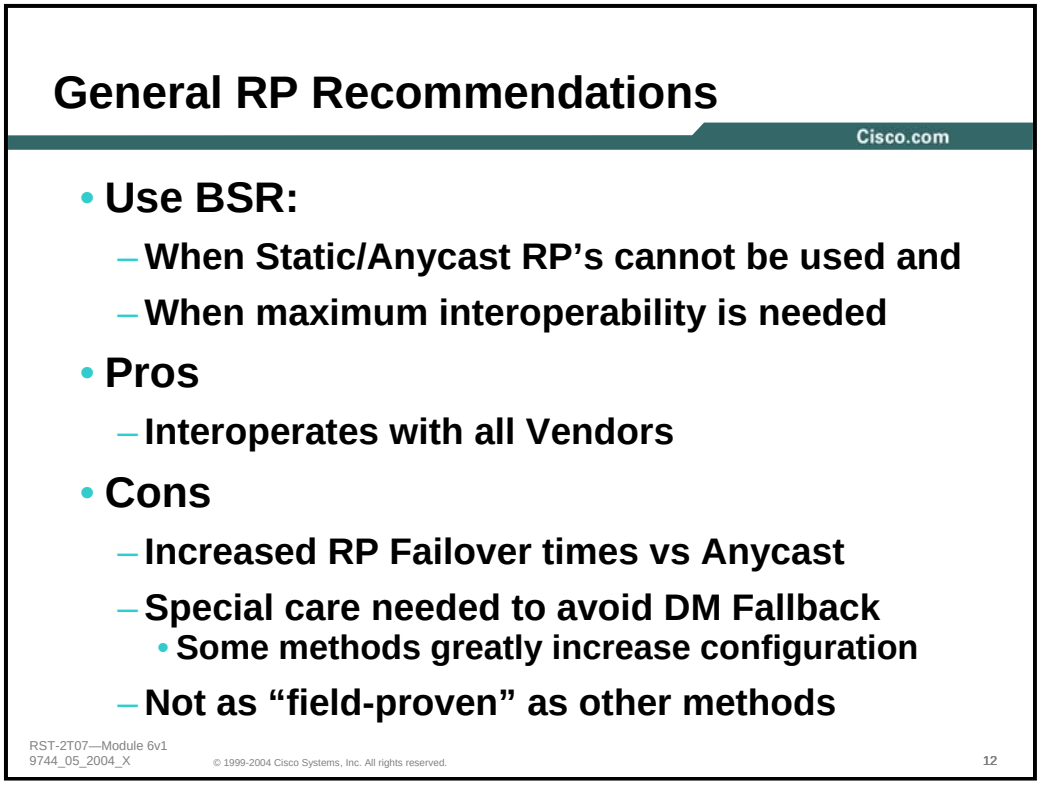
RP Engineering – General RP Recommendations



- **General RP Recommendations**
  - The use of Anycast RP's is recommended whenever the network must connect to the Internet to send/receive IP Multicast traffic or whenever rapid RP failover is critical to network operation.
  - Pros
    - Anycast RP's converge within seconds of the unicast routing protocol and provide the fastest RP convergence times currently possible.
    - The current methodology for sending/receiving multicast traffic to/from the Internet *requires* the use of MSDP to inform the Internet of active sources within the local network as well as learn of active sources in other portions of the Internet. Therefore, if Internet multicast connectivity is necessary, the use Anycast RP's should be considered as the primary RP configuration method.
  - Cons
    - Anycast RP's require more configuration on the individual routers in the network. Each router must be manually configured with the IP address of the Anycast RP. Furthermore, if there are numerous distinct group ranges each with a different Anycast RP address, each router in the network will need a separate configuration for each Anycast RP/Group-Range thus increasing the amount of configuration necessary.
    - Anycast RP's requires the use of MSDP to communicate active source information between all of the Anycast RP's. This means that MSDP must be configured on each of the Anycast RP's.

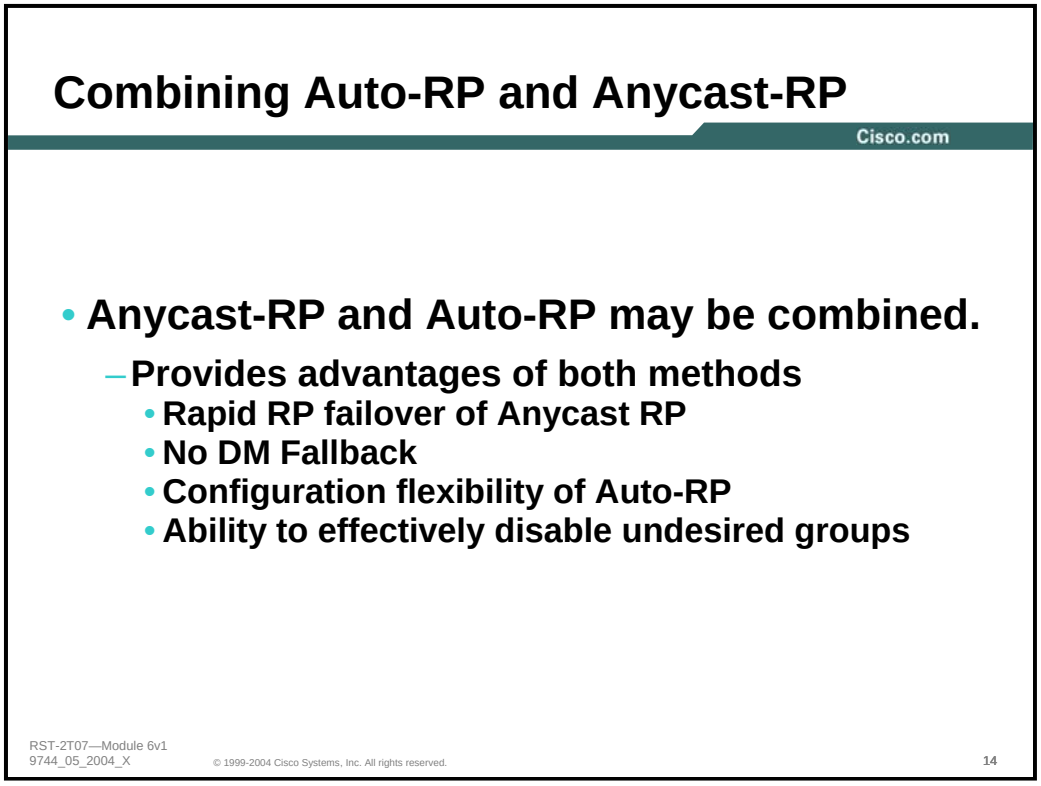


- **General RP Recommendations**
  - The use of Auto-RP is recommended whenever it is desired to minimize the amount of network configuration and/or where maximum flexibility in RP address election/assignment is desired.
  - Pros
    - Auto-RP is the most flexible method of RP configuration. If a new group-range is desired, it is only necessary to configure one or more routers to function as the Candidate RP (C-RP) for this new group range and all other routers in the network will learn which C-RP has been elected as active RP for the new group range.
    - Auto-RP is the easiest RP configuration mechanism when it comes to RP maintenance. C-RP's can be easily added, moved, modified or deleted as desired without having to reconfigure each router in the network.
  - Cons
    - The default Auto-RP failover times are in the order of several minutes. While this can be tuned to achieve failover times in under a minute, it is still tends to be slower than the Anycast RP failover times. Furthermore, reductions in failover time come at the expense of increased Auto-RP Announcement traffic in the network.
    - In the worst case scenario, all C-RP's for a group range could either fail or the Auto-RP mechanism can fail as a result of misconfiguration, network outage or network congestion. If this occurs, all RP information can be lost in some or all routers in the network resulting in the network (or portions thereof) falling back into Dense mode operation along with its associated Dense mode flooding.

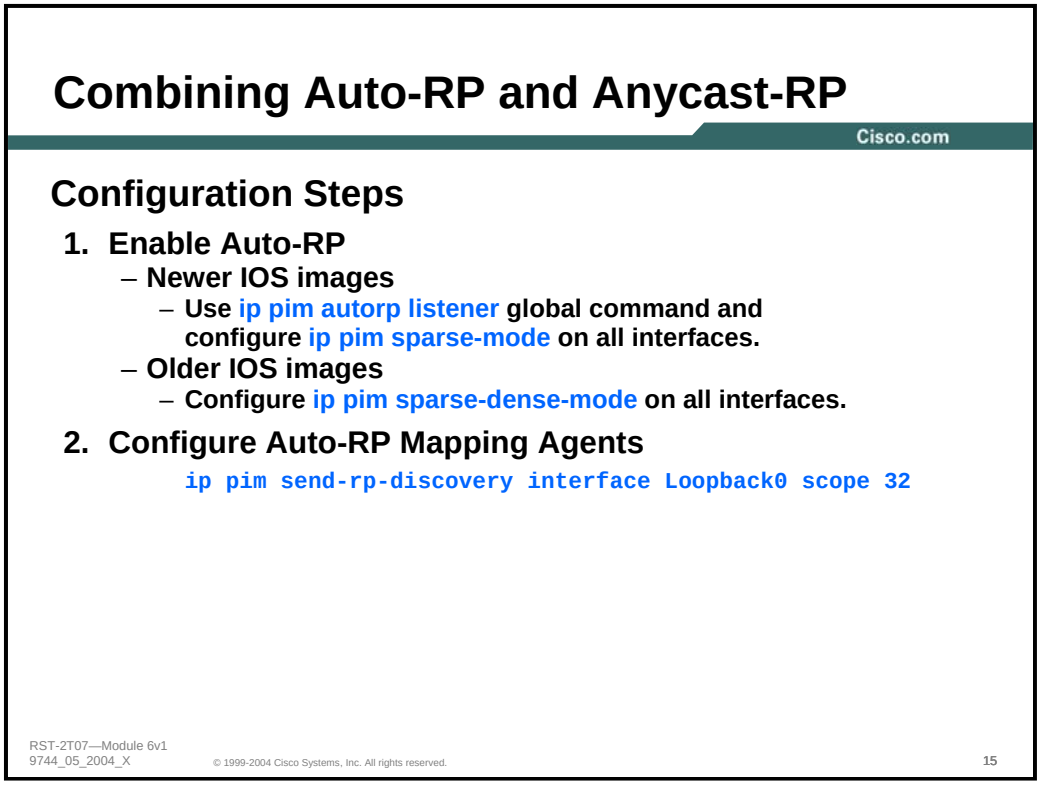


- **General RP Recommendations**
  - The use of BSR is recommended whenever Static/Anycast RP's cannot be used and it is desired to interoperate with non-Cisco routers that do not support Auto-RP. (Note: Some vendors have reversed engineered Auto-RP and offer an interoperable version on their routers.)
  - Pros
    - BSR is specified in the original PIMv2 specifications and hence is supported by virtually all PIMv2 compliant routers.
  - Cons
    - Like Auto-RP, the default BSR failover times are in the order of several minutes. While this can be tuned to achieve failover times in under a minute, it is still tends to be slower than the Anycast RP failover times. Furthermore, reductions in failover time come at the expense of increased Candidate-RP Announcement traffic in the network.
    - In the worst case scenario, all C-RP's for a group range could either fail or the BSR mechanism can fail as a result of misconfiguration, network outage or network congestion. If this occurs, all RP information can be lost in some or all routers in the network resulting in the network (or portions thereof) falling back into Dense mode operation along with its associated Dense mode flooding.
    - BSR has not been used as extensively as the Auto-RP, Static-RP and Anycast RP methods. As a result, the implementations of BSR may not be as robust the other methods.

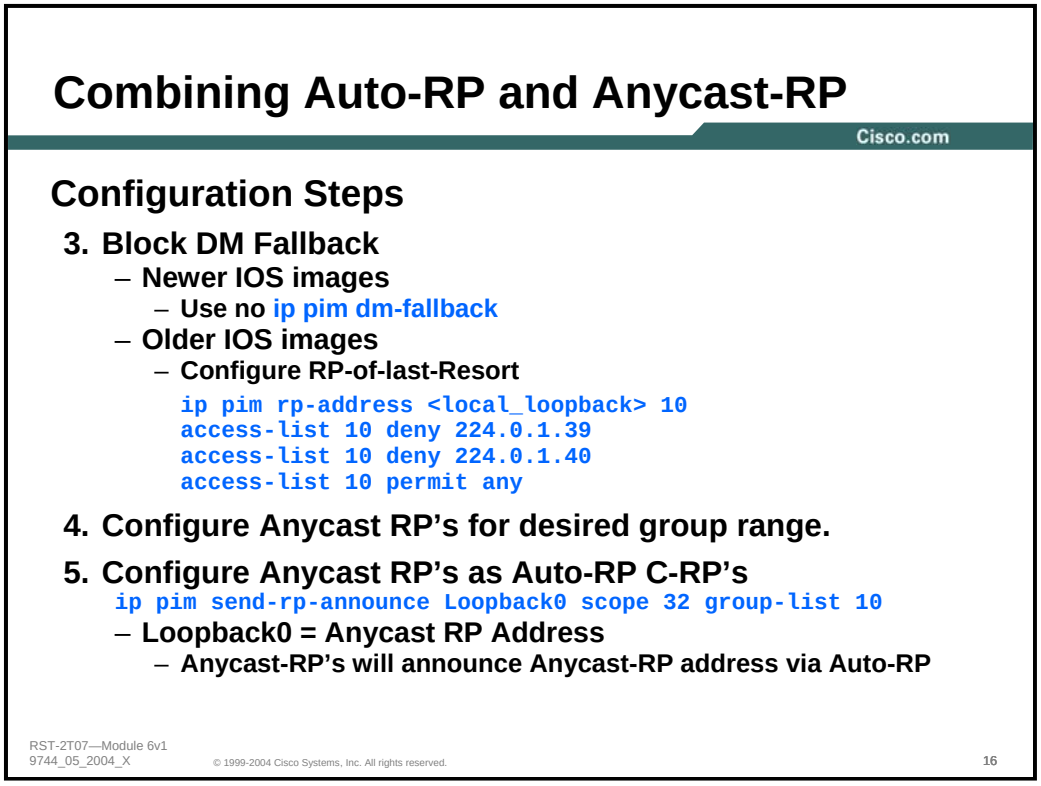
RP Engineering – Combining Anycast RP & Auto-RP



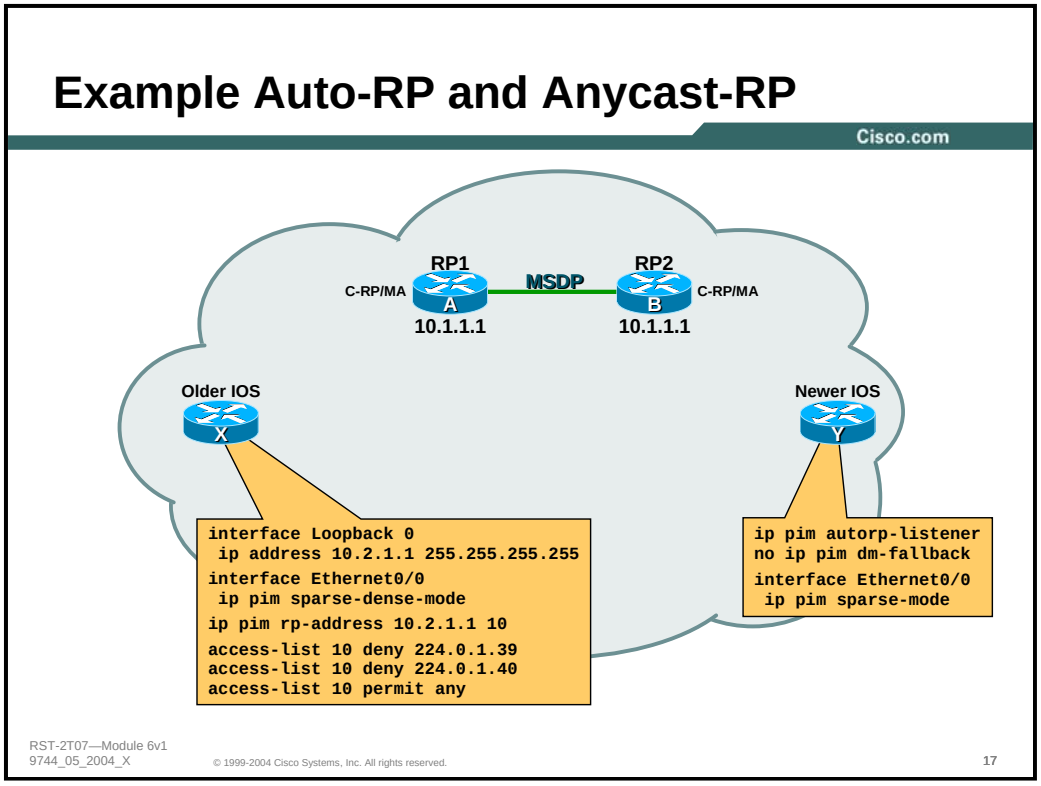
- **Combining Auto-RP and Anycast RP**
  - The techniques of Auto-RP and Anycast-RP are not mutually exclusive and may be combined in certain situations to meet a network administrator's needs for both configuration flexibility, rapid RP failover, no Dense mode fallback and the ability to effectively disable undesired multicast groups.



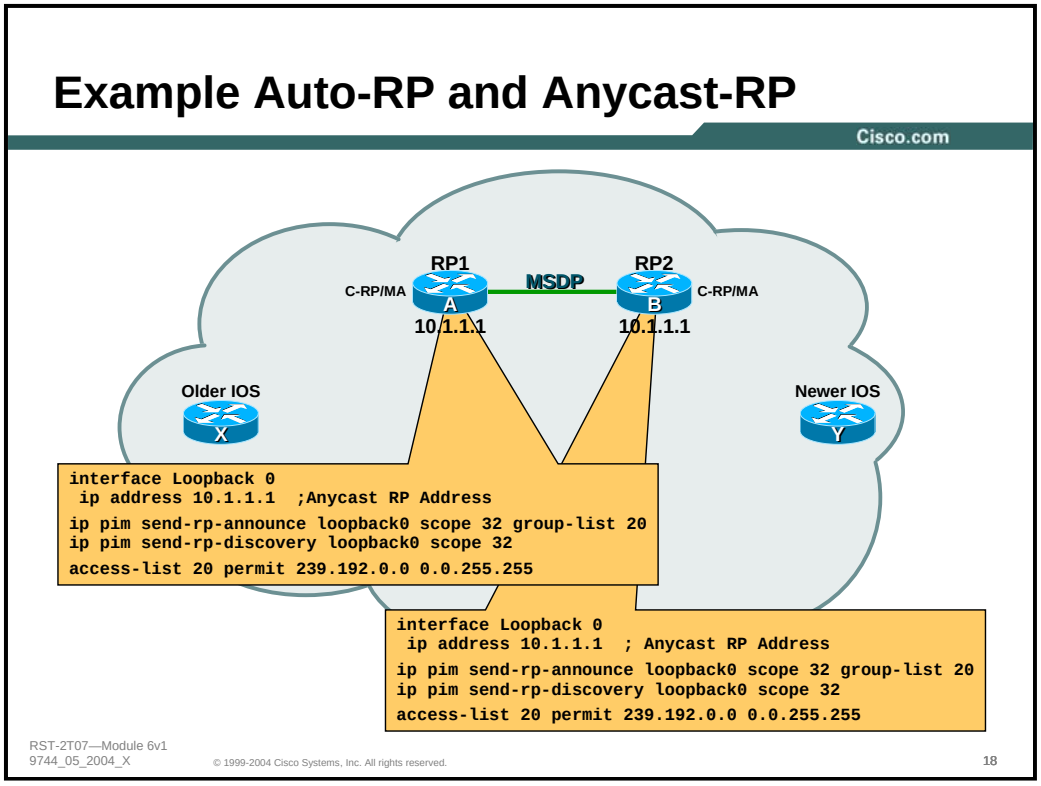
- **Combining Auto-RP and Anycast-RP**
  - Configuration Steps:
    - On newer IOS images, configure the **ip pim autorp listener** global command and also use **ip pim sparse-mode** on all interfaces. This will permit Auto-RP to operate in Dense mode while all other groups will operate in sparse mode.
    - On older IOS images that don't support **ip pim autorp listener** it is necessary to use the **ip pim sparse-dense-mode** command on all interfaces so that Auto-RP may operate in dense mode.
    - Configure Auto-RP Mapping Agents using the **ip pim send-rp-discovery** command on two or more routers in the PIM-SM domain.



- **Combining Auto-RP and Anycast-RP**
  - Configuration Steps:
    - Block DM Fallback by configuring the **no ip pim dm-fallback** command on newer IOS images that support this command. For older IOS images that don't support this command, configure an RP-of-last-Resort on every router in the network. Use the local loopback address of each router as the ip address of the RP-of-last-Resort. This will effectively prevent Dense mode fallback.
    - Configure Anycast-RP's for the desired group range. You may configure different routers as Anycast-RP's for different group ranges depending on your network requirements.
    - Configure the Anycast-RP's as Candidate RP's so that they will advertise the Anycast-RP address via Auto-RP. This is accomplished with the **ip pim send-rp-announce** command. Be sure that the interface specified in this command is the Loopback interface that is being used as the Anycast-RP address.

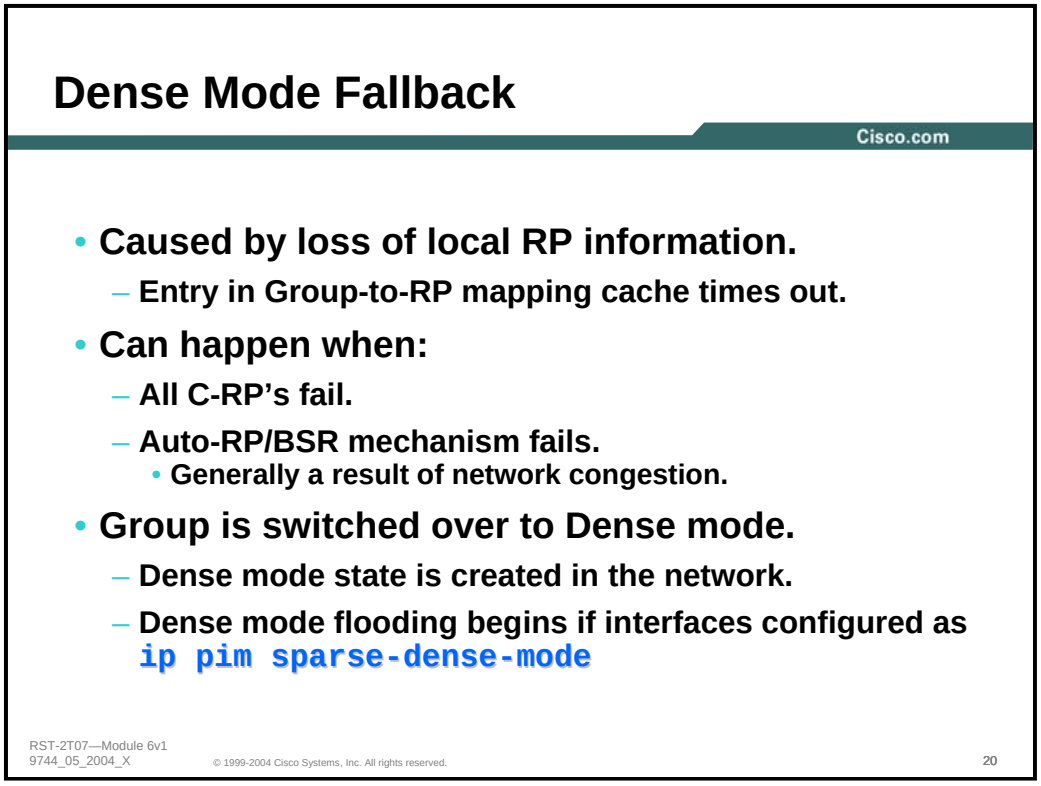


- **Example Auto-RP and Anycast-RP**
  - This diagram shows non-RP routers X and Y. Notice that router X is an older IOS version and requires the use of **ip pim sparse-dense-mode** on all interfaces as well as the RP-of-last-Resort technique with the local loopback interface being used as the RP-of-last-Resort address. Router Y is a newer IOS verson and makes use of the **ip pim autorp-listener, no ip pim dm-fallback** global commands while the **ip pim sparse-mode** command is configured on each interface.

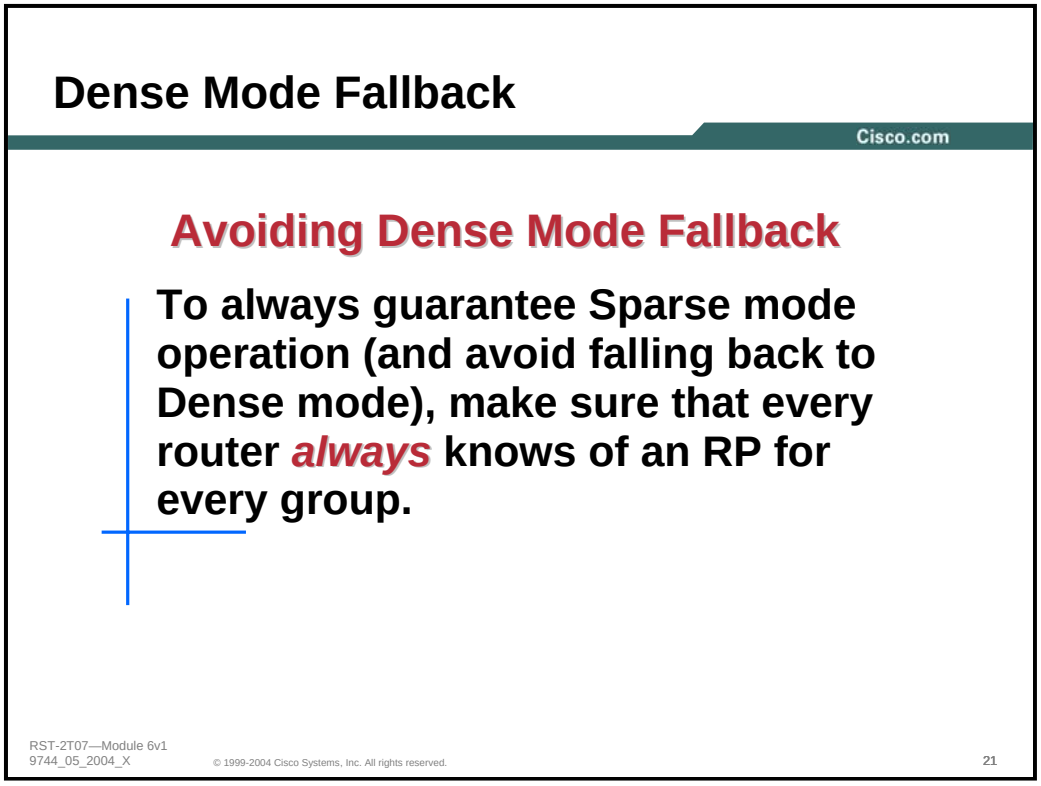


- **Example Auto-RP and Anycast-RP**
  - This diagram shows Anycast-RP routers A and B. Notice that these routers are serving as the Anycast-RP's for 239.192/16 multicast group range. Other routers can be easily configured to serve as Anycast-RP's for other group ranges. In order to combine Auto-RP and Anycast-RP, it is necessary for these two routers to also be configured as Auto-RP Candidate-RP's via the **ip pim send-rp-announce** command.
  - In this example, routers A and B are also configured as Mapping Agents via the **ip pim send-rp-discovery** command. This is not necessary of some other router(s) are configured as Mapping Agents within this PIM-SM domain.

RP Engineering – Avoiding Dense Mode Fallback



- **Dense Mode Fallback**
  - The switch from Sparse mode to Dense mode occurs automatically whenever a router looses RP information for a group. In other words, if for some reason the only entry in the Group-to-RP mapping cache for a group-range times out, the router will automatically change the state of any active entries in the mroute table from Sparse to Dense. This automatic switchover was part of the original design of IOS multicast. The thinking at that time was that if the RP's failed, the network would revert to Dense mode and traffic would continue to flow.
  - Loss of RP information can occur under a variety of situations including:
    - If all Candidate RP's fail
    - If the Auto-RP or BSR mechanisms fail due to network congestion or some other outage.
  - Since in most Auto-RP based network designs, the interfaces have been configured to operate in ip pim sparse-dense-mode so that the two Auto-RP groups will be flooded throughout the network in Dense mode, a switch to Dense mode will be followed by Dense mode flooding of all multicast traffic in this group range. This in turn will cause (S,G) state to be created in every router in the network for any active sources in the group range.
  - We now know that the non-deterministic behavior of PIM Dense mode is highly undesirable and has been know to actually cause network meltdowns under certain conditions. Therefore, steps should be taken to prevent the network from falling back into Dense mode.



- **Avoiding Dense Mode Fallback**
  - Currently, there is only one method that will absolutely insure that Dense mode Fallback doesn't occur and that is to configure the routers in the network so that there is always an RP defined for every group except the two Auto-RP groups, 224.0.1.39 and 224.0.1.40.

