



## Other Related Presentations

- **Multicast Sessions**

<u>Session #</u>	<u>Title</u>
2214	Introduction to IP Multicast
2215	PIM Multicast Routing
2216	Deploying IP Multicast
2217	Advanced IP Multicast Routing

- **MBGP Related Sessions**

<u>Session #</u>	<u>Title</u>
2209	Deploying BGP

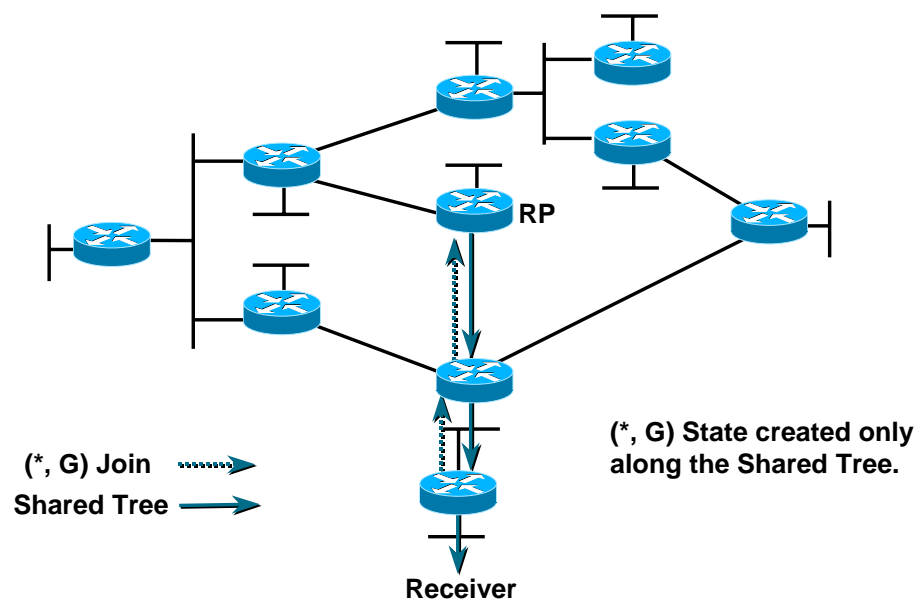
## Agenda

- **PIM-SM review (forwarding)**
- **Inter-Domain Multicast**
  - MBGP (routing)
  - MSDP (source discovery)
  - Topology Examples
- **SSM (Source Specific Multicast)**

