NETWORKERS 2004

Deploying IP Multicast

Session RST-2701

Agenda

Cisco.com

- Basic Multicast Engineering
  - PIM Configuration Steps
  - Which Mode: Sparse or Dense?
  - Basic RP Engineering
- Advanced Multicast Engineering
  - PIM Protocol Extensions
  - Combining Auto-RP and Anycast RP
  - Multicast Group Control
  - Using Admin. Scoped Zones

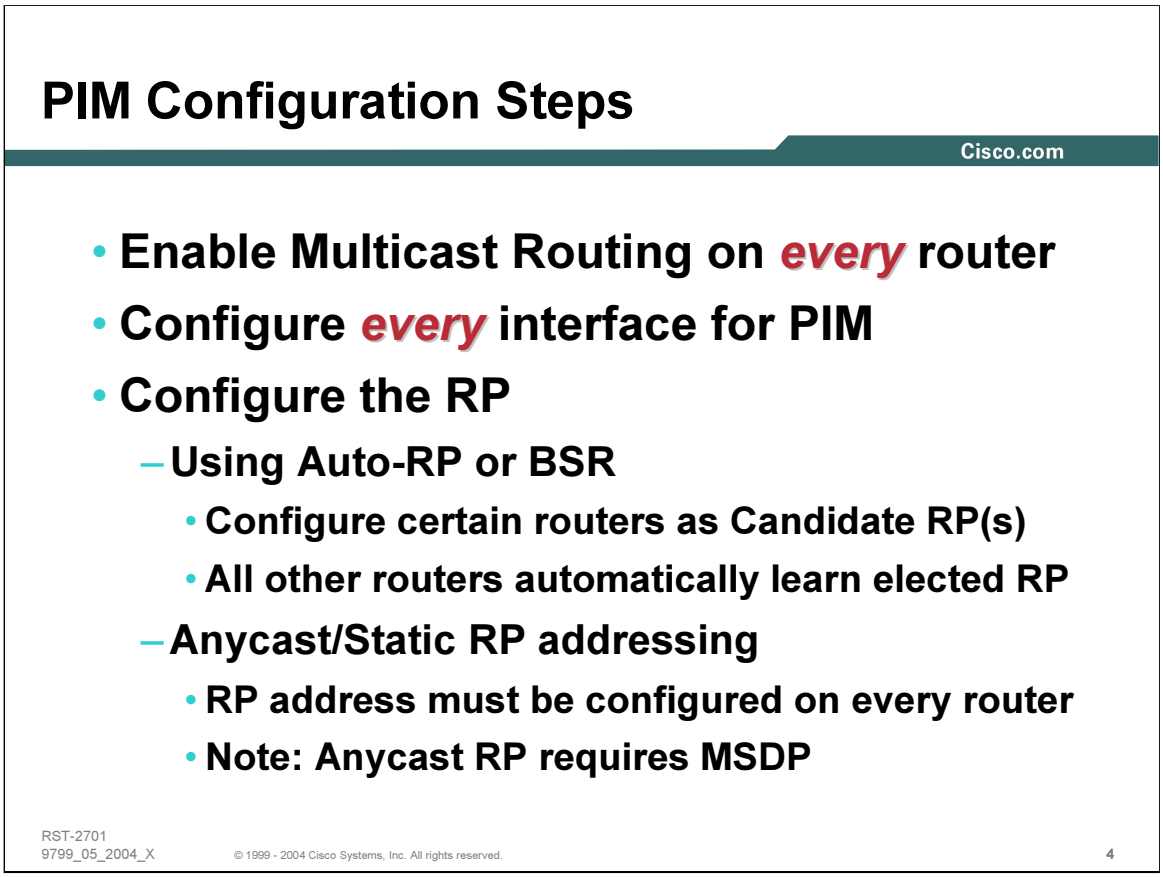
RST-2701
9799_05_2004_X

2

Basic Multicast Configuration – PIM Configuration Steps

RST-2701
9799_05_2004_X

3



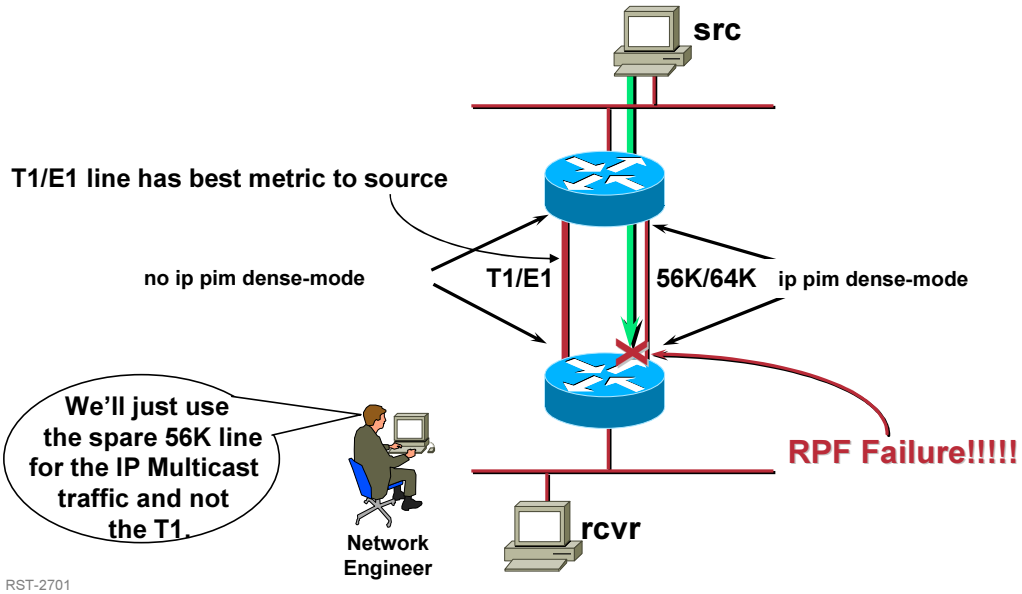
- **Enable Multicast on every router**
  - This is necessary because the PIM RPF (Reverse-Path Forwarding) check depends on the unicast route table to calculate the proper incoming interface for each multicast forwarding entry. Unless extraordinary measures have been taken in the network design (using 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.

Configure PIM on Every Interface

Cisco.com

Classic Partial Multicast Cloud Mistake #1



RST-2701
9799_05_2004_X

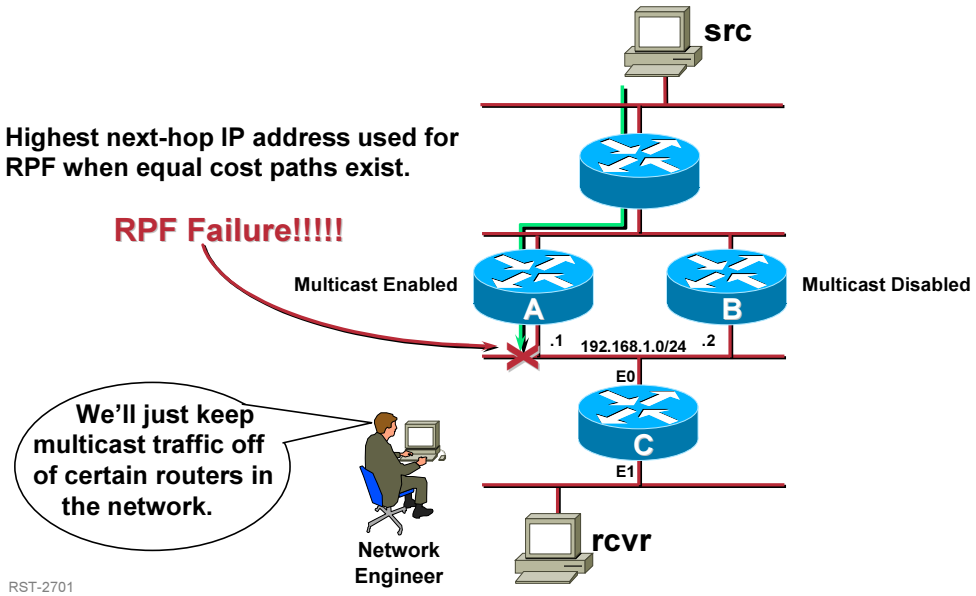
5

- **Classic Mistake #1**
  - In this example, our network engineer has either decided to "try before he buys" (i.e. he has decided to perform a limited test of multicast) or he has chosen to try to traffic engineer the flow of multicast traffic on the initial deployment of multicast.
  - Note that there are two links between the two routers, one being the primary link and the other a slower backup link. In an attempt to avoid possible impact to unicast traffic flowing over the primary link, the network engineer has only enabled PIM on the low-speed, backup link.
  - The problem is that the normal RPF calculation will (unless extraordinary steps are taken) use the unicast routing table to calculate the best path back to the source and use that as the RPF or "Incoming Interface" for the traffic. Because the primary link is the higher speed link, it will have a better metric than the low-speed link. Therefore the high-speed link will be the computed RPF interface and not the low-speed link thereby causing the multicast traffic arriving via the low-speed link to be discarded due to a failure of the RPF check. (I.e The traffic is arriving on the wrong interface.

Configure PIM on Every Router

Cisco.com

Classic Partial Multicast Cloud Mistake #2



RST-2701
9799_05_2004_X

6

- **Classic Mistake #2**
  - Once again our network engineer is attempting to "traffic engineer" the flow of multicast traffic through the network by only enabling multicast on a subset of the routers in the network. Note that in this example router A is multicast enabled but not router B.
  - In this case, we have equal cost metrics back to the source via routers A and B. Thus the RPF calculation on router C will certainly calculate interface **E0** as the RPF interface and would accept multicast traffic from either router A or router B. However, the RPF neighbor is also part of the RPF calculation and for this flow of multicast traffic it would resolve to router B. This is because multicast cannot have multiple "best paths" back to a source like unicast can. Instead, a single path must be chosen as the RPF path back to the source. However, since the metrics through routers A and B are equal, we cannot use routing metrics to resolve this tie. Therefore, the RPF calculation will use the highest of the IP addresses of the next hop neighbors to break the tie. In this case, that is router B.
  - The reason that this scenario causes problems is that any PIM control traffic that router C needs to send up the source tree will be sent to router B and not router A. Since router B is not multicast enabled, this can cause problems in building or maintaining the distribution tree.

Which Mode: Sparse or Dense?

RST-2701
9799_05_2004_X

7

Which Mode—Sparse or Dense

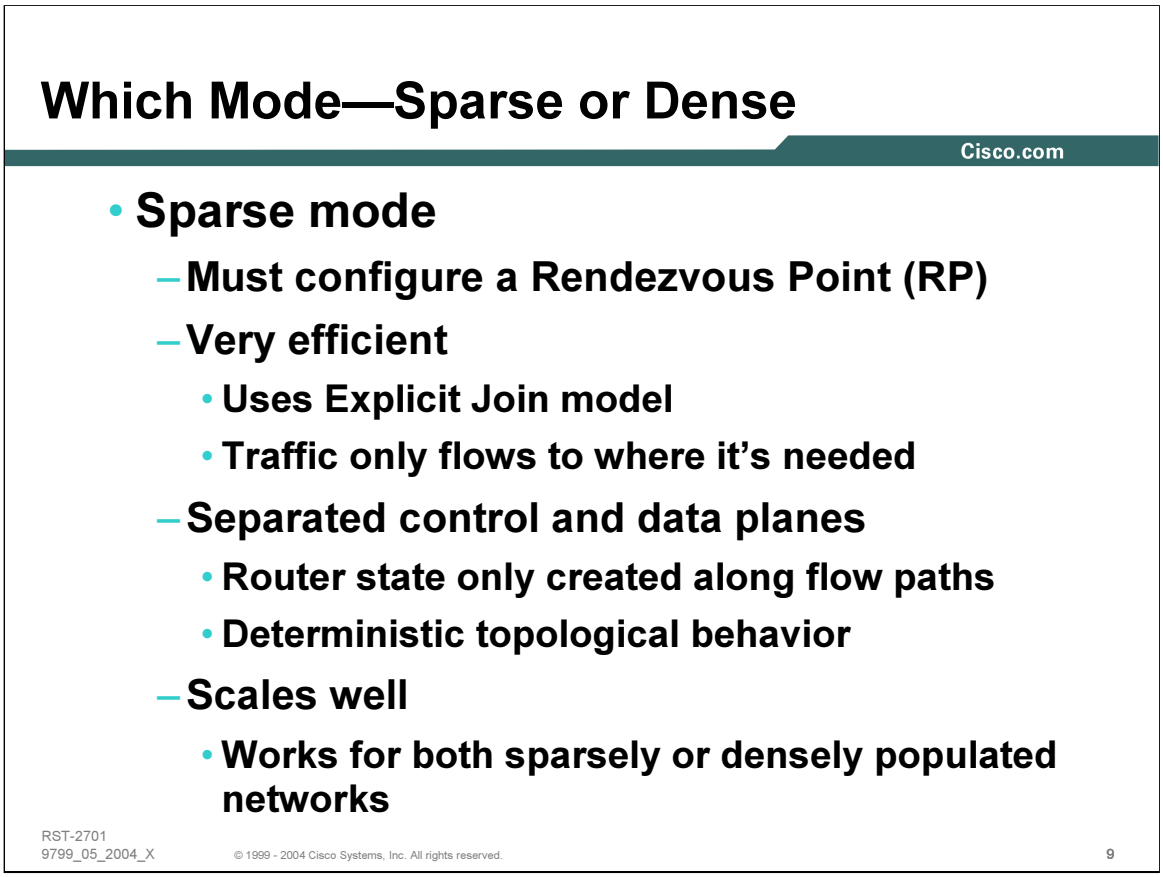
Cisco.com

- **Dense mode**
  - **Flood and Prune behavior very inefficient**
    - **Can cause problems in certain network topologies**
  - **Creates (S, G) state in EVERY router**
    - **Even when there are no receivers for the traffic**
  - **Complex Assert mechanism**
  - **Mixed control and data planes**
    - **Results in (S, G) state in every router in the network**
    - **Can result in non-deterministic topological behavior**
      Read: It can black-hole traffic and/or melt down your network!
  - **Primarily usage:**
    - **Testing a router's performance in the lab**

RST-2701
9799_05_2004_X

8

- **Dense Mode**
  - Dense mode uses a "Flood and Prune" behavior to build the multicast distribution trees. Basically, multicast traffic is flooded away from the source and (at least initially) flows down every link and reaches every point in the network. Routers with no downstream receivers will send back Prune messages to cut off the flow of unwanted traffic. These "prunes" shutoff the flow for roughly 3 minutes at which time they expire and the traffic is reflooded. Therefore, traffic is flooded throughout the network every 3 minutes resulting a periodic "Flood and Prune" cycle which is very inefficient.
  - Because the multicast traffic reaches every router in the network, it causes every router to create and maintain (S, G) multicast forwarding state regardless of whether there are any receivers in the network.
  - In order to prune off redundant links, the PIM Assert mechanism is used extensively by PIM Dense mode. The complexities of this mechanism and Dense mode's dependence on it can sometimes result in traffic being "black-holed".
  - Since the Dense mode distribution trees are built as a result of data arrival, there is no true Control Plane to the protocol. This results in non-deterministic network behavior which has been known to result in "black-holes" and multicast route-loops that have actually melted down a few networks.
  - As a result of the above, Dense mode typically only suited for testing the multicast forwarding performance of a single router in a test environment.



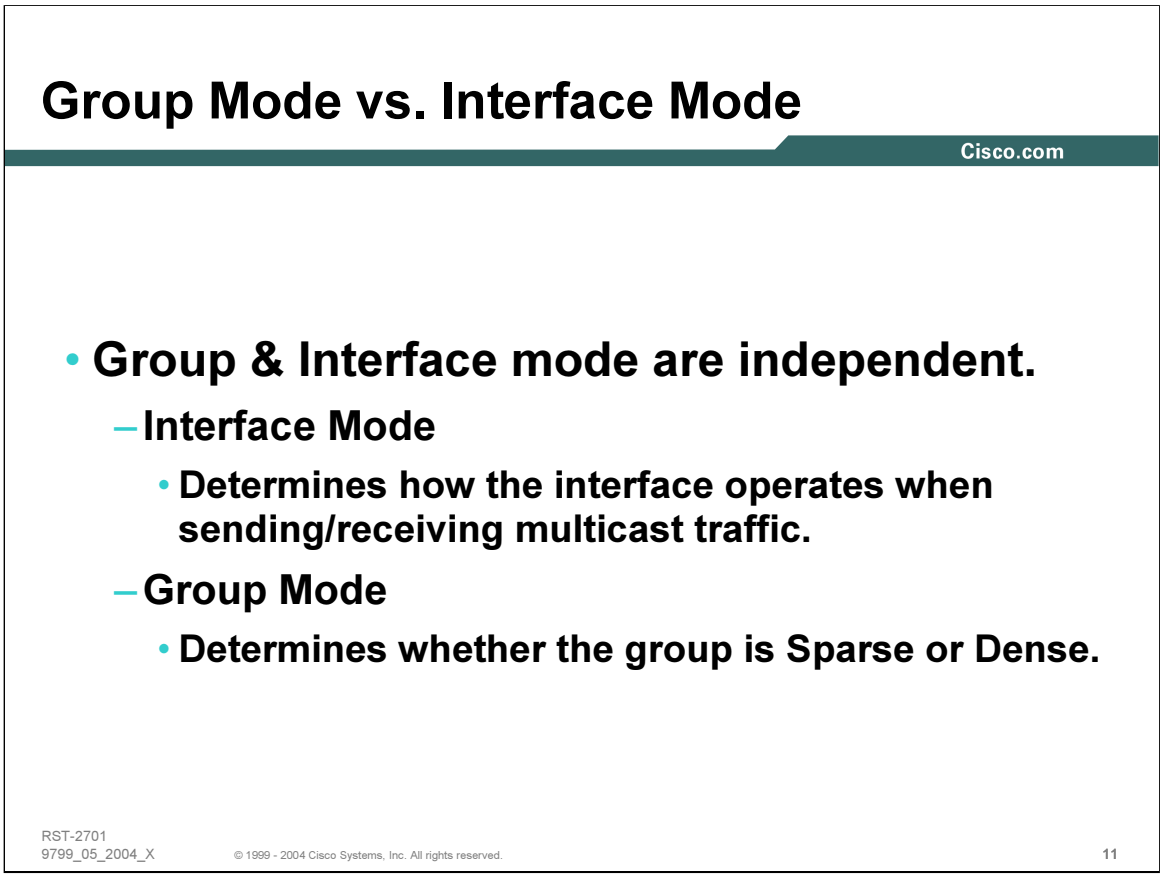
- **Sparse Mode**
  - Only requires one extra simple configuration step and that is to configure a Rendezvous Point (RP). This can be accomplish via a variety of means, all of which will be discussed in a later section. Therefore, it is basically as easy to deploy Sparse mode as it is to deploy Dense mode.
  - Sparse mode is much more efficient than Dense mode as it uses an explicit join model to request the flow of multicast traffic to the receivers. Using this model, PIM Joins are sent to build the branches of the distribution tree thus traffic is only forwarded along the paths necessary to reach the receivers. This also means that forwarding state (i.e. (*,G) and (S,G) mroute entries) are generally only created where multicast is actually flowing to a receiver and also nearly eliminates the use of the PIM Assert mechanism which can sometimes cause problems.
  - Because Spares mode uses explicit Join and Prune messages to build, maintain and tear-down the distribution trees, the tree maintenance is much more deterministic than Dense mode. This means that Sparse mode is much, much better at detecting network topology failures and immediately rerouting the branches of the distribution tree around failures.
  - Finally, because of its more deterministic behavior, lack of Asserts and multicast forwarding state reduction properties, Sparse mode scales much better than Dense mode.

    Note: Don't get caught by the classic myth that states, "If you have a dense population of receivers [i.e. a receiver on most subnets in the network] then you should use a Dense mode protocol." This is just not true. Sparse mode works well for both densely and sparsely populated networks.

CONCLUSION

"Sparse mode Good! Dense mode Bad!"