Avoiding DM Fallback – Old Workaround

Cisco.com

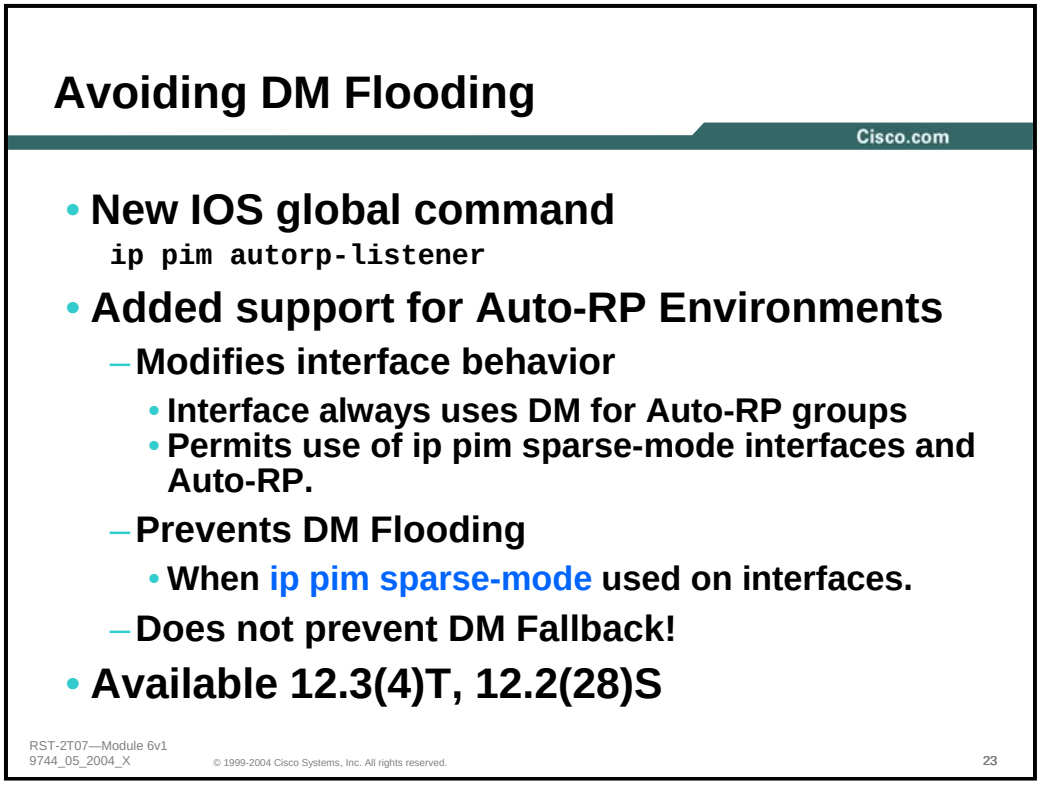
- **Define an "RP-of-last-resort"**
  - **Configure as a Static RP on every router**
    - **Will only be used if all Candidate-RP's fail**
    - **Can be a dummy address or local Loopback**
      - **Recommendation: Use local Loopback on each router**
  - ***MUST use ACL to avoid breaking Auto-RP!***

```
ip pim rp-address <RP-of-last-resort> 10
access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```

RST-2T07—Module 6v1
9744_05_2004_X

22

- **RP-of-last-resort**
  - An "RP-of-last-resort" is a statically defined RP that is just what the name implies; it is a last resort RP if all else fails.
  - Auto-RP and BSR learned RP's take precedence over statically defined RP's (unless the override qualifier is used). This means that any RP's that are statically defined in the router configuration will only be used if there is no Auto-RP or BSR learned RP's for a group. This would be the case should all Candidate RP's were to fail or if a network outage starved the routers in the network from important Auto-RP or BSR information. In that case, the router would "resort" to the statically defined RP for the group. Furthermore, because of this RP is statically defined, it will always exist in the Group-to-RP mapping cache and therefore the router will never entirely lose all RP information and revert to Dense mode.
  - The actual address specified as the "RP-of-last-resort" can be just about any address the network administrator wishes. One recommendation is to use the lowest priority Candidate RP address as the "RP-of-last-resort". This means that should the Auto-RP or BSR mechanism fail (or all C-RP's fail), the router will use the lowest priority Candidate RP as the active RP until the error condition has been resolved. This means that the multicast groups will remain in Sparse mode and any existing (S,G) state in the network will continue to remain active which in turn, will allow multicast traffic to flow over existing Shortest-Path trees. The only impact to the multicast network is that new receivers and sources will be unable to Join the Shared Tree or Register to the RP, respectively.
  - Care must be taken when defining the "RP-of-last-resort" since it can cause the two Auto-RP groups to switch to Sparse mode. This is because unlike Auto-RP learned RP's which never apply to the two Auto-RP groups, any statically define RP's that cover the Auto-RP groups will be used by the router. This would cause the two Auto-RP groups to try to run in Sparse mode which is normally not desired. In order to avoid this problem, it is necessary to specify an RP group range ACL that specifies the two Auto-RP groups with a deny clause to prevent these groups from switching to Sparse mode.



- **Avoiding DM Flooding**
  - One of the major problems with Auto-RP is that it requires all interfaces be configured with **ip pim sparse-dense-mode** so that the two Auto-RP groups can operate in Dense mode. Unfortunately, this can result in ALL groups reverting to Dense mode operation of all C-RP's fail or if there is a problem with the Auto-RP mechanism that results in RP information being timed out in each router's Group-to-RP Mapping Cache.
  - To help avoid the resulting dense mode flooding that can occur when this happens, the **ip pim autorp-listener** command has been introduced. This command allows all interfaces to configured with **ip pim sparse-mode** and yet continues to allow the Auto-RP groups to operate in dense mode on the interface. Because the interface is now operating in sparse mode for all other groups, no dense mode flooding can occur out this interface.
  - NOTE: This technique does not prevent Dense Mode FALLBACK which results in all group state switching to dense mode in the router. All it does is prevent dense mode FLOODING. To completely avoid the Dense mode Fallback problem, the use of another IOS feature is required. (See next section.)

Avoiding DM Flooding

Cisco.com

- **Deploying ip pim autorp-listener**
  - **Must be configured on every router.**
  - **Use RP-of-last-resort on older IOS versions until upgraded**
    - **Assign local Loopback as RP-of-last-resort on each router.**
    - **Example**

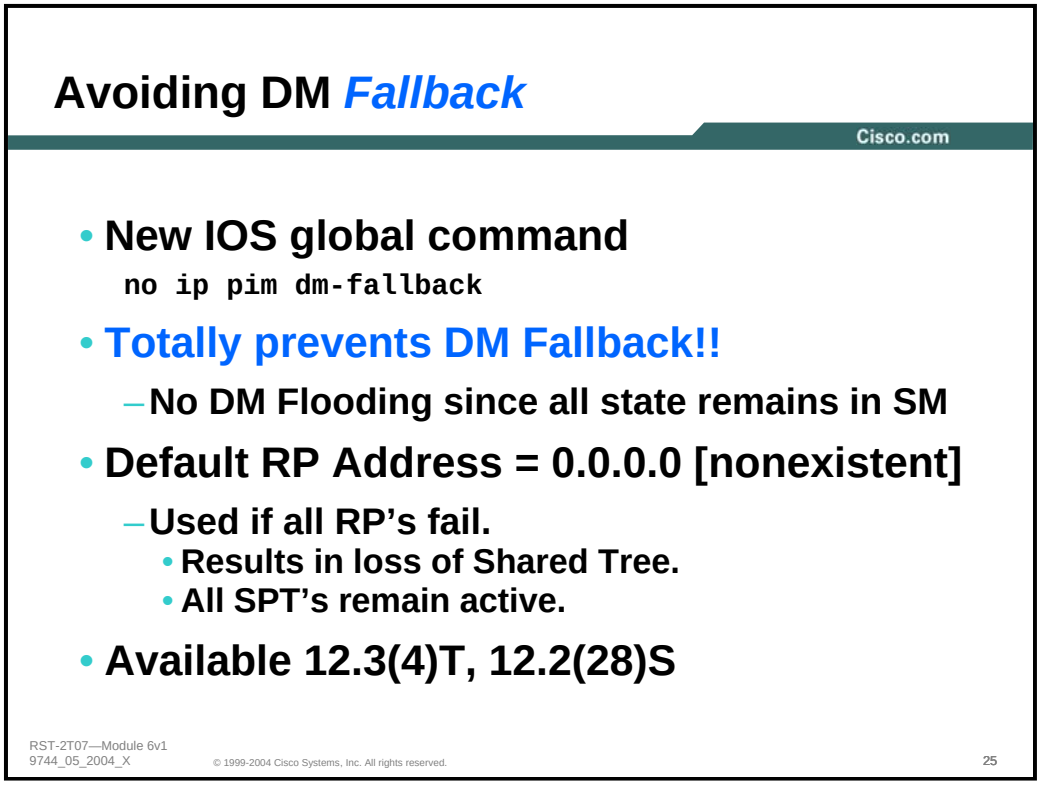
```
ip pim rp-address <local_loopback> 10
access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```

RST-2T07—Module 6v1
9744_05_2004_X

24

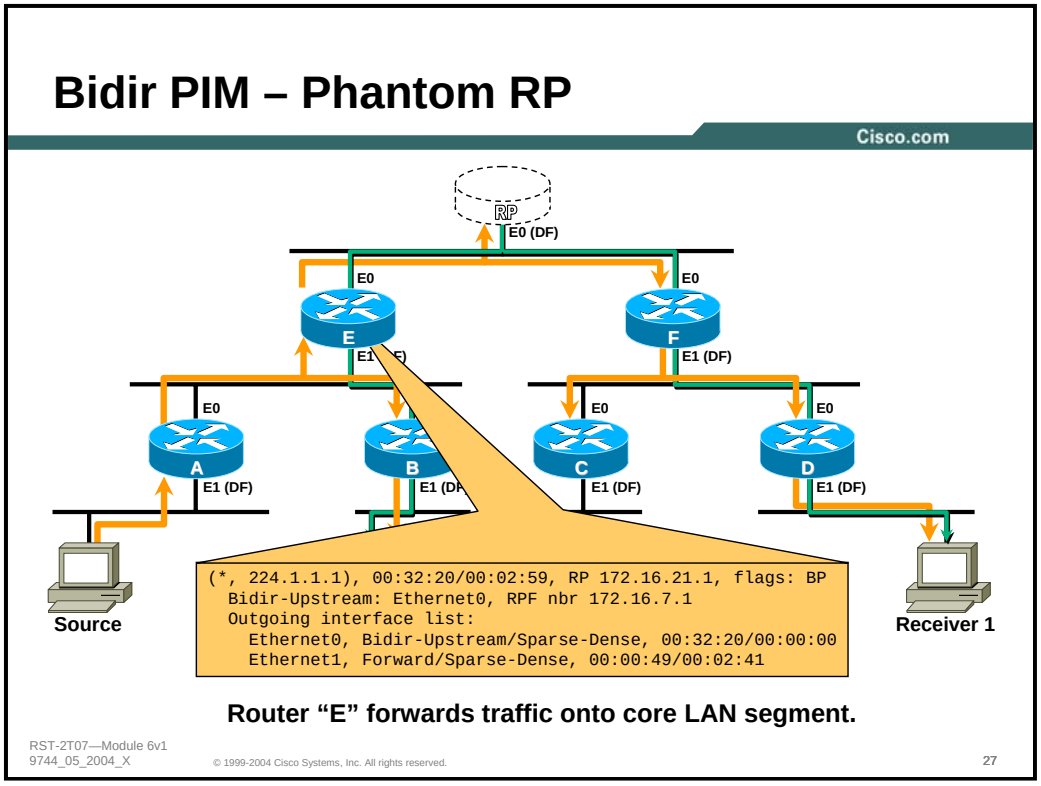
- **Avoiding DM Flooding**
  - Configure **ip pim autorp-listener** on every router in the network.
  - For older versions of IOS that do not support this command it is necessary to configure an RP-of-last-Resort. (It is recommended that the address of the RP-of-last-Resort be the address of the local loopback interface on each router.)
    - Example RP-of-Last Resort configuration:

```
ip pim rp-address <local_loopback> 10

   access-list 10 deny 224.0.1.39
   access-list 10 deny 224.0.1.40
   access-list 10 permit any
```

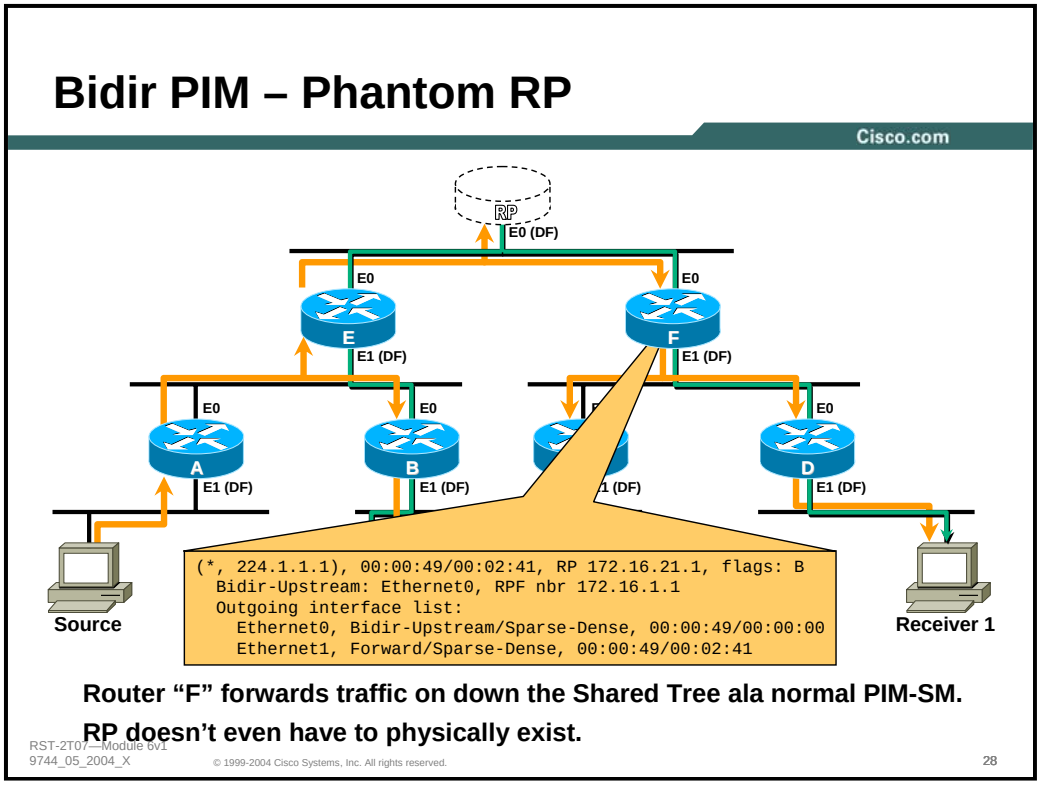


- **Avoiding DM Fallback**
  - In order to completely avoid Dense mode Fallback including the switch of all group state in the router to Dense mode when all RP's fail, a new IOS command, **no ip pim dm-fallback** has been introduced.
  - When the **no** form of this command is configured, a default RP address of 0.0.0.0 is assumed for all multicast groups (except for the Auto-RP groups). Because 0.0.0.0 is a non-existent RP, only the Shared Tree and RP will lost when the router reverts to this RP address in the case of total RP failure. This will prevent any new receivers or senders from being Joined or Registered. However, all existing Shortest-Path Trees will be maintained by the network and existing flows continued.

RP Engineering – Phantom Bidir RP's



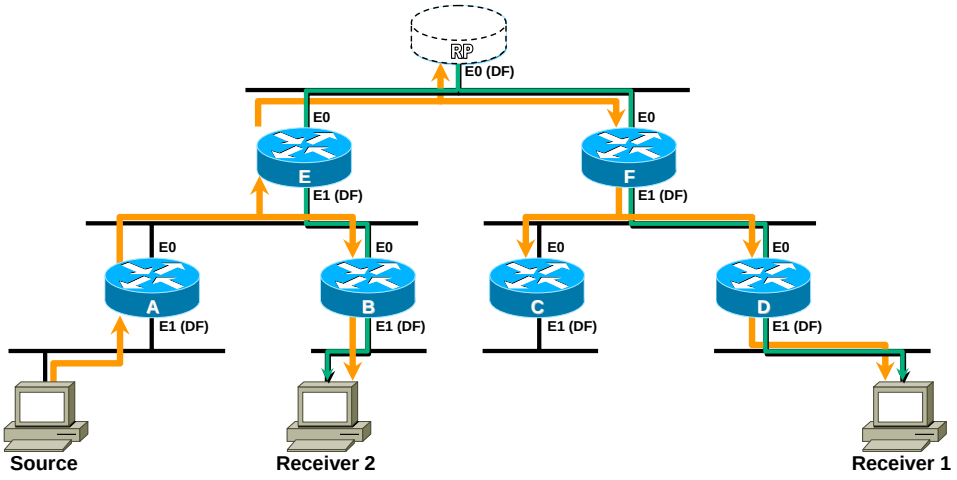
- **Bidir Phantom RP**
  - Bidir PIM differs from normal PIM-SM operation in that sources are not Registered with the RP and the RP does not join the SPT to a source to "pull" the traffic to the RP. Instead, source traffic is permitted to flow "up" the Shared Tree in the direction of the RP and from there back down the existing branches of the Shared Tree. (Because the traffic can flow both up or down the Shared Tree, this form of Shared Tree is referred to as a Bidirectional Shared Tree and hence the name "Bidir PIM".)
  - Consider the Bidir PIM network example in the above slide. Traffic from the source in the lower left corner is flowing up one branch of the Shared Tree and then back down the other branch to Receiver 2.
  - Examining the state on router "B" we see that the traffic is being forwarded up the Shared Tree onto the common multi-access LAN segment where the RP resides.



- **Bidir Phantom RP**
  - Note that when the traffic reaches the RP at the top of the hierarchy, it also automatically flows across the multicast access LAN segment and then back down the other branch of the Shared Tree to Receiver 1 without the RP having to take any action to forward the traffic.

Bidir PIM – Phantom RP

Cisco.com



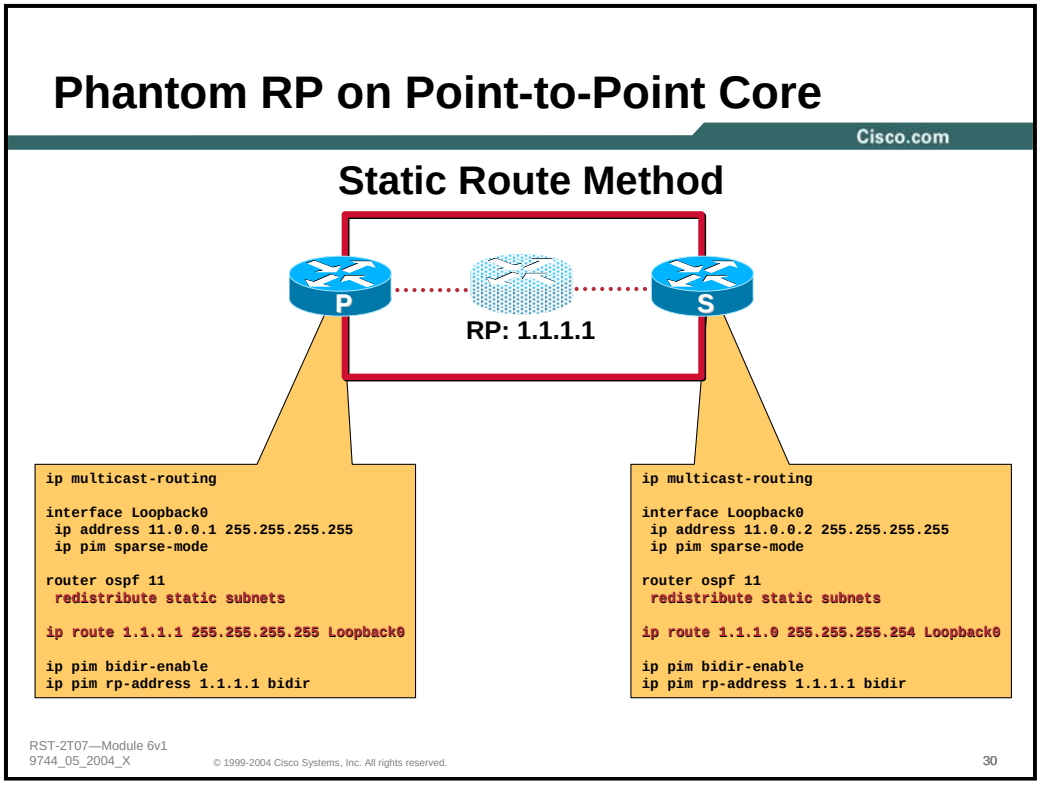
Question: Does a Bidir RP even have to physically exist?
Answer: No. It can just be a phantom address.

- **Bidir Phantom RP**
  - Consider the following questions:

    What did the RP contribute to the flow of traffic?

    If the RP failed, what affect would it have on network operation?

  - The answer to the both of the questions above is: **NOTHING.**
    - In fact, the RP for a Bidir network does not even have to be a physical device. It is only necessary for the network to agree as to the network location of root of the Shared Tree. This can actually be a phantom, nonexistent RP address that has been configured as the root of the Bidir Shared Tree.



- **Phantom RP on Point-to-Point networks**

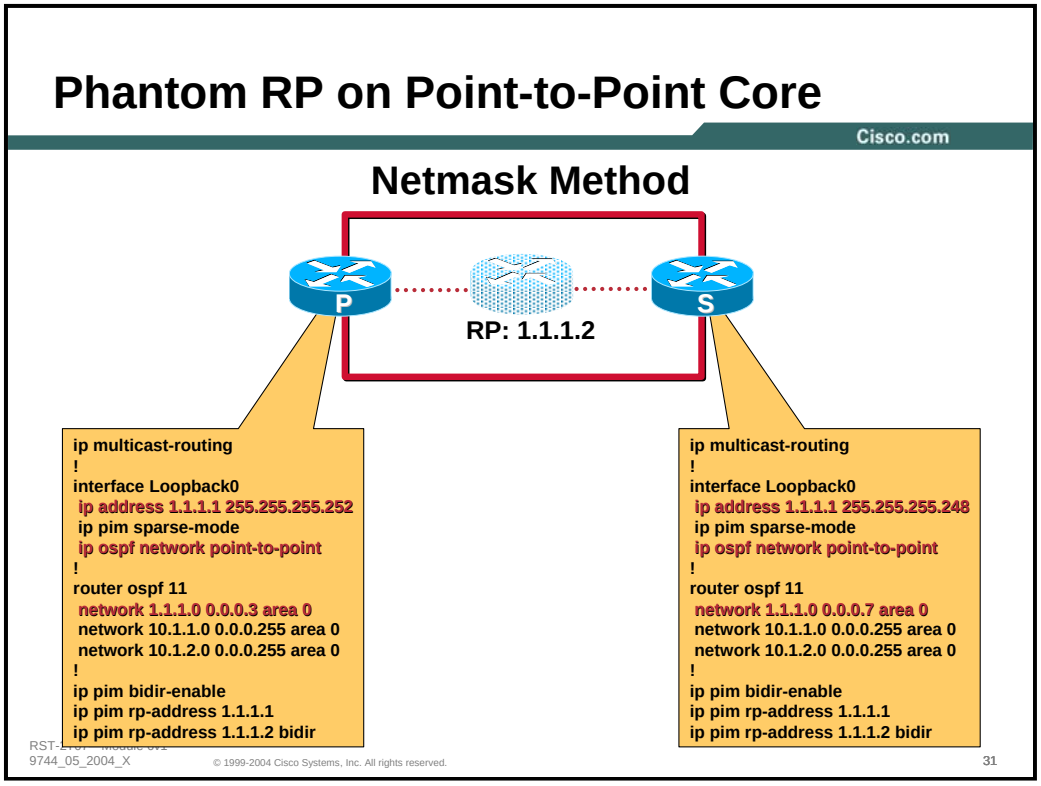
  The concept of Phantom RP can be extended to networks that use a point-to-point core instead of multi-access LANs.

  The fundamental concept is to advertise a route to the Phantom RP using different mask lengths; the longest of which is advertised from the Primary router and the shorter from the Secondary router. Because the normal routing behavior is to select the route with the longest match (all other metrics being equal), the Bidir Shared Tree is built toward the Phantom RP via the Primary router as long as it is up and advertising the shorter mask length route. Should the Primary router fail, only the route advertised by the Secondary router would remain (after unicast routing converges). This would cause the Bidir Shared Tree to be rebuilt through the Secondary router.

  There are are at least two methods to accomplish the above:

  - Static Route Method
  - Netmask Method

  – The Static Route method is shown in the example above.
    - The Primary router has defined a static route to the Phantom RP, 1.1.1.1 with a mask length of /32. This route is advertised to the rest of the network by the “redistribute static subnets” command.
    - The Secondary router has defined a static route to the Phantom RP, 1.1.1.1 with a mask length of /31. This route is advertised to the rest of the network by the “redistribute static subnets” command.
    - As long as the Primary router is up, it will function as the core of the Bidir Shared Tree. If it fails, the Bidir Shared Tree will be rerouted to the Secondary router which will then become the core of the Bidir Shared Tree.



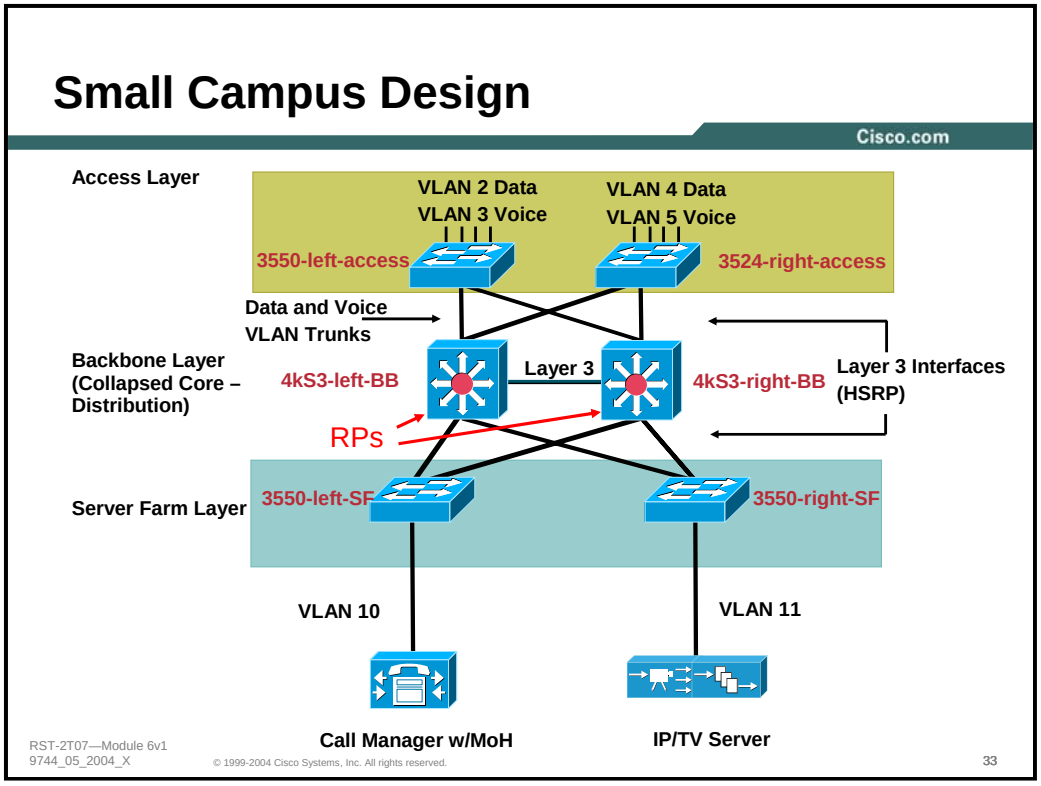
- **Phantom RP on Point-to-Point networks**

  The Netmask method is shown in the example above.

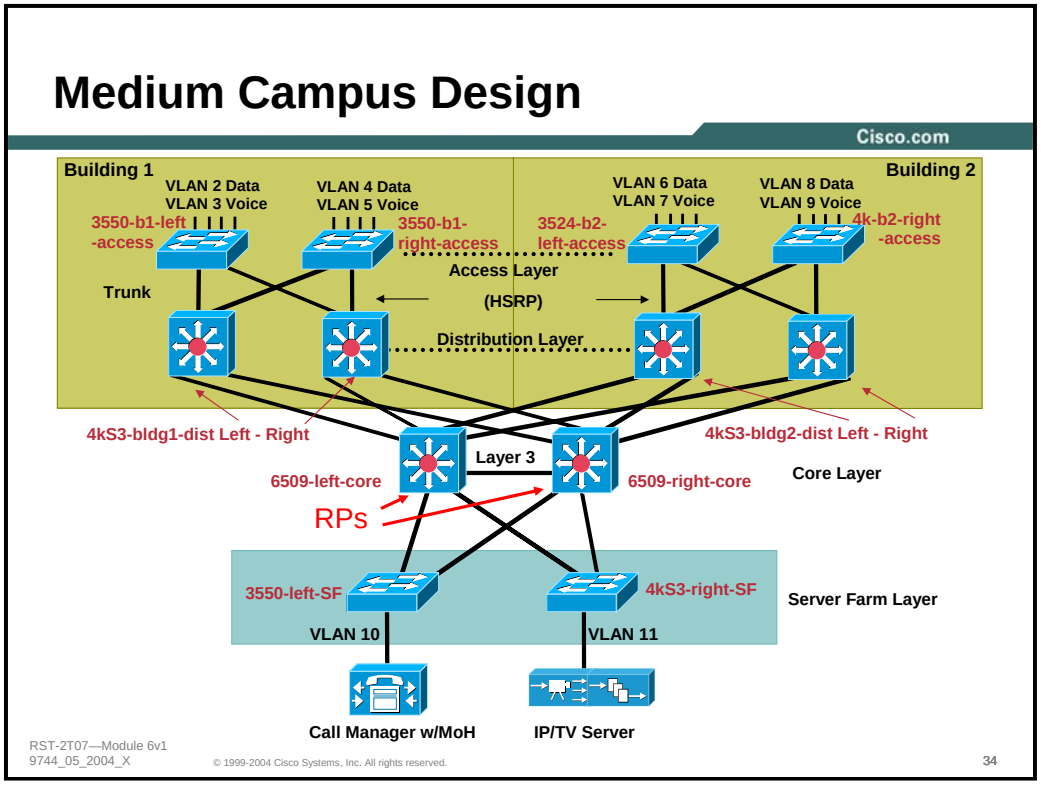
  - Both the Primary and Secondary Loopback interfaces are assigned identical addresses that would exist on the same virtual network as the Phantom RP whose address is 1.1.1.2.
  - In this configuration, a netmask length of /30 is used on the Loopback interface of the Primary router. This, coupled with the OSPF “network 1.1.1.0 0.0.0.3 area 0” command results in a /30 route to the 1.1.1.0 network being advertised to the rest of the network.

    Note: ‘ip ospf network point-to-point’ command is required on the Loopback interface to prevent the router from attempting to elect an OSPF DR and BDR on this subnet.. This route is advertised to the rest of the network by the “redistribute static subnets” command.
  - In the case of the Secondary router, a netmask length of /29 is used on the Loopback interface. This, coupled with the OSPF “network 1.1.1.0 0.0.0.7 area 0” command results in a /29 route to the 1.1.1.0 network being advertised to the rest of the network.
  - As long as the Primary router is up, it will function as the core of the Bidir Shared Tree. If it fails, the Bidir Shared Tree will be rerouted to the Secondary router which will then become the core of the Bidir Shared Tree.

Basic Campus Designs



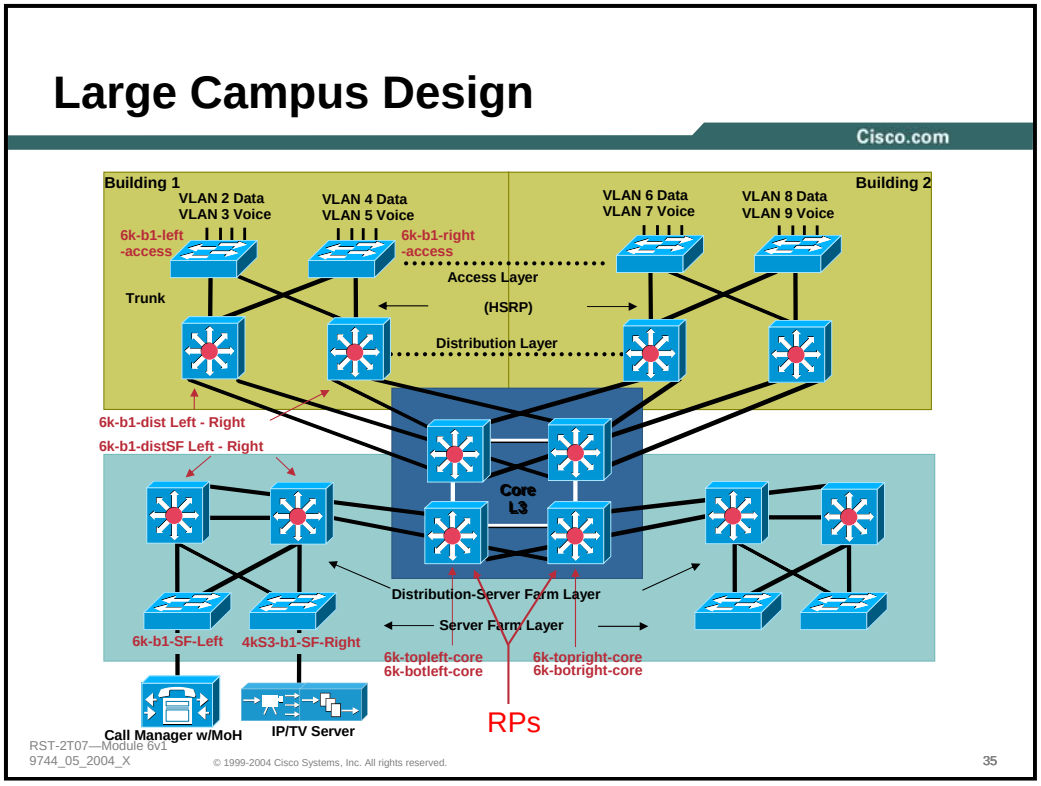
- **Small Campus Design**
  - A small campus is an isolated network consisting of a single building or site.
- **Access Layer**
  - We show two different Catalyst models here to illustrate that either IGMP-snooping or CGMP capable switches or both may be used. The 3550 supports IGMP-snooping, which is implemented via a specialized hardware ASIC. The 3524 (Typical of the Catalyst XL series) supports CGMP only. CGMP operates in software and requires an attached router known as the CGMP server.
- **Backbone Layer**
  - The two Layer 3 switches are acting as the RPs in the network (they are the only L3 devices capable of running PIM)
  - This network is very simple and, in this illustration, has no outside connection that would require an advanced deployment model like Anycast RP, MSDP for ISP peering and many other detailed features
  - Auto-RP can be used in this model as a good starting point.
    - Enabling both BB switches to be both Candidate-RPs and Mapping Agents allows for a fault-tolerant design (there will always be one RP on the network) and allows for further growth if the small campus one day connects to a larger network that also uses IP Multicast
- **Server-farm layer**
  - The two server-farm switches are Layer 2 only and are running IGMP-snooping (enabled by default)
  - In this example we have two active multicast sources on the network (Call Manager with Multicast Music-on-Hold (MMoH) and a IP/TV server streaming multicast media



- **Medium Campus Design**

  A medium campus is a network that can have 1-2 buildings and can be attached to a larger campus network via a WAN deployment. A medium campus will have a distinct core and distribution layer.

- **Access layer**
  - The access layer in the medium campus design is exactly like the small deployment. It is preferred to use IGMP-snooping enabled switches as there are no additional commands required and the L2 forwarding of multicast is performed in hardware.
- **Distribution layer**
  - The distribution layer will have multicast routing and PIM enabled. In this example we are using Anycast RP and have static RP entries defined on the distribution layer switches pointing to the RPs in the core.
  - Additional features will be deployed to control multicast traffic. Multicast boundaries will be used.
- **Core layer**
  - In this example the core layer is where the meat of the configuration will be.
  - This design is will use Anycast RP. The two core switches will use PIM-SM and MSDP as the mechanisms to facilitate the RP deployment.
  - Additional features will be used to help protect against illegal sources and RPs. Accept register ACLs and static RP configs will be used.
- **Server farm layer**
  - The SF design will use Layer 2 switches that use IGMP-snooping



- **Large Campus Design**

  A large campus consists of multiple buildings with 1000 users or more. A large campus may also have a multicast peering relationship with an ISP requiring a more advanced configuration to account for multiple multicast domains.