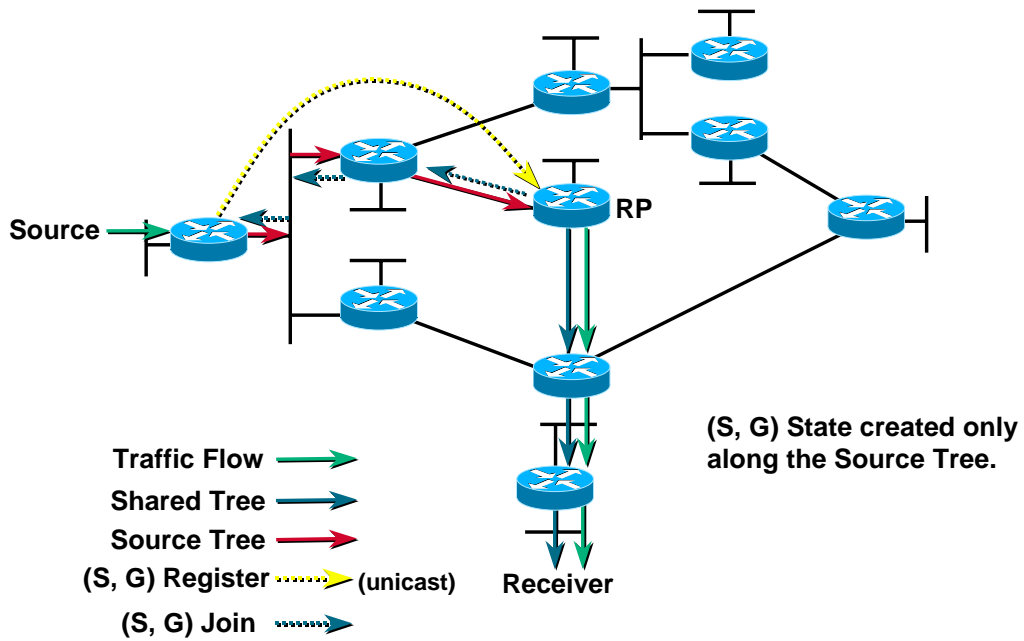
# PIM-SM Review

- Everyone here is a PIM-SM expert, right?
- So just for review...

# PIM-SM Shared Tree Join



# PIM-SM Sender Registration



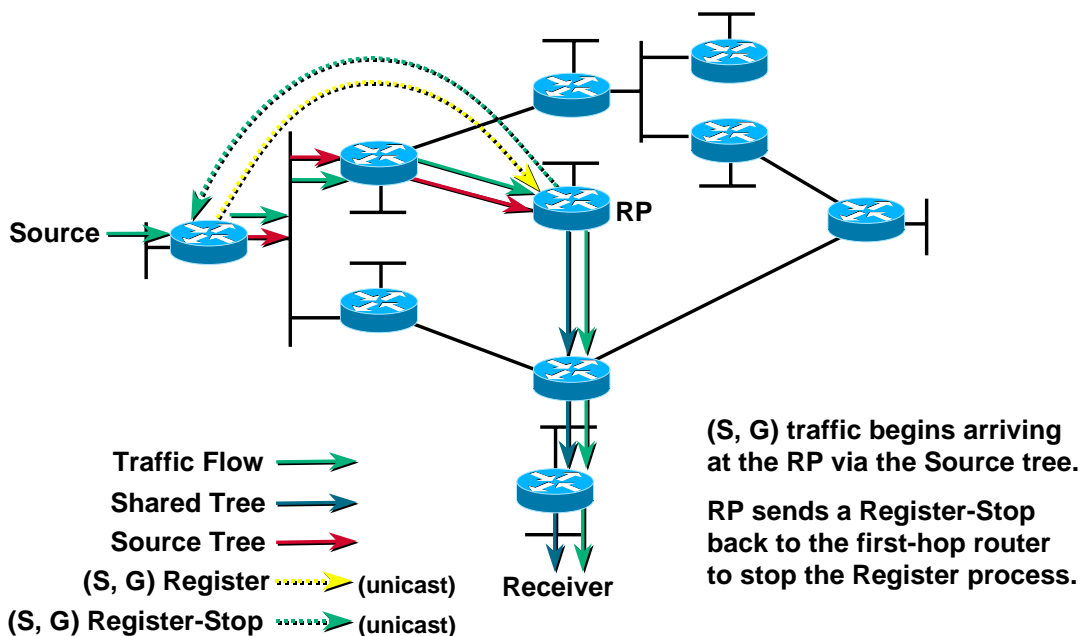
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