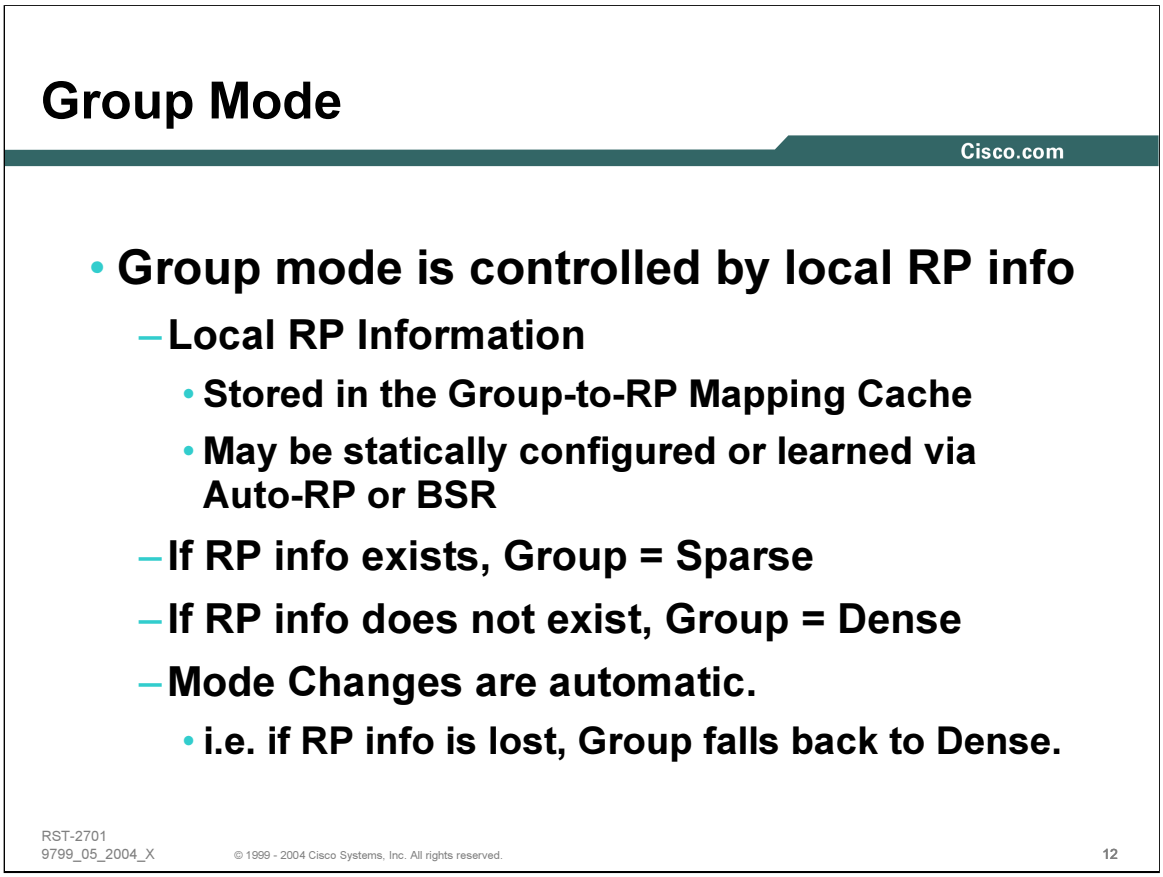
Source: "The Caveman's Guide to IP Multicast", ©2000, R. Davis



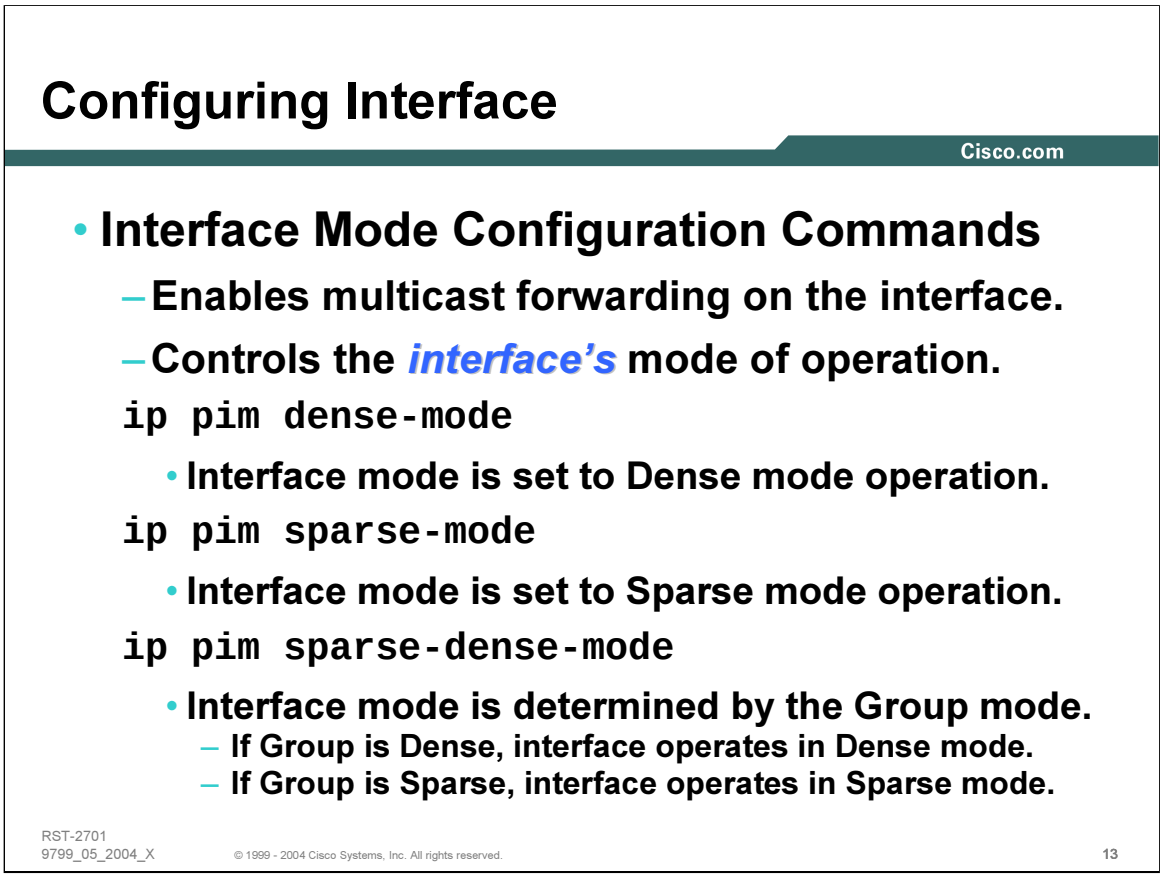
- **Group mode vs. Interface mode.**
  - These are totally independent concepts and often a source of confusion for newcomers to IP Multicast.
  - **Interface mode**

    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.



- **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



- **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.)

Basic RP Engineering – RP Configuration Methods

RST-2701
9799_05_2004_X

14

RP Configuration Methods

Cisco.com

- Static
- Auto-RP
- BSR
- Anycast-RP's

Static RP's

Cisco.com

- **Hard-coded RP address**
  - **When used, must be configured on every router**
  - **All routers must have the same RP address**
  - **RP fail-over not possible**
    - **Exception: If Anycast RPs are used. (More on that later.)**
  - **Group can never fall back into Dense mode.**

RST-2701
9799_05_2004_X

16

- **Hard-code RP Addresses**
  - Requires every router in the network to be manually configured with the IP address of a *single* RP.
  - If this RP fails, there is no way for routers to fail-over to a standby RP.

    The exception to this rule is if "Anycast-RP's" are in use. This requires MSDP to be running between each RP in the network.
- **Command**

  `ip pim rp-address <address> [group-list <acl>] [override]`
  - The 'group-list' allows a group range to be specified.

    The default is ALL multicast groups or 224.0.0.0/4

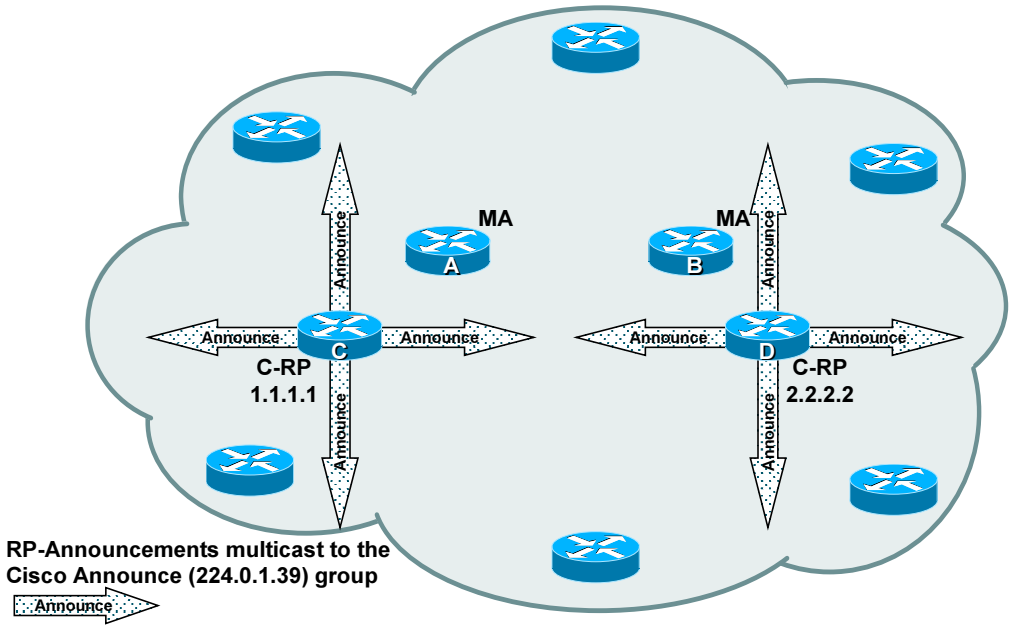
    DANGER, WILL ROBINSON!!!

    The default range includes the Auto-RP groups (224.0.1.39 and 224.0.1.40) which will cause this router to attempt to operate these groups in Sparse mode. This is normally not desirable and can often lead to problems where some routers in the network are trying to run these groups in Dense mode (which is the normal method) while others are trying to use Sparse mode. This will result in some routers in the network being starved of Auto-RP information. This in turn, can result in members of some groups to not receive multicast traffic.
  - The 'override' keyword permits the statically defined RP address to take precedence over Auto-RP learned Group-to-RP mapping information.

    The default is that Auto-RP learned information has precedence.

Auto-RP Overview

Cisco.com

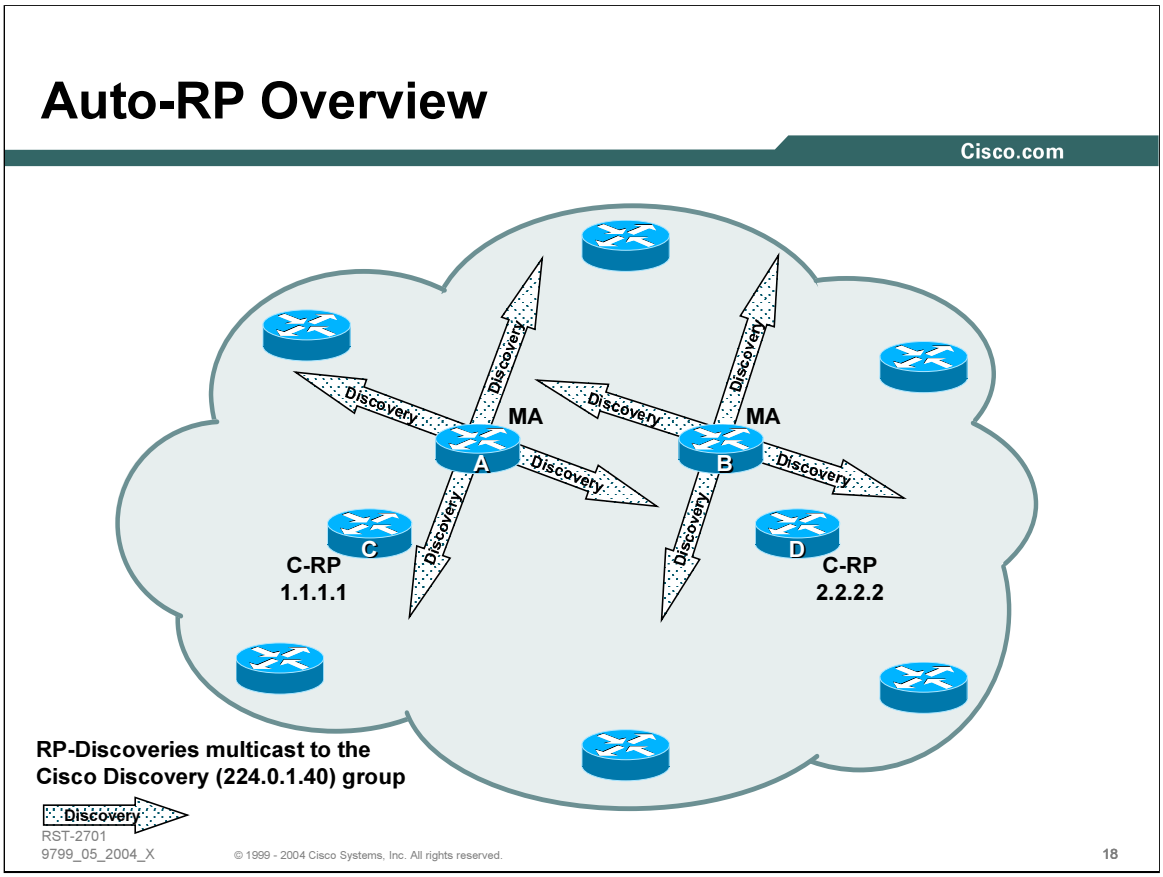


RST-2701
9799_05_2004_X

17

- **Auto-RP Overview**
  - The network administrator configures one or more routers in the network to serve as Candidate RP's (C-RPs).

    Candidate RPs "announce" their willingness to serve as RP for a particular group range by multicasting "Auto-RP Announce" messages to the Cisco Announce multicast group, 224.0.1.39. These messages are normally flooded throughout the network using Dense mode so that they reach every router in the network.



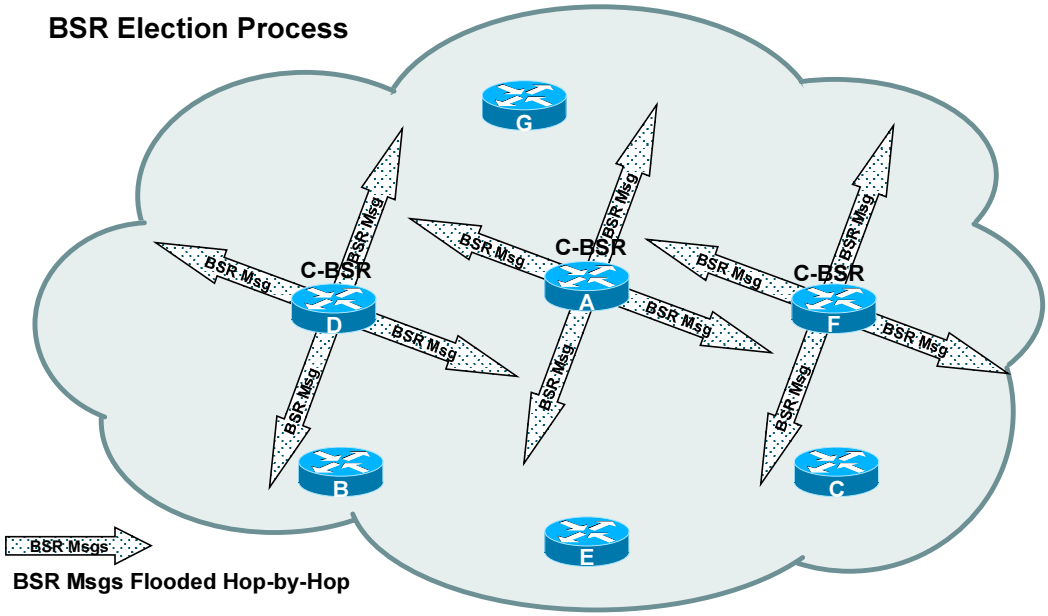
- **Auto-RP Overview**
  - The network administrator also configures one or more routers in the network to serve as Mapping Agents.

    Mapping Agents join the Cisco Announce multicast group so that they can receive the "Auto-RP Announce" messages from the Candidate RPs. Mapping Agents store the received announcements in their Group-to-RP mapping caches and elect the C-RP with the highest IP address as the currently active RP for each group range.

    Mapping Agents then announce the elected RP for each group range by multicasting "Auto-RP Discovery" messages to the Cisco Discovery multicast group, 224.0.1.40. These messages are normally flooded throughout the network using Dense mode so that they reach every router in the network.
  - All routers in the network automatically join the Auto-RP Discovery group so that they can receive the "Auto-RP Discovery" messages from the Mapping Agents. When they receive one of these messages, they store the elected RP information in their Group-to-RP mapping cache so that they know what RP is active for what group range.

BSR Overview

Cisco.com



RST-2701
9799_05_2004_X

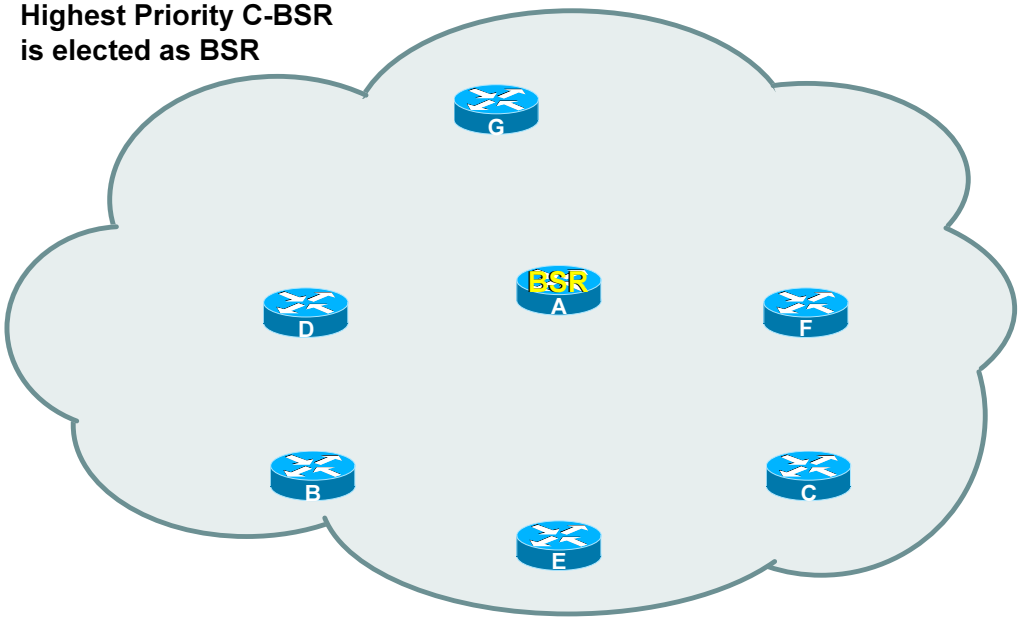
19

- **BSR Overview**
  - The network administrator configures one or more routers in the network to serve as Candidate BSR's (C-BSR).

    At network startup, all Candidate BSR's participate in the BSR election process by sending a PIM BSR message containing its BSR priority out all interfaces. These BSR messages are flooded hop-by-hop throughout the entire network.

BSR Overview

Cisco.com



RST-2701
9799_05_2004_X

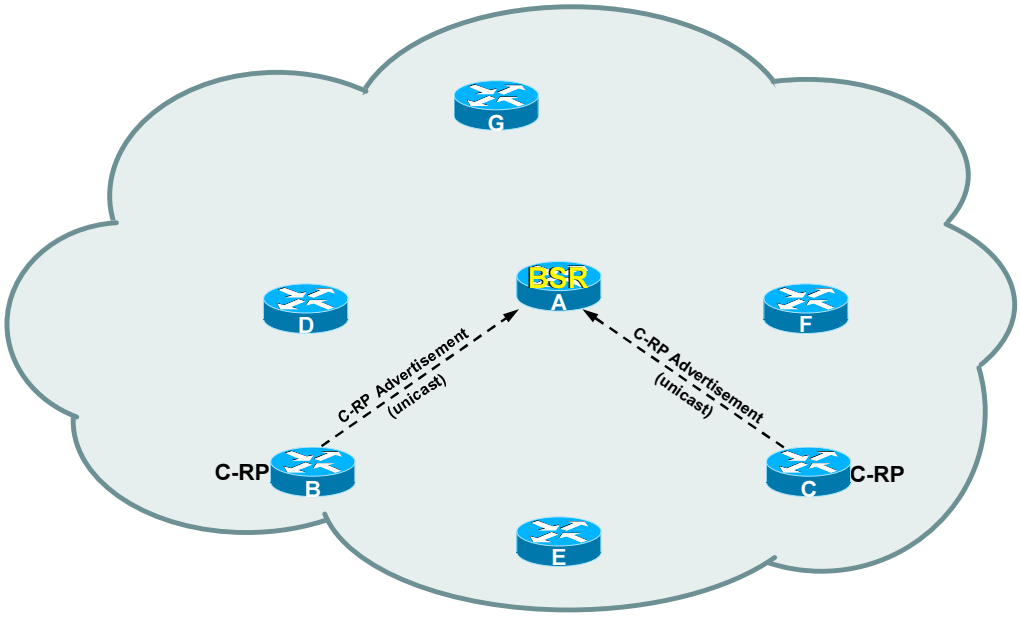
20

- **BSR Overview**
  - At the end of the “BSR-Election-Interval”, the BSR with the highest BSR priority is elected as the active BSR (Bootstrap Router).