- **Access layer**
  - Higher-end switches are typically used in a large campus simply to support a higher user population (more access ports). All medium to high-end Catalyst switches support IGMP-snooping

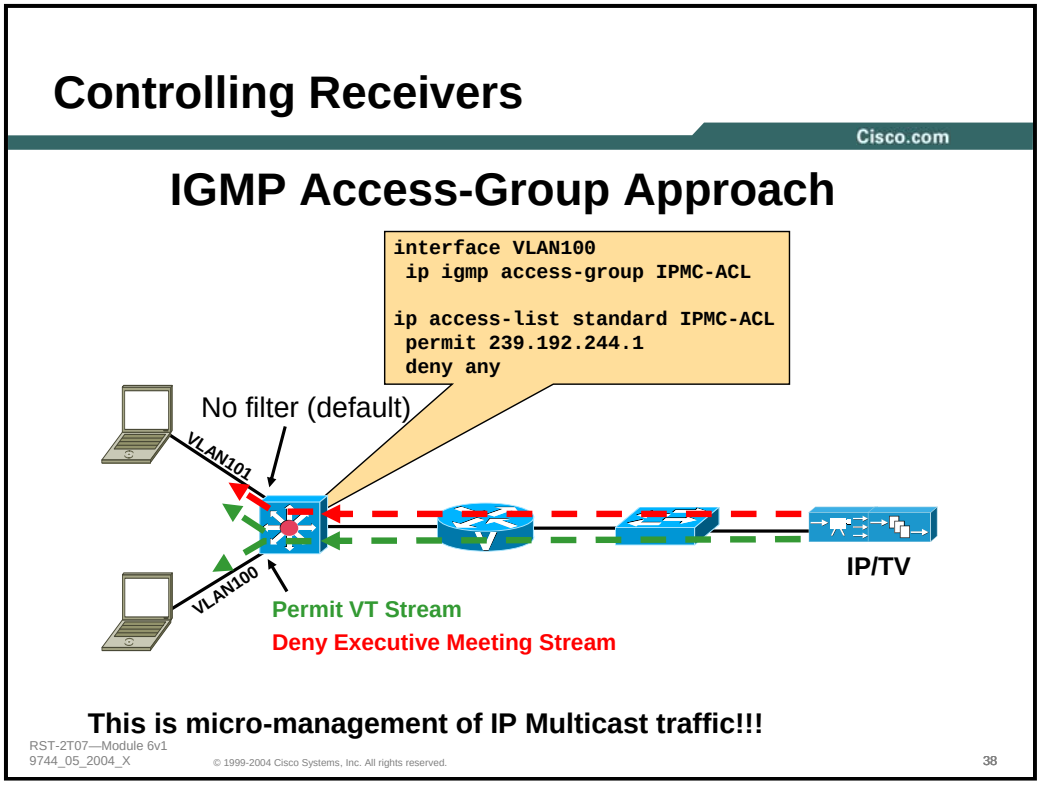
- **Distribution layer**
  - The distribution layer design is the same as in the medium campus network. There could be a need to support multiple multicast group ranges for varying types of multicast applications. For example, you may use Auto-RP to advertise globally accessible multicast applications and static-RP definitions (in conjunction with Anycast) for localized applications.
  - The distribution layer switches will utilize the RPs located in the core.
  - Multicast boundaries will be used to control the flow of multicast traffic.

- **Core layer**
  - Large networks will commonly have multiple core switches to cover various blocks in the network (WAN, ISP and Data Center blocks). Typically there would be no reason to use more than two of the core layer switches as RPs.

- **Server farm (Data Center)**
  - Multicast sources such as streaming media servers should be directly connected to the aggregation layer switches on a separate VLAN just for multicast. This permits multicast to be forwarded natively rather than tunneling multicast via GRE through DC devices (firewalls for example) that do not support multicast.

Advanced Multicast Engineering

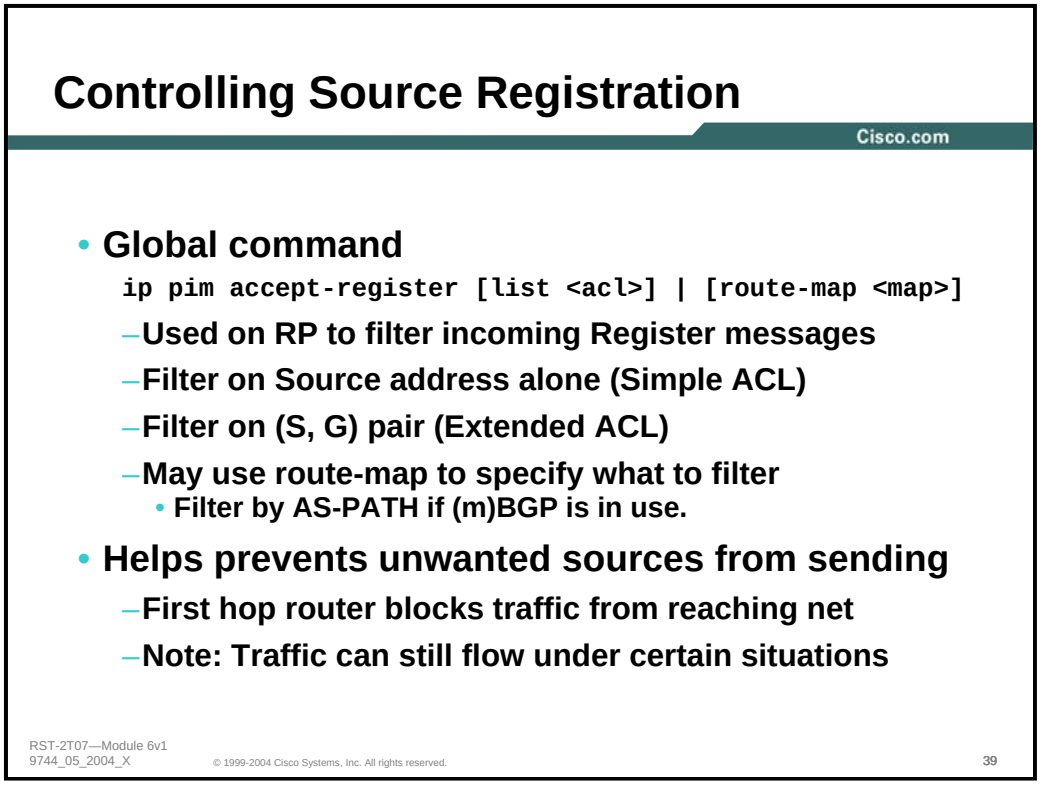
Security



- **IGMP Access-Group Approach**
  - Access control of received multicast groups is still an area of multicast technology that is “Under Development”. Unfortunately, this means that the tools available for Multicast Group Access Control are extremely primitive.
  - The IOS interface command

    `ip igmp access-group <ACL>`

    provides the ability to specify which groups can be Joined by hosts on a router interface. When this command is configured on an interface, IGMP Membership Reports for the groups that have been “denied” by the associated ACL will be ignored by the router on that interface.
  - In the example above, the Network Administrator has configured an ‘ip igmp access-group’ command on the router interface to prevent all groups except the “Virtual Team” IP/TV Stream from being joined by hosts on the interface. This results in these hosts being able to view the “Virtual Team” stream while at the same time blocking all other groups including the “Executive Meeting” stream.
  - It should be obvious that attempts to perform wholesale, Multicast Group Access Control using this technique does not scale.



- **Controlling Source Registration**

  In some cases, it may be desirable to control which hosts in the network can actually source traffic to a group. While there is currently no way to prevent a bogus source from transmitting traffic on its local segment, we can prevent it from being registered to the RP. This will, in most cases, prevent this traffic from going past the first-hop router and reaching other hosts in the network.

  A new IOS command, 'ip pim accept-register' was introduced which when configured on an RP, controls which (S, G) Register messages will be accepted and which will be rejected.

- **Global Command (IOS 12.0(6) or later)**

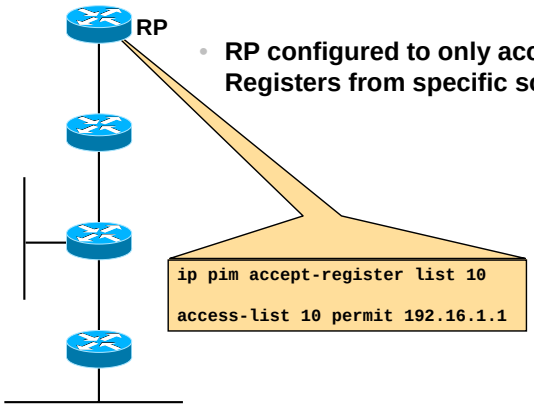
  `ip pim accept-register [list <acl>] | [route-map <map>]`

  – If the "list <acl>" is specified, the <acl> can either be a simple access list to control which hosts may send to any groups or an extended access list that specifies both source and group address combinations that are permitted or denied from sending.

  – If the "route-map <map>" is specified, then only matching (S, G) traffic will be accepted. (Note: This permits other matching criteria to be considered such as AS-PATH.)

Controlling Source Registration

Cisco.com



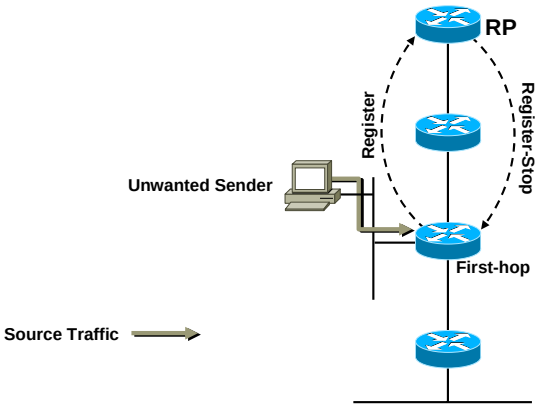
RST-2T07—Module 6v1
9744_05_2004_X

40

- **Accept-Register Operational Example**
  - In this example, the RP has been configured with the ‘ip pim accept-register’ command with an associated ACL whose goal is to disable all multicast sources in the network except for host 192.16.1.1. (Presumably the only authorized source of IP multicast traffic.)

Controlling Source Registration

Cisco.com



- **Unwanted source traffic hits first-hop router.**
- **First-hop router creates (S,G) state and sends Register.**
- **RP rejects Register, sends back a Register-Stop.**

RST-2T07—Module 6v1
9744_05_2004_X

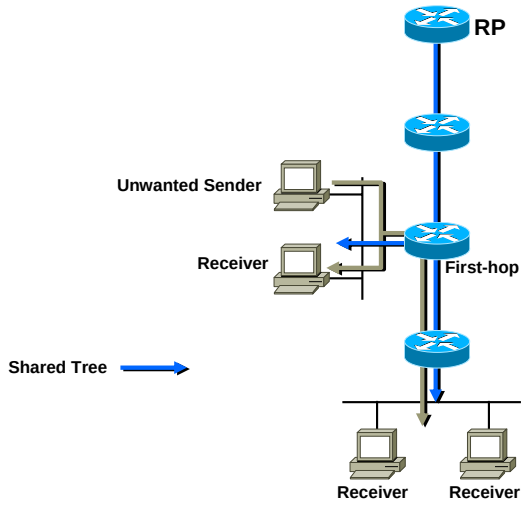
41

- **Accept-Register Operational Example**
  - When some unwanted/unauthorized source begins transmitting multicast traffic, the First-Hop router encapsulates the packets into PIM Register messages and transmits them to the RP.
  - Because this is an unwanted/unauthorized source that is “denied” by the accept-register ACL, the RP discards the PIM Register message and sends back an immediate “PIM Register-Stop” message to shutoff the flow of PIM Registers for this source.

Controlling Source Registration

Cisco.com

Weaknesses in 'accept-register' usage.

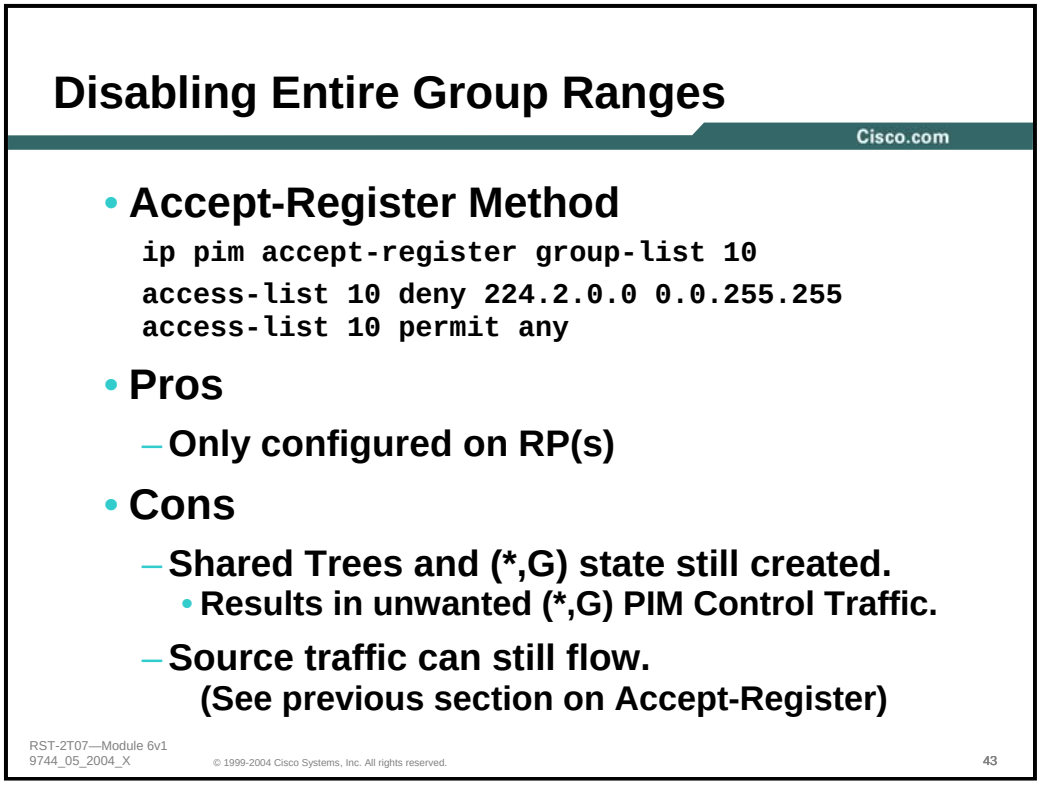


- **Traffic will flow on local subnet where source resides.**
- **Traffic will flow from first-hop router down any branches of the Shared Tree.**
  - **Results when (*,G) OIL is copied to (S,G) OIL at first-hop router.**
  - **Causes (S,G) traffic to flow down all interfaces in (*,G) OIL of first-hop router.**
  - **Fundamental limitation of PIM protocol.**

RST-2T07—Module 6v1
9744_05_2004_X

42

- **Weaknesses in the Accept-Register mechanism.**
  - While the 'ip pim accept-register' mechanism is a step in the right direction for multicast source control, it is not perfect and has some limitations due to the very nature of the way the PIM protocol operates.
  - The first issue is that the mechanism can do nothing to stop the flow of unwanted multicast traffic on the local subnet where the unwanted source resides.
  - The second (and more subtle) limitation is that unwanted source traffic ***can still flow*** down existing branches of the Shared Tree even though the RP has sent back a Register-Stop message to the First-Hop router. This is due to the following sequence of events in the normal PIM-SM forwarding mechanism:
    - As soon as the First-Hop router detects the arrival of the traffic from the directly connected unwanted source, it must create (S,G) state in order to perform the PIM Register function.
    - In the PIM forwarding model, whenever a new (S,G) mroute entry is created it "inherits" a copy of the Outgoing Interface List (OIL) from it's parent (*,G) entry. The OIL of the parent (*,G) entry contains a list of any active branches of the Shared Tree. This results in the initial OIL of the (S,G) containing a non-empty set of interfaces which point down the existing Shared Tree.
    - In addition to encapsulating and forwarding the initial unwanted source packets to the RP in PIM Register messages, unwanted source traffic begins flowing down the Shared Tree as a result of the newly created (S,G) entry and its "inherited" copy of the (*,G) OIL.
    - By the time the "Register-Stop" is received by the First-Hop router, the damage is already done as unwanted (S,G) traffic is already flowing down the Shared Tree which in turn, can create even more unwanted (S,G) state downstream if normal SPT-Switchover is in effect.



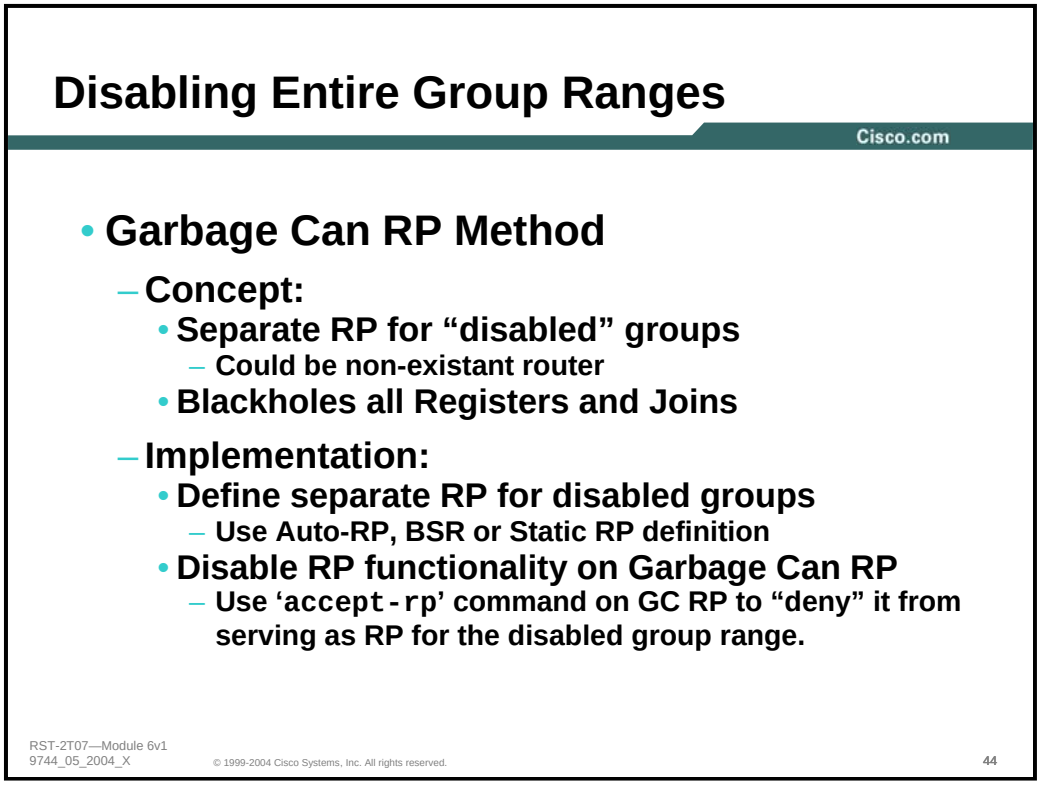
- **Disabling Entire Group Ranges – Accept Register Method**

  This method relies on the 'ip pim accept-register' command to completely disable the group range. This is accomplished by carefully crafting the ACL in the 'accept-register' command so that unwanted group ranges are denied.

  – Pros
    - This method is quite easy to implement as it only requires the RP(s) to configured.

  – Cons

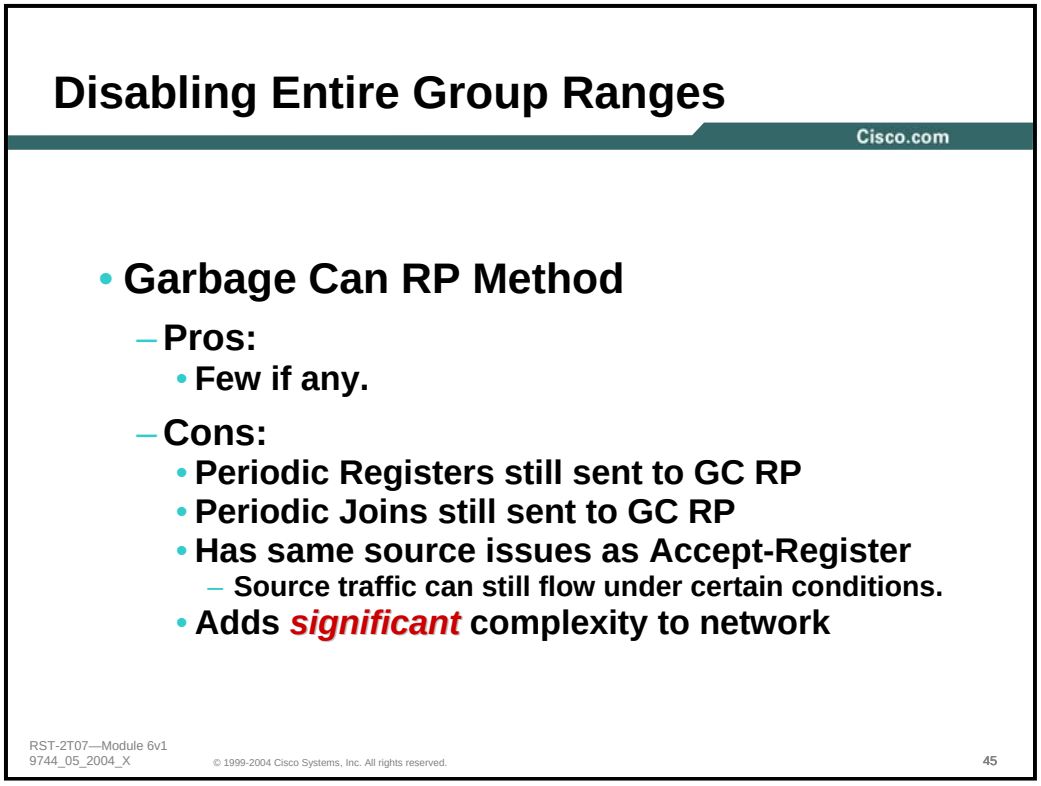
  This method suffers from all the issues previously discussed in the previous section on the Accept Register command. However, this solution may be sufficient for some cases where *absolute* control over disabled groups is not necessary. For example, if the desire is to simply "encourage" users to register their applications with IT for better muticast tracking and planning purposes, then this method may sufficiently "discourage" multicast to unauthorized or unregistered groups such that users would contact IT to register their applications.
    - Shared Trees can still be created for the unwanted multicast groups. This results in unwanted PIM control traffic in the network.
    - Source traffic can still flow to portions of the network as previously described in the section on the Accept Register.



- **Disabling Entire Group Ranges – Garbage-Can RP Method**

  This method relies on the definition of Garbage Can RP to completely disable the group range. The Garbage Can RP would “blackhole” all PIM Joins or Registers for the disabled groups.

  This is accomplished by using an RP address of either a non-existent router or a router that has been disabled from performing the Garbage Can RP function. This can be accomplished by using the ‘accept-rp’ command on the Garbage Can RP.



- **Disabling Entire Group Ranges – Garbage-Can RP Method**
  - Pros
    - Little if any.
  - Cons
    - PIM control traffic are still sent to the Garbage Can RP. This includes PIM Registers and (*,G) Joins.
    - (*,G) Joins for the unwanted groups will still create state on all the routers in the network which will, in turn, create a Shared Tree all the way to the Garbage Can RP. This results in the same issues as the Accept Register method as sources that exist along this unwanted Shared Tree will still forward traffic down the branches of the Shared Tree.
    - Given the above two items, the operational complexity of the network is significantly increased.

Disabling Entire Group Ranges

Cisco.com

- **Local Loopback RP Method**
  - **Concept:**
    - **Only Auto-RP-learned groups are authorized.**
    - **All other groups are considered *unauthorized.***
  - **Implementation:**
    - **Define local Loopback as RP for unauthorized groups on each router.**

```
ip pim rp-address <local_loopback> 10
access-list 10 permit 224.2.0.0 0.0.255.255
```

Note: The permit clause defines the unauthorized group.

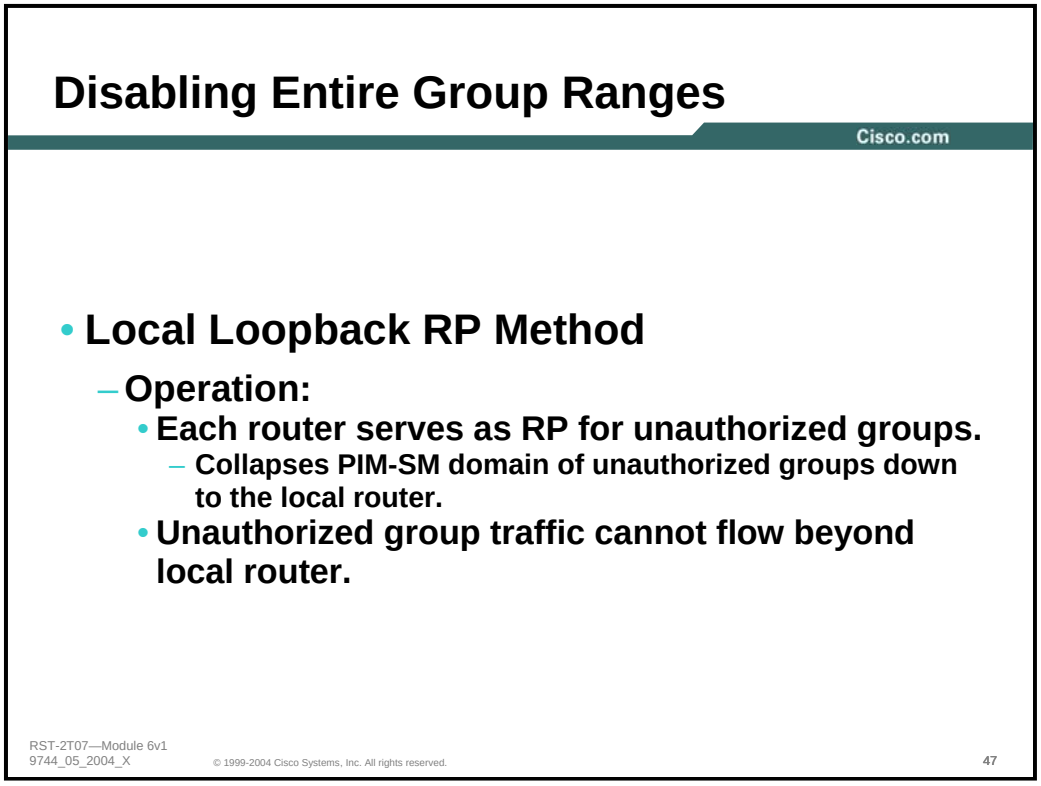
RST-2T07—Module 6v1
9744_05_2004_X

46

- **Disabling Entire Group Ranges – Local Loopback Method**

  This method relies on Auto-RP to distribute the list of RPs and group ranges that are authorized. Any group range that is not covered by an Auto-RP learned announcement is unauthorized and no Shared Tree is built for these groups.

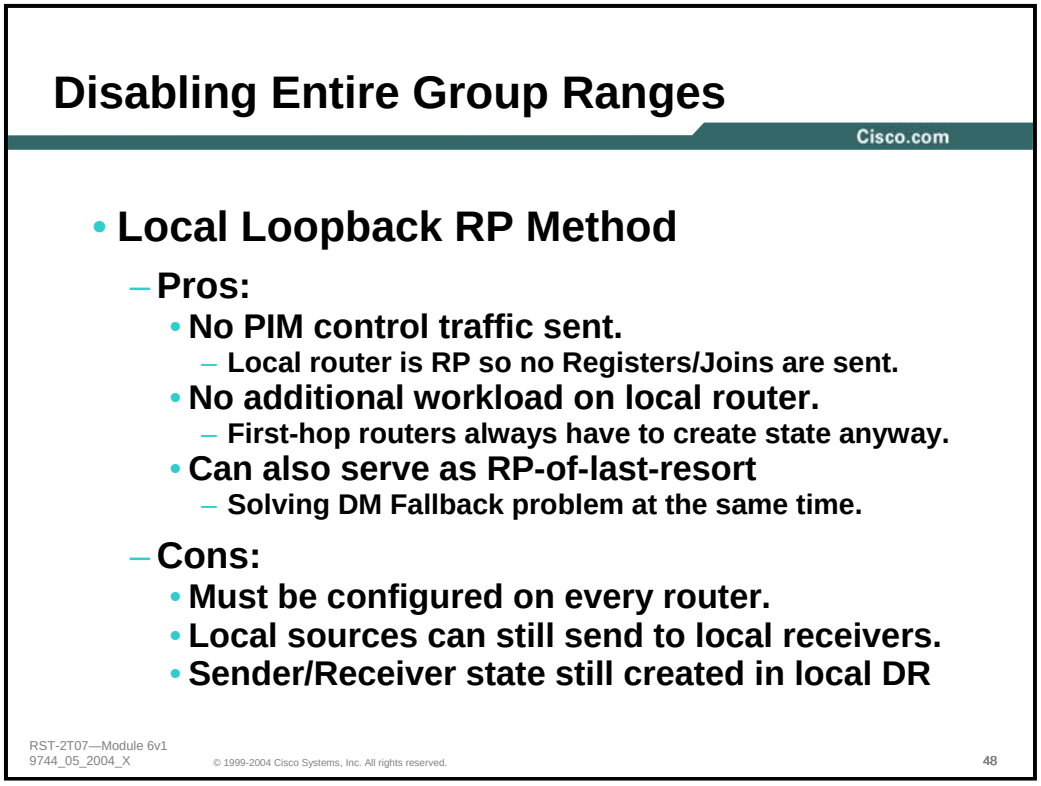
  In order to accomplish the above, the local Loopback address is configured as a static RP-of-last-resort on each router that covers the unauthorized groups.



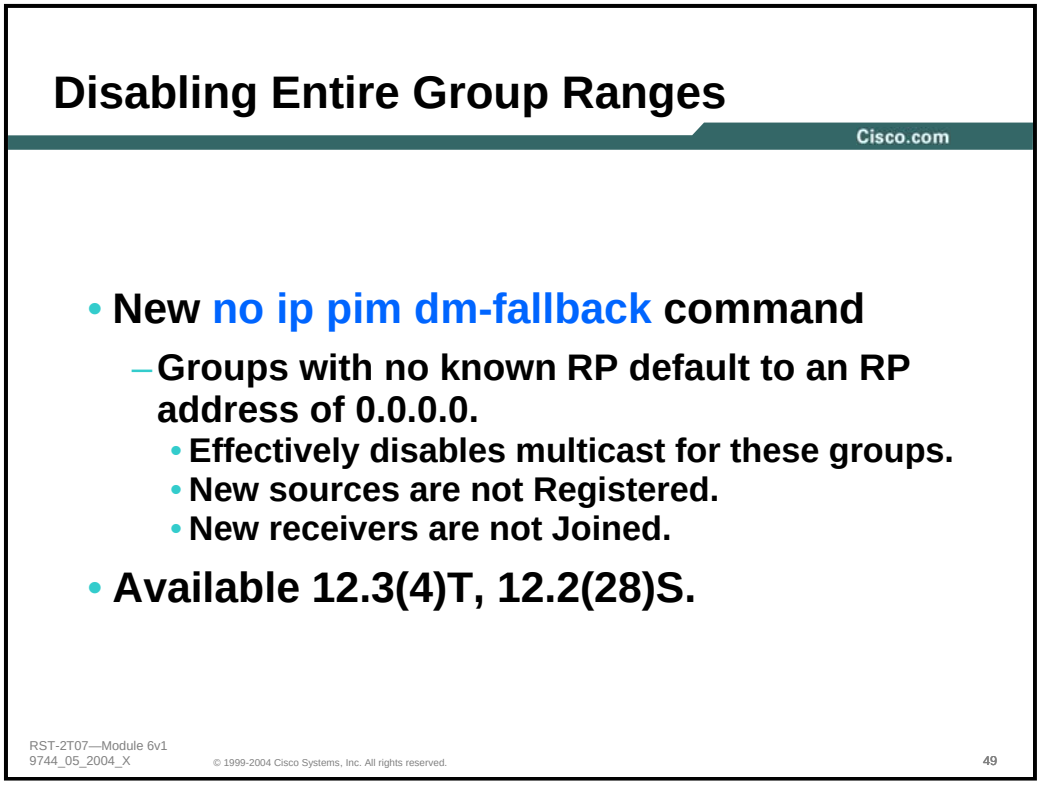
- **Disabling Entire Group Ranges – Local Loopback Method**
  - Operation

    The net affect is that each router in the network will assume that ***it*** is the RP for the unauthorized groups. As a result, the routers in the network will not send PIM Registers or (*,G) Joins for these groups since they all think that they are the RP. The Shared Tree for these groups actually becomes fragmented into a separate Shared Tree inside of each router. This effectively prevents sources and receivers from learning about each other and in turn, prevents multicast traffic in the unauthorized group range from flowing across the network.

    NOTE: This method is not perfect as multicast traffic can still flow between sources and receivers connected to a common router. However, this solution may be sufficient for cases where *absolute* control over disabled groups is not necessary. For example, if the desire is to simply "encourage" users to register their applications with IT for better multicast tracking and planning purposes, then this method may sufficiently "discourage" multicast to unauthorized or unregistered groups such that users would contact IT to register their applications.



- **Disabling Entire Group Ranges – Local Loopback Method**
  - Pros
    - No PIM control traffic is sent for the unauthorized groups. This is because each router thinks that it is the RP for these groups.
    - There is no additional workload on the routers as the First-Hop router must create (S,G) state for any directly connected sources anyway.
    - This technique can be combined with the RP-of-last-resort to prevent DM Fallback. This is accomplished by specifying a group range that includes all multicast addresses ***except*** the two Auto-RP groups 224.0.1.39 and 224.0.1.40.
  - Cons
    - This requires the RP-of-last-resort to be configured on every router.
    - Local sources can still send to local receivers. Therefore the method is not a perfect solution to group control.