7

# PIM-SM Sender Registration



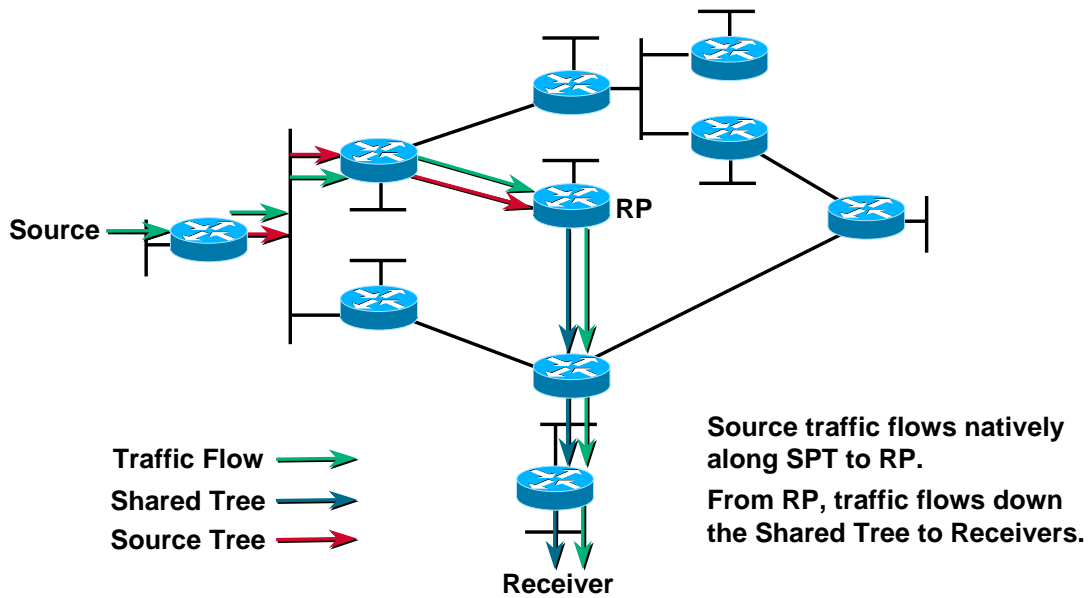
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

8

# PIM-SM Sender Registration



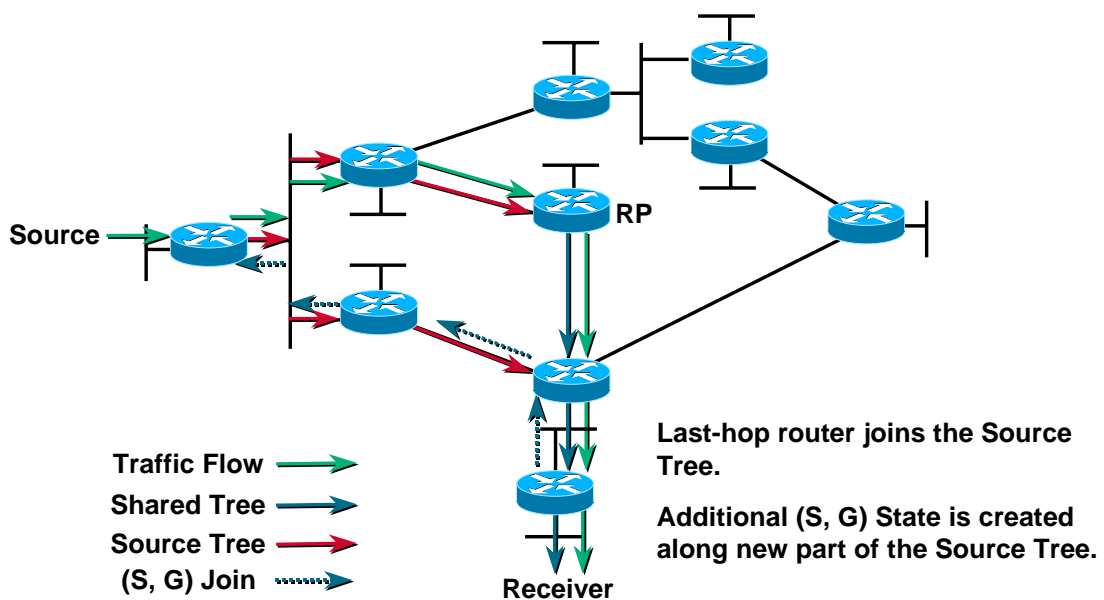
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