    (Note: The BSR election process is similar in nature to the Root-Bridge election mechanism in the Spanning-Tree protocol.)

BSR Overview

Cisco.com



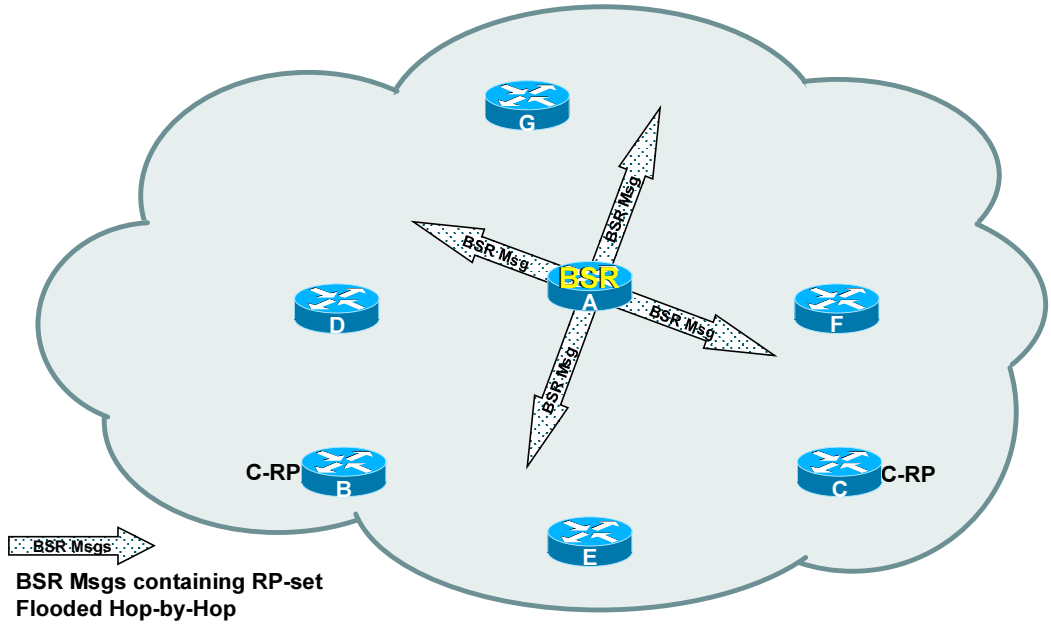
RST-2701
9799_05_2004_X

21

- **BSR Overview**
  - All routers in the network participate in the BSR Election process by forwarding BSR messages to their downstream neighbors. As a result, at the end of the BSR Election Interval, all routers in the network (including the routers that have been configured as Candidate RPs) know which C-BSR has been elected as the currently active BSR.
  - Since the Candidate RP's know the IP address of the currently active BSR, they can unicast their C-RP Announcement messages directly to the active BSR.

BSR Overview

Cisco.com



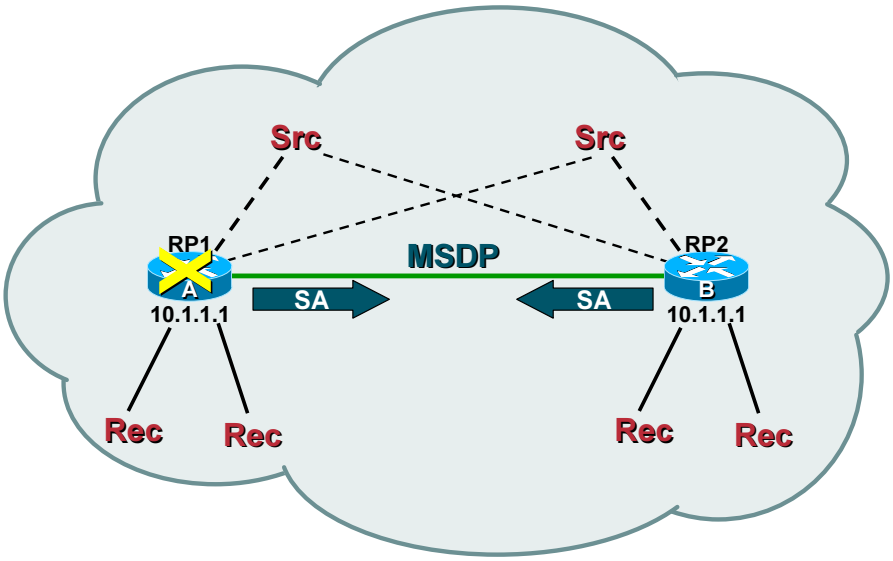
RST-2701
9799_05_2004_X

22

- **BSR Overview**
  - The active BSR stores all incoming C-RP Announcements in its Group-to-RP mapping cache. The BSR then sends the entire list of C-RP's from its Group-to-RP mapping cache in periodic BSR messages which are flooded hop-by-hop throughout the entire network. As each router receives a copy of these BSR messages, it updates the information in it's local Group-to-RP mapping cache so it knows the IP address of **all** C-RP's in the network.
  - However, unlike Auto-RP where the Mapping Agent "elects" the active RP for a group range and announces the election results to the network, the BSR does not "elect" the active RP for a group. Instead, it leaves this task to each individual router in the network.
  - Each router in the network will using a well-known hashing algorithm to elect the currently active RP for a particular group range. Since each router is running the same algorithm against the same list of C-RP's, then they will all elect the same RP for a particular group range.

Anycast RP—Overview

Cisco.com

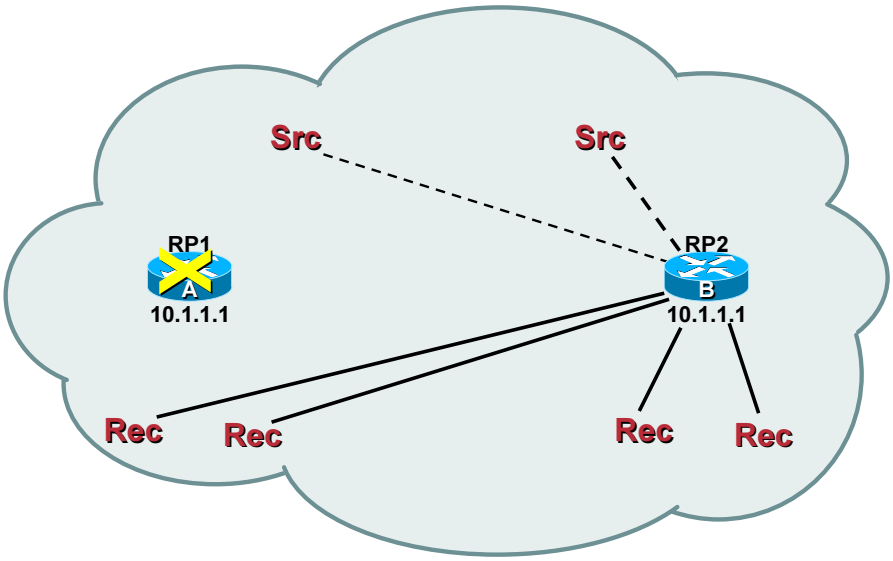


RST-2701
9799_05_2004_X

23

Anycast RP—Overview

Cisco.com



RST-2701
9799_05_2004_X

24

Basic RP Engineering – General RP Recommendations

General RP Recommendations

Cisco.com

- Use Anycast RP's:
  - When network must connect to Internet or
  - When rapid RP failover is critical
- Pros
  - Fastest RP Convergence method
  - Required when connecting to Internet
- Cons
  - Requires more configuration
  - Requires use of MSDP between RP's

General RP Recommendations

Cisco.com

- Use Auto-RP
  - When minimum configuration is desired and/or
  - When maximum flexibility is desired
- Pros
  - Most flexible method
  - Easiest to maintain
- Cons
  - Increased RP Failover times vs Anycast
  - Special care needed to avoid DM Fallback
    - Some methods greatly increase configuration

RST-2701
9799_05_2004_X

General RP Recommendations

Cisco.com

- Use BSR:
  - When Static/Anycast RP's cannot be used and
  - When maximum interoperability is needed
- Pros
  - Interoperates with all Vendors
- Cons
  - Increased RP Failover times vs Anycast
  - Special care needed to avoid DM Fallback
    - Some methods greatly increase configuration
  - Does not support Admin. Scoping

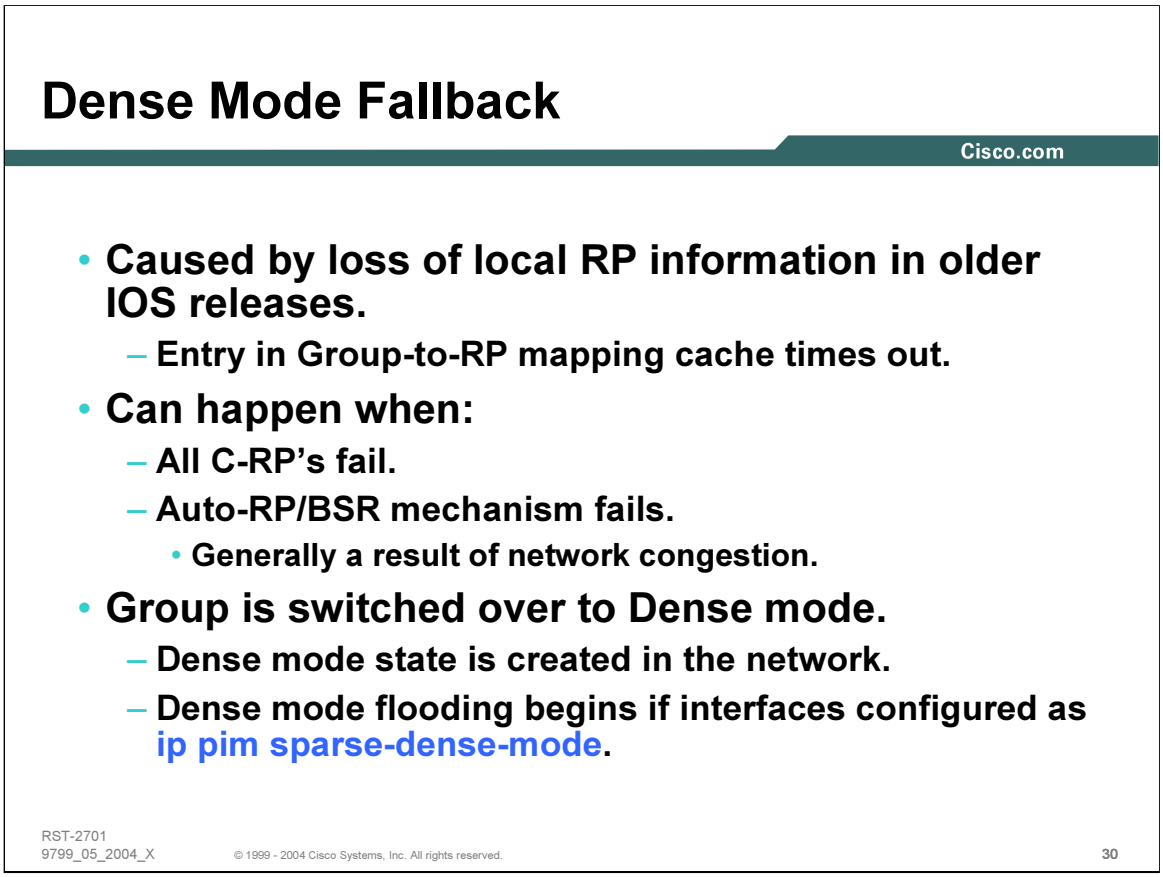
RST-2701
9799_05_2004_X

28

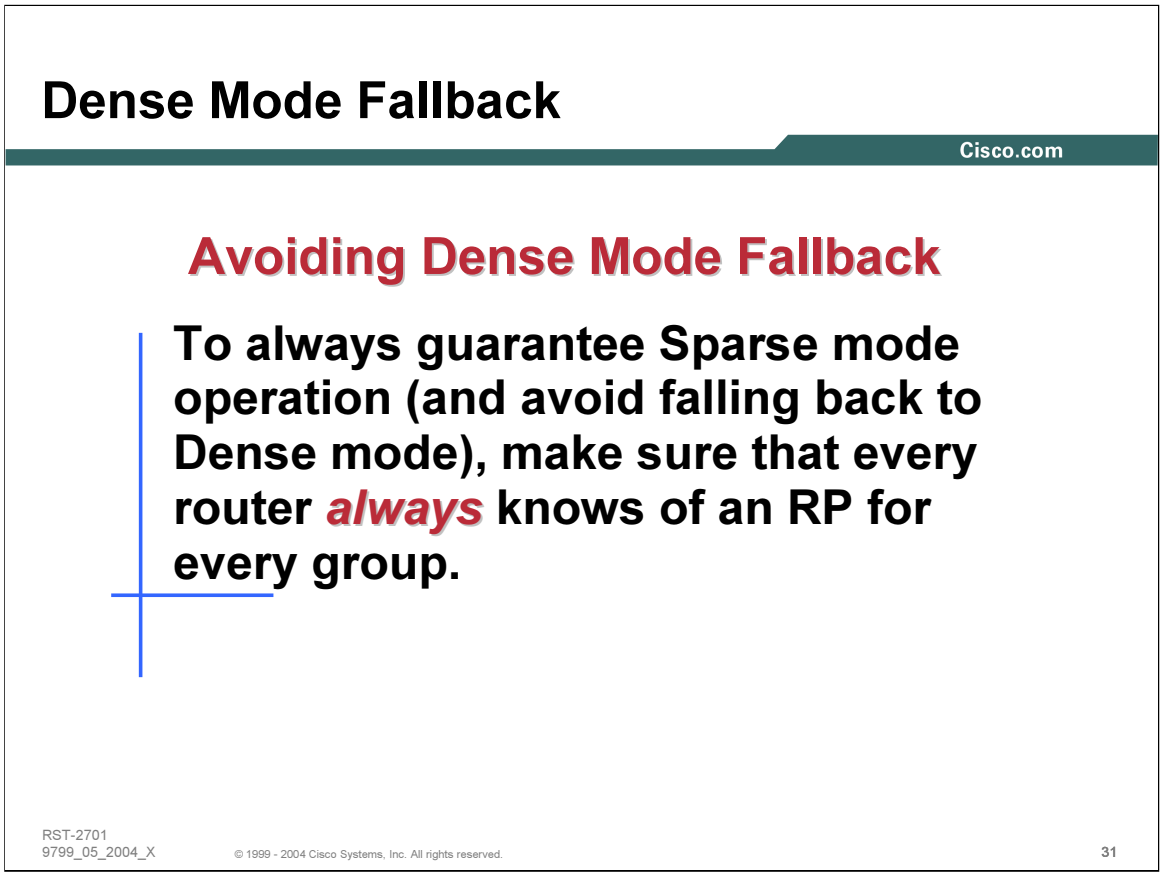
Basic RP Engineering – Avoiding Dense Mode Fallback

RST-2701
9799_05_2004_X

29



- **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:
    1. If all Candidate RP's fail
    2. 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-2701
9799_05_2004_X

32

- **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. (See example in the drawing above.)

Avoiding DM Flooding

Cisco.com

- New IOS global command

  `ip pim autorp-listener`

- Added support for Auto-RP Environments
  - Modifies interface behavior
    - Interface always uses DM for Auto-RP groups
    - Permits use of ip pim sparse-mode interfaces and Auto-RP.
  - Prevents DM Flooding
    - When ip pim sparse-mode used on interfaces.
  - Does not prevent DM Fallback!
- Available 12.3(4)T, 12.2(28)S

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-2701
9799_05_2004_X

Avoiding DM *Fallback*

Cisco.com

- New IOS global command

  `no ip pim dm-fallback`

- Totally prevents DM Fallback!!
  - No DM Flooding since all state remains in SM
- Default RP Address = 0.0.0.0 [nonexistent]
  - Used if all RP's fail.
    - Results in loss of Shared Tree.
    - All SPT's remain active.
- Available 12.3(4)T, 12.2(28)S