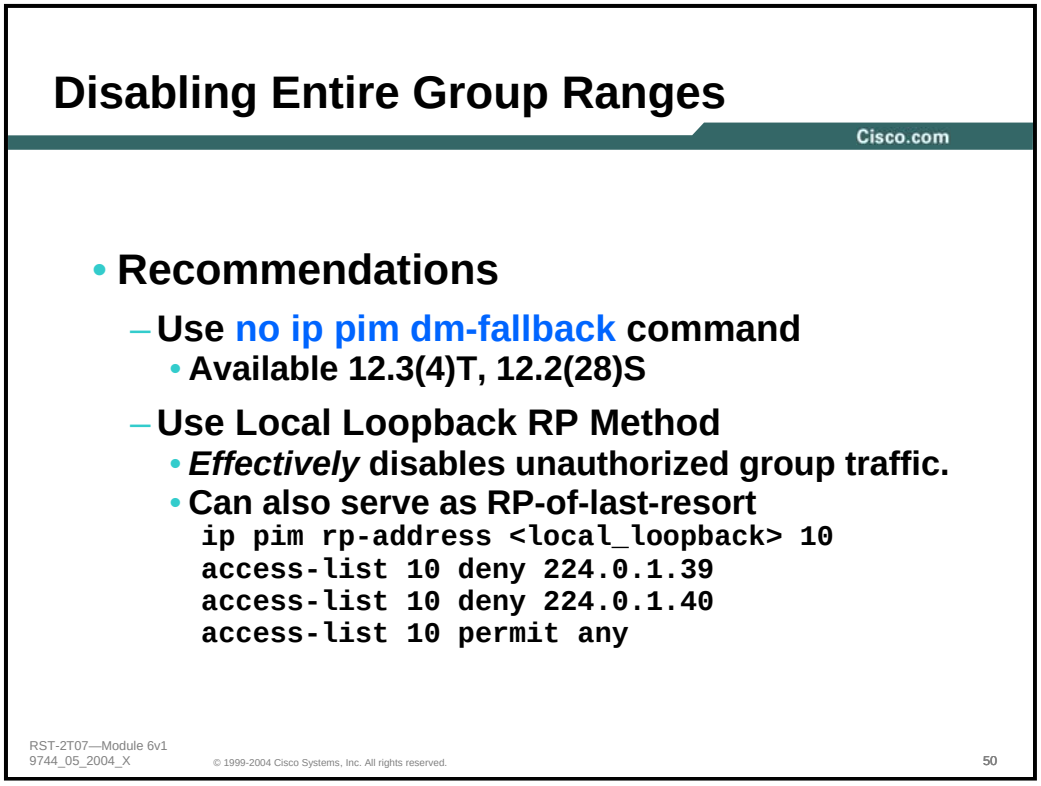
- **Disabling Entire Group Ranges – Future**

  While the previous solutions are sufficient to cover the requirements for many networks, it is obvious that it is not a perfect solution. A new IOS command

  ```
  no ip pim dm-fallback
  ```

  (previously discussed) can be used to disable unwanted groups.

  This command results in a default RP address of 0.0.0.0 which is a non-existent RP. If no Auto-RP or BSR RP information is learned, the router will default to using this RP address. An RP address of 0.0.0.0 [non-existent] prevents a Shared Tree from being built. This prevents multicast traffic from flowing for groups that have no Auto-RP or BSR RP definition.



- **Disabling Entire Group Ranges – Recommendation**

  The “Local Loopback” method is recommended when it is desired to disable a group range. While not completely fool-proof, it is the best method to date for *effectively* disabling groups.

  When combined with the RP-of-last-resort, this method can also prevent Dense mode Fallback via the configuration shown below.

```
ip pim rp-address <local_loopback> 10
access-list 10 deny 224.0.1.39
access-list 10 deny 224.0.1.40
access-list 10 permit any
```

  Furthermore, if the group-range of the RP-of-last-resort covers all groups except the Auto-RP groups, it becomes easy to administer which groups are authorized. All that is necessary to authorize a new group range is to (re)define a Candidate RP to advertise the new (extended) group-range. This will result in new Auto-RP learned information being distributed to all routers that defines an RP for the newly authorized group. This in turn, will override the RP-of-last-resort for the group range; thereby enabling a Shared Tree to be built for the group range.

Preventing RP-Spoofing DoS Attacks

Cisco.com

- **Global command**

  `ip pim rp-announce-filter rp-list <acl> [group-list <acl>]`

  `rp-list <acl>`
  - Specifies from which routers C-RP Announcements are accepted.

  `group-list <acl>`
  - Specifies which groups in the C-RP Announcement are accepted.
  - If not specified, defaults to deny all groups

- **Use on Mapping Agents to filter out bogus C-RP's**
  - Some protection from RP-Spoofing denial-of-service attacks
  - Multiple commands may be configured as needed

- **Filtering RP Announcements**

  Network Administrators may wish to configure Mapping Agents so that they will only accept C-RP Announcements from well-known routers in the network. This will prevent C-RP Announcements from bogus routers from being accepted and potentially being selected as the RP.

- **Global Command**

  `ip pim rp-announce-filter rp-list <acl> [group-list <acl>]`

  – The rp-list <acl> specifies the IP address(es) from which C-RP announcements will be accepted.

  – The option group-list <acl> specifies the group range(s) that are acceptable for the routers in the rp-list. If not specified, the default group-list <acl> is

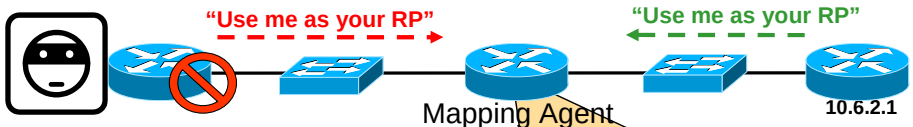
  deny all

  – Multiple instances of this command may be configured.

Preventing RP-Spoofing DoS Attacks

Cisco.com

Use ip pim rp-announce-filter on RP

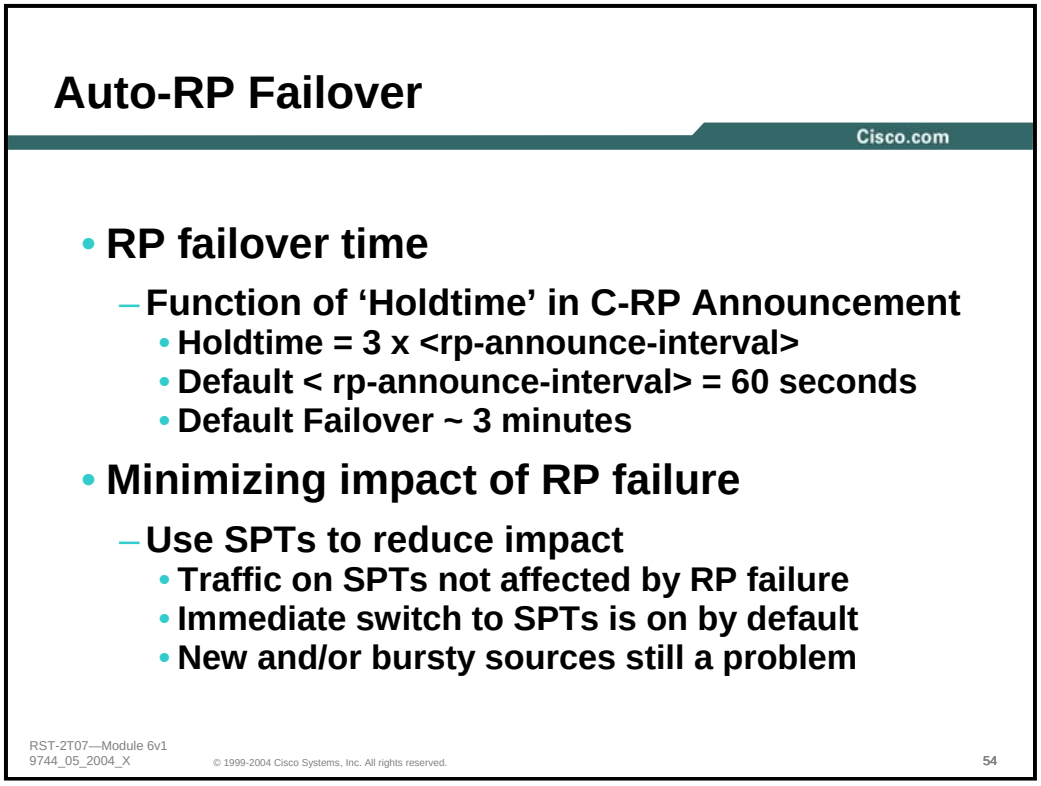



```
ip pim rp-announce-filter rp-list 11 group-list 12
access-list 11 permit 10.6.2.1                        !IP address of Permitted RP
access-list 12 permit 239.192.240.0 0.0.3.255         !Permit MoH
access-list 12 permit 239.192.244.0 0.0.3.255         !Permit Low Stream
access-list 12 permit 239.192.248.0 0.0.3.255         !Permit Medium Stream
access-list 12 permit 239.255.0.0 0.0.255.255         !Permit High Stream
access-list 12 deny 239.0.0.0 0.255.255.255           !Deny remaining Admin. Scoped range
access-list 12 permit 224.0.0.0 15.255.255.255        !Permit Link Local/Reserved Addr.
```

- **Using this configuration allows the router to accept RP announcements from only the RP in 'access-list 11' for group ranges described in 'access-list 12'**

- **Preventing RP-Spoofing DoS Attacks**
  - The "ip pim rp-announce-filter" command is used on Mapping Agents. The command filters Auto-RP announcement messages coming from the RP. Using this command, you can prevent maliciously configured routers from acting as a candidate RP.
  - By filtering which RPs and groups are authorized on the network, The MA will control what traffic non-RPs will be allowed to receive.

High Availability



- **Auto-RP Failover**
  - RP Failover
    - In the case of Auto-RP, the RP failover time is a direct function of the "Holdtime" value advertised in the Candidate RP Announcement. The value of the Holdtime is fixed at 3x the <rp-announce-interval> which has a default value of 60 seconds. This results in a default Auto-RP failover time of approximately 3 minutes.
  - Minimizing impact of RP Failure
    - Failure of the RP generally has no impact on existing Shortest-Path Trees. Even if the RP is on the SPT path, the routers in the network will try to rebuild the SPT around the failed router and restore the SPT. Since the default behavior of Cisco's PIM-SM implementation is to immediate cut-over to SPTs, the impact of an RP failure is generally limited to the inability of new sources and receivers to learn of each other and build new SPTs. In some networks, this is not an issue as long as the RP failure is transient and another RP takes over within a short period of time.

Tuning Auto-RP Failover

Cisco.com

- Tune Candidate RPs
- Use 'interval' clause to control failover times

```
ip pim send-rp-announce <intfc> scope <ttl>
                        [group-list acl]
                        [interval <seconds>]
```

- Allows rp-announce-interval to be adjusted
- Smaller intervals = Faster RP failover
- Smaller intervals increase amount Auto-RP traffic
  - Increase is usually insignificant
- Total RP failover time reduced
  - Min. failover ~ 3 seconds
- *Consider using Anycast RP for faster failover*

RST-2T07—Module 6v1
9744_05_2004_X

55

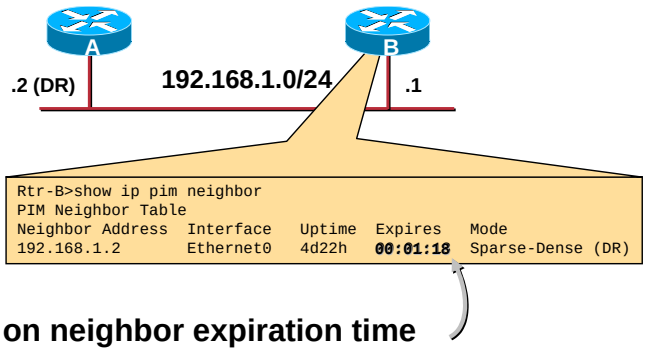
- **Tuning Auto-RP Failover**

  Auto-RP's failover time is dictated by the value of the "holdtime" in the Candidate RP Announcement being sent by the currently elected RP. This holdtime is 3x the <rp-announce-interval> which can be adjusted via the optional **interval** keyword in the **ip pim send-rp-announce** global command.

  The default value of the <rp-announce-interval> is 60 seconds. Tuning the **interval** value to values under 60 seconds can reduce the Auto-RP failover times. If the **interval** is set to 1 second, the theoretical worst case failover is approximately 3 seconds. However, this improvement in failover time comes at the expense of increased Auto-RP Announcement traffic.

DR Failover

Cisco.com



- **Depends on neighbor expiration time**
- **Expiration Time sent in PIM query messages**
  - **Expiration time = 3 x <query-interval>**
  - **Default <query-interval> = 30 seconds**
  - **DR Failover ~ 90 seconds (worst case) by default**

RST-2T07—Module 6v1
9744_05_2004_X

56

- **Designated Router (DR) Failover**
  - The DR failover mechanism is based the PIM Hello “Expiration” time that is sent with each PIM Hello message. This “Expiration” value is 3x the PIM Hello <query-interval> which has a default value of 30 seconds. This results in an advertised Expiration timer of 90 seconds (1:30).
  - Consider the example shown above. If PIM Router B fails and PIM Neighbor router A does not hear another PIM Hello message from B within the advertised “Expiration” time, then router A will assume router B has failed and will initiate a new DR election. Given the default <query-interval> values of 30 seconds, the worst case scenario of DR failover is approximately 90 seconds.

Tuning DR Failover

Cisco.com

- **Tune PIM query interval**
  - **Use interface configuration command**
    `ip pim query-interval <period> [msec]`
    - **Default <period> = seconds**
    - **"msec" keyword available beginning with 12.1(11b)E**
  - **Permits DR failover to be adjusted**
    - **Min. DR failover ~ 3 seconds (worst case)**
    - **Smaller intervals increase PIM query traffic**
      - **Increase is usually insignificant**

RST-2T07—Module 6v1
9744_05_2004_X

57

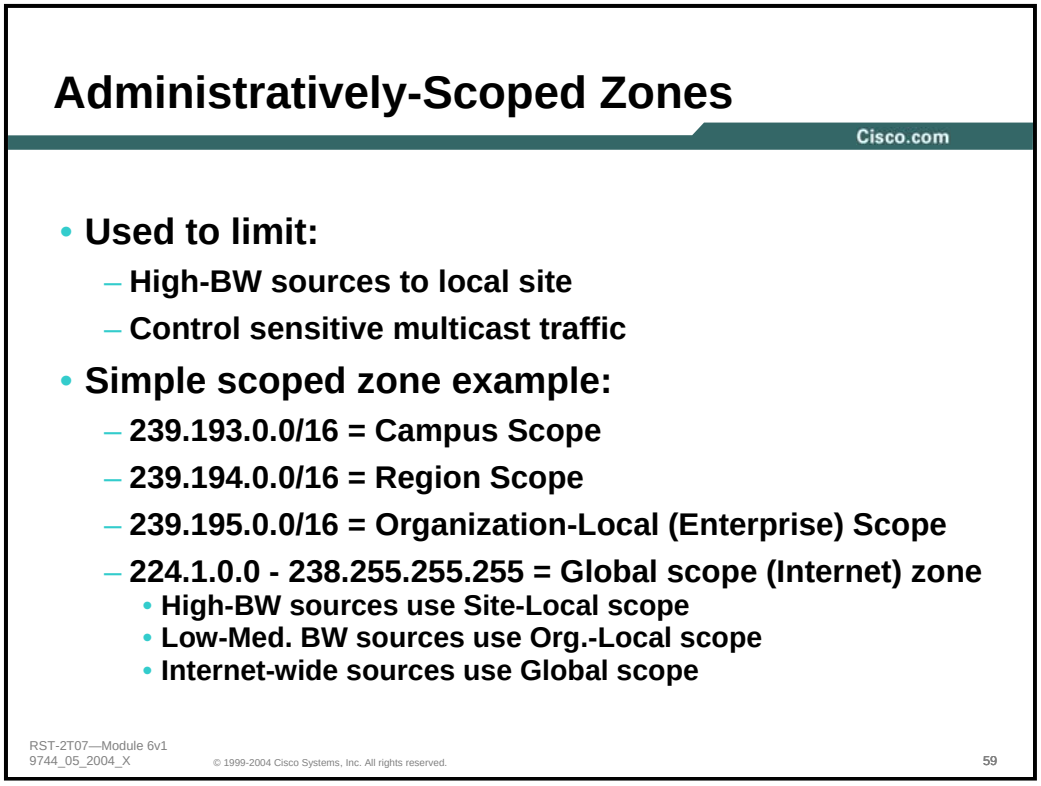
- **Tuning DR Failover**
  - The PIM <query-interval> value may be modified on an interface basis via the

    ```
    ip pim query-interval <period> [msec]
    ```

    IOS interface command. For example, if the <query-interval> period is set to 1 second, the worst case DR failover would be approximately 3 seconds. This improvement in DR failover comes at the expense of increased PIM Hello traffic on the local subnet. However, this is often insignificant when compared to the benefits and the amount of data traffic flowing on the subnet.

  - Beginning with IOS release 12.1(11b)E, a new **msec** keyword was added to the syntax of this command. This permits <query-interval> values to be set as low as 1 millisecond which in turn, leads to sub-second DR failovers.

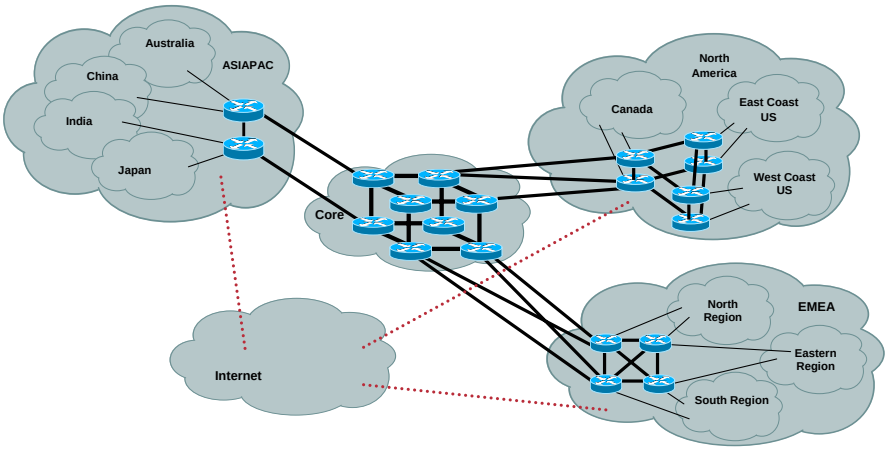
Using Admin. Scoped Zones



- **Administratively-Scoped Zones**
  - As your network makes more and more usage of IP Multicast there will come a time when you will wish to limit the “scope” of particular multicast flows. This can be as a result of:
    - The need to limit high-rate multicast flows to a local site where bandwidth is plentiful and to avoid congesting slower WAN links.
    - The need to control sensitive multicast traffic from leaving a particular area such as a building or from leaking into the Internet.
  - In the following pages an example of a three-tiered scoped zone example will be presented. These scopes are as follows:
    - The Campus Scope in the range of 239.193/16. This scope is used to prevent multicast traffic in this group range from leaving the local Campus.
    - The Region Scope in the range of 239.194/16. This scope is used to prevent multicast traffic in this group range from leaving a particular Region of the network (i.e. Americas Region, EMEA Region, etc.).
    - The Enterprise Scope (sometimes called the Organization-Local scope) in the range of 239.195/16. This scope is used to prevent multicast traffic in this group range from leaving the Enterprise network and leaking into the Internet or other adjacent PIM domains.

Administratively-Scoped Zones Example

Cisco.com



RST-2T07—Module 6v1
9744_05_2004_X

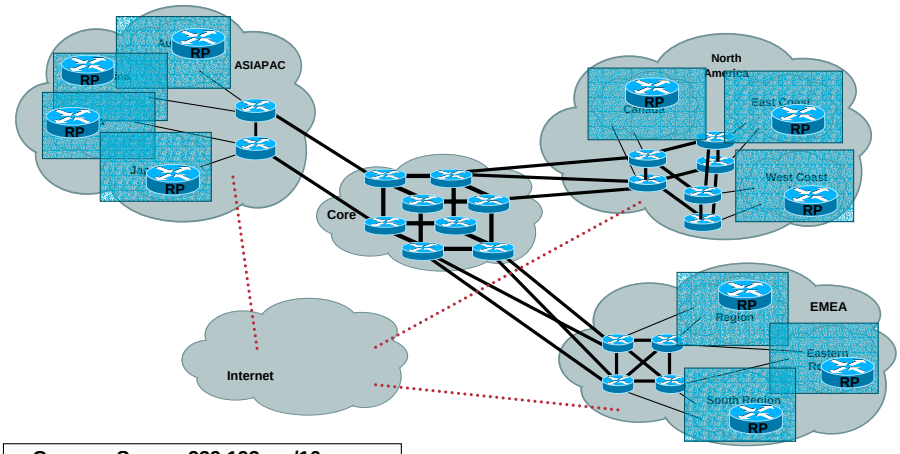
60

- **Administratively-Scoped Zone Example**
  - This diagram shows our example global network as well as the connections to the Internet.

Administratively-Scoped Zones Example

Cisco.com

Level1: Campus Scope



- **Campus Scope: 239.193.x.x/16**
- **RP per Campus**

RST-2T07—Module 6v1
9744_05_2004_X

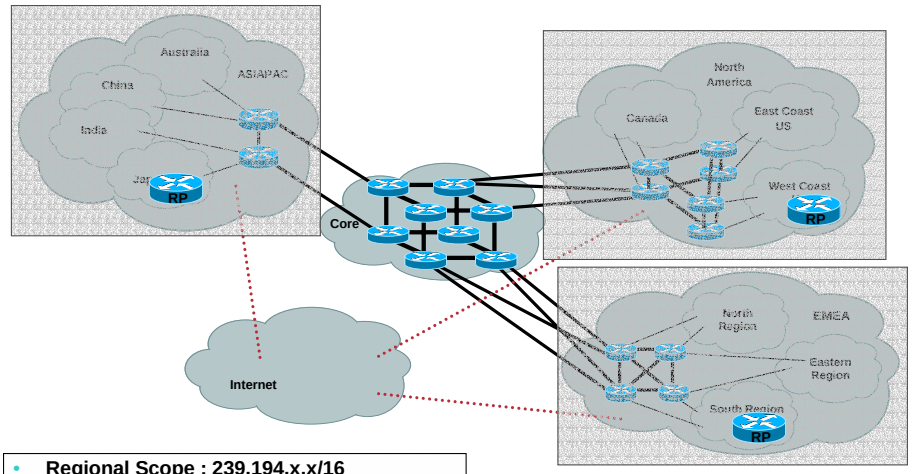
61

- **Administratively-Scoped Zone Example**
  - This diagram highlights the Campus scopes of our example network.
  - Notice that each Campus functions as an independent PIM-SM domain for the Campus scope range. Therefore, each Campus must have its own RP.

Administratively-Scoped Zones Example

Cisco.com

Level2: Regional Scope



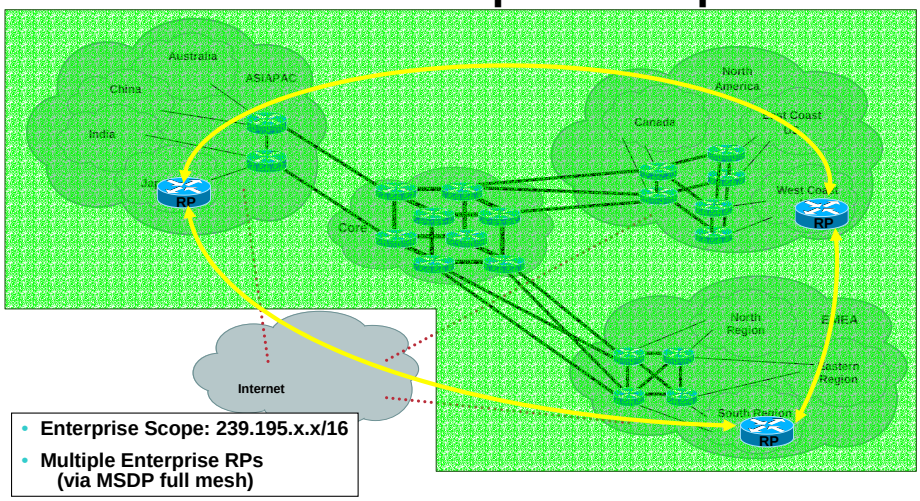
- **Regional Scope : 239.194.x.x/16**
- **RP per Region**

- **Administratively-Scoped Zone Example**
  - This diagram highlights the Region scopes of our example network.
  - Notice that each Region also functions as an independent PIM-SM domain for the Region scope range. Therefore, each Region must have its own RP.

Administratively-Scoped Zones Example

Cisco.com

Level3: Enterprise Scope

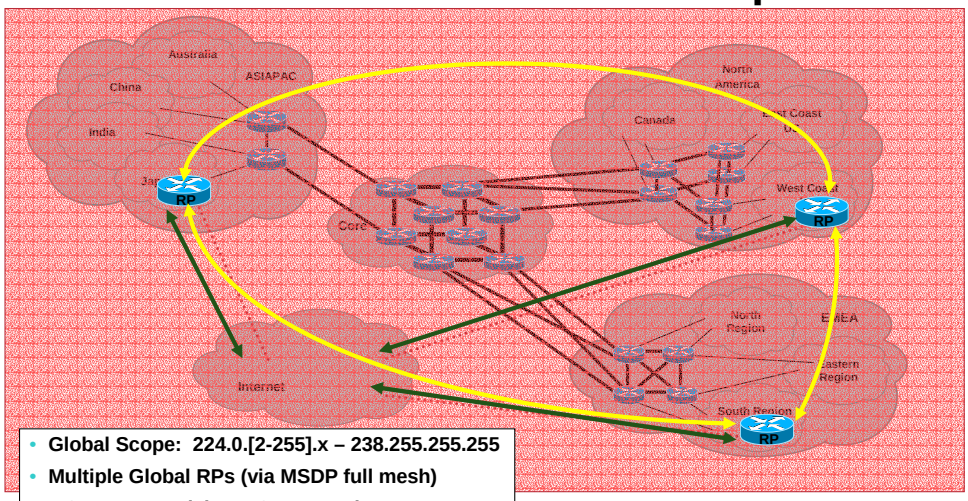


- **Administratively-Scoped Zone Example**
  - This diagram highlights the Enterprise scope of our example network.
  - In this case we have deployed multiple Anycast-RP's that serve as the RP's for the Enterprise scope.

Administratively-Scoped Zones Example

Cisco.com

Level 4: Internet Global Scope



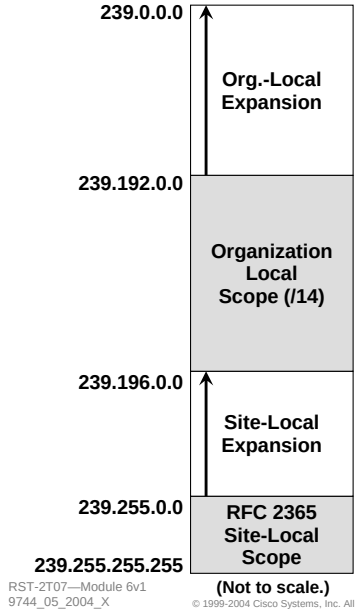
RST-2T07—Module 6v1
9744_05_2004_X

64

- **Administratively-Scoped Zone Example**
  - This diagram highlights the Internet Global scope of our example network.
  - In this case the multiple Anycast-RP's that serve as the RP's for the Enterprise scope also have connections to the Service Providers MSDP nodes to allow Interdomain multicast to and from the Internet.

Administratively Scoped Address Range

Cisco.com

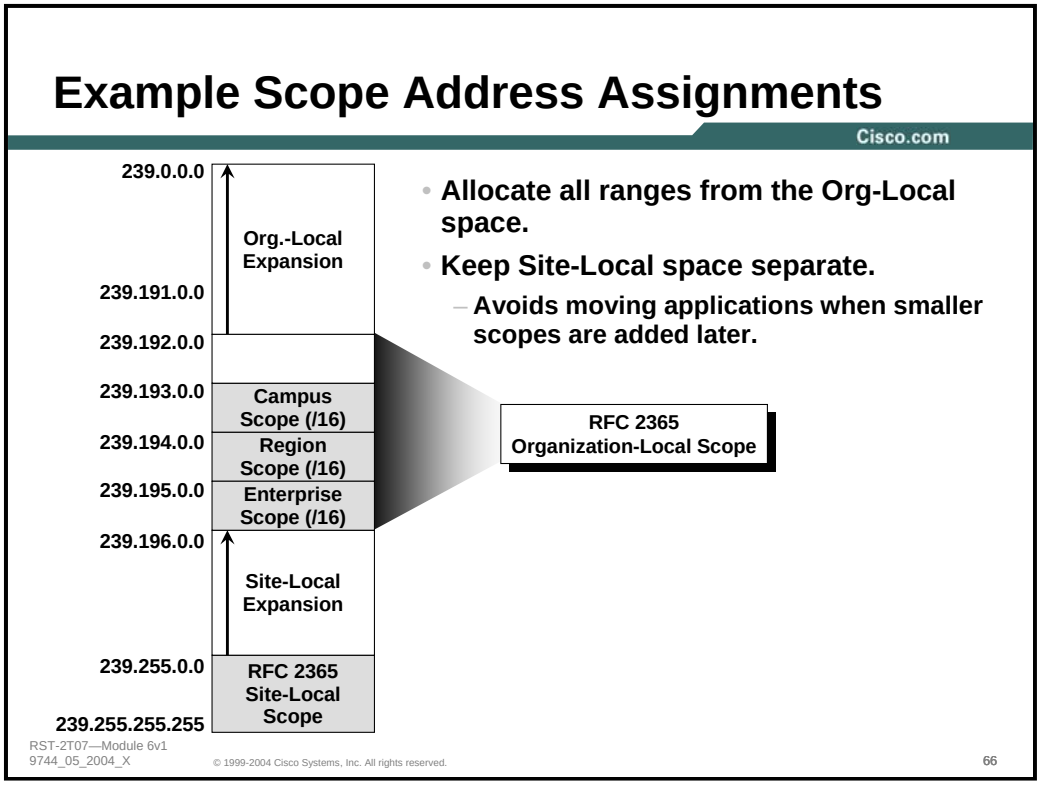


- **RFC 2365 Administratively Scoped Zones.**
  - **Organization-Local Scope (239.192/14)**
    - **Expands downward in address range.**
  - **Site-Local Scope (239.255/16)**
    - **Expands downward in address range.**
    - **Smallest possible scope.**
    - **Other scopes may be equal but not smaller.**

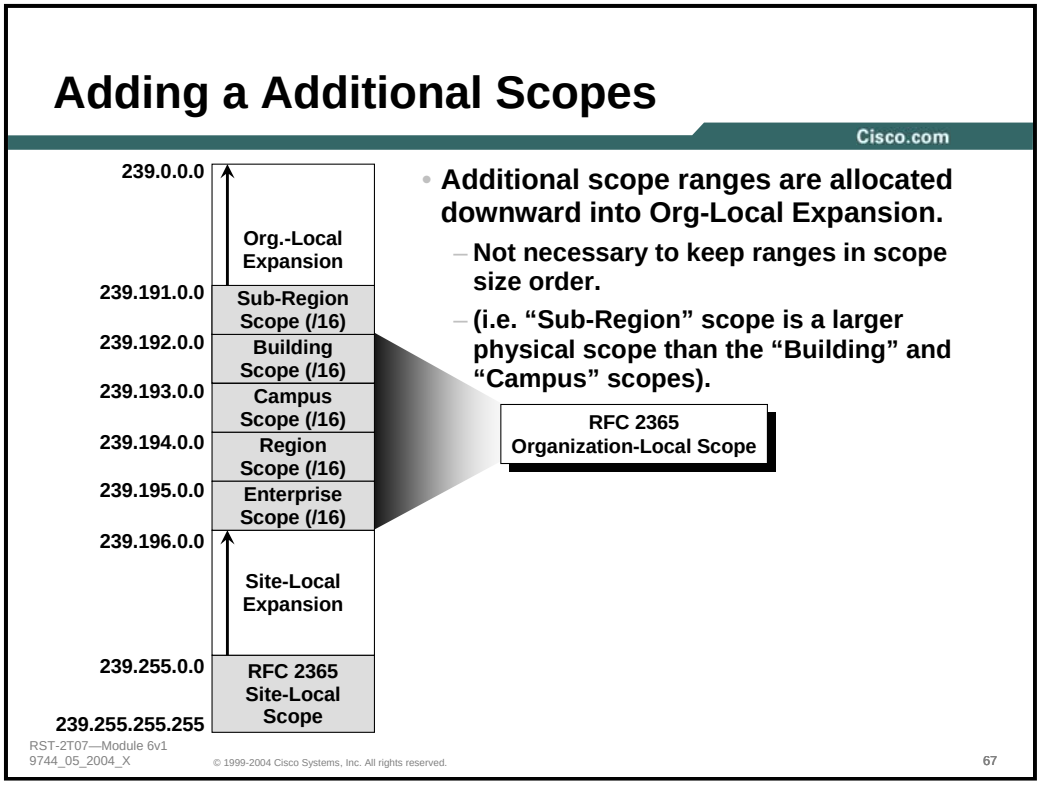
RST-2T07—Module 6v1
9744_05_2004_X

65

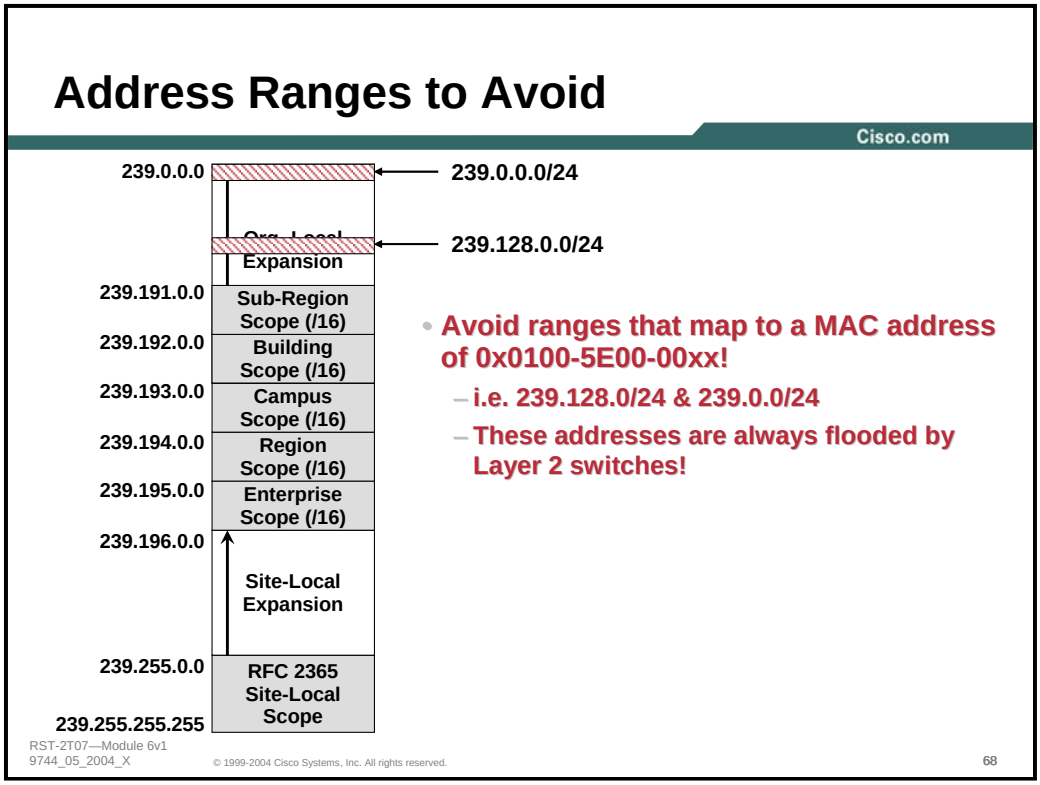
- **Administratively Scoped Address Range**
  - RFC 2365 *"Administratively Scoped Zones"* defines the most fundamental allocation of group ranges for Administratively Scoped Zone usage. Only two zones are predefined by RFC 2365. These are as follows:
  - Organization-Local Scope (239.192/14)
    - This range is the largest scope within an Enterprise network and is often referred to as the "Enterprise" scope.
    - In actual practice, this range may be subdivided into smaller scopes within the Enterprise.
    - The address space below this range may be used as expansion space should it become necessary to add additional scopes.
  - Site-Local Scope (239.255/16)
    - This is the smallest scope within an Enterprise network. NOTE: No other defined scope may be smaller in size than this scope.