9

# PIM-SM SPT Switchover



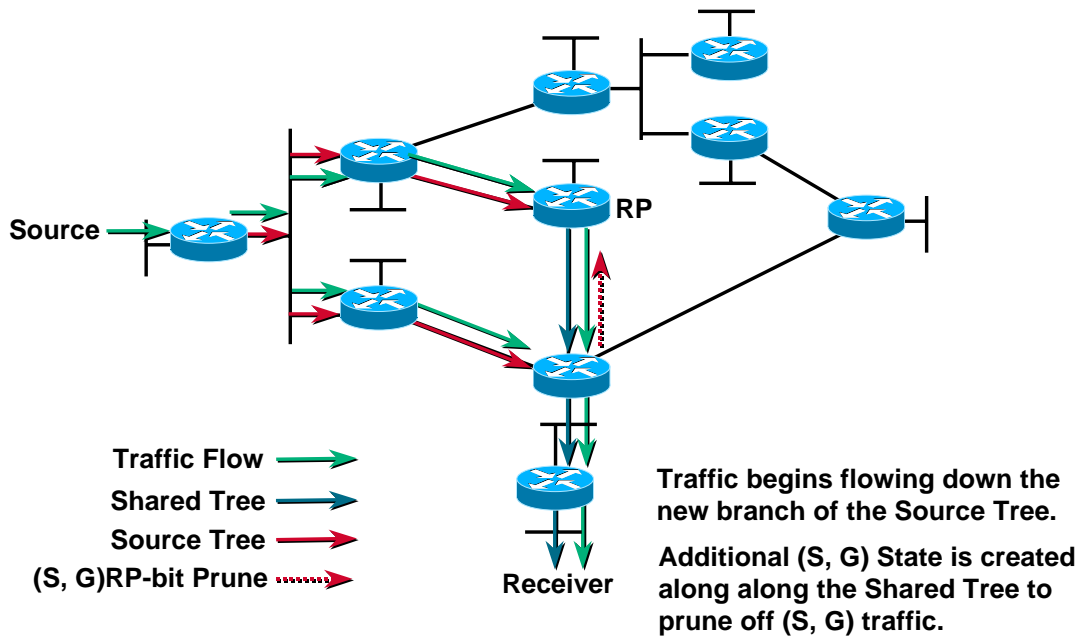
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