Advanced Multicast Engineering

RST-2701
9799_05_2004_X

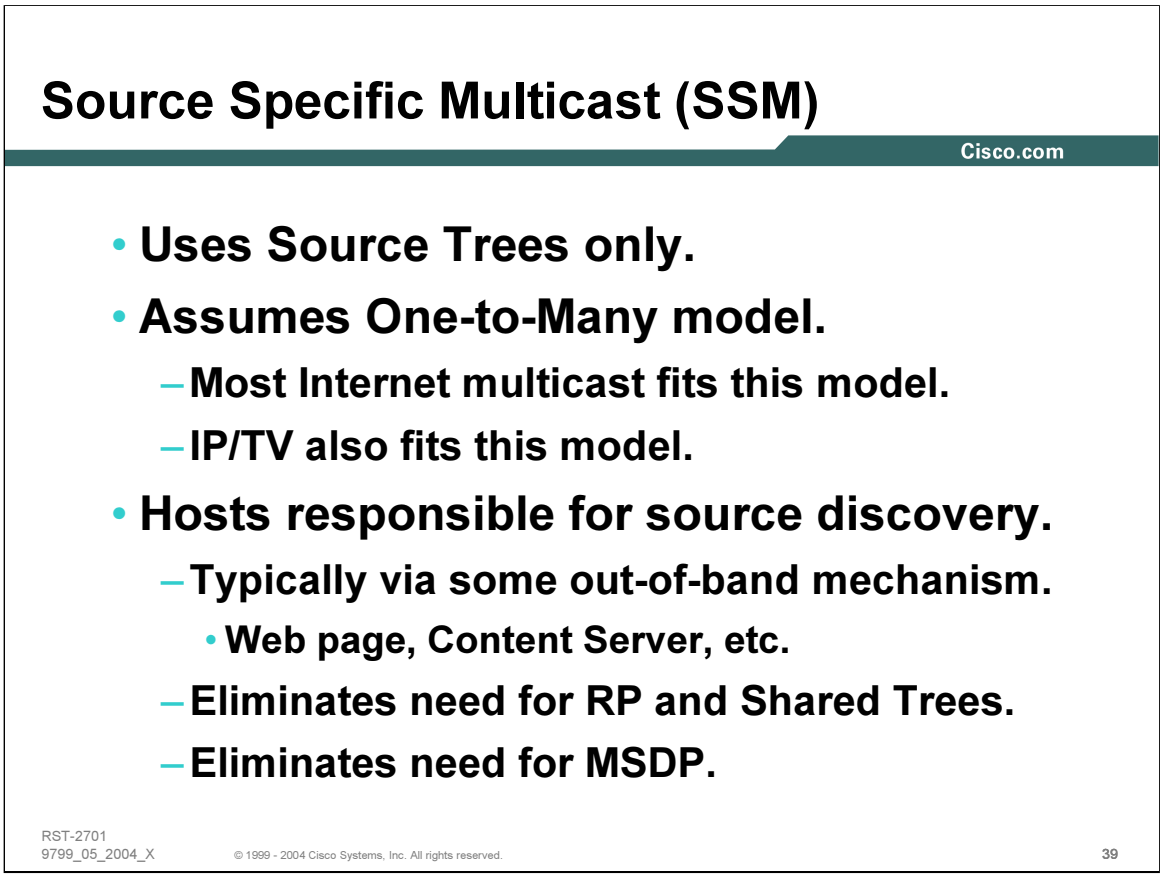
Source Specific Multicast

RST-2701
9799_05_2004_X

Barriers to Multicast Deployment

Cisco.com

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



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

SSM Overview

Cisco.com

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

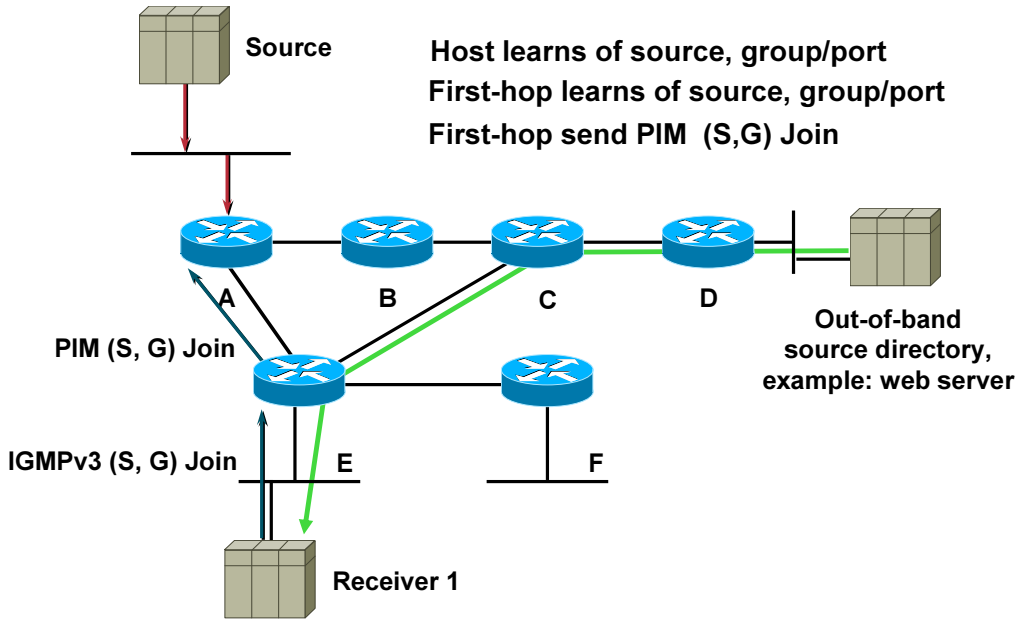
RST-2701
9799_05_2004_X

40

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

SSM Example

Cisco.com



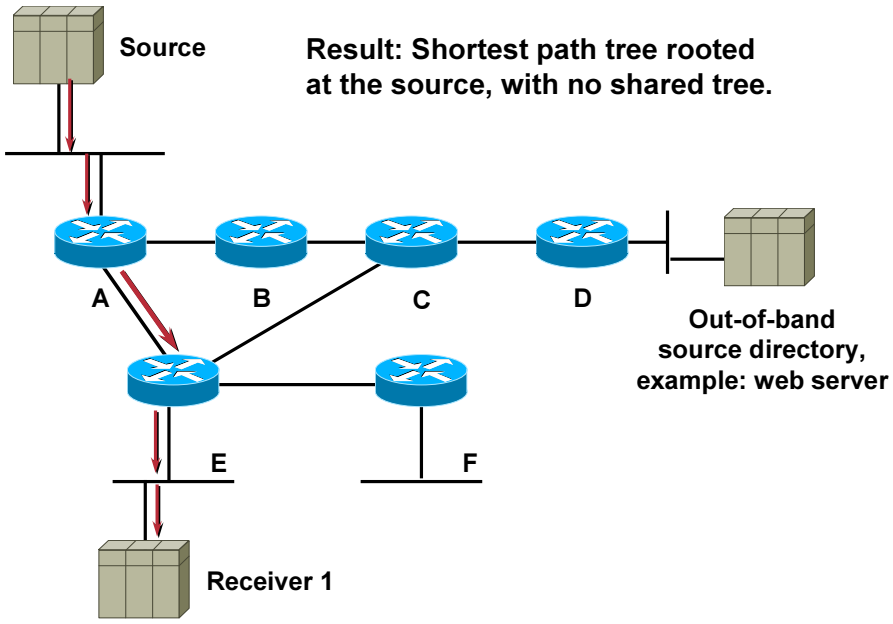
RST-2701
9799_05_2004_X

41

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

SSM Example

Cisco.com



RST-2701
9799_05_2004_X

42

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

SSM Configuration

Cisco.com

- **Global command**

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

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

SSM – Summary

Cisco.com

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

So where is SSM?

Cisco.com

- Dependant on IGMPv3
  - Microsoft supports IGMPv3 in Windows XP
- Workarounds
  - IGMPv3 lite
    - API/Library/DLL
    - Used by Cisco IP/TV 3.2 and later.
  - URL RenDezvous (URD)
    - Redirect from Web page with specific information intercepted by Router
  - Static Source Mapping
    - Router maps IGMPv2 Joins in SSM range to well-known sources via DNS or static configuration

IGMP v3lite

Cisco.com

- Source side:
  - No application changes required!
- Receiver side:
  - Application must use IGMPv3 API:
  - IGMP v3lite Library Component
    - Provides the IP SSM subset of IGMPv3 API
      - Applications must still filter out unwanted traffic.
  - IGMP v3lite Daemon Component
    - Sends special (S,G) Join to local router via UDP port 465

URD

Cisco.com

- **A content provider builds a web page that contains URD links.**
  - **List of sources willing to provide multicast content**
- **The user (receiver) clicks on one of the links**
- **Web Server sends back an HTTP redirect containing source and group info to TCP port 465**
- **Host sends the redirect via TCP port 465**
- **Local router intercepts TCP port 465 traffic**
  - **Uses source/group information in the redirect to identify the requested SSM flow.**

RST-2701
9799_05_2004_X

47

- **URD Overview**
  - The idea of URD as an interim solution for transition to IGMP v3 is that the content provider builds a web page that contains URD links. Those links contain information on sources that are willing to provide the multicast content for certain groups.
  - When a user clicks on such a link the browser of a host will try to open a TCP connection to the web server on port 659. If the last hop router is enabled for URD on the interface where the router receives the TCP packets from the host, it will intercept all packets for TCP connections destined to port 659 independent of the actual destination address of the TCP connection. From the information in URD the router learns about sources and groups.
  - Because normal IGMPv1/v2 group membership reports are still sent by the application, URD is compatible with IGMPv1/v2 snooping and CGMP in switches.

SSM Mapping

Cisco.com

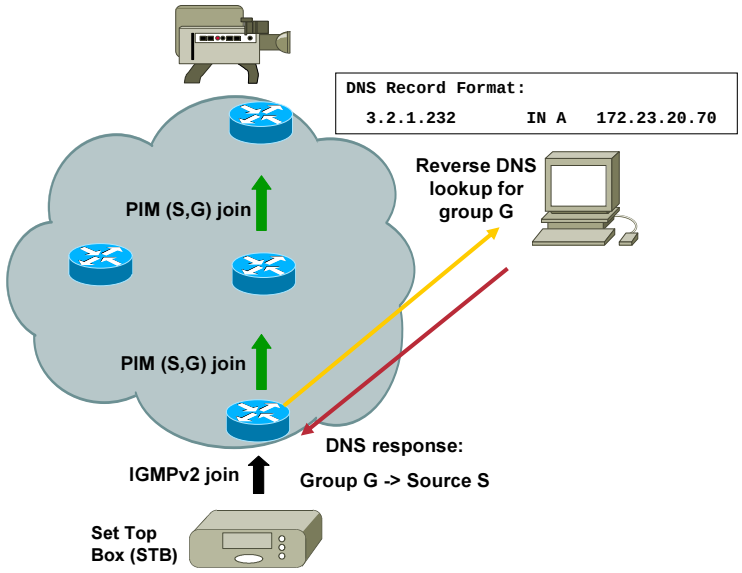
- Allows only for one source per Group
- Router maps group to a single source
  - Uses either DNS or static internal database
    - DNS method allows content providers to provide the mapping
    - DNS Method independent from network operators

RST-2701
9799_05_2004_X

48

SSM Mapping – DNS Example

Cisco.com



RST-2701
9799_05_2004_X

49

SSM Mapping Configuration

Cisco.com

Enabling SSM mapping on the router

```
ip igmp ssm-map enable
```

For static mapping:

```
ip igmp ssm-map static <acl-1> <source-1 IP address>
ip igmp ssm-map static <acl-2> <source-2 IP address>
```

For DNS mapping (existing commands):

```
ip domain-server <ip address>
ip domain-name <domain.com>
```

To disable DNS mapping

```
no ip igmp ssm-map query dns
```

```
DNS Record Format:   3.2.1.232       IN A   172.23.20.70
```

Where Is SSM?

Cisco.com

- Framework
  - draft-holbrook-idmr-igmpv3-ssm-06.txt
  - draft-ietf-ssm-arch-04.txt
- BCP proposal
  - draft-ietf-mboned-ssm232-08.txt
  - RFC 3569 Overview of SSM
- Supported in:

  IOS 12.X

  Windows XP, FreeBSD, Linux

  ftp://ftpeng.cisco.com/ipmulticast/ssm/index.html#Stacks

SSM – Summary

Cisco.com

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

Bidirectional (Bidir) PIM

RST-2701
9799_05_2004_X

Multicast Application Categories

Cisco.com

- One-to-Many Applications
  - Video, TV, Radio, Concerts, Stock Ticker, etc.
- Few-to-Few Applications
  - Small (<10 member) Video/Audio Conferences
- Few-to-Many Applications
  - TIBCO RV Servers (Publishing)
- Many-to-Many Applications
  - Stock Trading Floors, Gaming
- Many-to-Few Applications
  - TIBCO RV Clients (Subscriptions)

RST-2701
9799_05_2004_X

Multicast Application Categories
PIM-SM (S, G) State

Cisco.com

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

RST-2701
9799_05_2004_X

55

Multicast State Maintenance

Cisco.com

- CPU load factors
  - Must send/receive Registers
  - Must send periodic Joins/Prunes
  - Must perform RPF recalculations
    - Watch the total number of mroute table entries
    - Unicast route table size impacts RPF recalculation
- Memory load factors
  - (*, G) entry ~ 380 bytes + OIL size
  - (S, G) entry ~ 220 bytes + OIL size
  - Outgoing interface list (OIL) size
    - Each oil entry ~ 150 bytes

Many-to-Any State Problem

Cisco.com

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

RST-2701
9799_05_2004_X

Bidirectional (Bidir) PIM

Cisco.com

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

RST-2701
9799_05_2004_X

58

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

Bidirectional (Bidir) PIM

Cisco.com

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

RST-2701
9799_05_2004_X

59

Bidirectional (Bidir) PIM

Cisco.com

- **Designated Forwarders (DF)**
  - **On each link the router with the best path to the RP is elected to be the DF**
    - **Note: Designated Routers (DR) are not used for bidir groups**
  - **The DF is responsible for forwarding traffic upstream towards the RP**
  - **No special treatment is required for local sources**

RST-2701
9799_05_2004_X

60

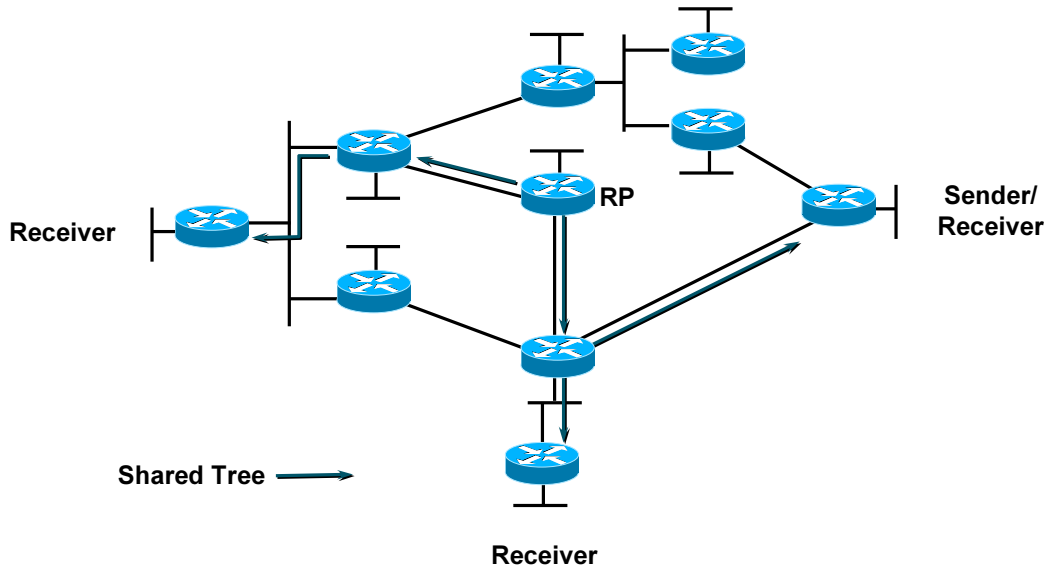
- **PIM Modifications for Bidir Operation**
  - The major modification of PIM Sparse Mode to support bidirectional mode is an addition of a Designated Forwarder, which takes over the role of a Designated Router (DR) and has the following responsibilities:

    It is the only router that forwards packets travelling downstream (towards receiver segments) onto the link

    It is the only router that picks-up upstream traveling packets (away from the source) off the link and forwards them towards the RP
  - There is one DF per RP for bidirectional group(s) on each link. One and only one election is performed at RP discovery time. There is no constant control traffic and control messages appear only on changes. The election is robust and enforces consistent view on all routers on link. The router with the best unicast route to the RP is elected as a DF.
  - There is no effect of this election on local sources – their traffic reaches locally attached receivers directly and special treatment is no longer required when the sources are directly connected to a router. Data from those sources will automatically be picked up by the DF and forwarded towards the RP.

Bidirectional PIM — Example

Cisco.com

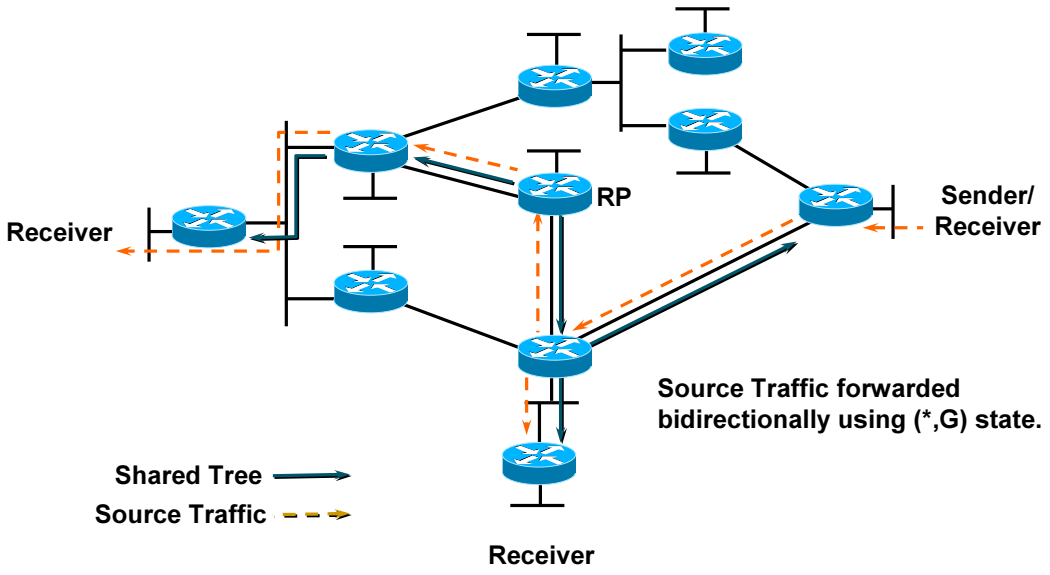


RST-2701
9799_05_2004_X

61

Bidirectional PIM — Example

Cisco.com



RST-2701
9799_05_2004_X

62

Configuring Bidir PIM
(Auto-RP Example)

Cisco.com

- **Define Candidate RP and groups / modes it is willing to serve**

```
ip pim send-rp-announce Loopback0 scope 10 group-list 45 bidir
ip pim send-rp-announce Loopback1 scope 10 group-list 46
! Two loopbacks needed due to a nature of ACLs (permit, deny)
ip pim send-rp-discovery scope 10

access-list 45 permit 224.0.0.0 0.255.255.255
access-list 45 permit 227.0.0.0 0.255.255.255
! 224/8 and 227/8 will be PIM Bidir groups
access-list 45 deny 225.0.0.0 0.255.255.255
! 225/8 will be a PIM Dense Mode group

access-list 46 permit 226.0.0.0 0.255.255.255
! 226/8 will be a PIM Sparse Mode group
```

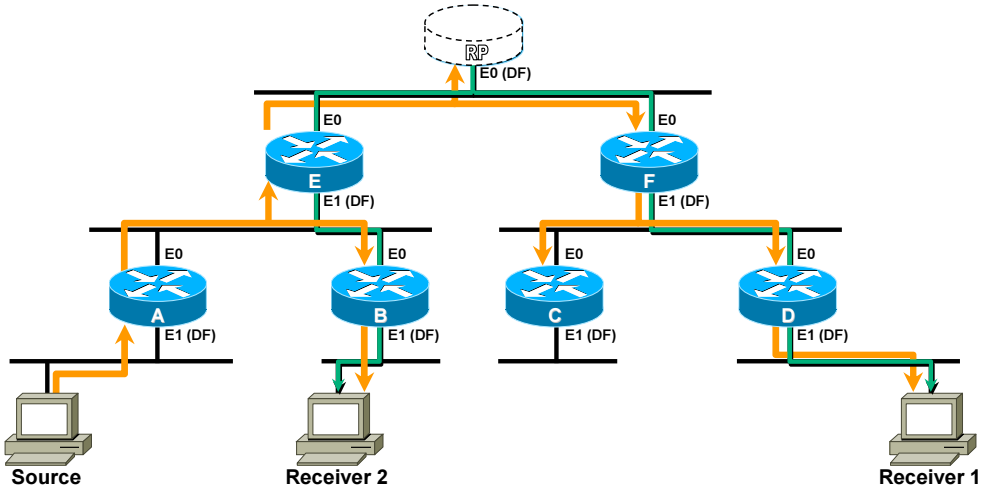
RST-2701
9799_05_2004_X

63

- **Configuring Bidir PIM (BSR Example)**
  - A bidirectional PIM capable router can run in bidirectional mode, sparse mode, dense mode or any combination of them. If a router is configured for bidirectional mode but does not learn of a bidirectional capable RP it will operate in sparse mode. If a bidirectional capable router learns of a bidirectional RP then the group range advertised by the RP will operate in bidirectional mode. If the RP advertises any groups with a negative prefix they will operate in dense mode.
  - By default a bidirectional RP advertises all groups as bidirectional. An access group on the RP can be used to specify a list of groups to be advertised as bidirectional. Groups with the "deny" clause will operate in dense mode.
  - A different (non bidirectional) RP address needs to be specified for groups that need to operate in sparse mode. This is because a single access-list allows only "permit" or a "deny" clause.
  - The example shows how to configure a bidirectional RP to run all 3 modes. 224/8 and 227/8 are bidirectional groups, 225/8 is dense mode and 226/8 is sparse mode. Both the bidirectional RP and the sparse mode RP are configured on one router using two different loopback interfaces.

Bidir PIM – Phantom RP

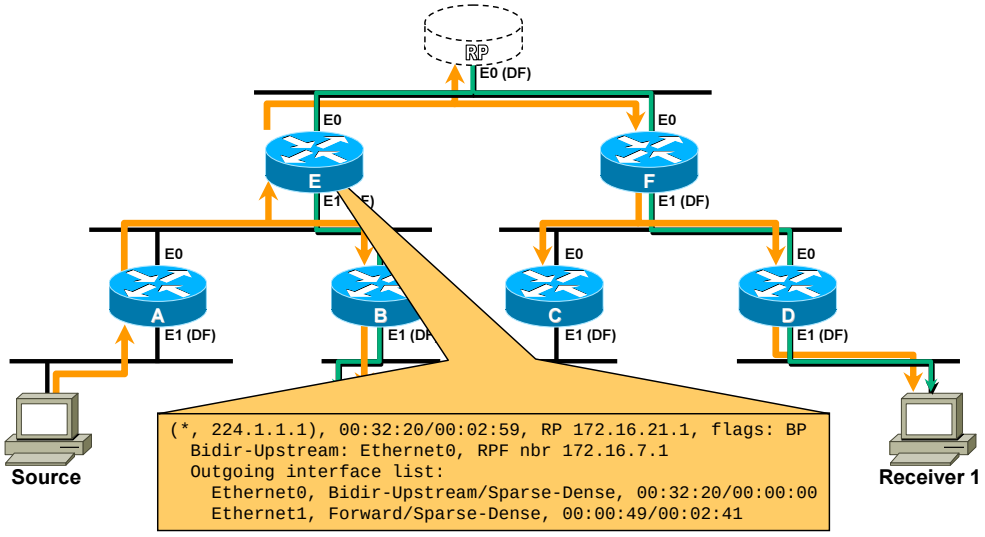
Cisco.com



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

Bidir PIM – Phantom RP

Cisco.com



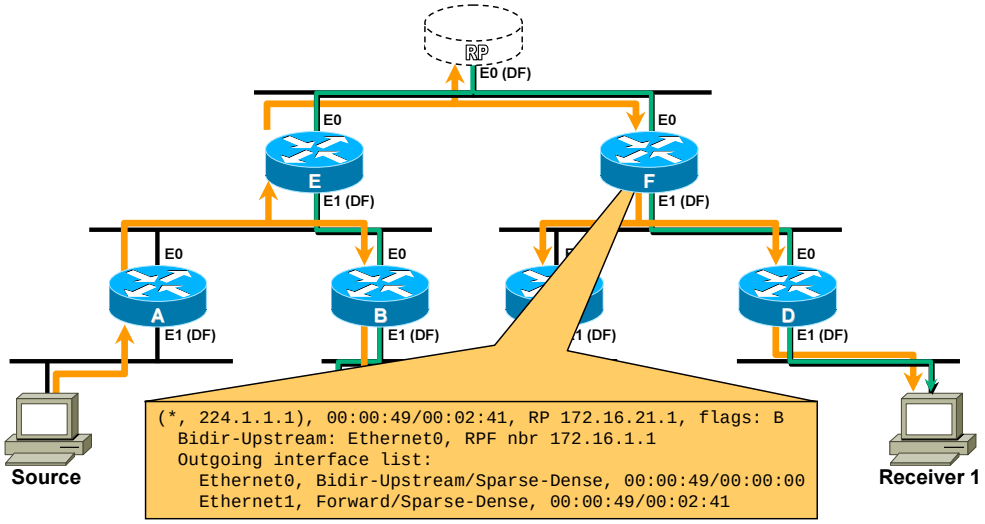
Router “E” forwards traffic onto core LAN segment.