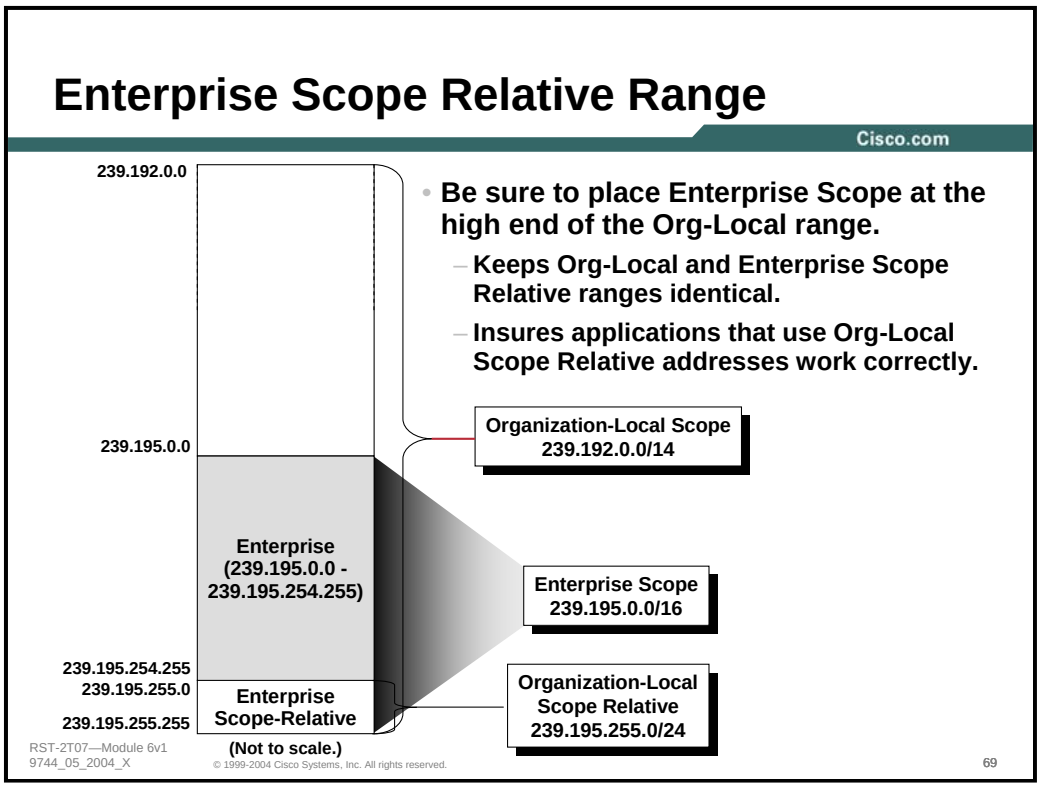
- **Example Scope Address Assignments**
  - Using our example network we allocate non-overlapping group address ranges for each of the Enterprise, Region and Campus scopes from the Organization-Local range as follows:
    - Enterprise Scope (239.195/16)
    - Region Scope (239.194/16)
    - Campus Scope 239.193/16)
  - Notice that the Site-Local range is not used for the Campus scope and is kept as a separate scope/group range. Instead, the Site-Local scope will be (at least initially) assigned the same physical boundaries as the Campus scope. The reason for this separation is that it allows for future definition of smaller scopes than Campus without violating the rule that the Site Local scope must be the smallest physical scope.



- **Adding Additional Scopes**
  - Address space for additional scopes are defined downward into the Organization-Local space and on into its expansion space. Note that is not necessary to maintain the address ranges in scope size order.
    - NOTE: The exception to this rule is that the Enterprise scope must be assigned to the top of the Organization-Local space to preserve Scope Relative addressing for the Organization-Local scope defined in RFC 2365.
  - In this example, new "Building" and "Sub-Region" scopes have been defined. While the physical size of the "Sub-Region" scope is larger than the "Building" and "Campus" scopes, it is not necessary to reorder the group range assignments. (This would require all applications to be reassigned to new addresses.)

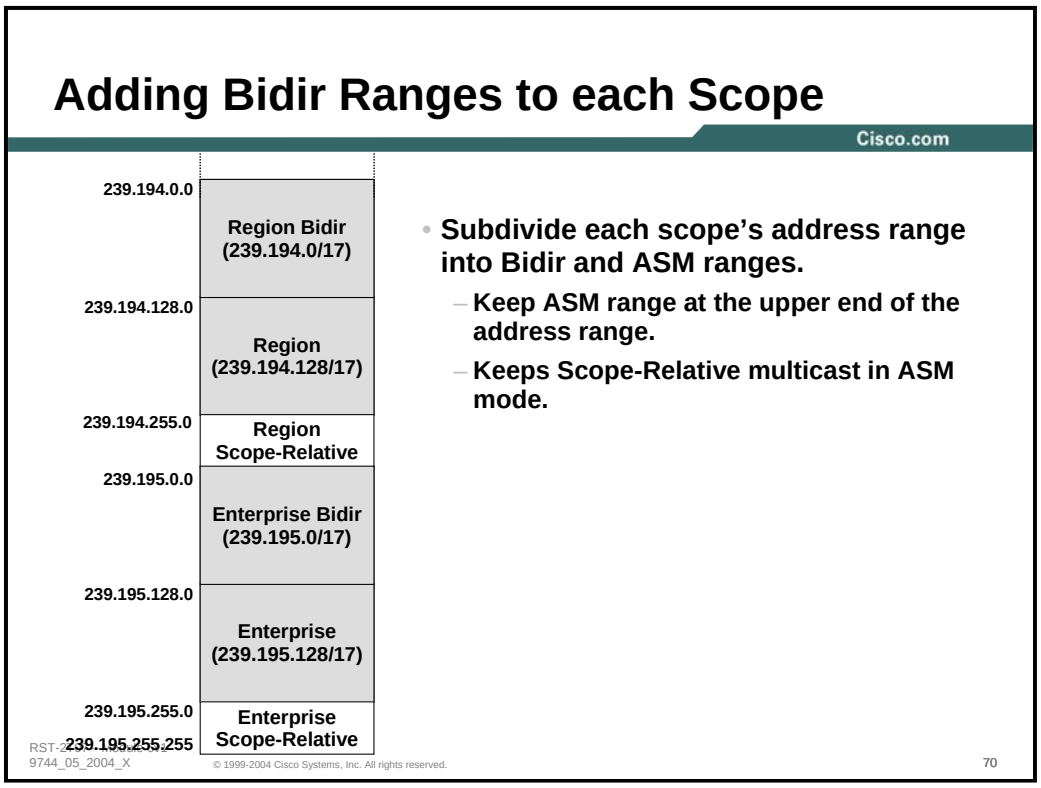


- **Address Ranges to Avoid**
  - As more address space is allocated for additional scopes, two ranges must be avoided. These are 239.128.0/24 and 239.0.0/24. This is because these multicast group ranges map to the 0x0100-5E00-00xx MAC address range which is always flooded by Layer 2 switches.

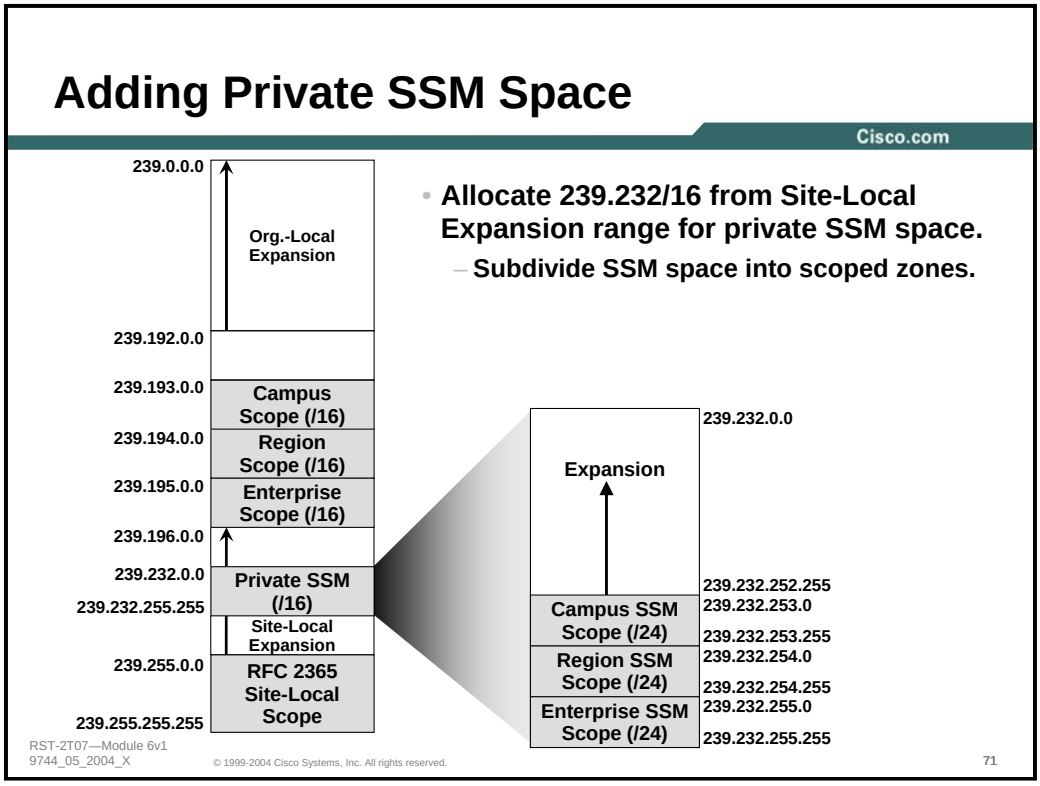


- **Enterprise Scope Relative Range**
  - In order to maintain proper Scope Relative addressing for the Organization-Local scope, it is necessary to allocate the address range for the Enterprise Scope (which is functionally equivalent with the Organization-Local scope) at the top of the Organization-Local address range.
  - In our network example, we have accomplished this by assigning the Enterprise scope to the group range of 239.195/16.

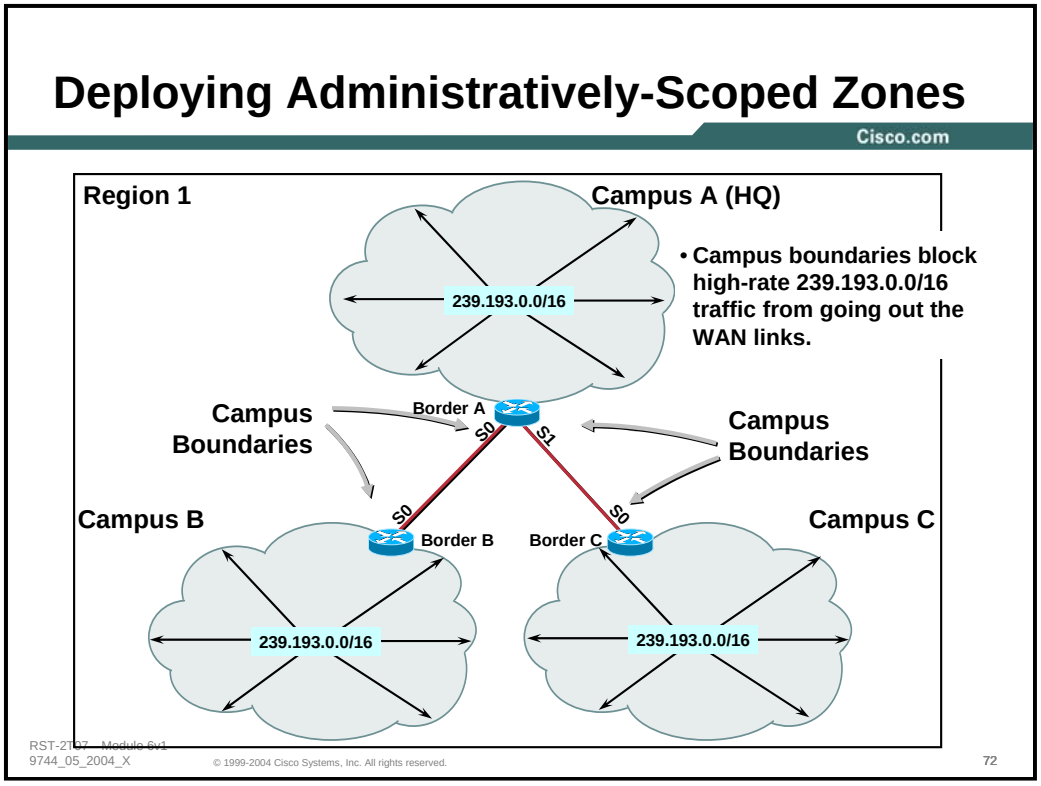
    Note: Scope Relative addressing uses the top 256 group addresses of each scope's multicast group range as shown in the example above.



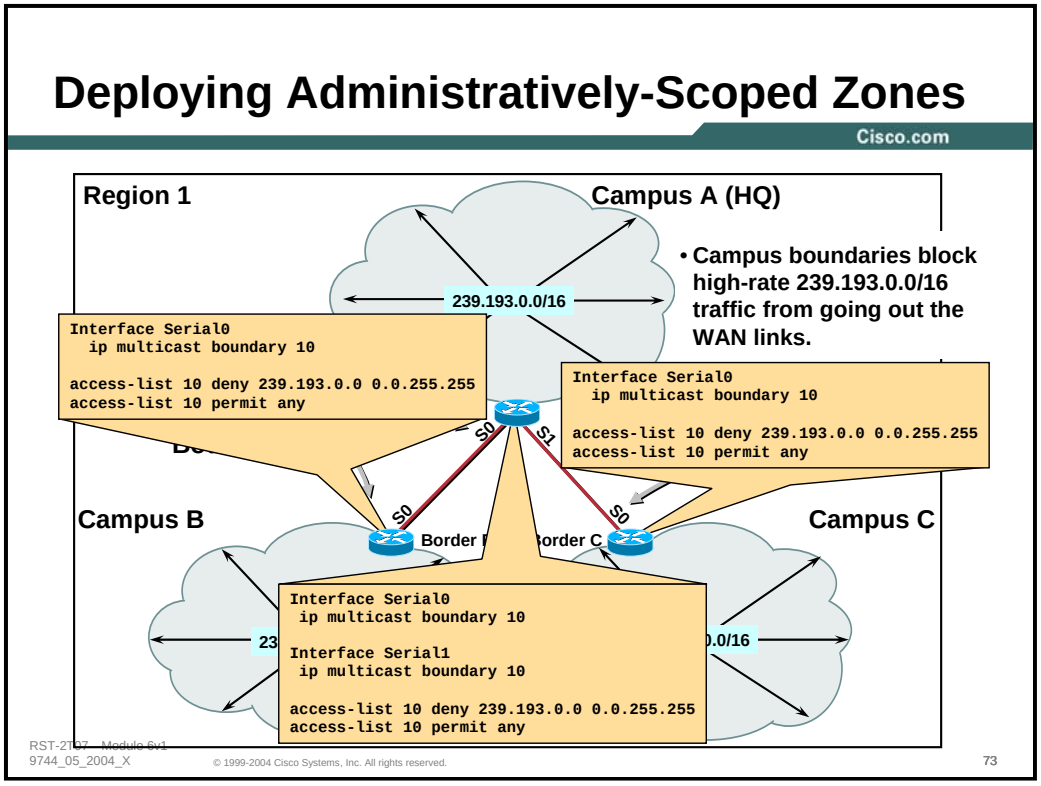
- **Adding Bidir PIM Ranges to each Scope.**
  - The previous examples assumed classic PIM-SM (aka Any-Source Muticast or ASM) operation. However, if we wish to also support Bidir PIM in our network it will be necessary to define address ranges for Bidir PIM. Furthermore, it may be the case that we wish to apply scoping to Bidir multicast as well.
  - In order to support Bidir PIM scoping in our example network we have subdivided each address range into a classic PIM-SM range and a Bidir PIM range as shown above. (Only Enterprise and Region ranges shown due to a lack of space.)
    - Note that the ASM range is kept at the top of the range while the Bidir range is below it. This maintains the important Scope Relative range in its normal position.



- **Adding Private SSM Space**
  - While IANA has allocated the 232/8 multicast address space for Source Specific Multicast (SSM), it can be argued that this space is intended for global Internet SSM operation. Having said that, it *is* possible to use the 232/8 address space inside of the Enterprise. However, it becomes increasingly difficult to keep track of which SSM ranges are intended for Enterprise and which are intended for Internet Global usage. If we furthermore desire to scope SSM within the Enterprise to smaller scopes such as Campus, Region, etc., this problem quickly gets out of hand. Therefore, we have allocated a private SSM space for use inside of the Enterprise at the 239.232/16 range. (This range was picked because the second octet quickly reminds us that this is SSM since it is the same as the first octet for global SSM.)
  - Extending our scoping model of Campus, Region and Enterprise, we subdivide the 239.232/16 space into smaller address ranges for SSM scopes.

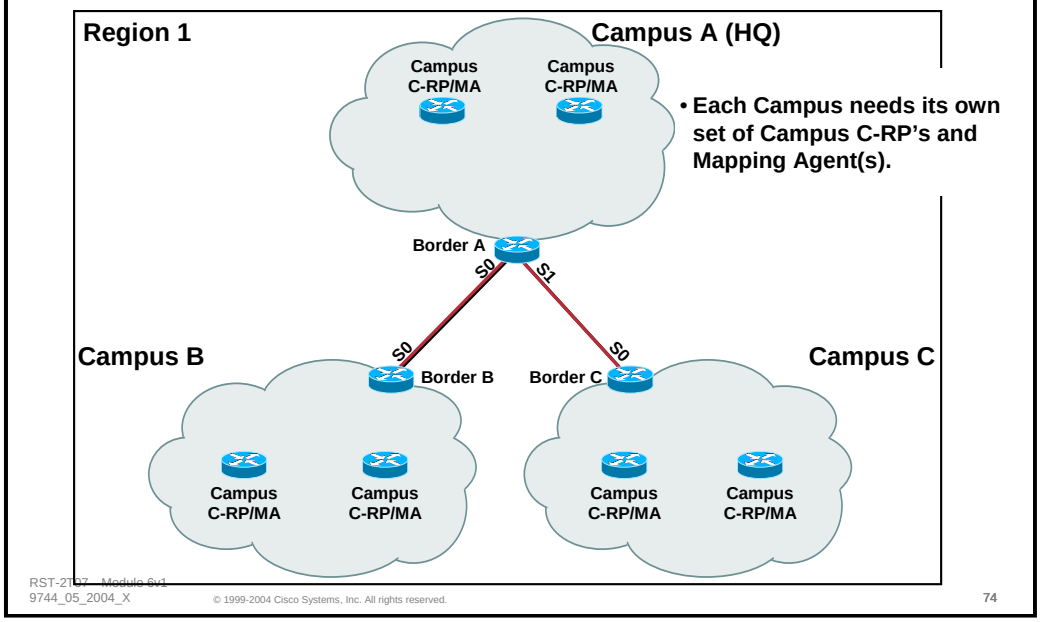


- **Administratively-Scoped Zone**
  - The slide above shows a deployment of Admin-Scoped Zones based on the address scheme shown on the previous page.
    - In the example above, a Headquarters site is connected to two other remote sites: one in Los Angeles and another in Atlanta. Note that each of these sites (including the HQ site) have site local boundaries configured to prevent the flow of 239.255.0.0/16 multicast traffic from leaving the site.

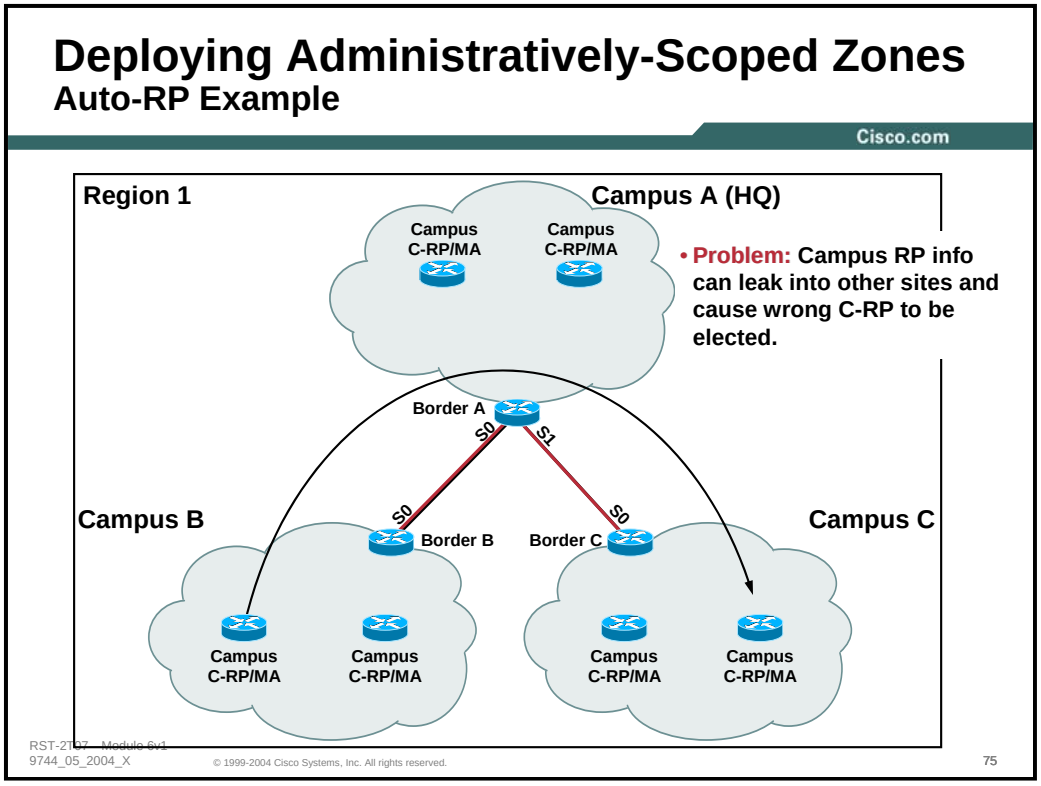


- **Administratively-Scoped Zone**
  - The slide above shows the configuration commands necessary to establish the "Site-Local" Admin-Scoped Zones.
    - Notice that the **ip multicast boundary** command is used with the appropriate ACL to deny any high-bandwidth multicast traffic in the 239.255.0.0/16 multicast group range from entering/leaving the sites and possibly congesting the WAN links.

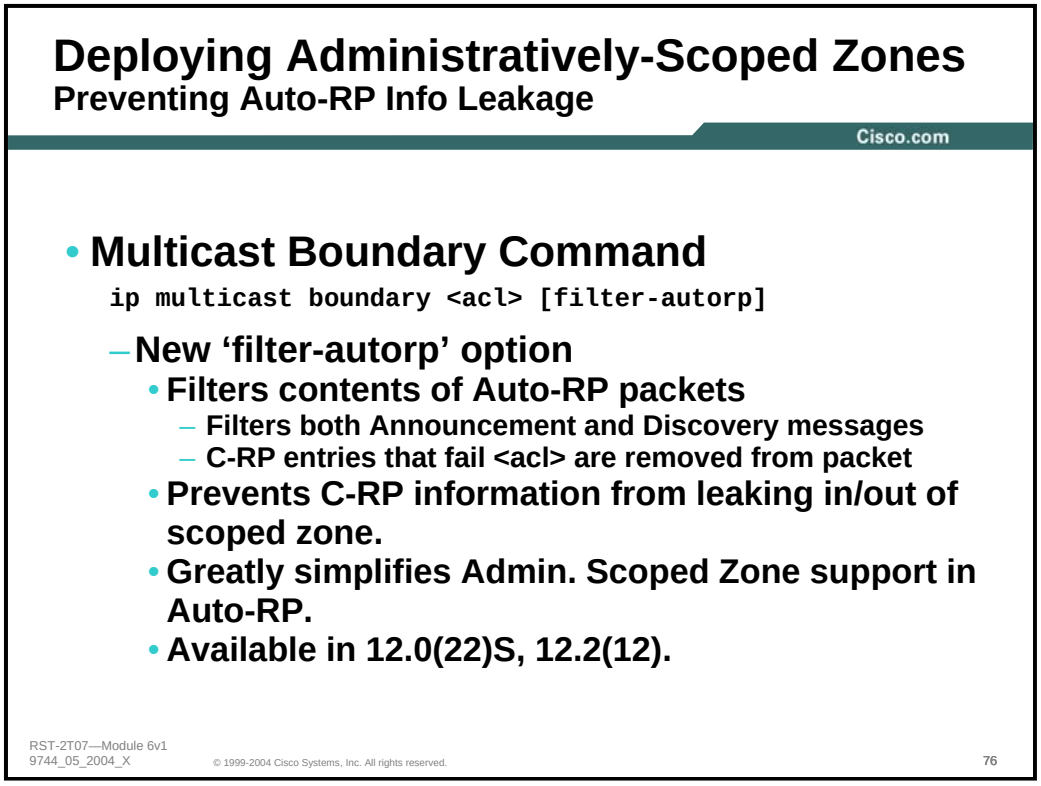
Deploying Administratively-Scoped Zones
Auto-RP Example



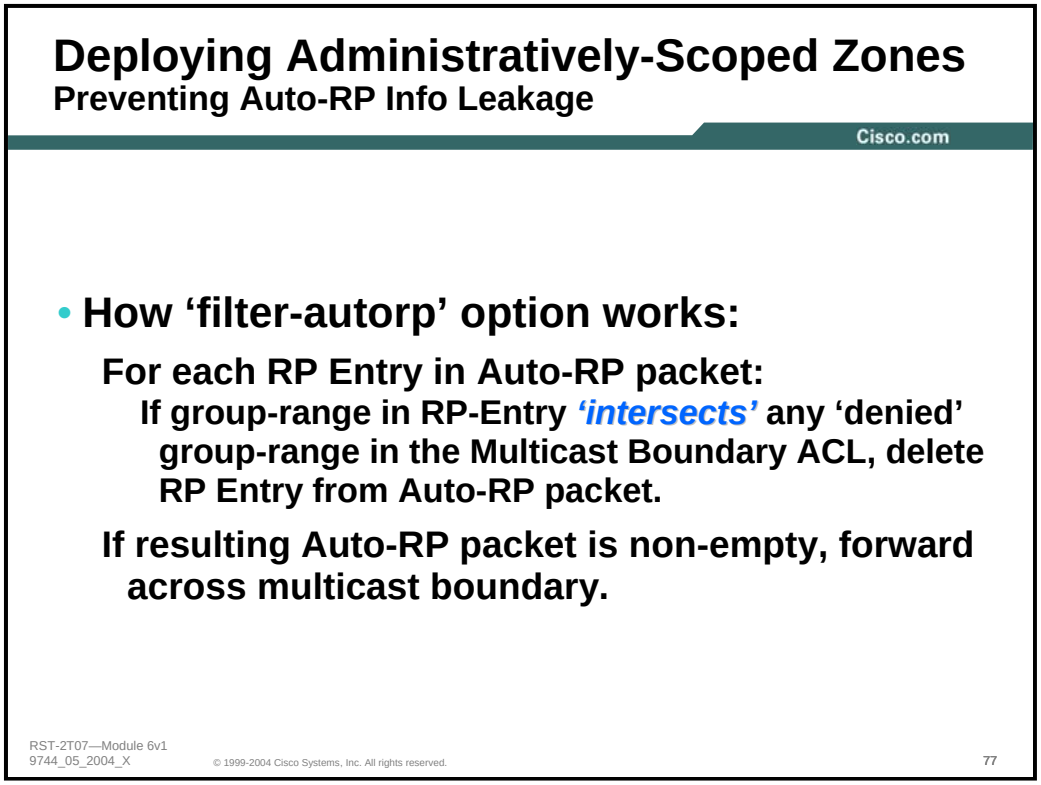
- **Administratively-Scoped Zone**
  - As a result of the **multicast boundary** commands being placed on the WAN links, each site effectively becomes an independent Sparse Mode domain for the 239.255.0.0/16 "Site-Local" group range. This means that each site must have it's own RP for the "Site-Local" group range.
  - In this example, we are using Auto-RP to configure RP's at each site. Two Candidate RP's and two Mapping Agents are configured in each site in order to provide RP redundancy within the site for the "Site-Local" group range. (In this example we've placed the Mapping Agent and C-RP router functions on the same two routers within each site to simplify the drawing. This is not a requirement, however, as these functions could just as easily be placed on separate routers within the site.)



- **Administratively-Scoped Zone**
  - The problem here is that Auto-RP Announcement and Discovery traffic can “leak” between sites. If this is allowed to happen, the site in Atlanta, for example, could erroneously “elect” a Site-Local RP in Los Angeles. This would result in a Campus multicast failure in the Atlanta site.
  - While it may seem that the simple solution would be to block ***all*** Auto-RP traffic between sites, we cannot take this approach. The reason is that we will need to distribute other Admin-Scope RP information (e.g. Organization-Local RP information) between the sites. If we block all Auto-RP multicast traffic in the 224.0.1.39 and 224.0.1.40 range, we will not be able to distribute this information and hence multicast for these group ranges would break somewhere in the network.
  - What is needed is a special filter function that will selectively filter the contents of Auto-RP Announcement and Discovery messages and remove the Site-Local advertisements from the messages so that Site-Local information does not leak between sites.