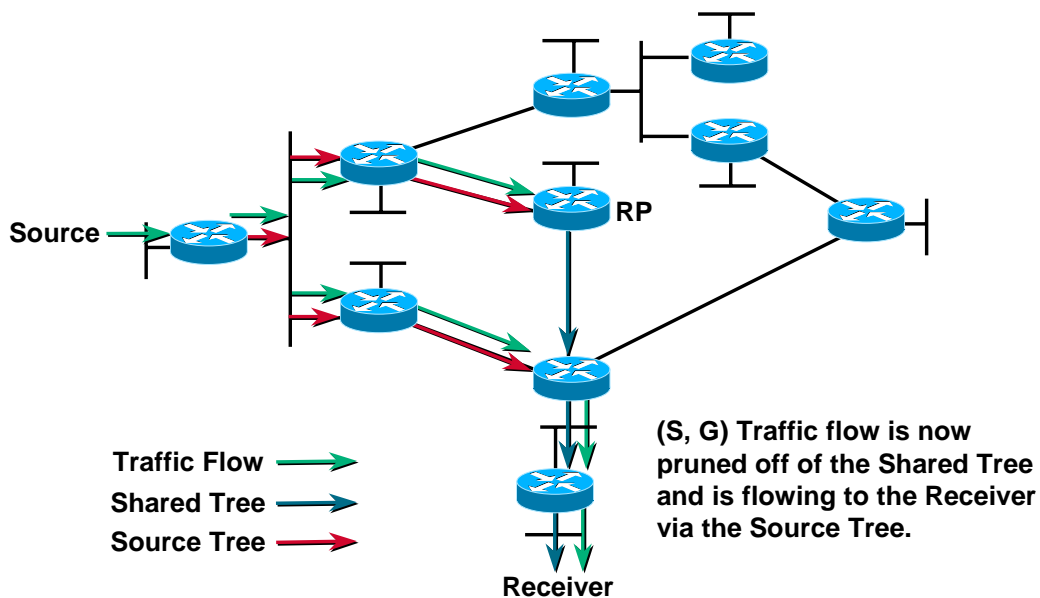
cisco.com

10

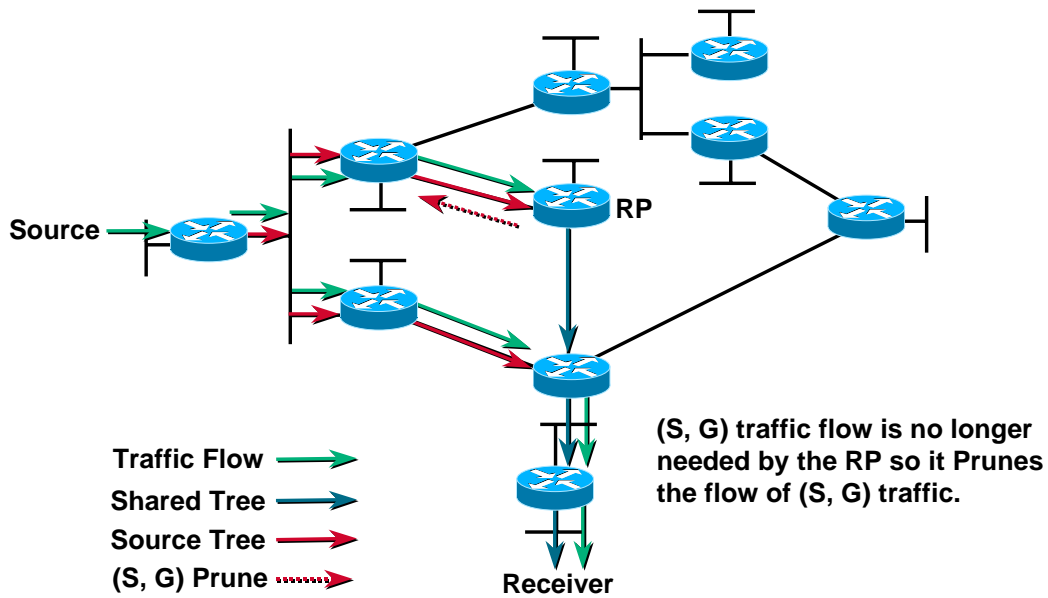
# PIM-SM SPT Switchover



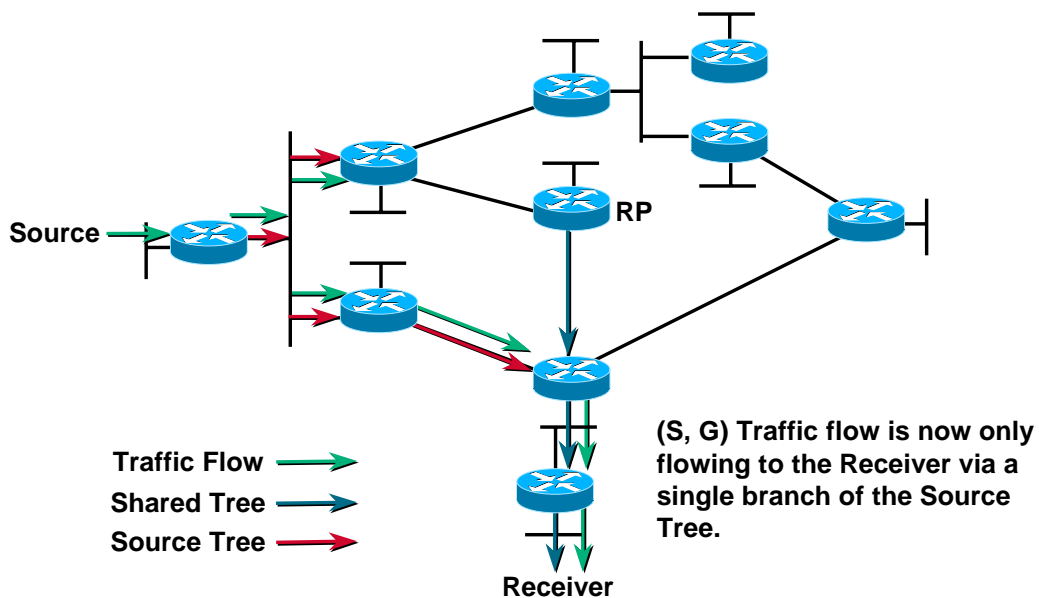
# PIM-SM SPT Switchover



# PIM-SM SPT Switchover



# PIM-SM SPT Switchover



# Agenda

- PIM-SM review (forwarding)
- **Inter-Domain Multicast**
  - MBGP (routing)
  - MSDP (source discovery)
  - Topology Examples
- SSM (Source Specific Multicast)

# Inter-Domain Multicast

- Past history
- In the future
- **ISP requirements to deploy now**
  - PIM-SM
  - MBGP
  - MSDP
  - Topology Examples

## Past History

**A long time ago,  
in a galaxy  
far, far away...**

## Past History

**MBONE...**

## Past History

**MBONE... vat, nv, wb, sd,...**

## Past History

- **DVMRP MBONE**
  - **Virtual network overlaid (tunneled) on the unicast Internet infrastructure**
  - **DVMRP MBONE uses RIP-like routing**
  - **Flood and Prune technology**
  - **Initially instantiated by MROUTED, and later implemented by various router vendors**
  - **Very successful in academic circles**

# Past History

## Problem

- **DVMRP can't scale to Internet sizes**
  - Distance vector-based routing protocol
  - Periodic updates
    - Full table refresh every 60 seconds
  - Table sizes
    - Internet > 40,000 prefixes
  - Stability
    - Hold-down, count-to-infinity, etc.

# In the Future?

- **BGMP (Border Gateway Multicast Protocol)**
  - Shared tree of domains
    - Bidirectional trees
    - Explicit join-model
    - Joins sent toward root domain
  - Single root domain per group
    - Multicast group prefixes assigned by domain
    - MASC proposed as assignment method
  - Requires BGP4+ (aka MBGP)