RST-2701
9799_05_2004_X

65

Bidir PIM – Phantom RP

Cisco.com



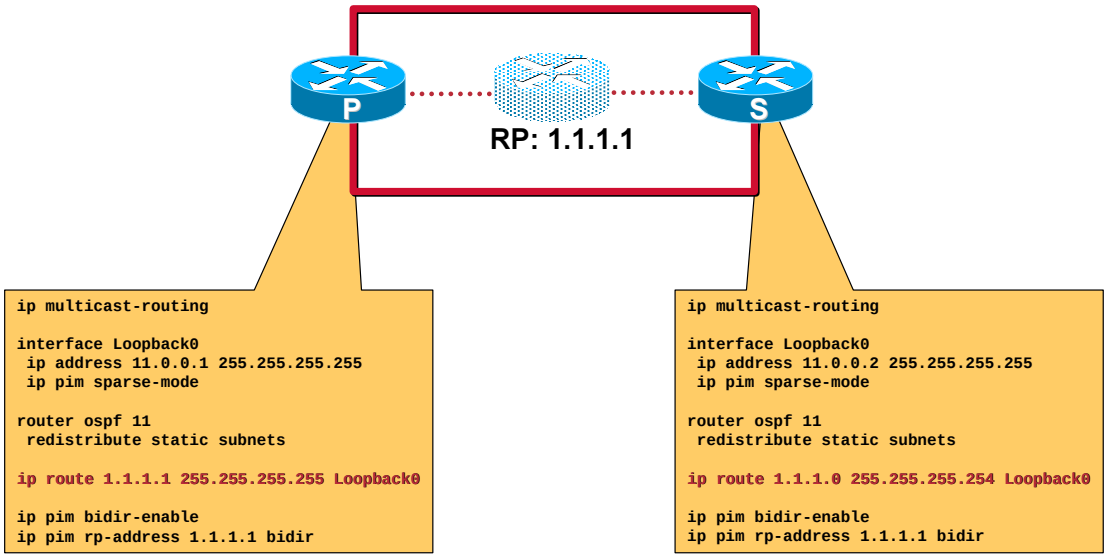
Router “F” forwards traffic on down the Shared Tree ala normal PIM-SM.

RP doesn’t even have to physically exist.

Phantom RP on Point-to-Point Core

Cisco.com

Static Route Method




```
ip multicast-routing

interface Loopback0
 ip address 11.0.0.1 255.255.255.255
 ip pim sparse-mode

router ospf 11
 redistribute static subnets

ip route 1.1.1.1 255.255.255.255 Loopback0

ip pim bidir-enable
ip pim rp-address 1.1.1.1 bidir
```

```
ip multicast-routing

interface Loopback0
 ip address 11.0.0.2 255.255.255.255
 ip pim sparse-mode

router ospf 11
 redistribute static subnets

ip route 1.1.1.0 255.255.255.254 Loopback0

ip pim bidir-enable
ip pim rp-address 1.1.1.1 bidir
```

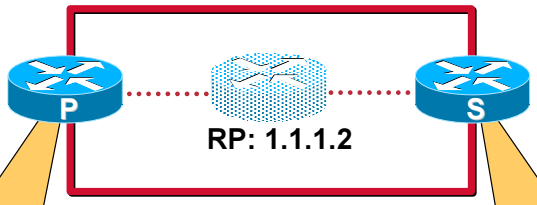
RST-2701
9799_05_2004_X

67

Phantom RP on Point-to-Point Core

Cisco.com

Netmask Method




```
ip multicast-routing
!
interface Loopback0
 ip address 1.1.1.1 255.255.255.252
 ip pim sparse-mode
 ip ospf network point-to-point
!
router ospf 11
 network 1.1.1.0 0.0.0.3 area 0
 network 10.1.1.0 0.0.0.255 area 0
 network 10.1.2.0 0.0.0.255 area 0
!
ip pim bidir-enable
ip pim rp-address 1.1.1.1
ip pim rp-address 1.1.1.2 bidir
```

```
ip multicast-routing
!
interface Loopback0
 ip address 1.1.1.1 255.255.255.248
 ip pim sparse-mode
 ip ospf network point-to-point
!
router ospf 11
 network 1.1.1.0 0.0.0.7 area 0
 network 10.1.1.0 0.0.0.255 area 0
 network 10.1.2.0 0.0.0.255 area 0
!
ip pim bidir-enable
ip pim rp-address 1.1.1.1
ip pim rp-address 1.1.1.2 bidir
```

RST-2701
9799_05_2004_X

68

Bidir PIM—Summary

- Drastically reduces network mroute state
  - Eliminates *ALL* (S,G) state in the network
    - SPT's between sources to RP eliminated
    - Source traffic flows both up and down Shared Tree
  - Allows Many-to-Any applications to scale
    - Permits virtually an unlimited number of sources

Multicast Group Control

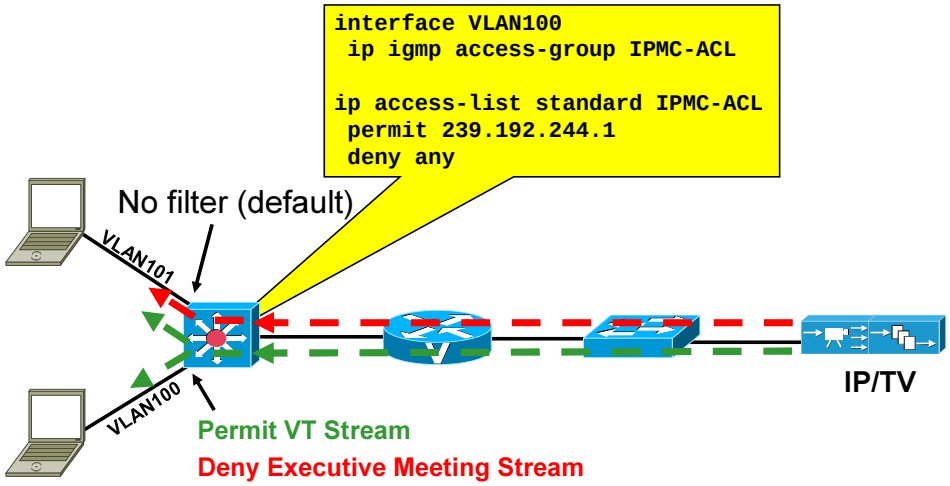
RST-2701
9799_05_2004_X

70

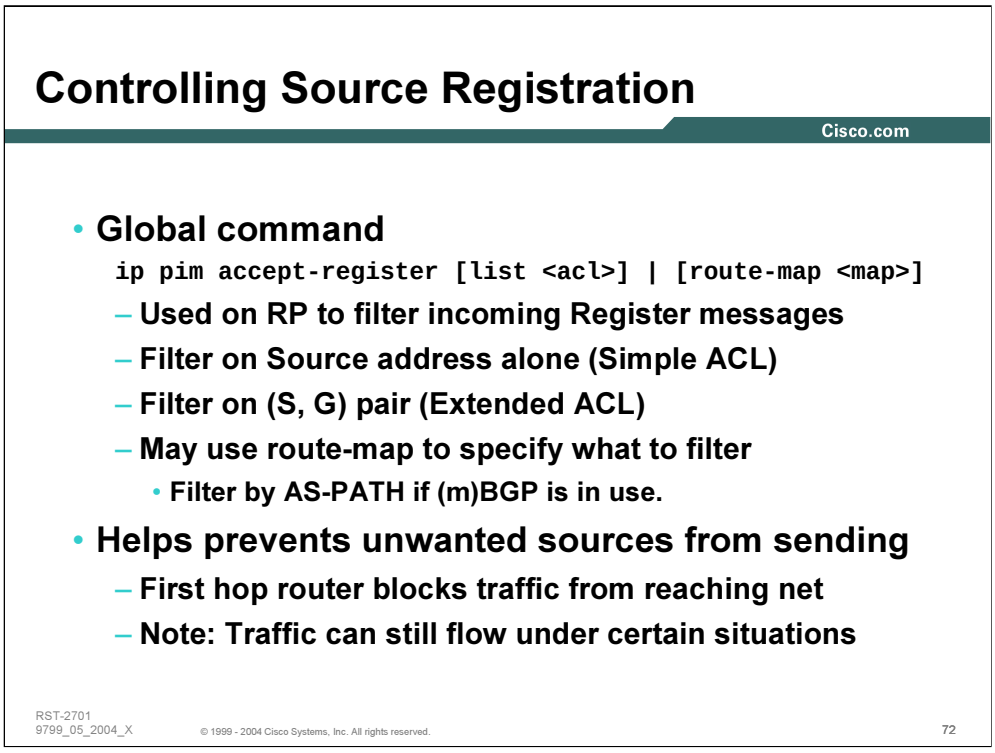
Controlling Receivers

Cisco.com

IGMP Access-Group Approach



This is micro-management of IP Multicast traffic!!!



- **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)**

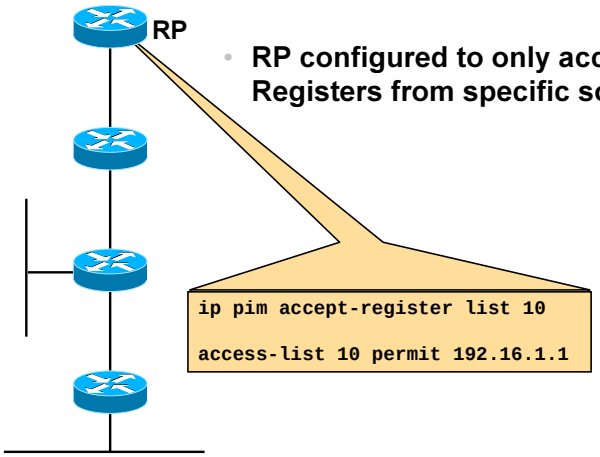
  ip pim accept-register [list <acl>] | [route-map <map>]

  – 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



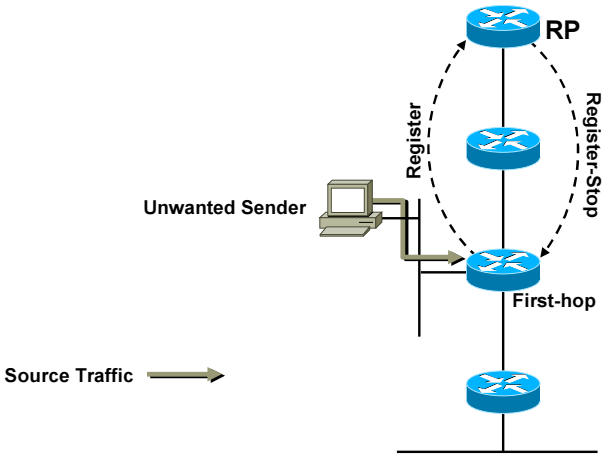
- RP configured to only accept Registers from specific source.

RST-2701
9799_05_2004_X

73

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.

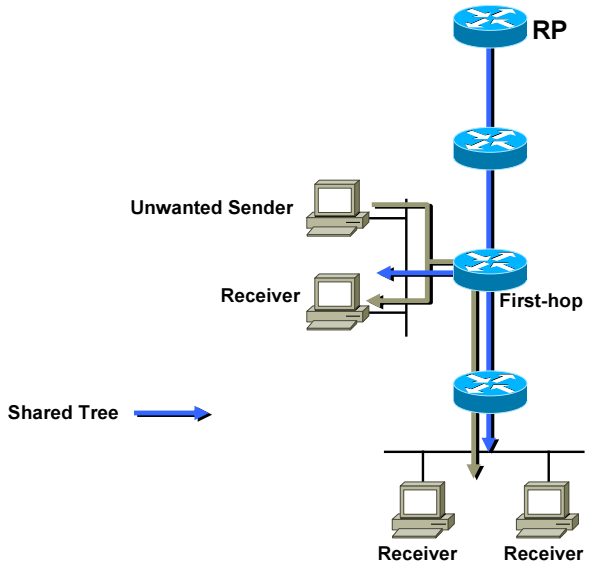
RST-2701
9799_05_2004_X

74

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-2701
9799_05_2004_X

75

Disabling Entire Group Ranges

Cisco.com

- **Accept-Register Method**

```
ip pim accept-register group-list 10
access-list 10 deny 224.2.0.0 0.0.255.255
access-list 10 permit any
```

- **Pros**
  - Only configured on RP(s)
- **Cons**
  - Shared Trees and (*,G) state still created.
    - Results in unwanted (*,G) PIM Control Traffic.
  - Source traffic can still flow.

    (See previous section on Accept-Register)

Disabling Entire Group Ranges

Cisco.com

- Garbage Can RP Method
  - Concept:
    - Separate RP for “disabled” groups
      - Could be non-existant router
    - Blackholes all Registers and Joins
  - Implementation:
    - Define separate RP for disabled groups
      - Use Auto-RP, BSR or Static RP definition
    - Disable RP functionality on Garbage Can RP
      - Use ‘accept-rp’ command on GC RP to “deny” it from serving as RP for the disabled group range.

Disabling Entire Group Ranges

Cisco.com

- Garbage Can RP Method
  - Pros:
    - Few if any.
  - Cons:
    - Periodic Registers still sent to GC RP
    - Periodic Joins still sent to GC RP
    - Has same source issues as Accept-Register
      - Source traffic can still flow under certain conditions.
    - Adds *significant* complexity to network

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-2701
9799_05_2004_X

Disabling Entire Group Ranges

Cisco.com

- **Local Loopback RP Method**
  - **Operation:**
    - **Each router serves as RP for unauthorized groups.**
      - **Collapses PIM-SM domain of unauthorized groups down to the local router.**
    - **Unauthorized group traffic cannot flow beyond local router.**

Disabling Entire Group Ranges

Cisco.com

- Local Loopback RP Method
  - Pros:
    - No PIM control traffic sent.
      - Local router is RP so no Registers/Joins are sent.
    - No additional workload on local router.
      - First-hop routers always have to create state anyway.
    - Can also serve as RP-of-last-resort
      - Solving DM Fallback problem at the same time.
  - Cons:
    - Must be configured on every router.
    - Local sources can still send to local receivers.

Disabling Entire Group Ranges

Cisco.com

- New no ip pim dm-fallback command
  - Groups with no known RP default to an RP address of 0.0.0.0.
    - Effectively disables multicast for these groups.
    - New sources are not Registered.
    - New receivers are not Joined.
- Available 12.3(4)T, 12.2(28)S.

RST-2701
9799_05_2004_X

82

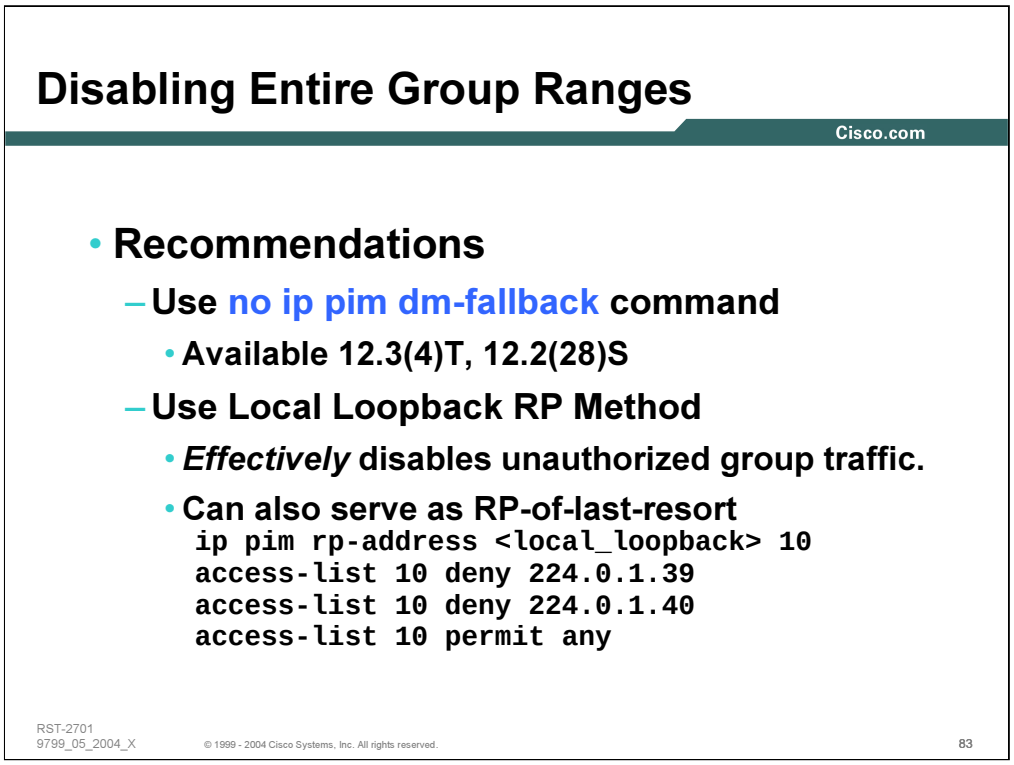
- **Disabling Entire Group Ranges – Future**

  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.