- **Preventing Auto-RP Info Leakage**
  - In order to selectively filter the *contents* of Auto-RP packets, a new **filter-autorp** option was added to the **ip multicast boundary** interface command.
    - When configured, this feature will filter the contents of both Auto-RP Announcement and Discovery messages. RP entries that are "denied" by the ACL are removed from the Auto-RP packet thereby preventing Auto-RP information from leaking across the multicast boundary.
  - This new option greatly simplifies the configuration steps necessary to deploy Admin. Scoped Zones.
    - All that is necessary to deploy a Scoped Zone is to configure a **multicast boundary** on an interface with the **filter-autorp** keyword and with an ACL that "denies" the Admin. Scoped range.
    - Note: Care must be taken to insure that the C-RP group-range definitions do not overlap. This can result in larger range scopes being filtered by accident which in turn, will result in loss of critical Auto-RP information.



- **How it works**
  - When an Auto-RP packet is to be received or sent on an interface configured with a **multicast** boundary command with the **filter-autorp** option enabled, the router intercepts the packet and applies the following logic:
    - for <each RP-Entry in the Auto-RP packet>
    - if <the RP-Entry Group-Range *intersects* any "denied" group-range
    - in the multicast boundary ACL> then
    - delete the RP-Entry from the Auto-RP packet;
    - endif
    - endfor
    - if <remaining Auto-RP packet is non-empty>
    - forward across multicast boundary
    - else
    - discard Auto-RP packet.
    - endif
  - Note that the function *intersects* in the above algorithm is true if any address in the RP-Entry Group-Range falls within a "denied" multicast boundary group range. This is why it is ***critical*** to make sure RP group-ranges do not overlap. (This means don't use 224.0.0.0/4 as a group range!!!!!)

Deploying Administratively-Scoped Zones
Preventing Auto-RP Info Leakage

Cisco.com

- **Using Multicast Boundary 'filter-autorp'**
  - **Avoid Auto-RP Group-Range Overlaps**
    - **Overlapping ranges can "intersect" denied ranges at multicast boundaries.**
      - **Can cause unexpected Auto-RP info filtering at multicast boundaries.**
      - **Results in loss of Auto-RP info to other parts of network.**
  - **Rule of Thumb:**
    - **Make sure Auto-RP Group-Ranges match exactly any Multicast Boundary Ranges!**
      (i.e. don't use overlapping Auto-RP group ranges.)

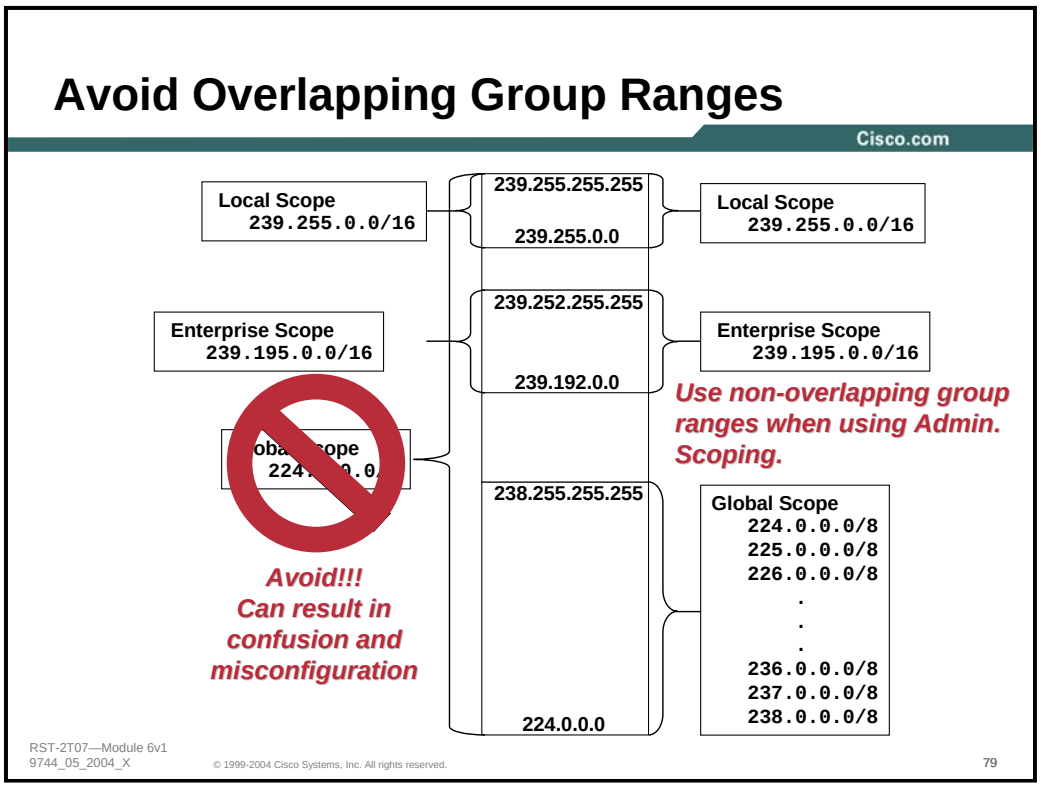
RST-2T07—Module 6v1
9744_05_2004_X

78

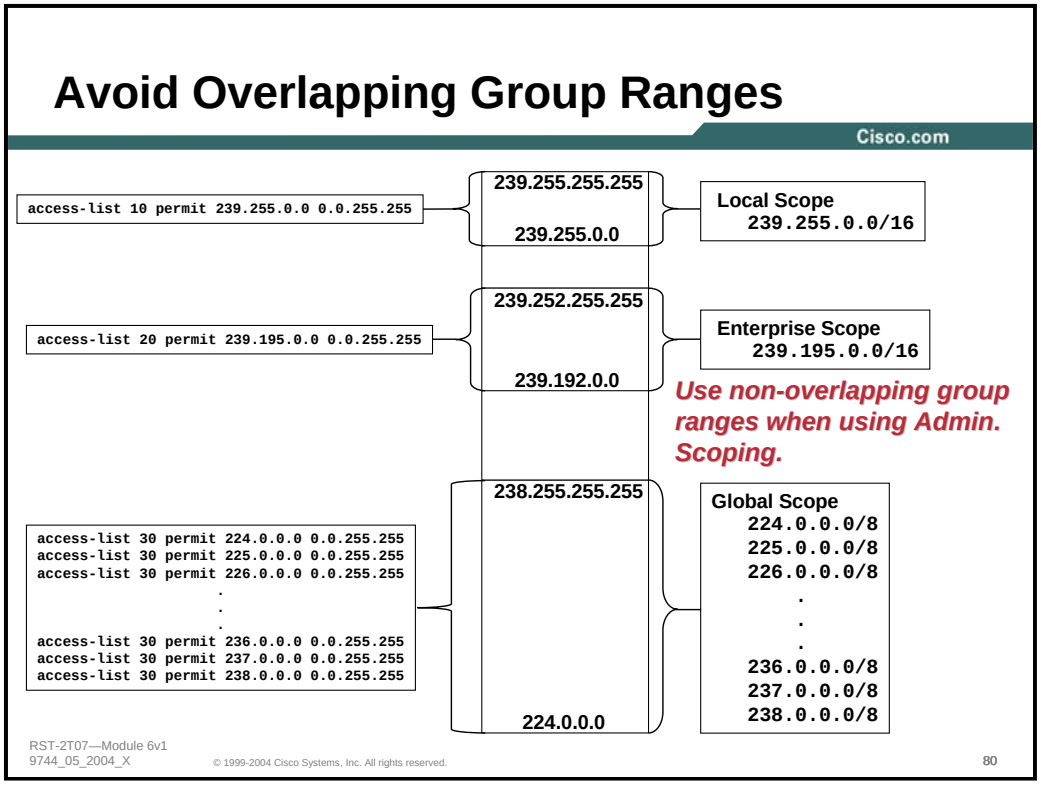
- **Using Multicast Boundary 'filter-autorp'**
  - It is *crucial* that one avoids overlapping RP group ranges when using Admin. Scoped Zones. The classic example of this is the use of a "catch-all" RP to cover everything *except* the Site-Local zone (or any other zone for that matter.) The catch-all C-RP definitions are often configured as

    ip pim send-rp-announce Loopback0 scope 32

    This will result in the group range of RP-Entry in the Auto-RP Announcement being 224.0.0.0/4. This overlaps the Site-Local range which will be "denied" by the multicast boundary ACL. The net result will be that the RP-Entry for the catch-all RP will be filtered at the multicast boundary.
    - This is why it is ***critical*** to make sure RP group-ranges do not overlap. (This means don't use 224.0.0.0/4 as a group range!!!!!)
  - Rule of Thumb
    - When using Admin. Scoped Zones, make sure that the RP group-ranges specified in the group-list ACL ***match exactly*** the multicast boundary group-ranges.



- **Avoiding Overlapping Group Ranges**
  - It is important to avoid overlapping group ranges especially when using Auto-RP and Administrative Scoping. This is because large overlapping group range information can accidentally be filtered at multicast boundary points if the **filter-autorp** feature is in use.



- **Avoiding Overlapping Group Ranges**
  - This example shows how the group ranges have been redefined to avoid any overlap.

Avoiding Overlapping Group Ranges

Cisco.com

- **Avoiding Overlapping Group Ranges**
  - **Can't use "deny" clause in C-RP ACL's**
    - **Implies "Dense-mode Override"**

```
ip pim send-rp-announce loopback0 scope 16 group-list 10
access-list 10 deny 239.0.0.0 0.255.255.255
access-list 10 permit 224.0.0.0 15.255.255.255
```

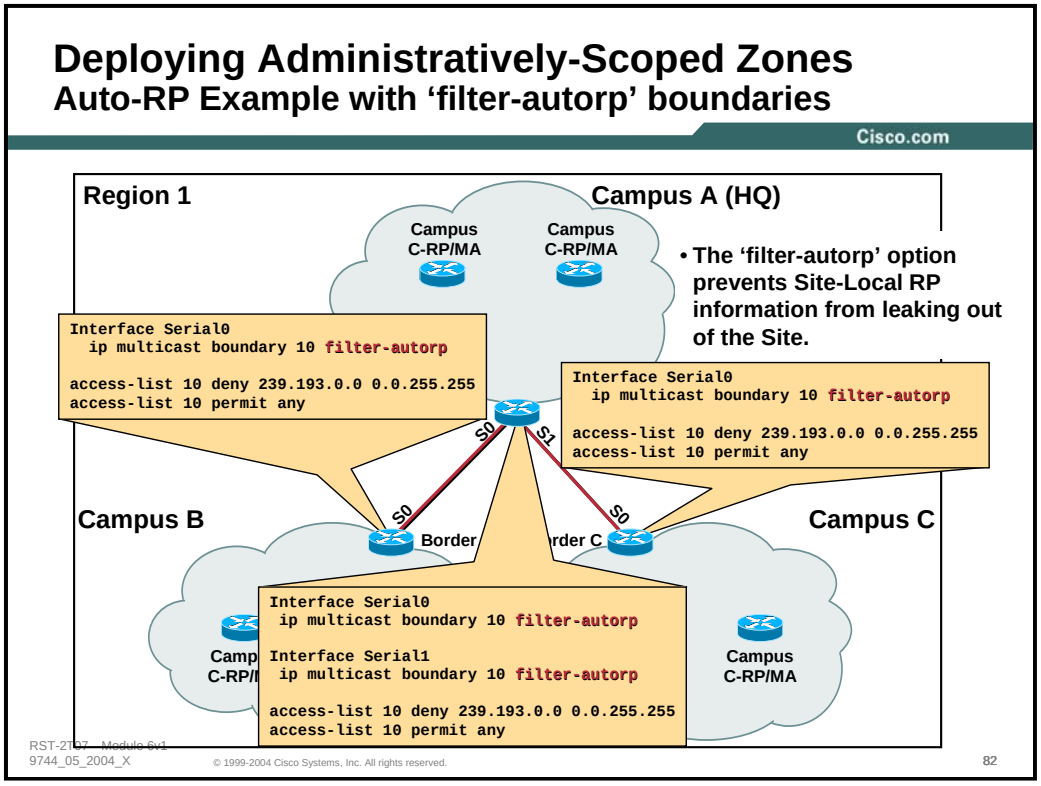
  - **Must only use "permit" clauses**

```
ip pim send-rp-announce loopback0 scope 16 group-list 10
access-list 10 permit 224.0.0.0 0.255.255.255
access-list 10 permit 225.0.0.0 0.255.255.255
                       .
                       .
                       .
access-list 10 permit 238.0.0.0 0.255.255.255
```

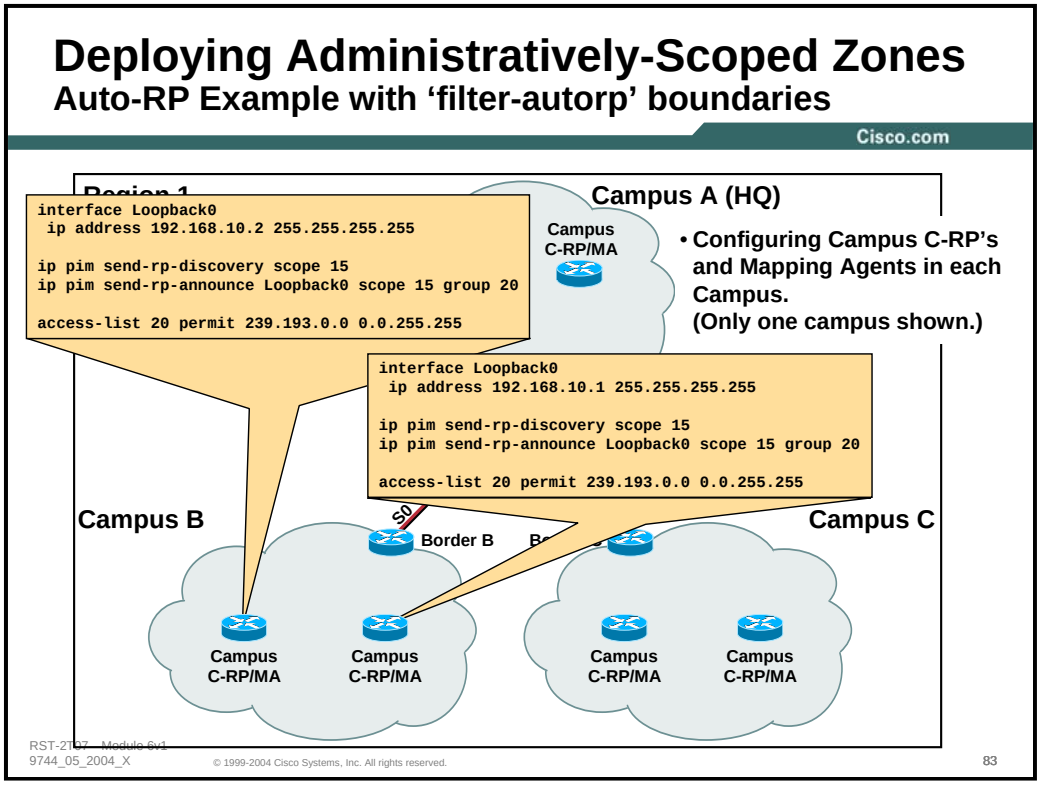
RST-2T07—Module 6v1
9744_05_2004_X

81

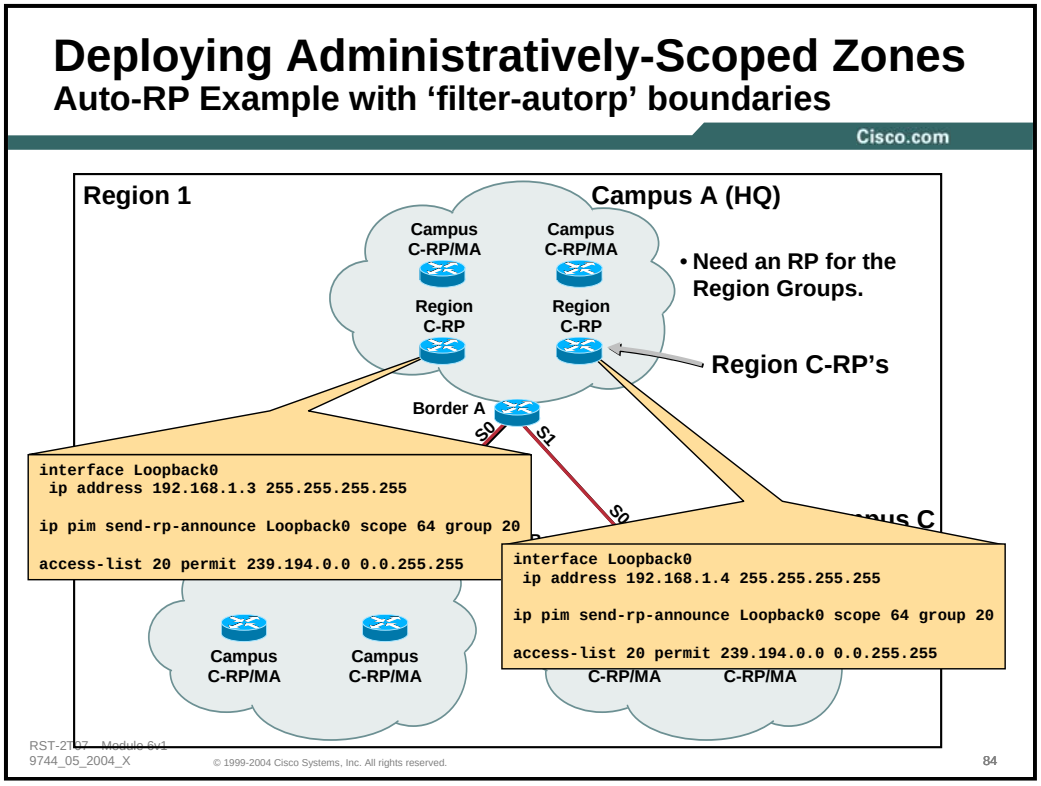
- **Avoiding Overlapping Group Ranges**
  - The problem with defining group-range ACL's that don't overlap is that we can not use the **deny** clause in our ACL's as we normally do in other ACL's. This is because the **deny** clause has a special *"Dense-mode Override"* meaning when used in RP group-range ACL's.
  - Therefore, when defining RP group-range ACL's, only **permit** clauses should be used. The impact of this is that we often have to use more statements to define the group range.



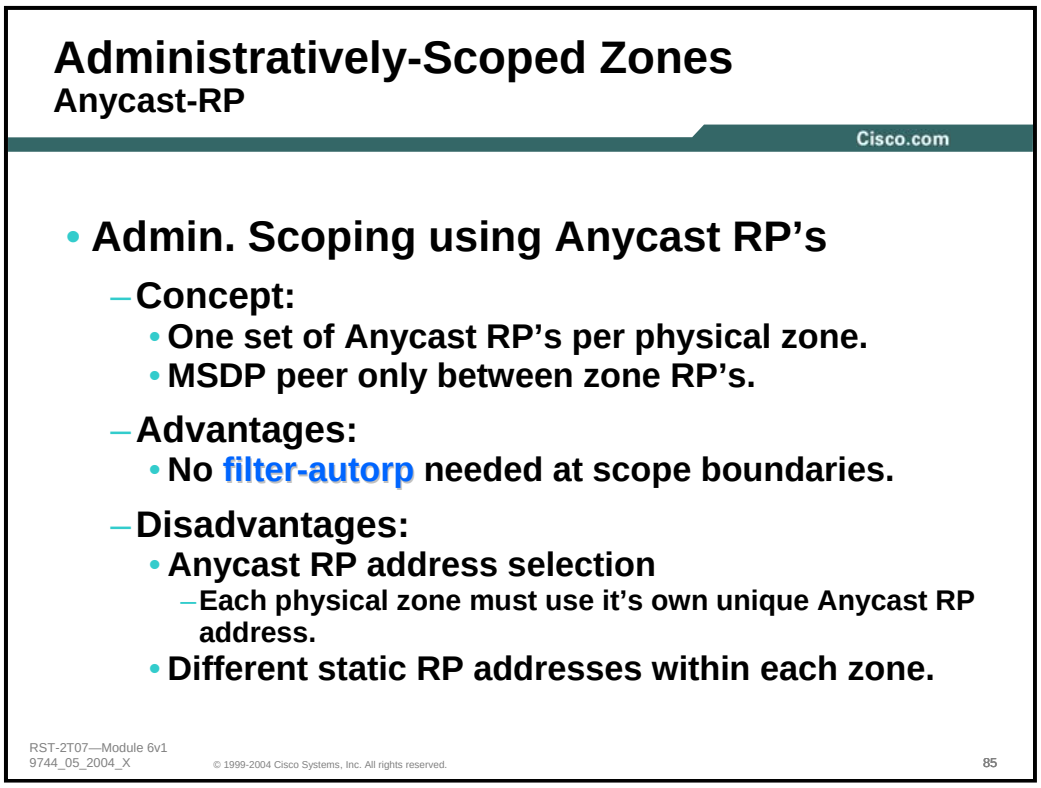
- **Auto-RP Example with 'filter-autorp' boundaries**
  - The example above shows our example network configured with the new **filter-autorp** feature on the **ip multicast boundary** command.
    - The ACL used in the boundary command will prevent multicast traffic in the 239.255.0.0/16 range from flowing across the boundary
    - The **filter-autorp** keyword will also filter the contents of any Auto-RP packets (Discovery and Announcement messages) and remove any RP-Entries from the packet whose group-range *intersects* with the denied range of 239.255.0.0/16.



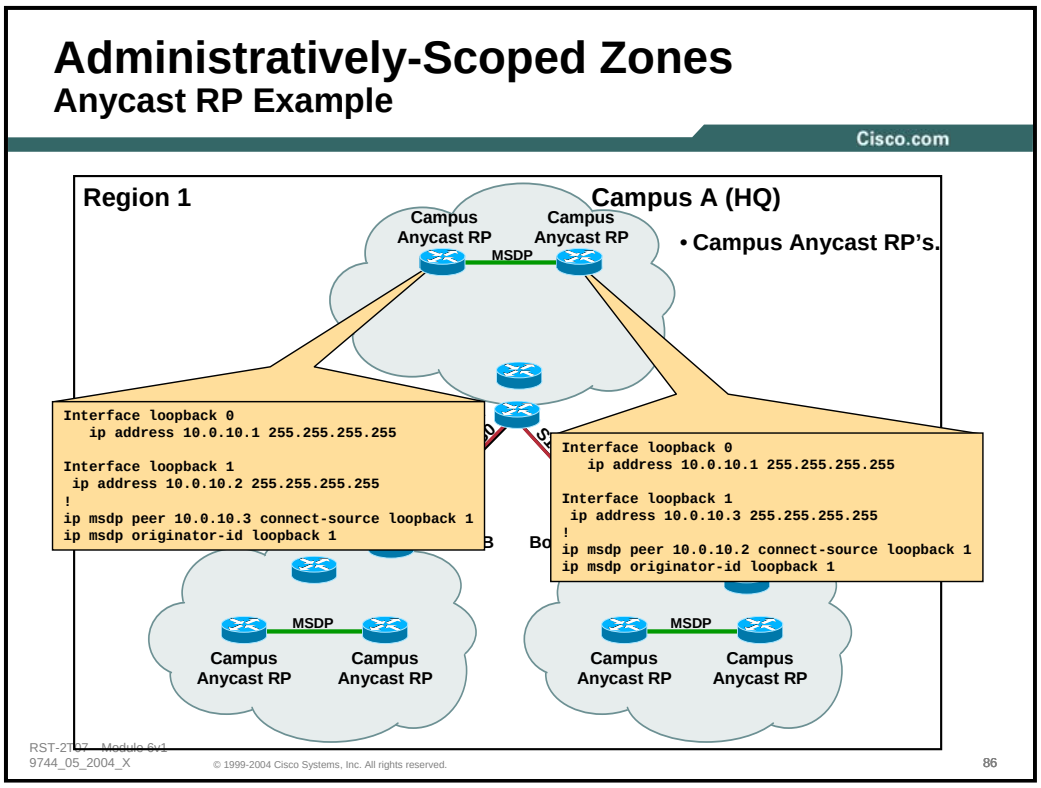
- **Auto-RP Example with 'filter-autorp' boundaries**
  - Continuing our example, it is necessary to configure Site-Local Candidate RPs and Mapping Agents inside of each site. The slide above shows the necessary configuration for the Los Angeles site. Notice that the following:
    - The group-list ACL used in the **ip pim send-rp-announce** command that defines the Site-Local Candidate RP *matches exactly* the group range used in the **multicast boundary** ACL on the previous page.



- **Auto-RP Example with 'filter-autorp' boundaries**
  - Finally, it is necessary to configure Organization-Local Candidate RPs that will serve as the RP for all other groups *except* the Site-Local group range. In this case, we have chosen to place all C-RP's for this group range at the HQ site although it would be just as easy to place C-RP's for this range at other sites as well. The slide above shows the necessary configuration for the C-RPs for the Organization-Local group range as well as all other remaining groups. (These are the catch-all C-RP's.) Notice that the following:
    - •The group-list ACL used in the **ip pim send-rp-announce** command that defines the catch-all Candidate RP's *does not overlap* the Site-Local group range used in the **multicast boundary** ACL. This is necessary so that the corresponding RP-Entry in the Auto-RP Announcement for these C-RP's ***will not*** intersect the "denied" Site-Local multicast boundary range of 224.0.0.0/16.
    - •Notice that the definition of the **group-list** ACL for the C-RP's uses only "permit" clauses. This is necessary as any "deny" clause in this ACL would force the specified group range into Dense mode for the *entire* network. This means we have to take the "long way around" to defining the catch-all group range that cover everything *but* the Site-Local group range.



- **Administrative Scoping using Anycast RP's**
  - The idea is to provide one set of Anycast-RP's per scoped zone. Since it is not desirable to have sources (sending traffic to the scoped range) inside a scope learned by another scope, there is no need to MSDP peer with other Anycast-RP's in other scopes. (In fact, this is would be a mis-configuration.)
  - The advantage of this approach is that there is no need for special 'filter-autorp' mechanisms at the scope boundaries. (Although adding this clause wouldn't hurt anything and would permit a combination of Auto-RP and Anycast-RP techniques to be used.)
  - The disadvantage is that each *physical* zone should use a unique Anycast RP address to avoid routers at the edge of a boundary to attempt to Register/Join source/receivers to an Anycast-RP in the adjacent scope. This also means that different static Anycast-RP addresses that must be statically configured in each zone as opposed to having a "cookie-cutter" configuration for every router in the network.
    - Note: This can be overcome by using a combination of Auto-RP and Anycast-RP techniques where Auto-RP is used to advertise the local Anycast-RP address to the routers within the scope. This considerably reduces the amount of configuration overhead. However, the down side is that you are running Auto-RP in your network and some network administrators prefer to not run Auto-RP for security reasons.

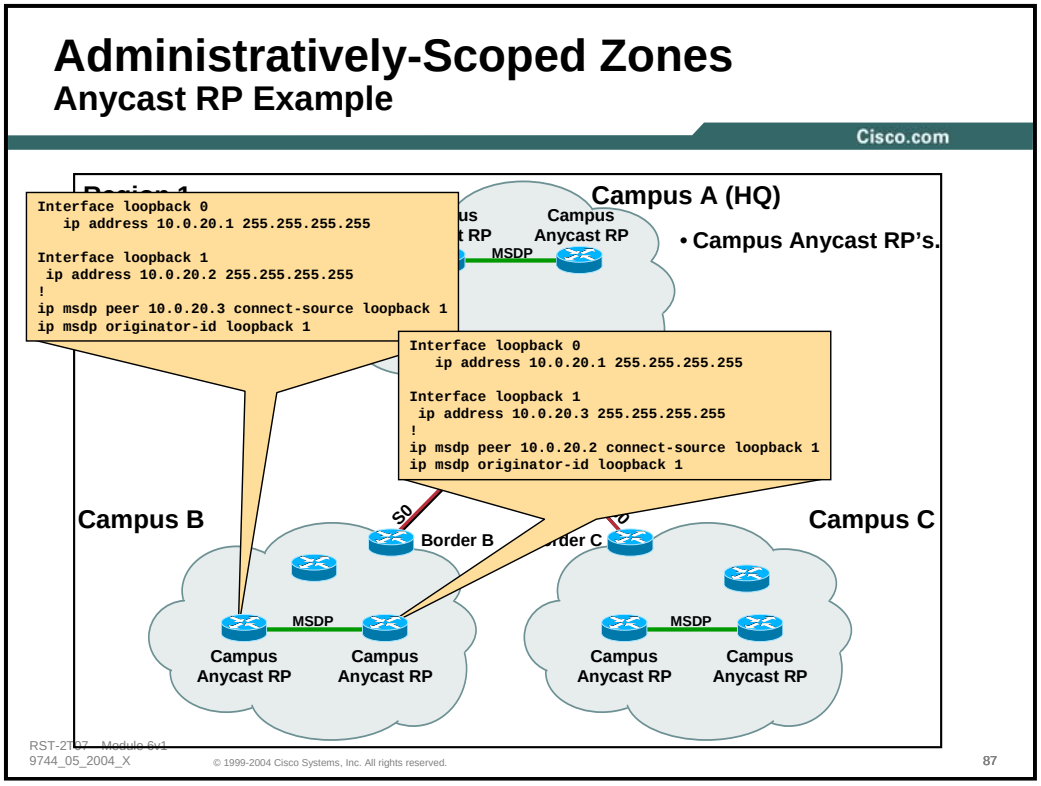


- **Example – Administrative Scoping using Anycast RP's**

  This example follows the Campus, Region, Enterprise examples seen previously. Here we see Region 1 which is comprised of three Campus scopes. (Note that we only show a single Region and not the entire Enterprise due to space limitations.)

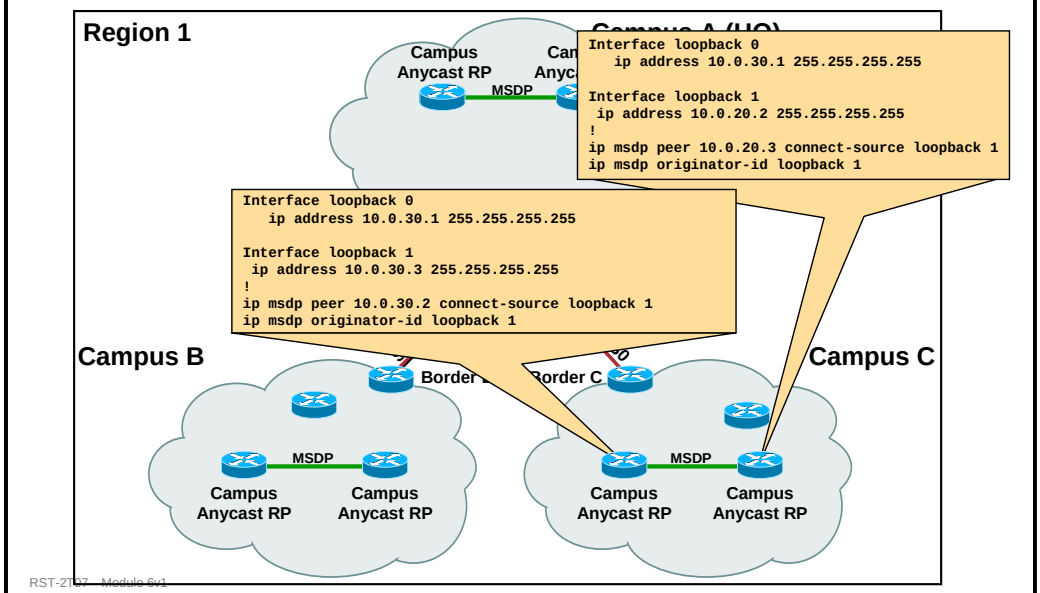
  – Each Campus scope is configured with two routers serving as Anycast-RP's for the Campus group range. Notice that these Anycast-RP's do not MSDP peer with any other routers in any other scopes. This is because we do not want Campus scope sources in one Campus being learned by Anycast-RP's in other Campus scopes.

  – Also make note that the Anycast-RP address for Campus A is **10.0.10.1**. This will be unique to all routers within Campus A.



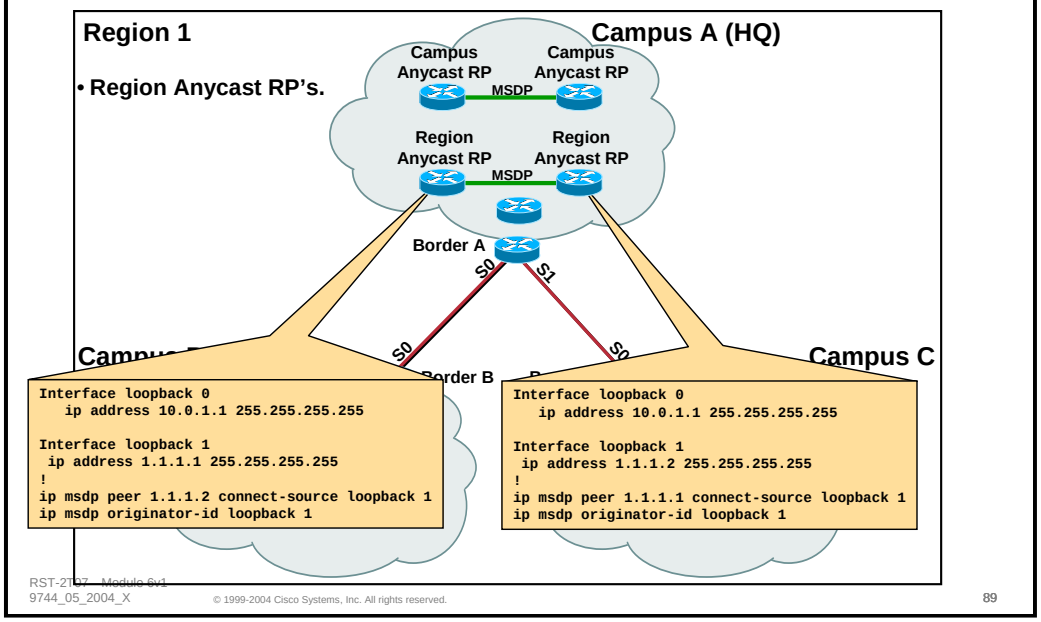
- **Example – Administrative Scoping using Anycast RP's**
  - Note that the Anycast-RP address for Campus B is **10.0.20.1**. This will be unique to all routers within Campus B.