Combining Anycast RP & Auto-RP

RST-2701
9799_05_2004_X

84

Combining Auto-RP and Anycast-RP

Cisco.com

- Anycast-RP and Auto-RP may be combined.
  - Provides advantages of both methods
    - Rapid RP failover of Anycast RP
    - No DM Fallback
    - Configuration flexibility of Auto-RP
    - Ability to effectively disable undesired groups

RST-2701
9799_05_2004_X

Combining Auto-RP and Anycast-RP

Cisco.com

Configuration Steps

1. Enable Auto-RP
   - Newer IOS images
     - Use ip pim autorp listener global command and configure ip pim sparse-mode on all interfaces.
   - Older IOS images
     - Configure ip pim sparse-dense-mode on all interfaces.
2. Configure Auto-RP Mapping Agents

```
ip pim send-rp-discovery interface Loopback0 scope 32
```

RST-2701
9799_05_2004_X

Combining Auto-RP and Anycast-RP

Cisco.com

Configuration Steps

3. Block DM Fallback
   - Newer IOS images
     - Use no ip pim dm-fallback
   - Older IOS images
     - Configure RP-of-last-Resort

```
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
```

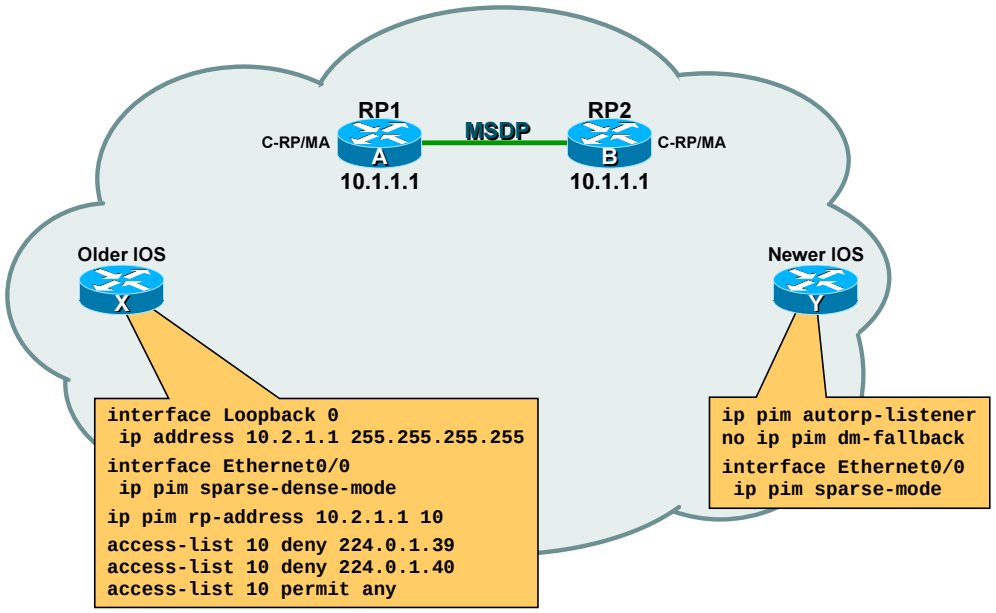
4. Configure Anycast RP's for desired group range.
5. Configure Anycast RP's as Auto-RP C-RP's

```
ip pim send-rp-discovery Loopback0 scope 32 group-list 10
```

   - Loopback0 = Anycast RP Address
     - Anycast-RP's will announce Anycast-RP address via Auto-RP

Example Auto-RP and Anycast-RP

Cisco.com

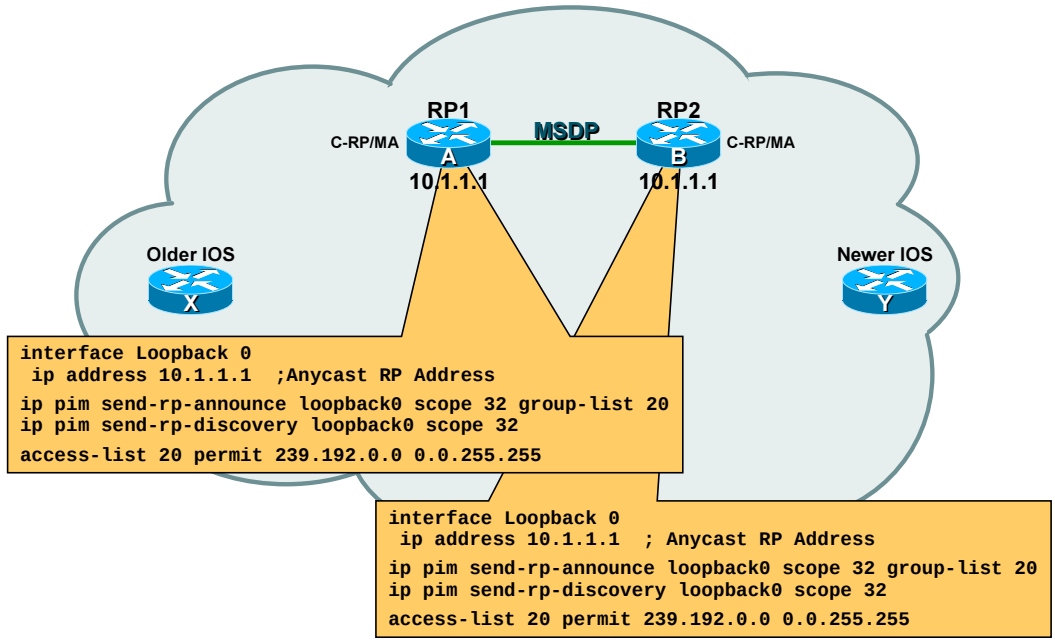


RST-2701
9799_05_2004_X

88

Example Auto-RP and Anycast-RP

Cisco.com



RST-2701
9799_05_2004_X

89

Using Admin. Scoped Zones

RST-2701
9799_05_2004_X

Administratively-Scoped Zones

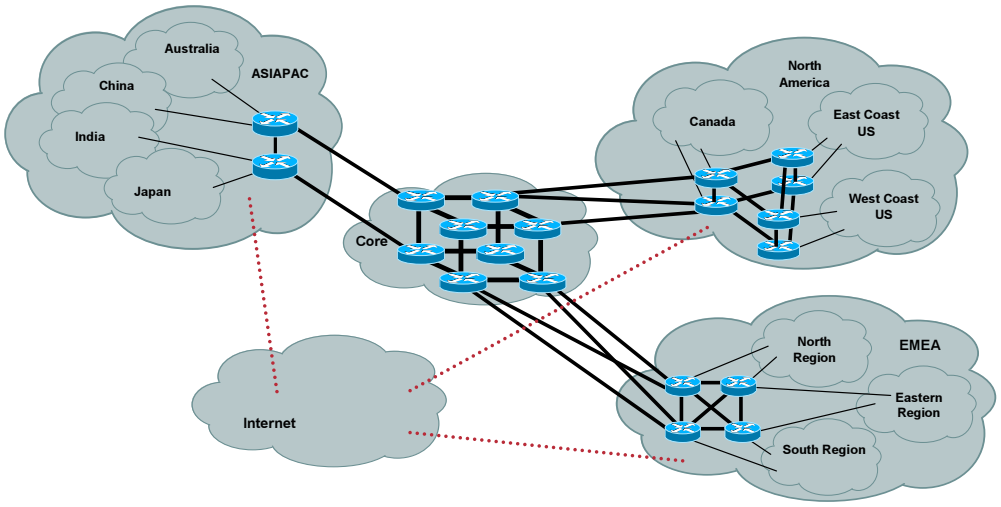
Cisco.com

- Used to limit:
  - High-BW sources to local site
  - Control sensitive multicast traffic
- Simple scoped zone example:
  - 239.193.0.0/16 = Campus Scope
  - 239.194.0.0/16 = Region Scope
  - 239.195.0.0/16 = Organization-Local (Enterprise) Scope
  - 224.1.0.0 - 238.255.255.255 = Global scope (Internet) zone
    - High-BW sources use Site-Local scope
    - Low-Med. BW sources use Org.-Local scope
    - Internet-wide sources use Global scope

RST-2701
9799_05_2004_X

Administratively-Scoped Zones Example

Cisco.com



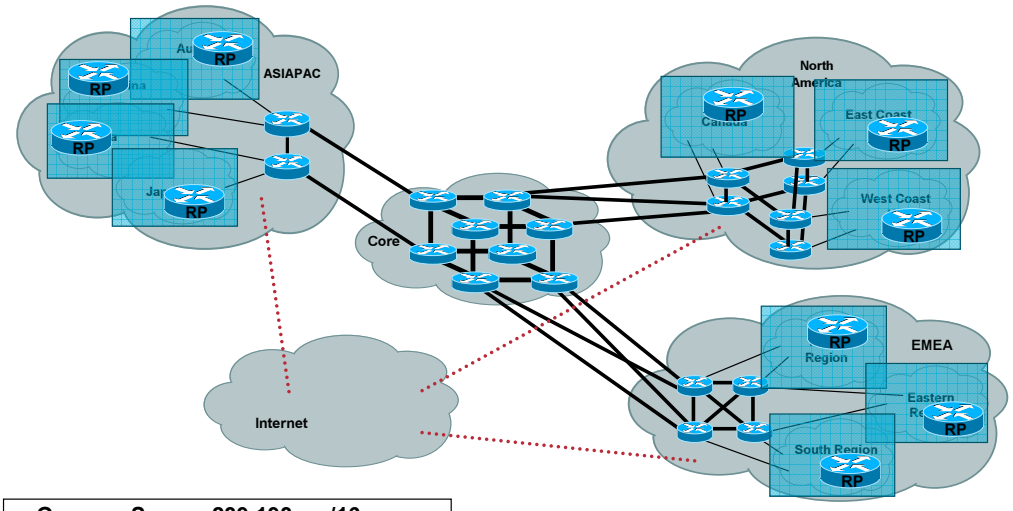
RST-2701
9799_05_2004_X

92

Administratively-Scoped Zones Example

Cisco.com

Level1: Campus Scope



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

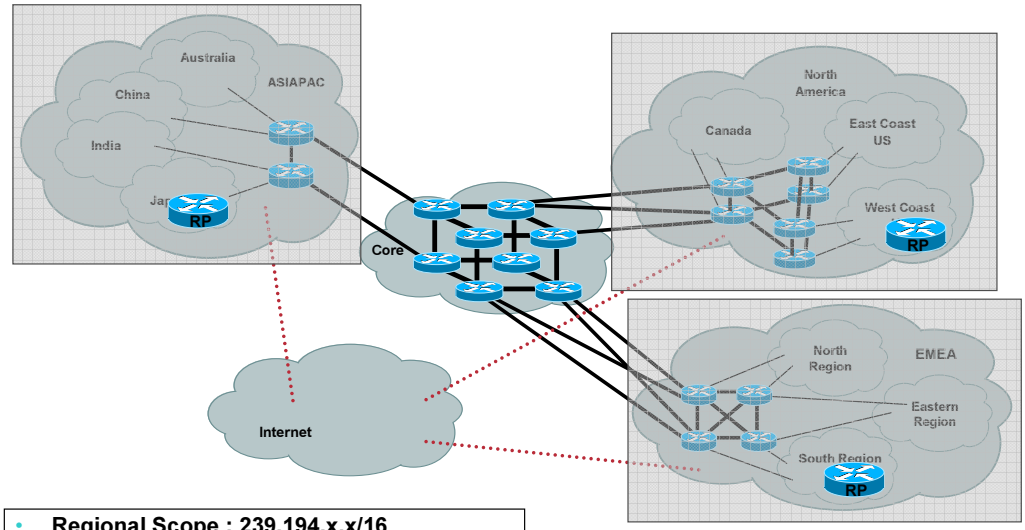
RST-2701
9799_05_2004_X

93

Administratively-Scoped Zones Example

Cisco.com

Level2: Regional Scope



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

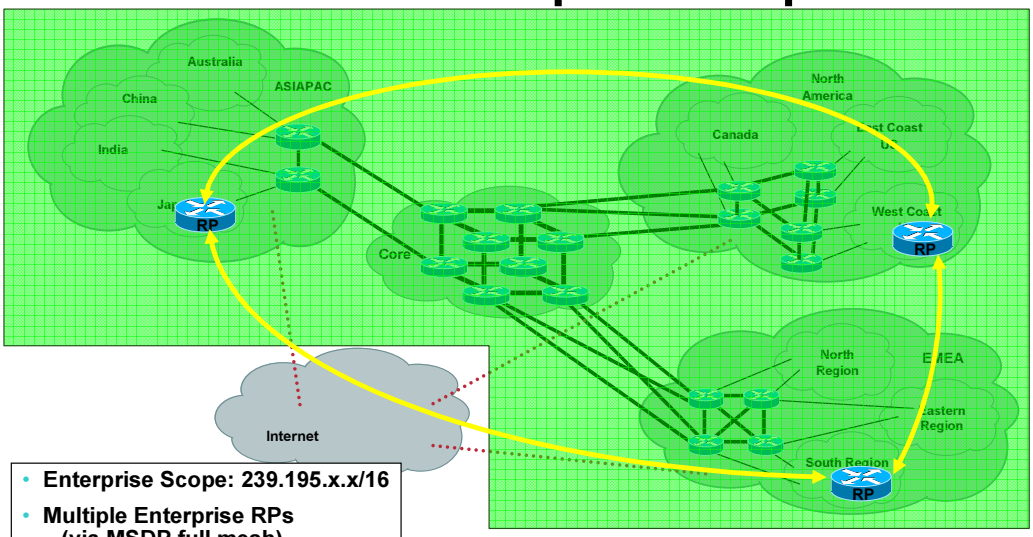
RST-2701
9799_05_2004_X

94

Administratively-Scoped Zones Example

Cisco.com

Level3: Enterprise Scope

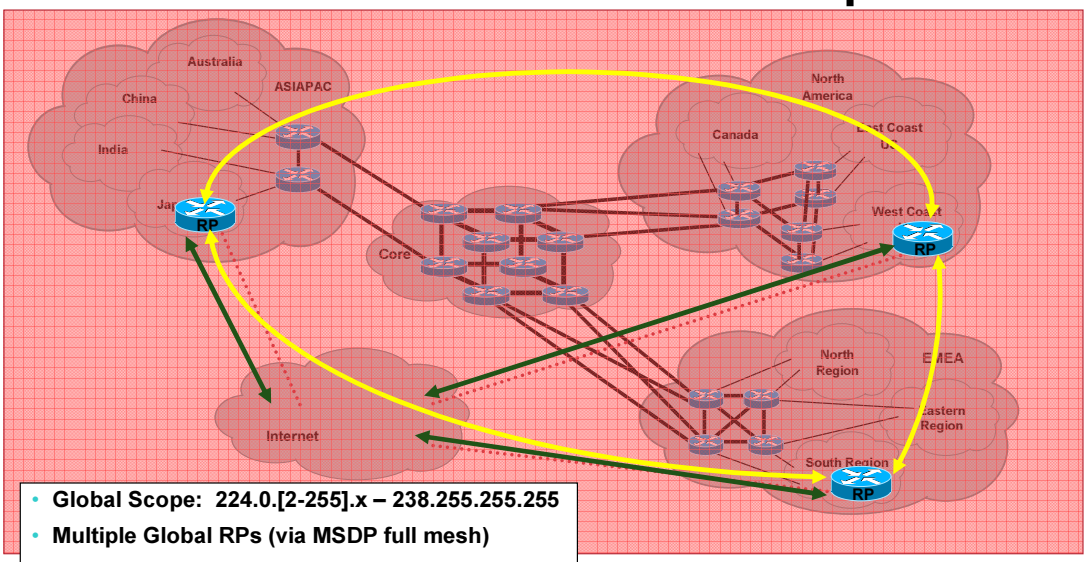


- Enterprise Scope: 239.195.x.x/16
- Multiple Enterprise RPs (via MSDP full mesh)

Administratively-Scoped Zones Example

Cisco.com

Level 4: Internet Global Scope



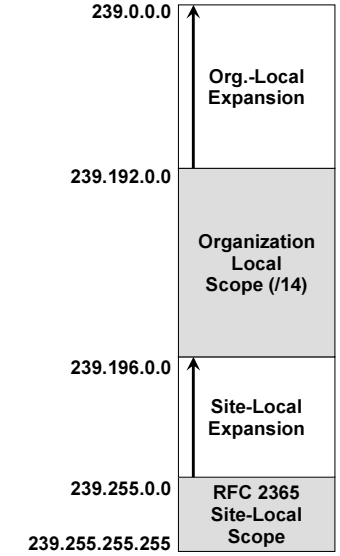
- Global Scope: 224.0.[2-255].x – 238.255.255.255
- Multiple Global RPs (via MSDP full mesh)
- MSDP connectivity to SP network

RST-2701
9799_05_2004_X

96

Administratively Scoped Address Range

Cisco.com



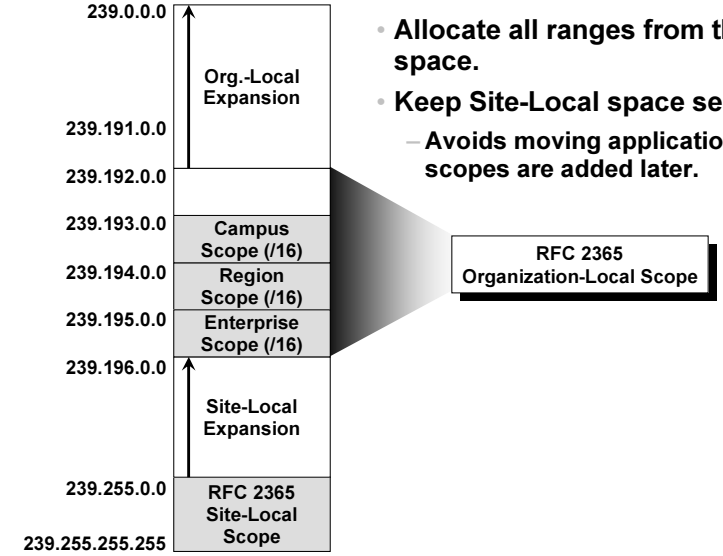
- RFC 2365 Administratively Scoped Zones.
  - Organization-Local Scope (239.192/14)
    - Expands downward in address range.
  - Site-Local Scope (239.255/16)
    - Expands downward in address range.
    - Smallest possible scope.
    - Other scopes may be equal but not smaller.

RST-2701
9799_05_2004_X

97

Example Scope Address Assignments

Cisco.com



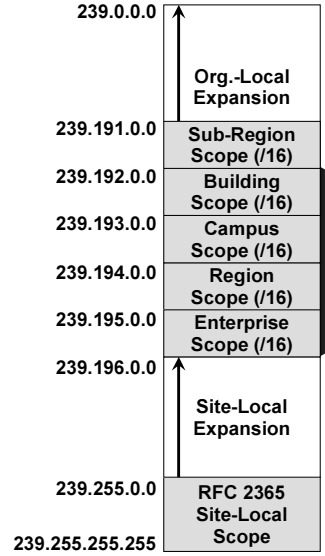
- Allocate all ranges from the Org-Local space.
- Keep Site-Local space separate.
  - Avoids moving applications when smaller scopes are added later.

RST-2701
9799_05_2004_X

98

Adding a Additional Scopes

Cisco.com



- Additional scope ranges are allocated downward into Org-Local Expansion.
  - Not necessary to keep ranges in scope size order.
  - (i.e. “Sub-Region” scope is a larger physical scope than the “Building” and “Campus” scopes).

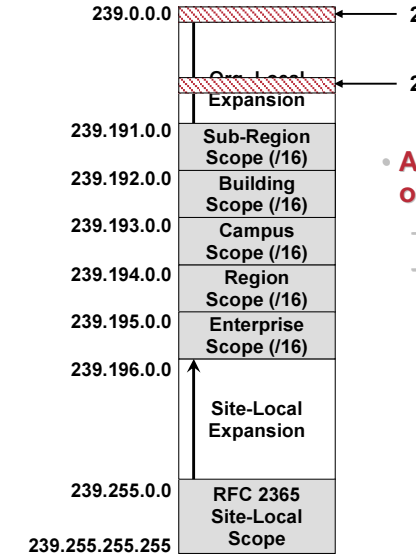
RFC 2365
Organization-Local Scope

RST-2701
9799_05_2004_X

99

Address Ranges to Avoid

Cisco.com



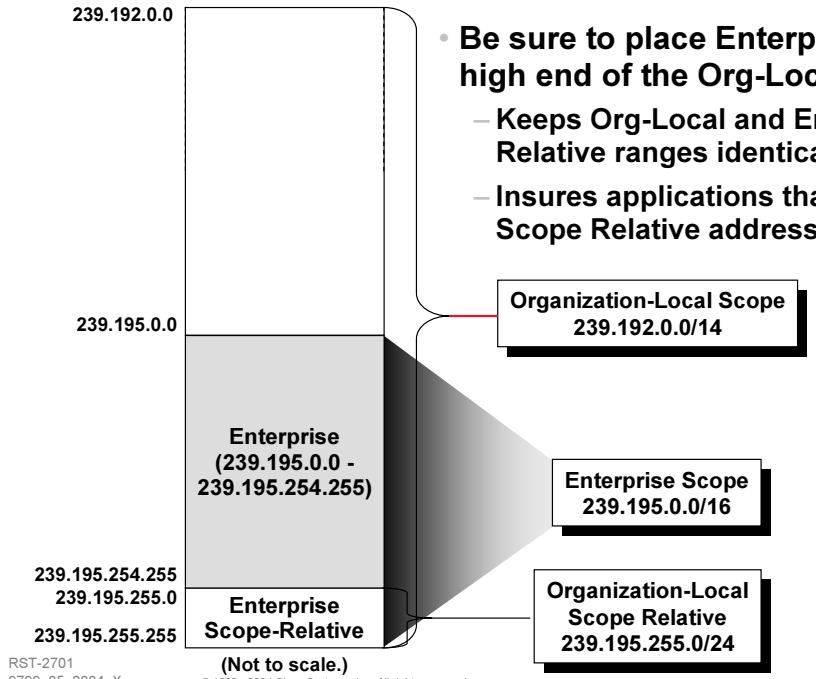
- Avoid ranges that map to a MAC address of 0x0100-5E00-00xx!
  - i.e. 239.128.0.0/24 & 239.0.0.0/24
  - These addresses are always flooded by Layer 2 switches!