Administratively-Scoped Zones
Anycast RP Example



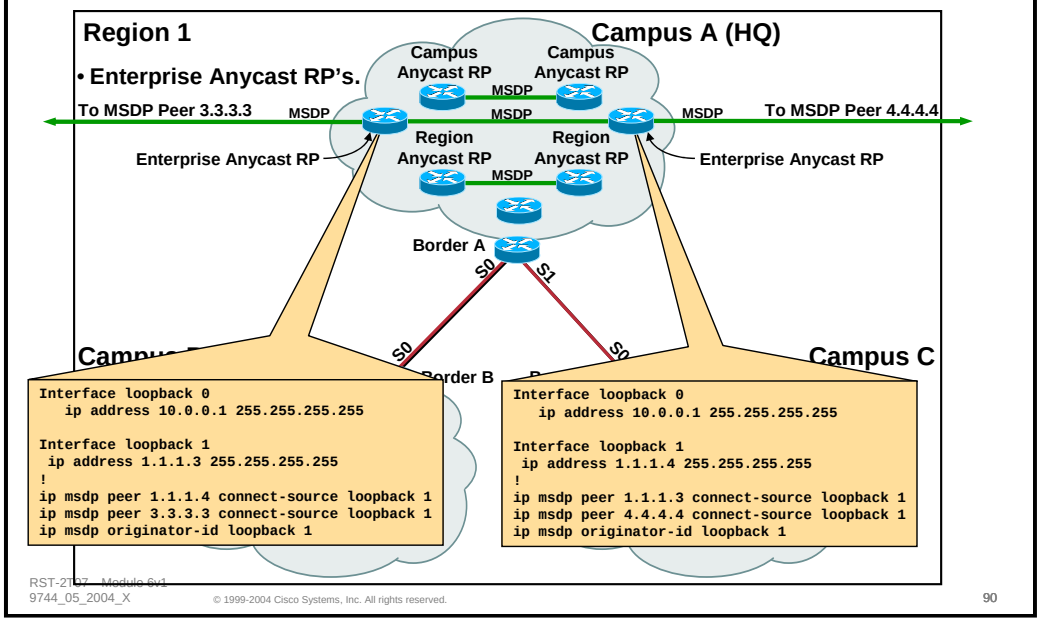
- **Example – Administrative Scoping using Anycast RP's**
  - And finally, note that the Anycast-RP address for Campus C is **10.0.30.1**. This will be unique to all routers within Campus C.

Administratively-Scoped Zones
Anycast RP Example

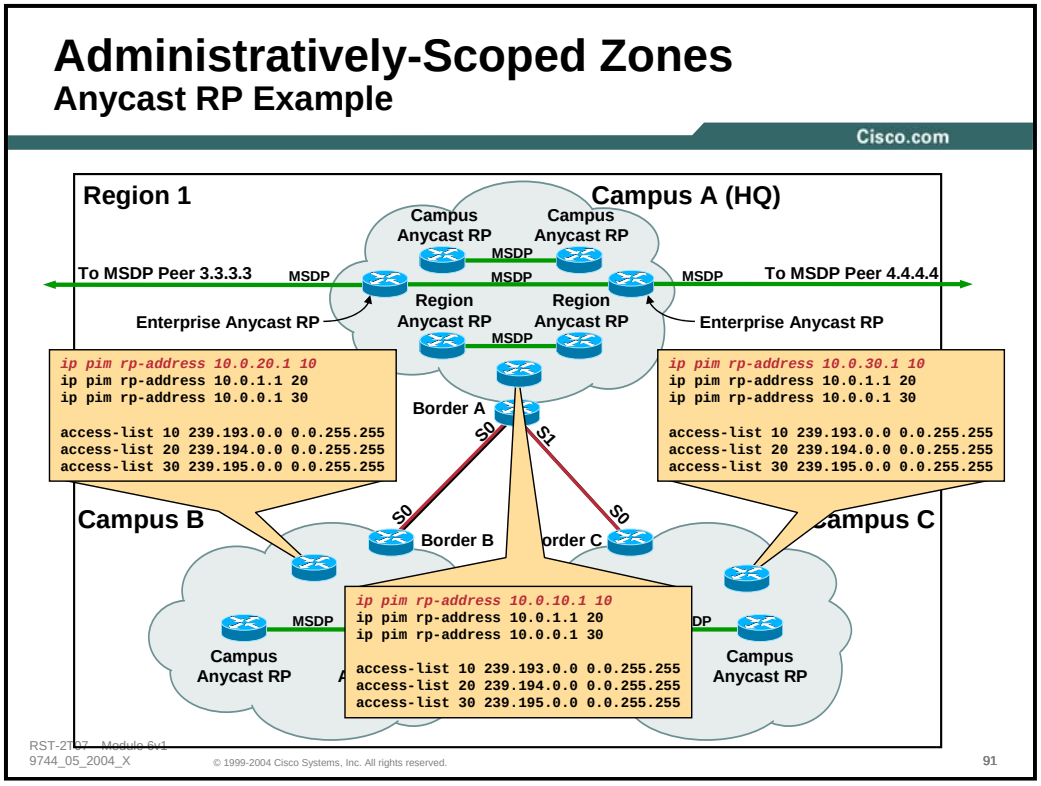


- **Example – Administrative Scoping using Anycast RP's**
  - Each Region scope is also configured with two routers serving as Anycast-RP's for the Region group range. In this example we have chosen to put both of these routers in the HQ Campus, Campus A. However, these routers could be distributed in any manner inside of the Region. In fact, we could use one router in each of the three Campus locations and MSDP peer each of them together to build our Region Anycast-RP's.
  - Notice too that these Anycast-RP's do not MSDP peer with any other routers in any other scopes. This is because we do not want Region scope sources in one Region scope being learned by Anycast-RP's in other Region scopes.
  - Also make note that the Anycast-RP address for Region 1 is **10.0.1.1**. This will be unique to all routers within Region 1.

Administratively-Scoped Zones
Anycast RP Example

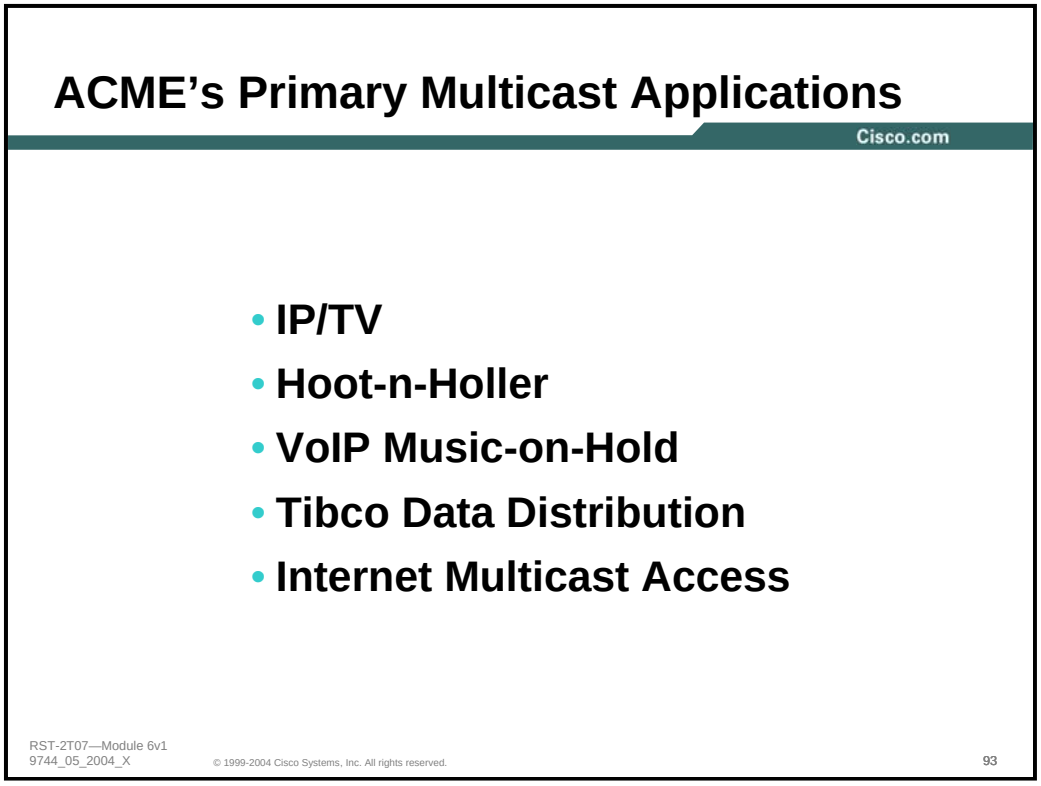


- **Example – Administrative Scoping using Anycast RP's**
  - Finally, the Enterprise scope is also configured with two routers serving as Anycast-RP's for the Enterprise group range. In this example we have again chosen to put two routers in the HQ Campus, Campus A. However, we have also put Anycast-RP's in other Regions (not shown). The MSDP peering connections to these Anycast-RP's (MSDP peers 3.3.3.3 and 4.4.4.4) are also shown above along with the corresponding configuration.



- **Example – Administrative Scoping using Anycast RP's**
  - The last step is to configure each router with the proper Anycast-RP addresses for the Campus, Region and Enterprise scope in which it resides. Notice that in the example above, the Campus Anycast-RP's are unique in each campus. This means that different static Anycast-RP addresses must be statically configured in each Campus as opposed to having a "cookie-cutter" configuration for every router in the network.
    - Note: This can be overcome by using a combination of Auto-RP and Anycast-RP techniques where Auto-RP is used to advertise the local Anycast-RP address to the routers within the scope. This considerably reduces the amount of configuration overhead. However, the down side is that you are running Auto-RP in your network and some network administrators prefer to not run Auto-RP for security reasons.

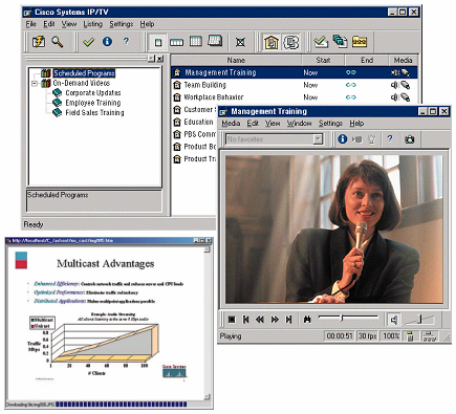
Case Study – ACME Financials



- **Case Study – ACME Financials**
  - ACME Financials has deployed several IP Multicast applications in their network. Each of which requires some special considerations for the proper deployment of the applications. These applications include:
    - IP/TV
    - Hoot-n-Holler
    - VoIP Music-on-Hold
    - Tibco Data Distribution
    - Internet Multicast Access

IP/TV

Cisco.com

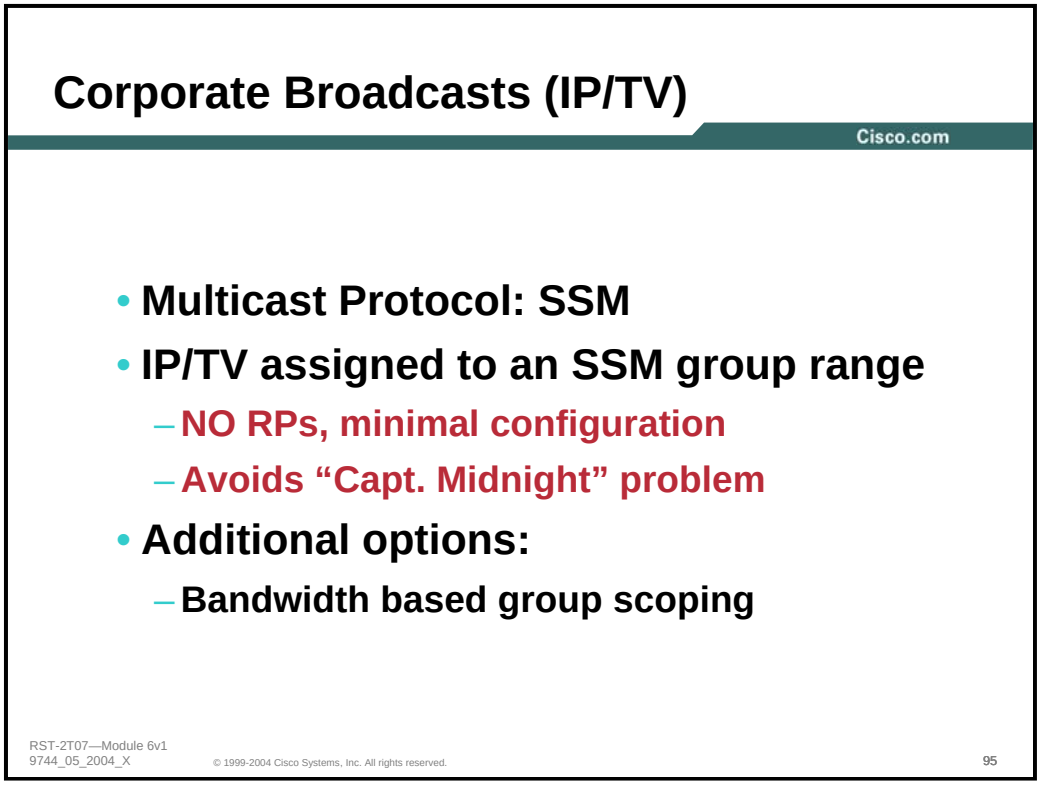


- **One-to-many video multicast**
  - **Live or Rebroadcast Content**
  - **Synchronized Presentations**
  - **Integrated “Question Manager”**
  - **Supports “Source Specific Multicast” (SSM)**
  - **Video-on-Demand (VoD) (Unicast only)**

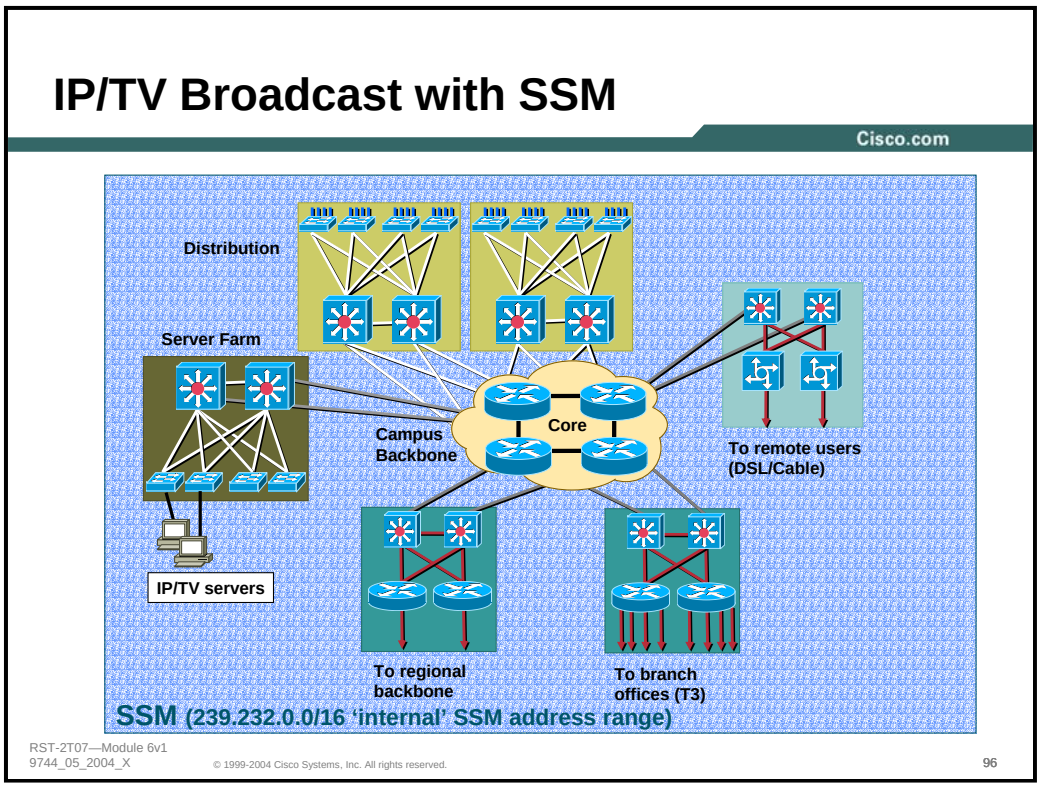
RST-2T07—Module 6v1
9744_05_2004_X

94

- **IP/TV**
  - This application is a one-to-many multicast application that multicasts Live or Rebroadcast video. It supports synchronized presentations and allows the viewers to pose questions to the presenter via an integrated Question Manager.
  - IP/TV now supports “Source Specific Multicast” (SSM) which provides an improved delivery model for one-to-many applications and avoids certain security and DoS attacks that can sometimes be a problem with standard multicast streams.



- **IP/TV Corporate Broadcasts**
  - ACME Financials has deployed IP/TV to do corporate broadcasts from their CEO to all employees as well as to provide regular “Distance-Learning” broadcasts for their Traders so that they can keep up to date on market trends.
  - In order to deploy IP/TV, ACME has chosen to use the new SSM model for IP/TV. This will eliminate the need for RP’s in this group range (along with the simplified network operational model) as well as avoiding the possibility of broadcast “jamming” from disgruntled employees.
  - Additionally, by using multiple servers for the same content source, ACME is able to stream the same content at two different rates. By combining this with Admin. Scoping techniques, ACME can offer remote sites with lower-speed streams that don’t congest the WAN links while at the same time allowing employees at the HQ site to watch a high-speed stream that is limited to the local HQ campus where bandwidth is plentiful.

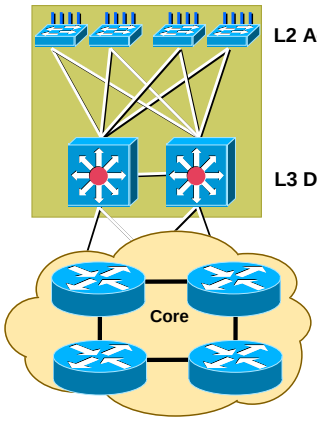


- **IP/TV Broadcast with SSM**
  - In order to provide for SSM multicast inside of the ACME Enterprise network, the address range of 239.232.0.0/16 was chosen by ACME engineers as the "internal" SSM range. This would be configured on the routers in the network using the following configuration commands:

```
ip pim ssm range 11

access-list 11 permit 239.232.0.0 0.0.255.255
```

IP/TV – SSM on Campus

Cisco.com



- **IGMPv3 Snooping on Access where available on clients and switches**
- **'static ssm mapping' when IGMPv3 is not available (L3 distribution)**
- **Use bandwidth scoping for SSM groups with different rates (239.232.QOS.x)**

RST-2T07—Module 6v1
9744_05_2004_X

97

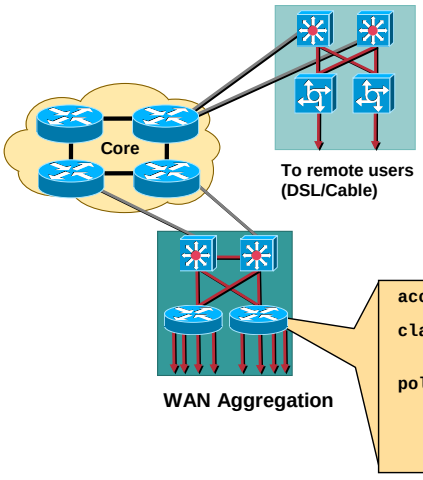
- **IP/TV – SSM on Campus**
  - Normal SSM operation requires hosts and their upstream switches and routers to support SSM. This means that IGMPv3 Snooping needs to be enabled on the switches and the host platforms need to have an OS stack that supports IGMPv3 (such as Microsoft XP or later).
  - In some cases, it may be desirable to employ SSM multicast even in portions of the network where the switches and/or the hosts have not yet been upgraded to fully support IGMPv3. In this case, the use of a new IOS feature, "SSM Static Mapping" may be used as an interim SSM deployment strategy.
    - Static SSM Mapping provides the ability to configure a static mapping of a single source to a single group address on the last-hop routers. The following is an example of a Static SSM Mapping configuration:

      ```
      ip igmp ssm-map enable

      ip igmp ssm-map static 10 192.162.20.20

      access-list 239.232.10.1
      ```

      The configuration above will statically map source 192.162.20.20 to SSM group 239.232.10.1. If the router receives an IGMPv2 Membership Report from a host connected to the router to join this group, the router will interpret the IGMPv2 Membership Report as if an "Include (192.168.20.20, 239.232.10.1) IGMPv3 Membership Report was received and will respond by sending a PIM (S,G) Join to join this SSM flow.
  - In order to provide bandwidth scoping for SSM flows, ACME engineers have followed the 239.232.<qos>.0 addressing scheme for the different QoS bandwidth flows of the same video content transmitted at different rates.

IP/TV – SSM over WAN

Cisco.com



- **QOS protection necessary for remote users and WAN aggregation**
- **Use bandwidth scoping to deny high rate streams**
- **Use MQC to guarantee bandwidth to IP/TV:**

```
access-list 101 permit ip any 239.232.224.0 0.0.31.255

class-map match-all iptv-qos-lowbandwidth
    match access-group 101

policy-map IPTV-over-T1
    class iptv-qos-lowbandwidth
       bandwidth 512
    class default
       fair-queue
```

RST-2T07—Module 6v1
9744_05_2004_X

98

- **IP/TV – SSM over the WAN**
  - ACME engineers have implemented QoS protection on WAN aggregation routers to guarantee bandwidth to IP/TV flows being transmitted to remote sites.
    - The configuration shown in the drawing above uses **access-list 101** to identify low-bandwidth IP/TV SSM flows in the 239.232.224.0/20 address range.
    - The **class-map** commands in the configuration above, maps all flows in the 239.232.224.0/20 address range to the **iptv-qos-lowbandwidth** class.
    - The **policy-map** commands in the configuration above assigns all low-bandwidth IP/TV SSM flows in the iptv-qos-lowbandwidth class a bandwidth of 512Kbps to insure that these flows have a guaranteed amount of bandwidth on the T1 WAN links. All other traffic flows fall into the **default** class which uses Fair Queuing.

Hoot-n-Holler

Cisco.com

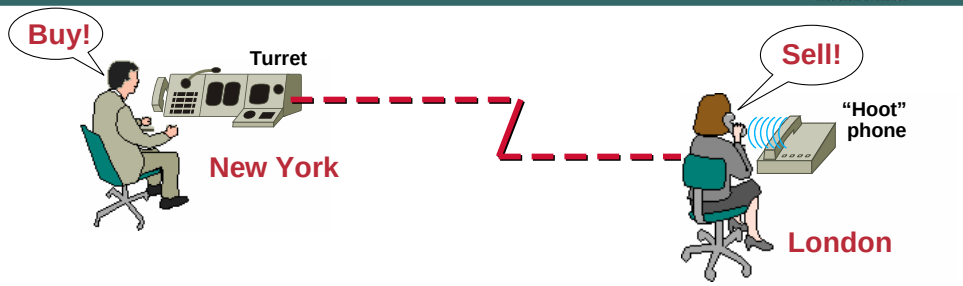
RST-2T07—Module 6v1
9744_05_2004_X

99

- **Hoot-n-Holler**
  - All ACME Financial Traders have Hoot-n-Holler turrets at their trading station. These turrets permit the traders to communicate as a group by simply selecting the Hoot-n-Holler channel button on their turret and talking. Their voice is then carried to all other Traders in the network that have selected this Hoot-n-Holler channel.

Hoot 'n' Holler

Cisco.com

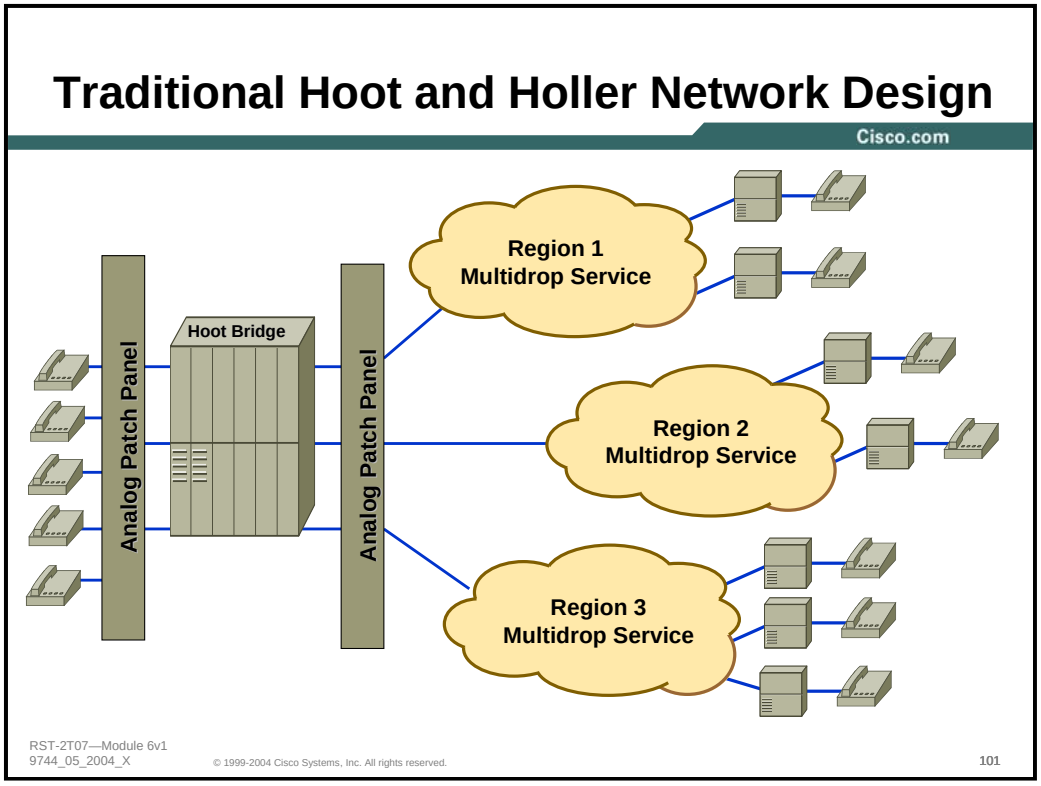


- Broadcast audio network
- Typically point to multipoint
- Uses specialized analog 4-wire phones (Hoot phones) and digital turrets
- Brokerages, utilities, media companies, mass transit, publishing, etc.

RST-2T07—Module 6v1
9744_05_2004_X

100

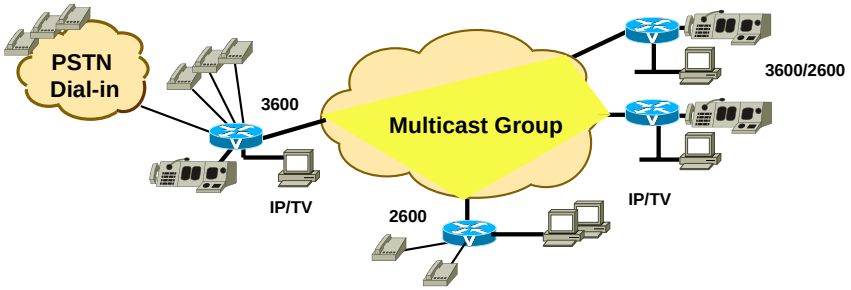
- **Hoot-n-Holler**
  - Traditional Hoot-n-Holler (Hootie, for short) networks have been used in Brokerage Houses, utilities, media companies, etc. for many years. It is basically an audio broadcast network (typically point-to-multipoint) that is carried over special 4-wire phones and digital turrets.



- **Hoot-n-Holler Networks**
  - A traditional Hoot-n-Holler network design might appear as shown above. Analog phones are tied into a special Hoot Bridge that allows all members on a Hootie channel to talk.

Hoot 'n' Holler over IP Multicast

Cisco.com



- **Leverages VoIP, IP Multicast (IPmc) & QoS**
- **Bridging & audio mixing occurs in router voice DSPs**
- **Dynamic bandwidth sharing & cost savings**
- **Existing analog end systems & procedures retained**

RST-2T07—Module 6v1
9744_05_2004_X

102

- **Hoot-n-Holler over IP Multicast**
  - ACME has moved away from the traditional analog Hoot-n-Holler network infrastructure and have migrated their Hoot-n-Holler application to IP Multicast. This permits them to eliminate the expense of maintaining multiple networks.
  - The basic idea of Hoot-n-Holler of IP Multicast is to combine Voice-of-IP (VoIP), QoS and IP Multicast technologies to accomplish the same goals.
  - The bridging and mixing of the audio streams of the participants on a Hootie channel takes place on the Digital Signal Processors (DSP) normally installed in the Cisco router/switch equipment for VoIP.
  - Because IP Multicast is being used to distribute the audio of the Hootie channel, only a single multicast stream is distributed in the network. This provides bandwidth and cost savings over point-to-point methods.
  - By using FXS cards on the 3600 routers, ACME is able to make a smooth transition to the IP Multicast Hoot-n-Holler environment by permitting the use of the existing analog end systems.

Hoot&Holler over IP Multicast – Considerations

Cisco.com

- **Multicast Protocol: Bidir PIM**
  - **Scales well for many-to-many applications**
- **Additional options:**
  - **LLQ (Low latency queue) for Hoot&Holler traffic (voice traffic)**
  - **CRTP Header Compression for low speed links**

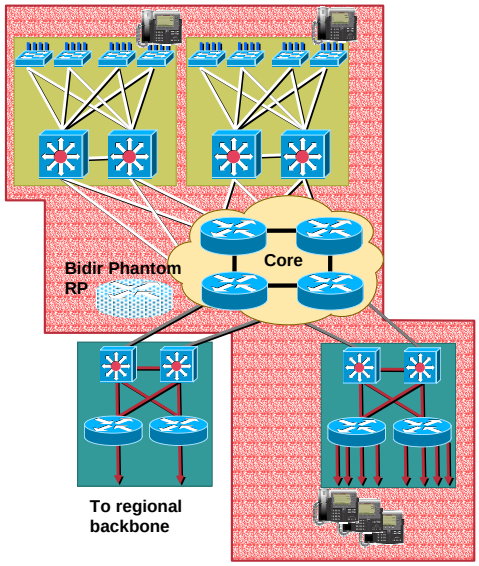
RST-2T07—Module 6v1
9744_05_2004_X

103

- **Hoot-n-Holler – Considerations**
  - Bidir PIM is a perfect match for the many-to-many nature of Hoot-n-Holler. Therefore, the ACME engineers have designed their network to carry the Hoot-n-Holler channels over Bidir PIM Shared Trees.
    - The use of Bidir PIM requires only a single Bidir (*,G) multicast routing entry in each router to carry the audio for the Hootie channel, regardless of how many stations and turrets are using the channel. This results in improved scalability for Hoot-n-Holler over standard PIM-SM techniques where every participant on the Hootie channel would require another (S,G) multicast routing entry in the routers.
  - Since Hoot-n-Holler flows are basically voice traffic, the use of Low Latency Queuing (LLQ) is often employed to insure the quality of the voice traffic is not impacted by other data flows in the network causing delay and jitter of the voice streams.
  - The CRTP IOS feature provides for the compression of the Real-Time Protocol (RTP) headers. This can be used to reduce the bandwidth demands of the Hoot-n-Holler flows over low speed WAN links.

Hoot and Holler on Campus and over WAN

Cisco.com



- **Choose groups from Campus range for H&H (example 239.193.255.x)**
- **Bidir enabled on all (red zone) routers**
- **H&H is voice traffic, so treat it accordingly with Low Latency Queues (LLQ)**

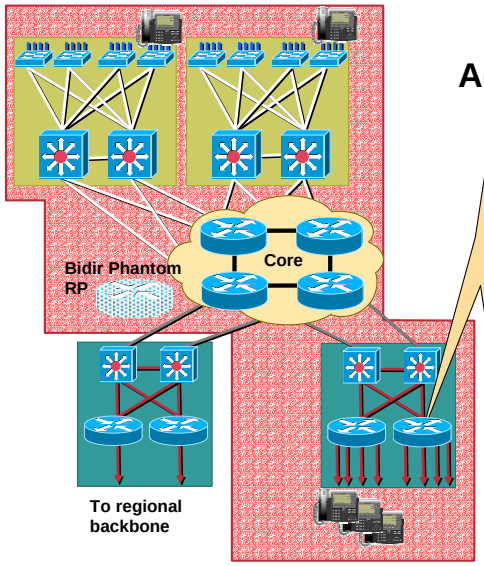
RST-2T07—Module 6v1
9744_05_2004_X

104

- **Adding Hoot-n-Holler**
  - The drawing above shows the implementation of the Hoot-n-Holler application in the ACME Enterprise network.
    - The concept of the Phantom Bidir RP is used to eliminate a single point of failure in the Bidir network.
    - It is not necessary for all routers in the network to be upgraded to support Bidir PIM since it was not necessary to provide Hoot-n-Holler to some parts of the ACME network.
    - Low-Latency Queuing (LLQ) has been used in the network to insure that Hoot-n-Holler and other voice traffic are not impacted by the data streams.