RST-2701
9799_05_2004_X

100

Enterprise Scope Relative Range

Cisco.com



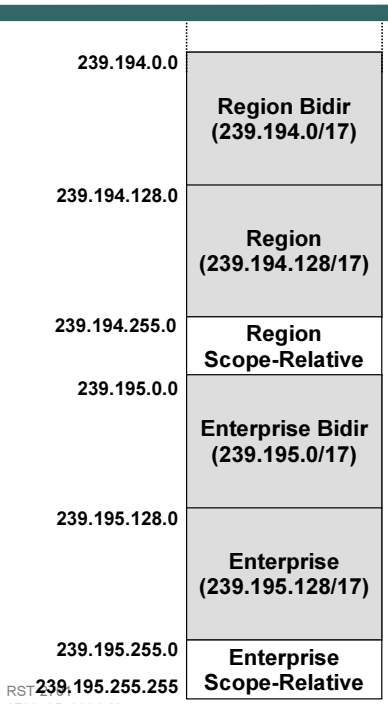
- **Be sure to place Enterprise Scope at the high end of the Org-Local range.**
  - **Keeps Org-Local and Enterprise Scope Relative ranges identical.**
  - **Insures applications that use Org-Local Scope Relative addresses work correctly.**

RST-2701
9799_05_2004_X

101

Adding Bidir Ranges to each Scope

Cisco.com

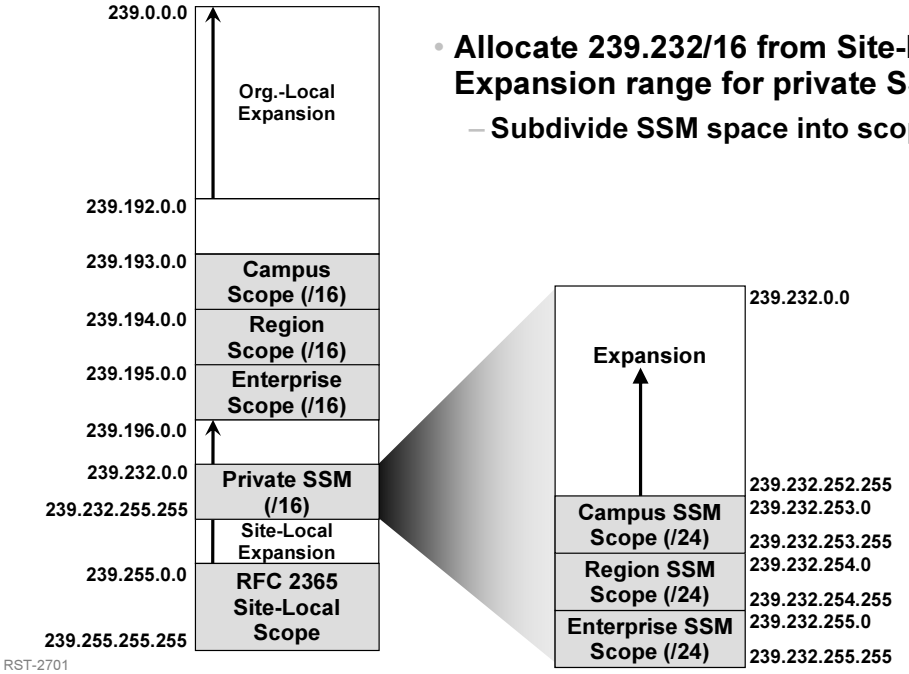


- **Subdivide each scope's address range into Bidir and ASM ranges.**
  - **Keep ASM range at the upper end of the address range.**
  - **Keeps Scope-Relative multicast in ASM mode.**

RST-2300, 9799_05_2004_X

Adding Private SSM Space

Cisco.com

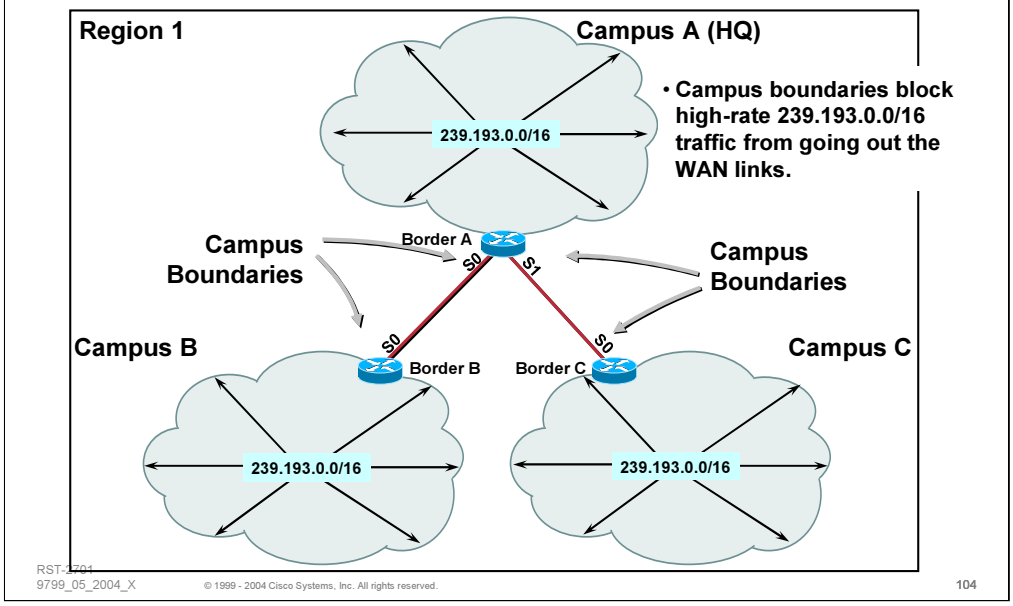


- **Allocate 239.232/16 from Site-Local Expansion range for private SSM space.**
  - **Subdivide SSM space into scoped zones.**

RST-2701
9799_05_2004_X

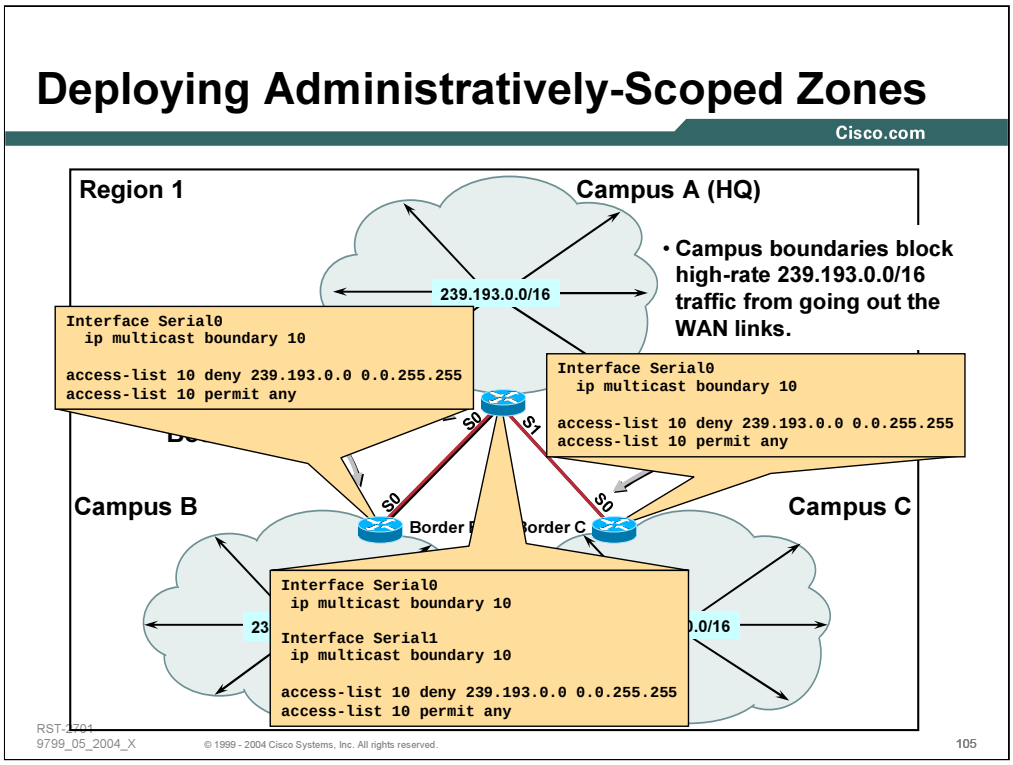
103

Deploying Administratively-Scoped Zones



- **Administratively-Scoped Zone**
  - The slide above shows a deployment of Admin-Scoped Zones based on the address scheme shown on the previous page.

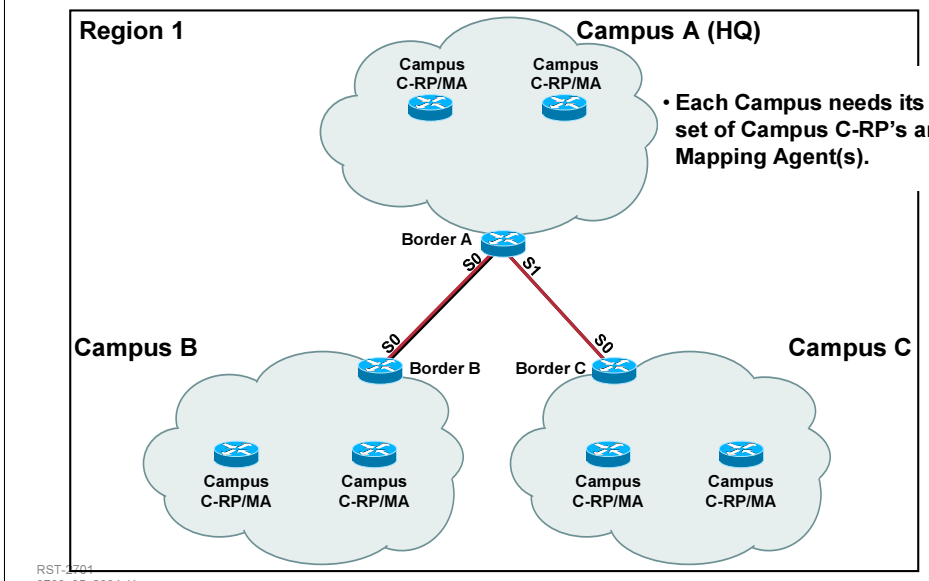
    In the example above, a Headquarters site is connected to two other remote sites: one in Los Angeles and another in Atlanta. Note that each of these sites (including the HQ site) have site local boundaries configured to prevent the flow of 239.255.0.0/16 multicast traffic from leaving the site.



- **Administratively-Scoped Zone**
  - The slide above shows the configuration commands necessary to establish the "Site-Local" Admin-Scoped Zones.

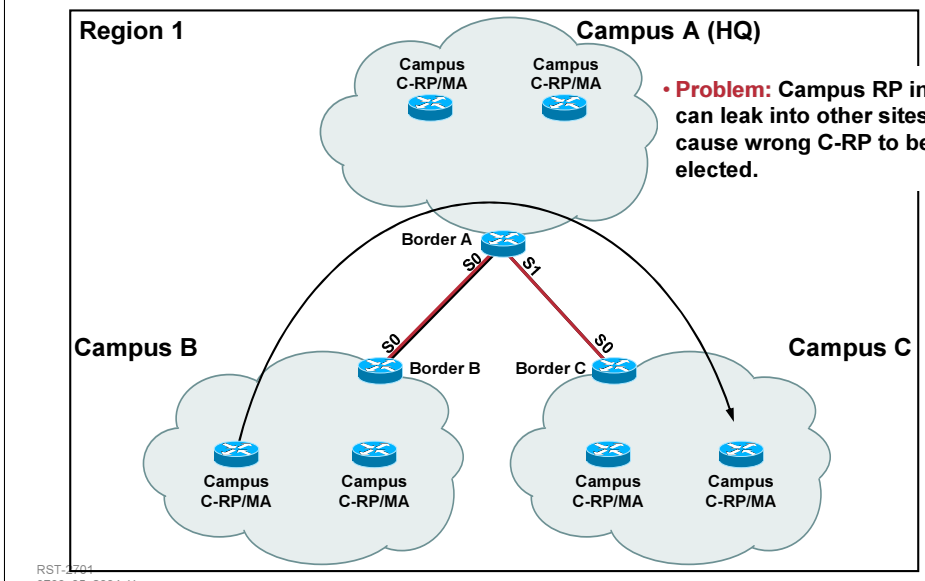
    Notice that the **ip multicast boundary** command is used with the appropriate ACL to deny any high-bandwidth multicast traffic in the 239.255.0.0/16 multicast group range from entering/leaving the sites and possibly congesting the WAN links.

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.)

Deploying Administratively-Scoped Zones
Auto-RP Example



- **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.

Deploying Administratively-Scoped Zones
Preventing Auto-RP Info Leakage

Cisco.com

- **Multicast Boundary Command**

  `ip multicast boundary <acl> [filter-autorp]`

  - **New 'filter-autorp' option**
    - **Filters contents of Auto-RP packets**
      - **Filters both Announcement and Discovery messages**
      - **C-RP entries that fail <acl> are removed from packet**
    - **Prevents C-RP information from leaking in/out of scoped zone.**
    - **Greatly simplifies Admin. Scoped Zone support in Auto-RP.**
    - **Available in 12.0(22)S, 12.2(12).**

RST-2701
9799_05_2004_X

108

- **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.

Deploying Administratively-Scoped Zones
Preventing Auto-RP Info Leakage

Cisco.com

- **How 'filter-autorp' option works:**
  - **For each RP Entry in Auto-RP packet:**
    - **If group-range in RP-Entry *'intersects'* any 'denied' group-range in the Multicast Boundary ACL, delete RP Entry from Auto-RP packet.**
  - **If resulting Auto-RP packet is non-empty, forward across multicast boundary.**

RST-2701
9799_05_2004_X

109

- **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-2701
9799_05_2004_X

110

- **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

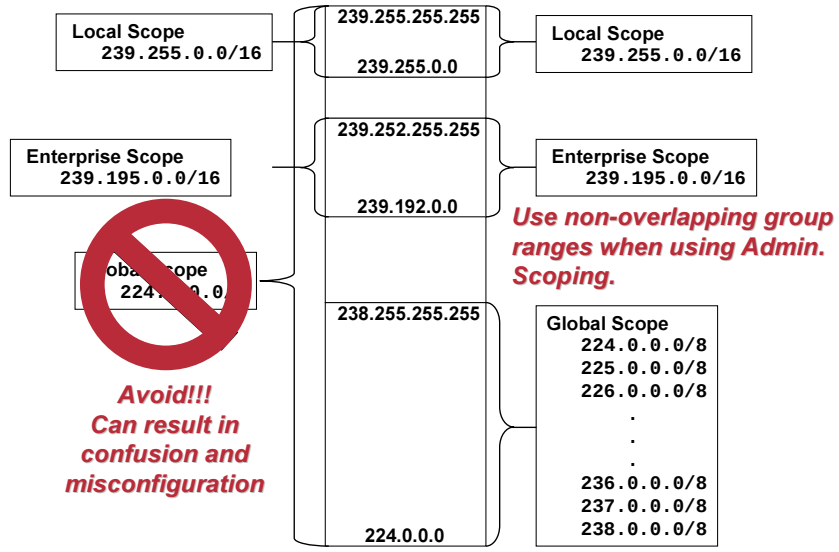
    This will result in the group range of RP-Entry in the Auto-RP Announcement being 224.0.0.0/4. This overlaps the Site-Local range which will be "denied" by the multicast boundary ACL. The net result will be that the RP-Entry for the catch-all RP will be filtered at the multicast boundary.

    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.

Avoid Overlapping Group Ranges

Cisco.com

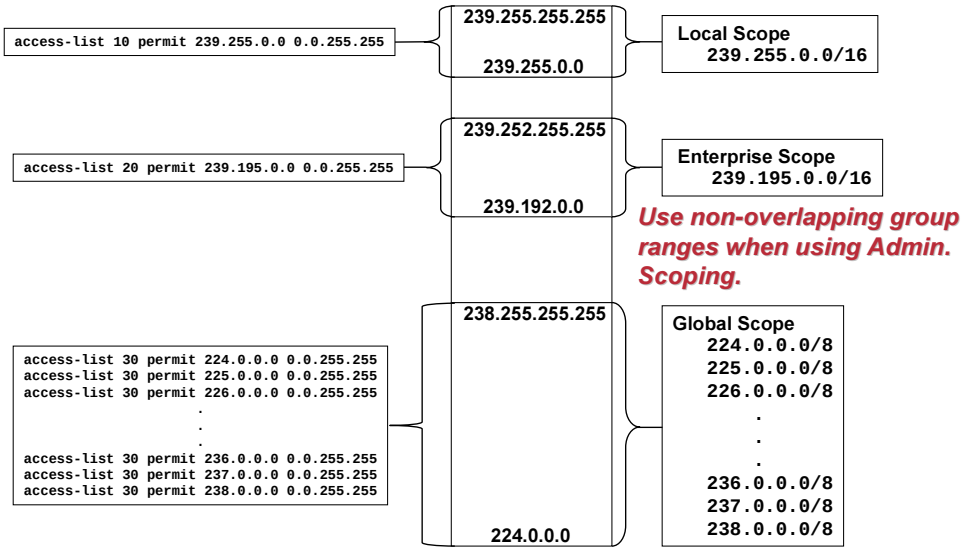


RST-2701
9799_05_2004_X

111

Avoid Overlapping Group Ranges

Cisco.com



RST-2701
9799_05_2004_X

112

Avoid 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
```

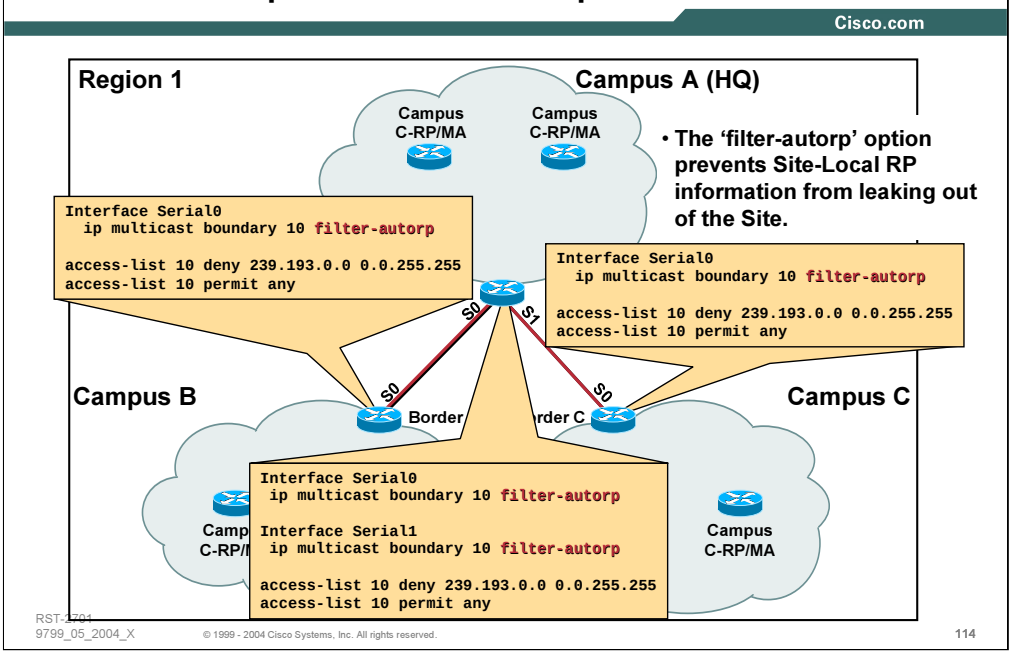
  - **Must only use "permit" clauses**

```
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-2701
9799_05_2004_X

113

Deploying Administratively-Scoped Zones
Auto-RP Example with 'filter-autorp' boundaries



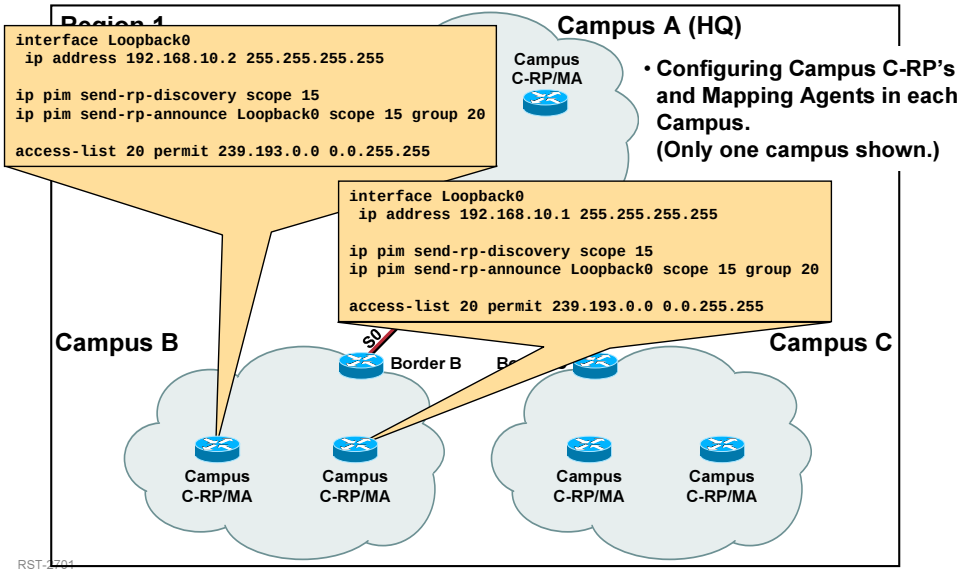
- **Auto-RP Example with 'filter-autorp' boundaries**
  - The example above shows our example network configured with the new **filter-autorp** feature on the **ip multicast boundary** command.

    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.