Hoot and Holler on Campus and over WAN

Cisco.com



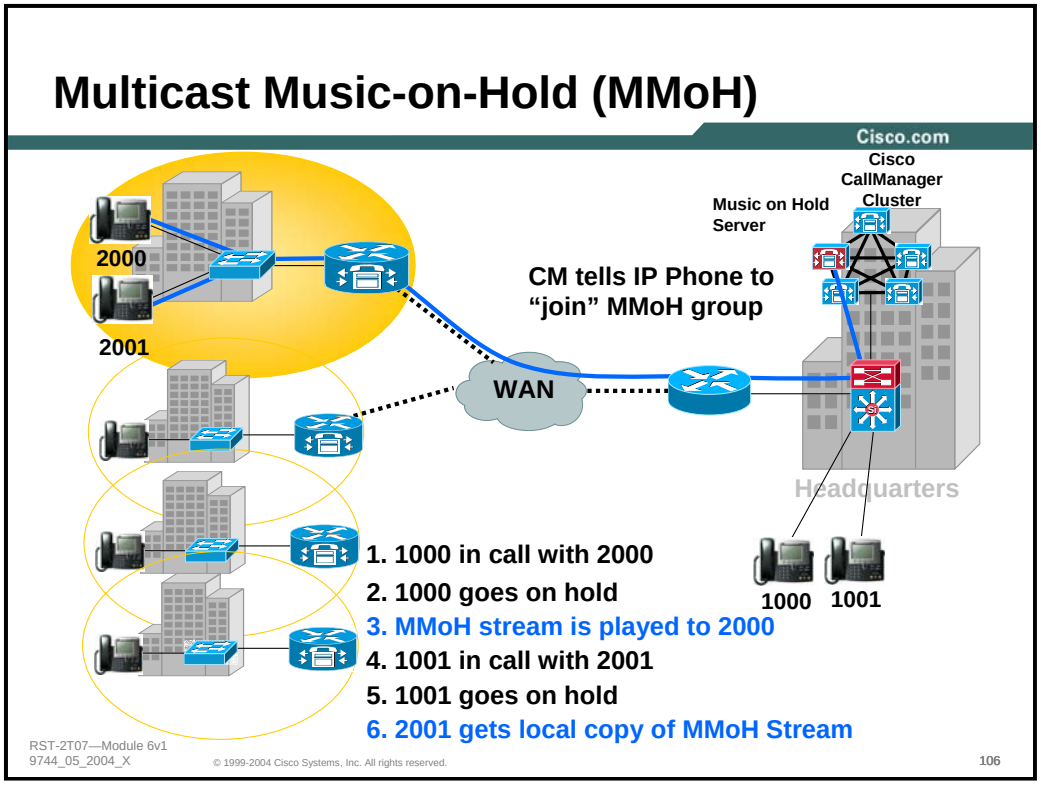
Adding QoS for Hoot and Holler

```
access-list 101 permit ip any 239.232.224.0 0.0.31.255
access-list 102 permit ip any 239.193.255.0 0.0.0.255

class-map match-all iptv-qos-lowbandwidth
     match access-group 101

class-map match-all hootie
     match access-group 102

policy-map Mcast-over-T1
     class hootie
          priority 495
     class iptv-qos-lowbandwidth
          bandwidth 512
     class default
          fair-queue
```

RST-2T07—Module 6v1
9744_05_2004_X

105

- **Adding Hoot-n-Holler**
  - The drawing above shows the configuration details associated with adding QoS for Hoot-n-Holler traffic to the existing QoS that was defined for IP/TV flows.
    - The configuration shown in the drawing above uses **access-list 102** to identify the Hoot-n-Holler flows in the 239.255.255.0/24 address range.
    - A new set of **class-map** commands have been added that maps all flows in the 239.232.224.0/12 address range to the **hootie** class.
    - The **policy-map** commands have been updated and now also assigns all Hoot-n-Holler flows in the **hootie** class a bandwidth of 495Kbps to insure that these flows have a guaranteed amount of bandwidth on the T1 WAN links.



- **Multicast Music-on-Hold**
  - Multicast Music on Hold (MMoH) is an extension to the Cisco Call Manager that permits music on hold audio programs to be delivered from the Music on Hold Server via IP Multicast instead of individual point-to-point connections.

    The drawing above demonstrates how this feature works.
    - Extensions 1000 and 2000 have an established call when extension 1000 goes on hold.
    - When this happens the Call Manager sends a message to the IP Phone at extension 2000 letting it know that the other end of the call has gone on Hold and that it should join the MMoH multicast group in order to play the music on hold stream. The IP Phone at extension 2000 sends an IGMP join for this multicast group and the music on hold program is played to the caller at this extension.
    - At the same time, extensions 1001 and 2001 have an established call when extension 1001 goes on hold.
    - The Call Manager informs the IP Phone at extension 2001 that the other end has gone on hold and instructs the phone to join the MMoH group. The IP Phone at extension 2001 also joins the MMoH multicast group and begins to play the music on hold multicast stream.
  - In the example above, considerable bandwidth savings can be gained as only a single music-on-hold stream is being delivered to the remote site.

Multicast Music-on-Hold (MMoH)

Cisco.com

- **Increment on IP Addresses as opposed to Port Numbers**
- **Modify the "Max Hops" (TTL) default value of 2**
  - **Adjust according to the network topology and hop count to the receiver**
- **Use administratively scoped addressing for the MMoH address range**
  - Note: CSCdv01308 - Default Multicast MoH IP address should not be 239.0.0.0 – Fixed in 3.1(3)
- **Use G.729 for low-bandwidth sites – WARNING - LOW QUALITY**
- **Know how many audio sources have been configured for IP Multicast**

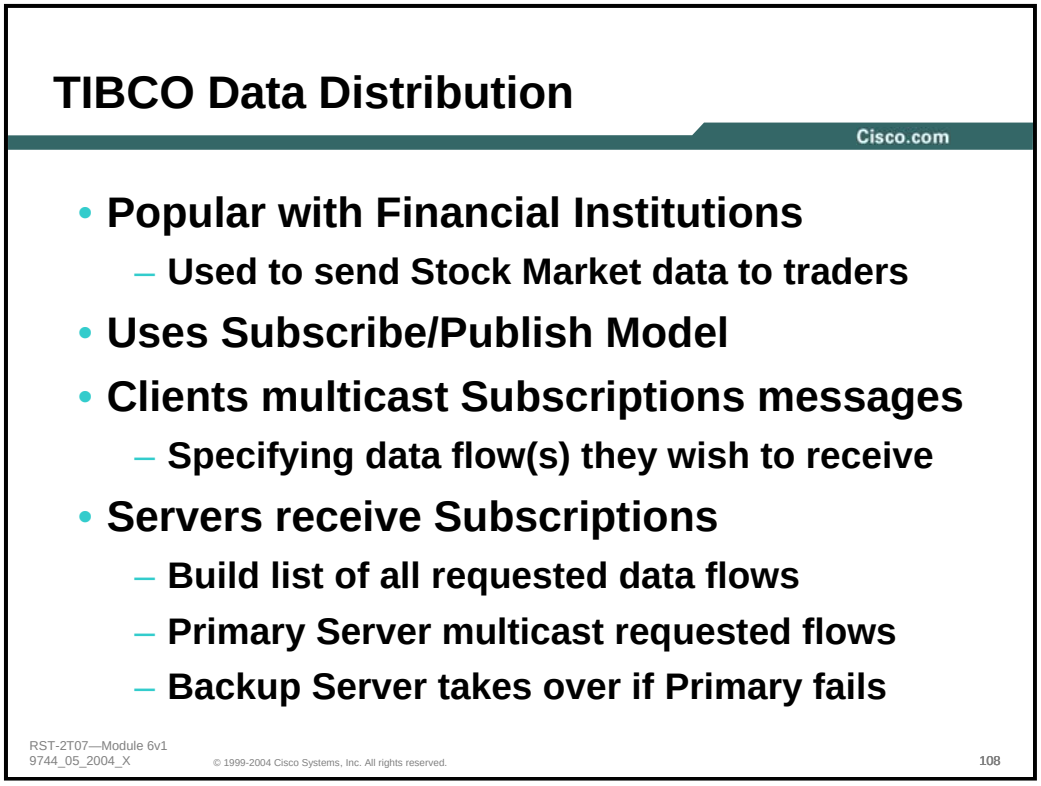
CODEC	Multicast Address
G.711ulaw	239.192.240.1
G.711alaw	239.192.240.2
G.729	239.192.240.3
Wideband	239.192.240.4

51 possible audio sources
X
4 Multicast addresses per source
=
204 multicast addresses consumed

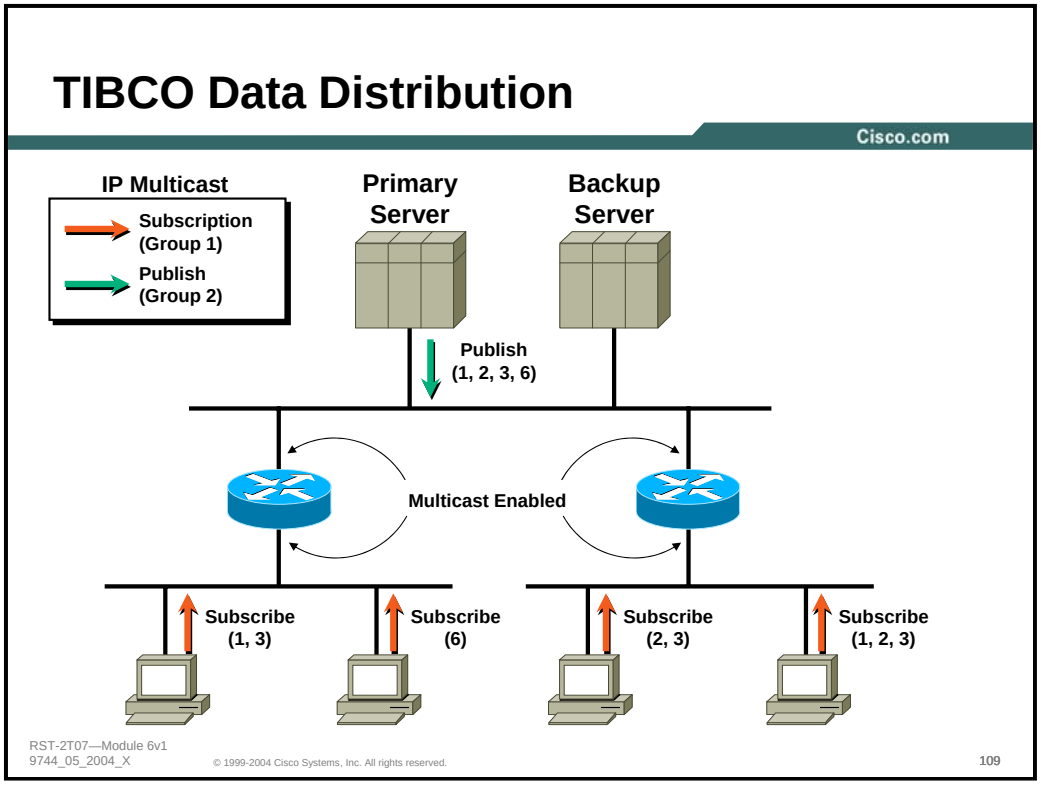
RST-2T07—Module 6v1
9744_05_2004_X

107

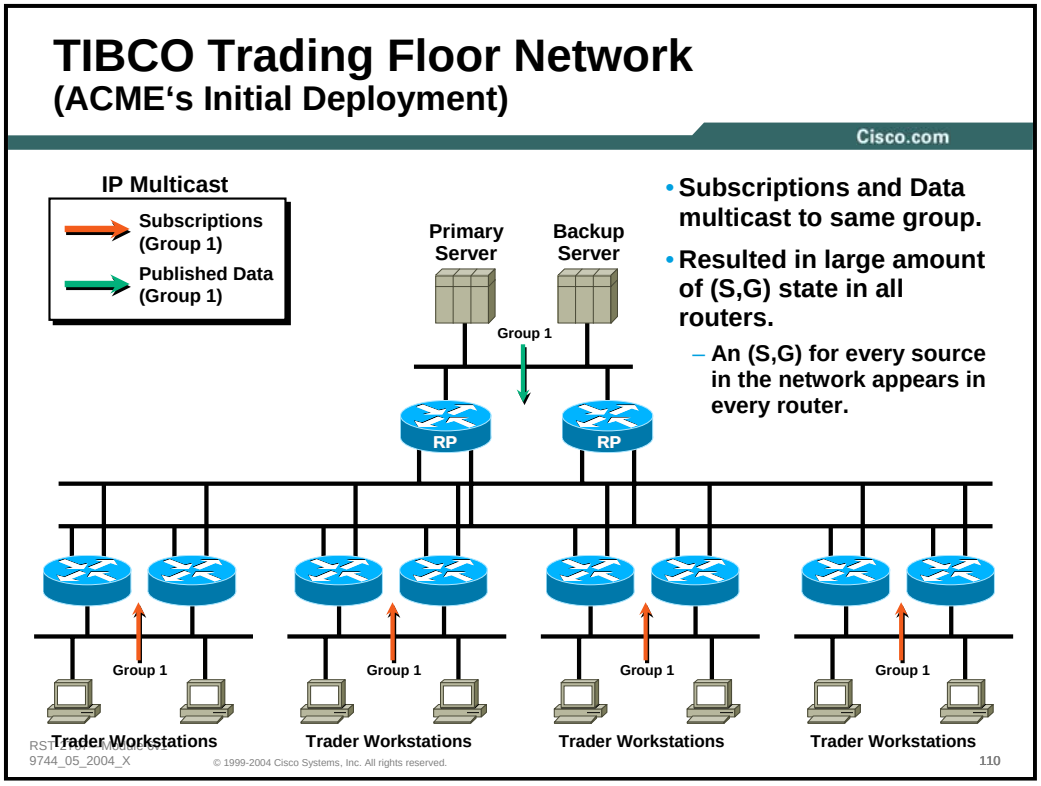
- When you configure the MoH server to increment multicast on IP address, the starting address is for the G.711ulaw CODEC and the addresses are incremented by 1.
- It is important to know how large the address requirements can get. For every file that is added as a multicast audio source, four addresses are consumed (one for every CODEC). The current maximum for audio sources on a single MoH server is 51 (50 file sources and 1 fixed source). So you could indeed have a requirement for 204 multicast addresses (51 files x 4 CODECs).
- Another consideration for making sure you increment multicast on IP address and not port number is that IP Multicast routers understand (S,G) notations and not port numbers.
- An example of why you should not use increment on port number is as follows.
- Let's say you use the increment on port number option for your audio sources and you want to use G.729 for the branch office phones. When the branch office IP Phone issues an IGMPv2 Join for the group associated with the audio source, for example 239.192.240.1, all four streams would be sent to the branch.
- Since we have four streams with the same group address (different port numbers ignored), the router at the branch pulls all four streams even though we wanted only one stream
- The result would be ugly given that you could have up to 51 sources streaming 480Kbps each.
- Change the base multicast IP address to the range you have identified from the administratively scoped range
- Change the default max hops (TTL) to reflect your routed network



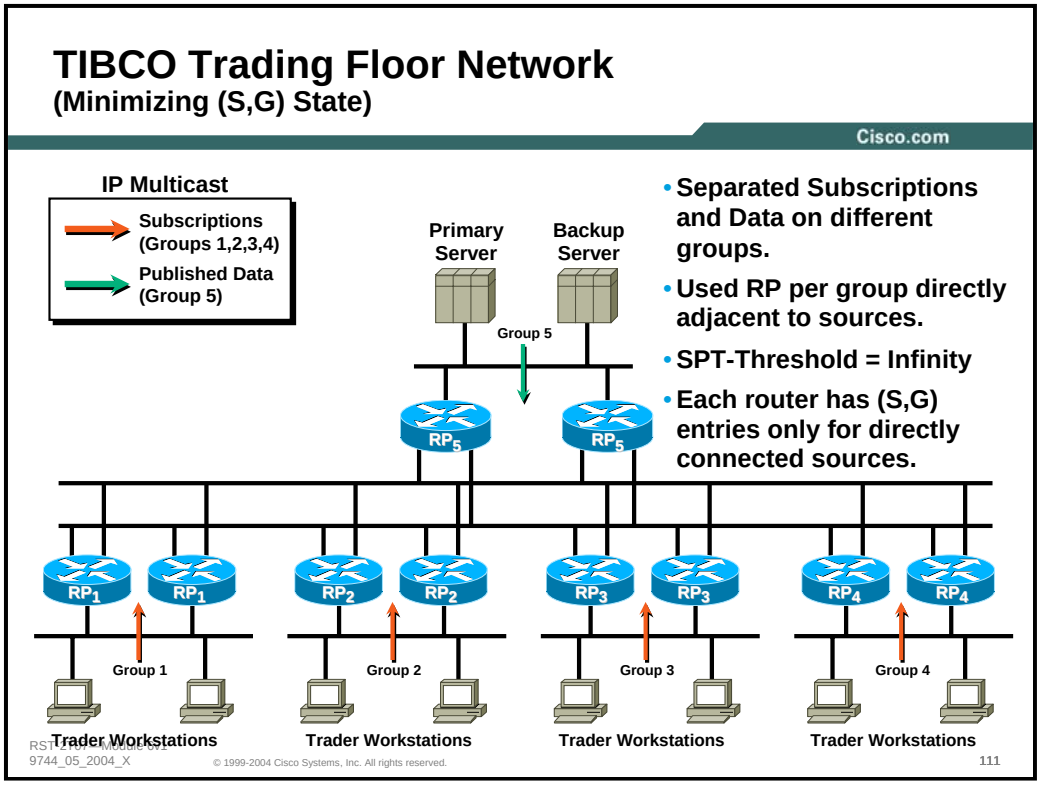
- **Tibco Data Distribution**
  - The Tibco “Rendezvous” (RV) product is often used at Financial Institutions to deliver Stock Market “ticker” data feeds to the traders.
    - Tibco Rendezvous employs a “Subscribe/Publish” model where client workstations multicast “Subscription” messages to the Rendezvous Data Server. These subscriptions tell the Server what data is desired by the clients in the network.
    - The Tibco Data Server then collects and merges all data subscriptions and builds a data set that includes all requested data. This data set is then multicast by the Server and received by all the clients in the network.
    - The use of multicast allows a backup Data Server to be operating in standby mode at all times and to also detect when the Primary Server has failed and begin transmitting the critical data flows to the clients.



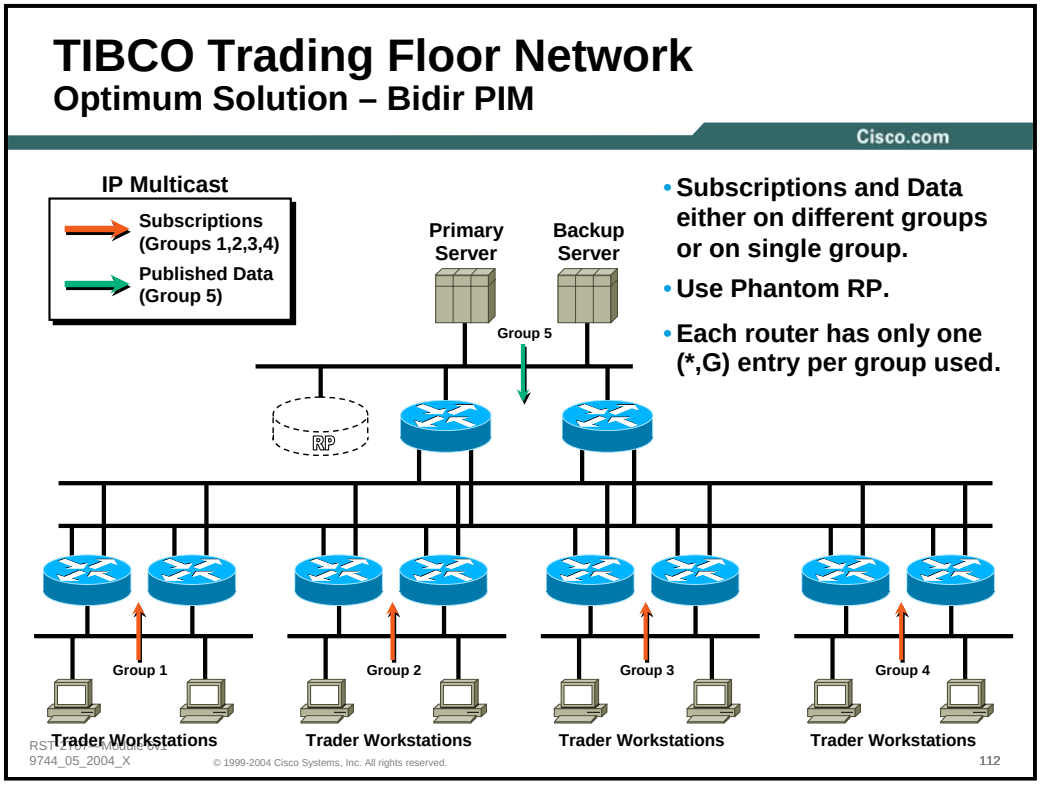
- **Tibco Data Distribution**
  - The drawing above shows an example of the Tibco Subscribe/Publish model in action.
    - Tibco client workstations at the bottom of the drawing are multicasting their Subscription messages to the Data Servers. Each subscription contains the set of requested data that the client wishes to receive. (These data are denoted in the drawing above by the use of numbers to represent different groups of data.)
    - The Subscription messages are received by the Primary and Backup data servers and collected into a complete set of all requested data. (In this case, data groups 1-3 and 6 have been requested by the clients in the network.)
    - The Primary server “Publishes” the requested data via IP Multicast back to the clients.
    - Notice that in this example, the Tibco configuration has been set up so that all Subscription messages are transmitted on one multicast group while the data are published on a separate multicast group.



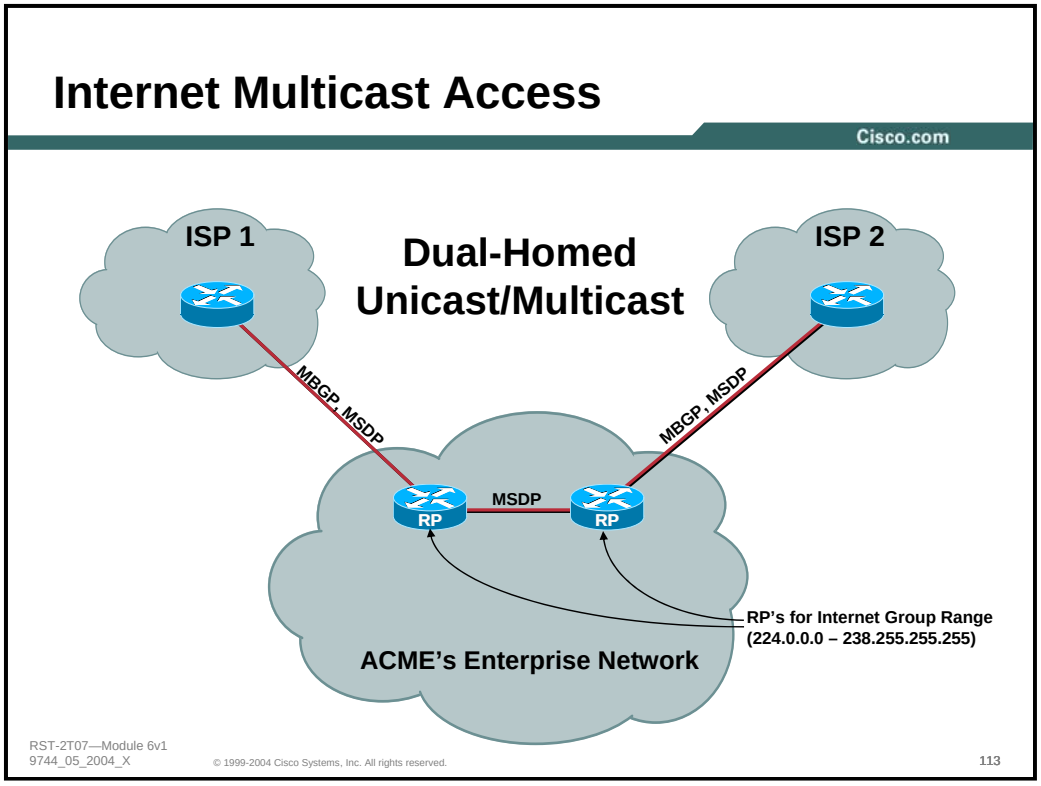
- **ACME's Initial Deployment**
  - When ACME first deployed their Market Feeds using the Tibco RV application, they used a simple two group model through-out the network.
    - All Subscriptions were multicast to Group 1.
    - All Data was published to Group 2.
    - One set of RP's (Primary and Backup) were used for all multicast groups.
  - This worked fine in the initial pilot program but as the number of Traders using this new application grew, so did the amount of (S,G) multicast state through-out the network. (Each trader workstation became a multicast source.) At some point, the router performance began to suffer as the routers worked harder and harder to maintain this growing amount of (S,G) state in this many-to-few multicast environment.



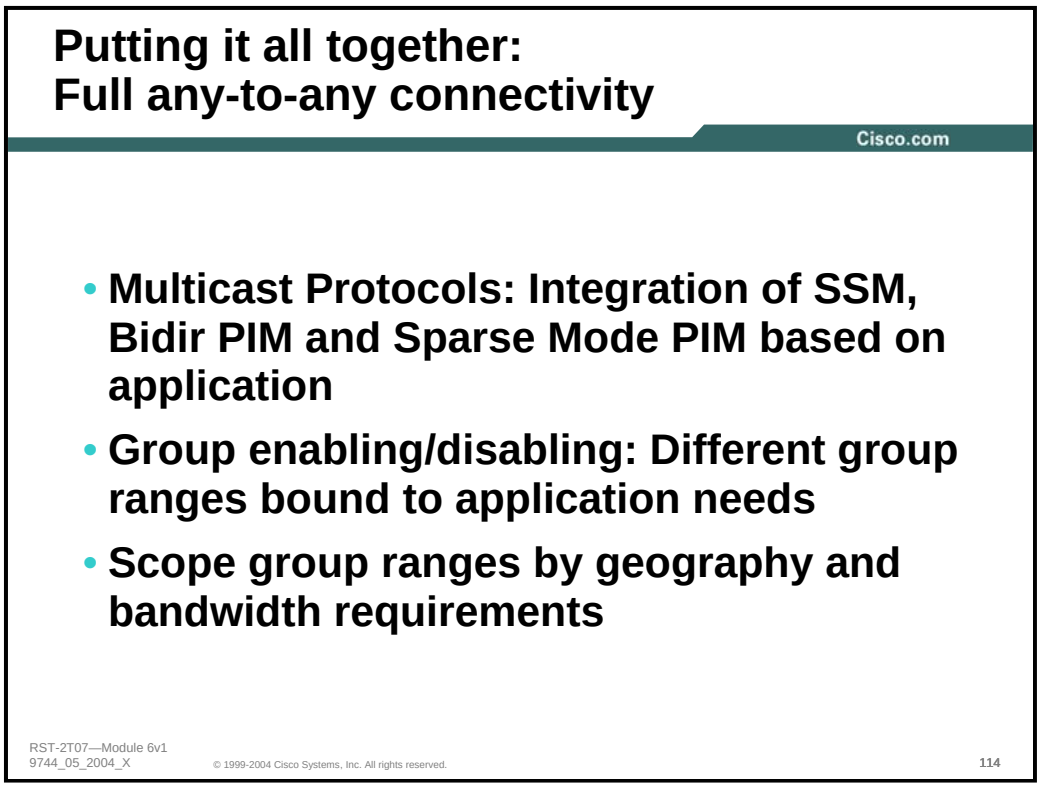
- **Minimizing (S,G) State**
  - In order to reduce the workload on any one router in the network, the ACME network engineers modified their Tibco RV configuration as well as the network operational model so that the RP for any multicast group was directly connected to all sources that sent to that group. The goal here was to reduce the amount of (S,G) state along the SPT to the RP. This was accomplished as follows:
    - Trader workstations on each building floor were configured to multicast their Subscription messages to a unique multicast group. The RP function for this group was assigned to the routers that directly connected the workstations on that floor to the network core. This model was replicated for every floor which resulted in four Subscription multicast groups.
    - The configuration of the data servers was modified so that they would join all 4 Subscription groups in order to receive all of the client subscription messages.
    - The data “publish” multicast group was left as it was previously and the RP function for this group was assigned to the routers directly connected to the data servers.
    - The SPT-Threshold on all routers was set to “Infinity” to prevent routers from joining the Shortest-Path Tree to the sources which would dramatically increase the amount of (S,G) state in the network.
  - This optimization resulted in a significant reduction in the amount of (S,G) state in the network. Each RP in the network would only have the number of (S,G) entries that corresponded to the number of trader workstations directly connected to it.



- **Optimum Solution – Bidir PIM**
  - Eventually, even the previous optimization failed to scale for the ACME network as the number of ACME traders at locations around the world increased into the 10,000's. The solution was to migrate the Tibco RV application to use Bidir PIM.
  - Once the migration to Bidir was complete, ACME's Tibco RV application was initially able to reduce the thousands of (S,G) multicast route entries distributed around the network down to 5 (*,G) multicast routing entries.
    - NOTE: In the future, ACME network and application engineers may reduce this even further by placing ALL Tibco RV traffic (both Subscriptions and Published data) on a single Bidir Shared tree. This will reduce the total state in the network down to a single (*,G) Bidir multicast routing entry.



- **Internet Multicast Access**
  - ACME Financials recently began offering their Stock Market data flows to their commercial customers via Internet multicast. This was accomplished as follows:
    - Since ACME is dual-home to the Internet via two different Service Providers that are multicast capable, they configured two RP's using the Anycast RP method for the 224.0.1.0 – 238.255.255.255 Internet multicast group range. This required the two Anycast RP's to be interconnected via the Multicast Source Discovery Protocol (MSDP).
    - These two RP's were then connected to MSDP peers, one in each Service Provider network as shown above. This permitted ACME to receive "Source Active" (SA) messages regarding other multicast sources in the Internet as well as to distribute SA messages regarding the address of the Stock Market data server that multicasts stock ticker information to ACME's customers elsewhere in the Internet.
    - Finally, the ACME border routers along with these two RP's, were configured to run MBGP so that multicast routing information could be sent/received to/from the Service Providers.



- **Putting it all Together**
  - ACME network engineers have made judicious use of all forms of PIM multicast routing including Bidir PIM, Source Specific Multicast and standard PIM-SM.
    - Applications are bound to the appropriate multicast group range depending on which of the above models best suits the application needs.
  - Admin Scope Ranges are applied by geography and bandwidth requirements to optimize network and application performance.

Recommended Reading

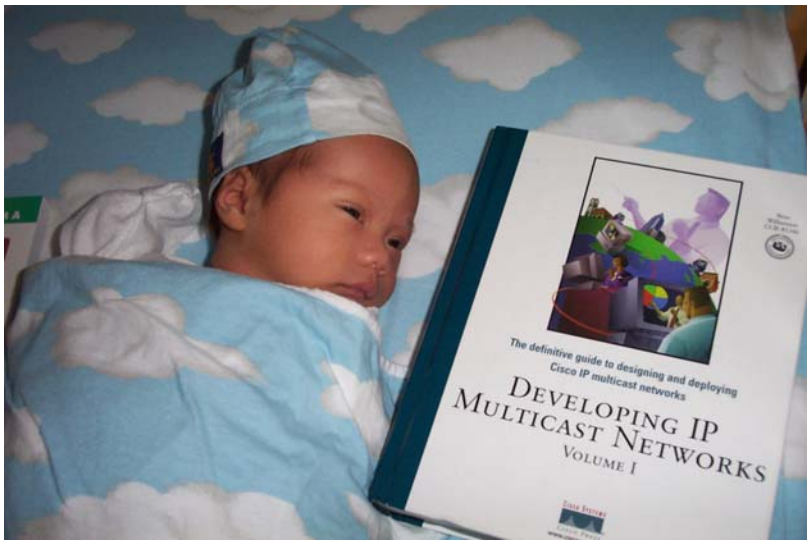
Cisco.com

- Continue your Networkers learning experience with further reading for this session from Cisco Press.
- Check the Recommended Reading flyer for suggested books.

Available on-site at the Cisco Company Store

Wonderful Bedtime Stories

Cisco.com



Complete Your Online Session Evaluation!

Cisco.com

WHAT: Complete an online session evaluation and your name will be entered into a daily drawing

WHY: Win fabulous prizes! Give us your feedback!

WHERE: Go to the Internet stations located throughout the Convention Center

HOW: Winners will be posted on the onsite Networkers Website; four winners per day