Deploying Administratively-Scoped Zones
Auto-RP Example with 'filter-autorp' boundaries

Cisco.com

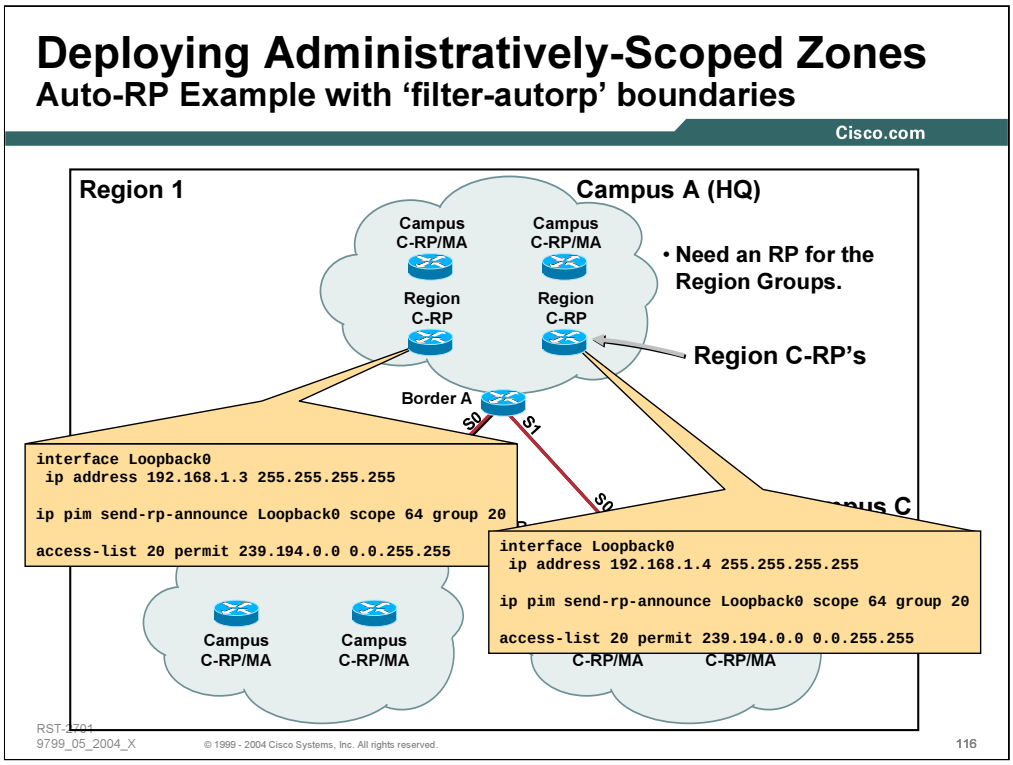


RST-2701
9799_05_2004_X

115

- **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.

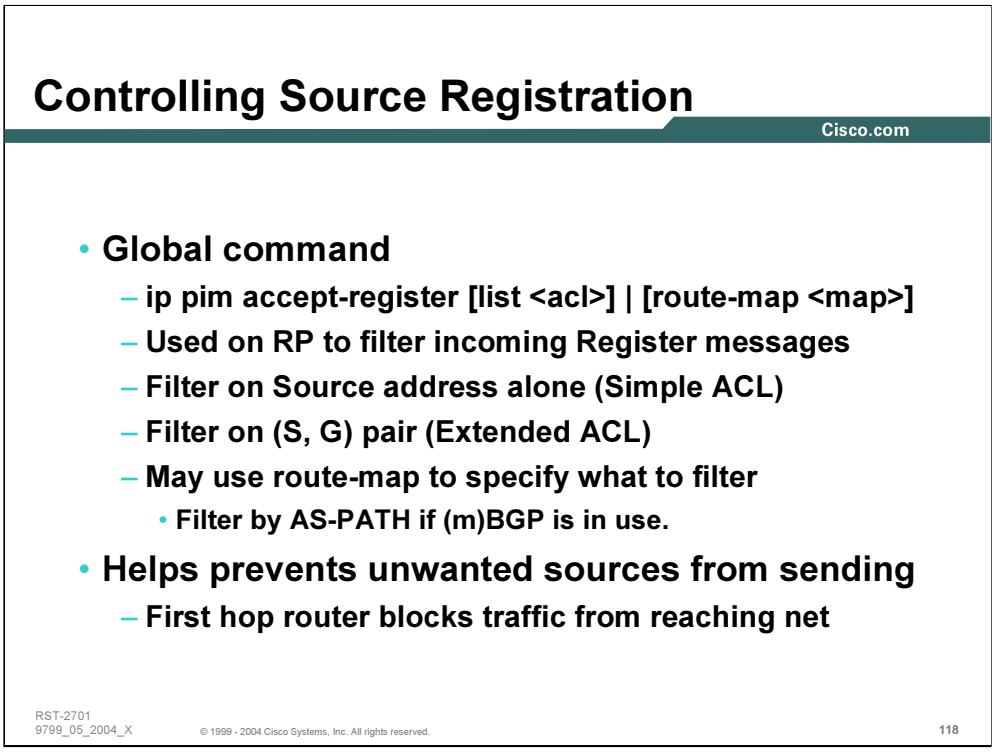


- **Auto-RP Example with 'filter-autorp' boundaries**
  - Finally, it is necessary to configure Organization-Local Candidate RPs that will serve as the RP for all other groups *except* the Site-Local group range. In this case, we have chosen to place all C-RP's for this group range at the HQ site although 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.

SECURITY



- **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)**

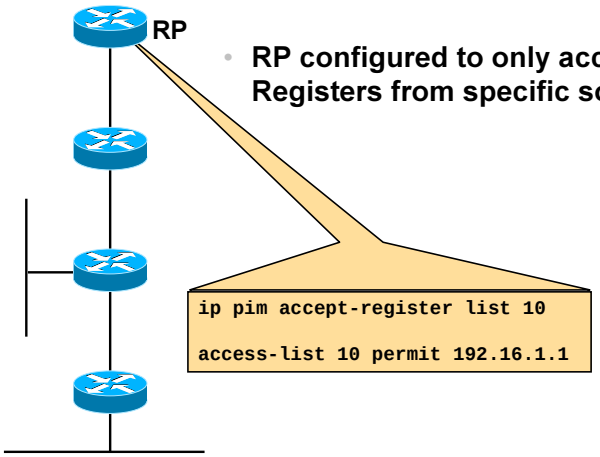
  ip pim accept-register [list <acl>] | [route-map <map>]

  – 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



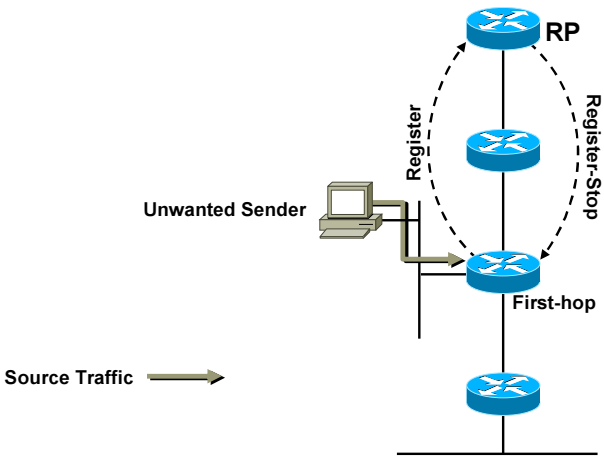
- RP configured to only accept Registers from specific source.

RST-2701
9799_05_2004_X

119

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-2701
9799_05_2004_X

120

Worms

Cisco.com

- **To propagate, the Sasser worm sent a packet to TCP port 445 of random IP addresses, including multicast. Filter TCP destined to all multicast addresses:**

  access-list 115 deny tcp any 224.0.0.0 15.255.255.255

- **Pings (ICMP) used to scan for hosts to infect. Filter ICMP packets towards 224/4:**

  access-list 115 deny icmp any 224.0.0.0 15.255.255.255

Commands used to Protect

Cisco.com

- **ip multicast route-limit** <routes>

  Use this command to limit the impact of Denial of Service attacks based on creating useless IP multicast routing state.
- **ip pim rp-address** <ip-address> [<group-access-list>] [override]

  A single RP will be used only for certain defined groups.
- **ip pim accept-rp** {<address> | auto-rp} [<acl>]

  To configure a router to accept Joins or Prunes destined for a specified RP and for a specific list of groups.
- **ip msdp sa-filter** in|out <ip-address-or-name> [list <acl>]

  Filters incoming/outgoing SA messages to/from a peer.
- **ip msdp sa-limit** <peer-address-or-name> <limit>

  Introduced as a mean of protection against (distributed) denial of service attacks. Limits the overall number of SA messages the router will accept from a peer.

RST-2701
9799_05_2004_X

122

Commands used to Protect

Cisco.com

- **ip pim neighbor-filter** <acl>

  Used to administratively deny a misconfigured PIM neighbor from participating in PIM
- **ip pim bsr-border**

  Bootstrap messages will not be able to pass through this border in either direction.
- **ip multicast boundary** <acl>

  No multicast data packets, defined in acl, will be allowed to flow across the boundary from either direction. For example, to configure a boundary for all administratively scoped addresses, do:

```
access-list 1 deny 239.0.0.0 0.255.255.255
access-list 1 permit 224.0.0.0 15.255.255.255
interface ethernet 0
ip multicast boundary 1
```

RST-2701
9799_05_2004_X

123

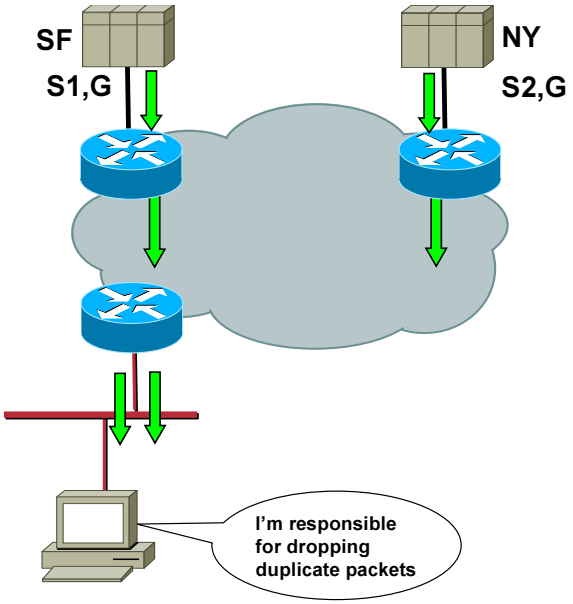
SOURCE REDUNDANCY

RST-2701
9799_05_2004_X

124

Source Redundancy (Duplicate Streams)

Cisco.com

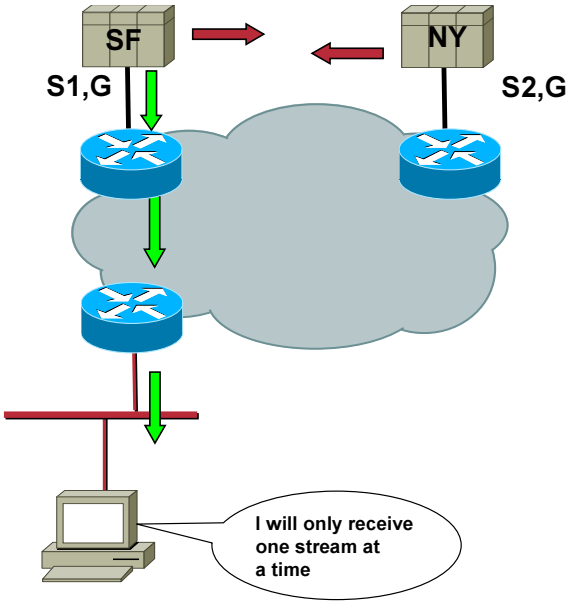


RST-2701
9799_05_2004_X

125

Source Redundancy (Server Heartbeat)

Cisco.com

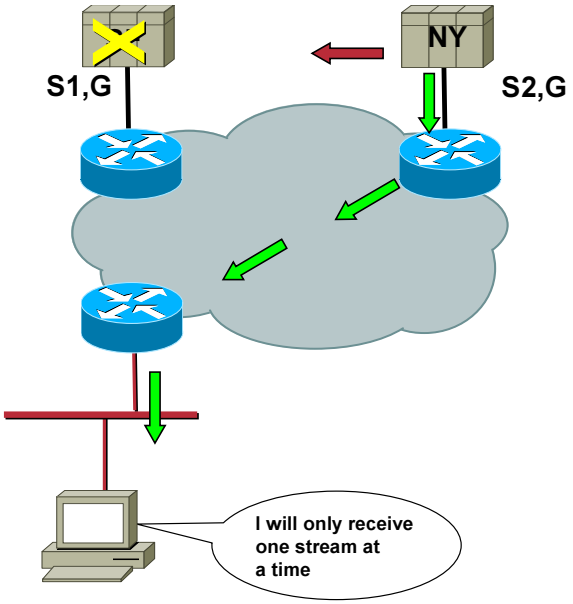


RST-2701
9799_05_2004_X

126

Source Redundancy (Server Heartbeat)

Cisco.com

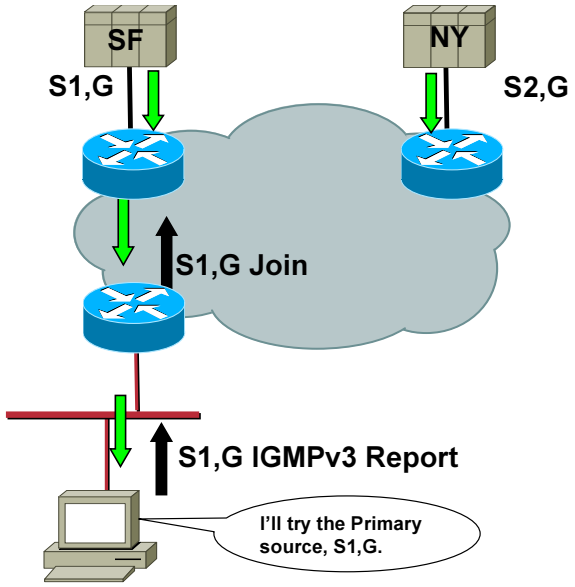


RST-2701
9799_05_2004_X

127

Source Redundancy (SSM)

Cisco.com

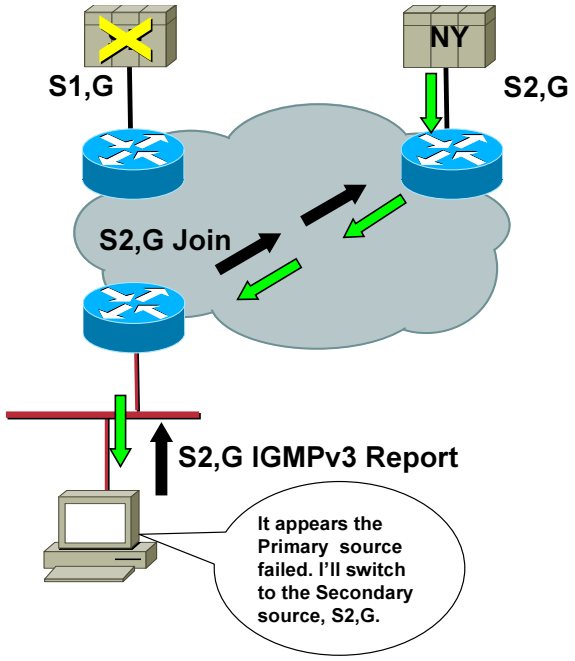


RST-2701
9799_05_2004_X

128

Source Redundancy (SSM)

Cisco.com

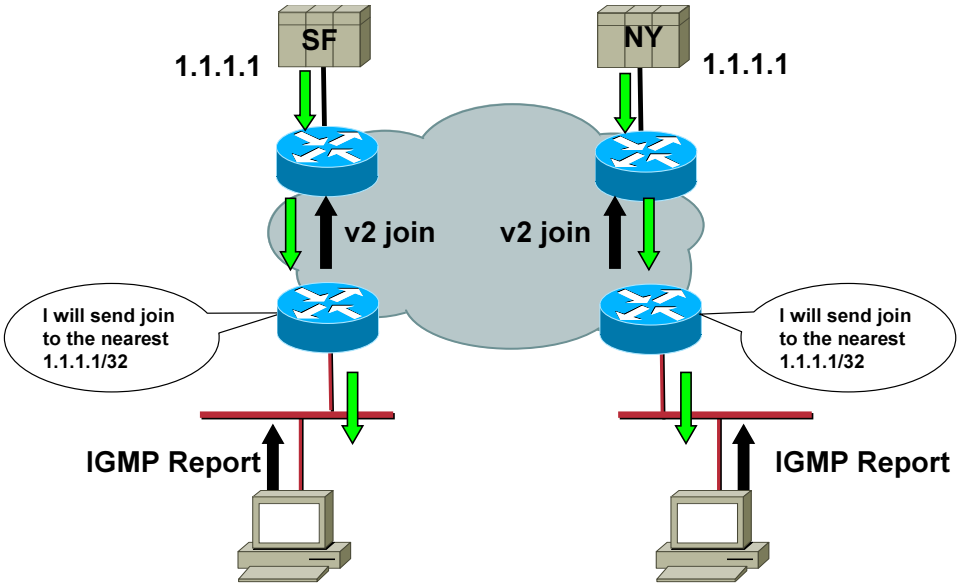


RST-2701
9799_05_2004_X

129

Anycast Sources

Cisco.com

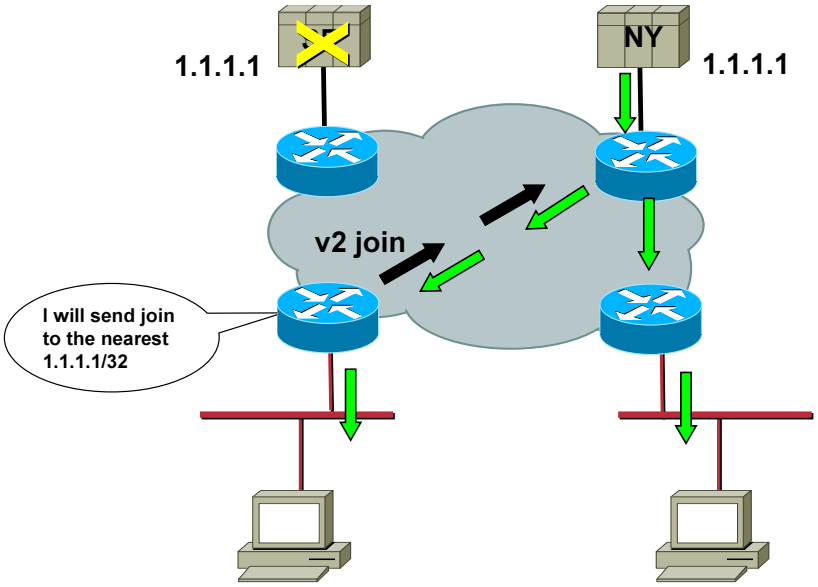


RST-2701
9799_05_2004_X

130

Anycast Sources

Cisco.com

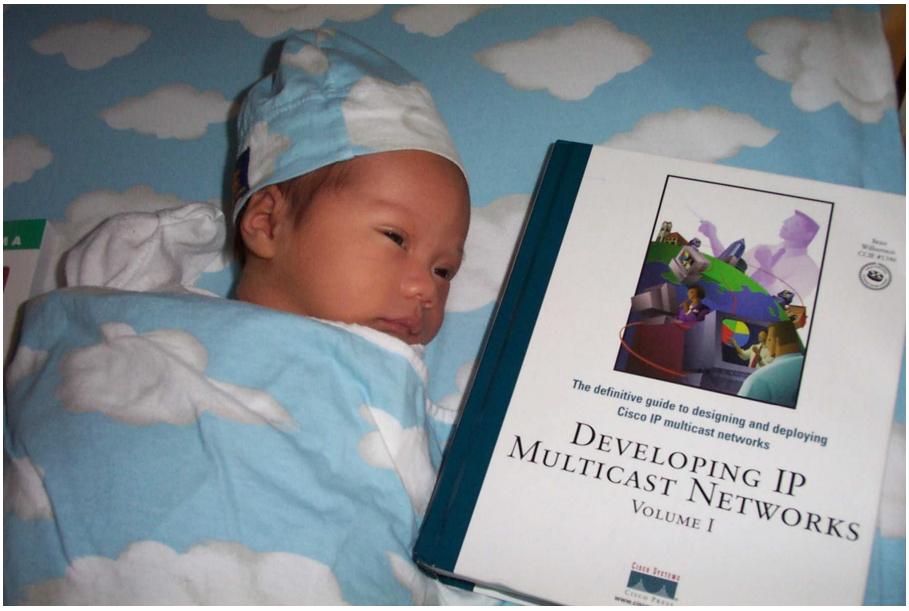


RST-2701
9799_05_2004_X

131

Wonderful Bedtime Stories

Cisco.com



RST-2701
9799_05_2004_X