    - Must carry group prefixes in NLRI field
    - Needed to build bidirectional trees



## In the Future?

- **MASC (Multicast Address Set-Claim)**
  - **Multicast address space is hierarchical**
    - Top of hierarchy is at an Internet exchange
    - Children get address space from parent
    - Results in aggregateable multicast address space
  - **Allocation has a lifetime**
    - Children must renew address allocation
    - May not receive same space at renewal time
    - Parent may reclaim space at renewal time
    - Permits reallocation of space
    - Complex “garbage collection” problem

## In the Future

- **BGMP and MASC are a long ways off**
  - Both are quite complex to implement
  - Still only in draft proposal stages
- **ISP's are deploying multicast now**
  - What are their minimum requirements?

# ISP Requirements to Deploy Now

- **Want an explicit join protocol for efficiency**
  - ✓ **PIM-SM (forwarding)**
- **Use existing (unicast) routing model**
  - Hmmm
- **Need interdomain source discovery**
  - **Separate PIM domain RP's need to share source information**
  - Hmmm

# ISP Requirements to Deploy Now

- **Use existing (unicast) routing model**
  - **Need same tool-set for multicast as unicast**
    - **Robust set of peering and policy controls**
    - **Ability to separate unicast and multicast topologies**
    - **Use familiar configuration, operation and terminology model**
  - **Something like BGP but for multicast**
- **Solution: Multiprotocol BGP (MBGP)**

# MBGP—Multiprotocol BGP

- **MBGP overview**
- **MBGP capability negotiation**
- **MBGP NLRI exchange**
- **MBGP-DVMRP redistribution**
- **Unicast to Multicast NLRI Translation**

## MBGP Overview

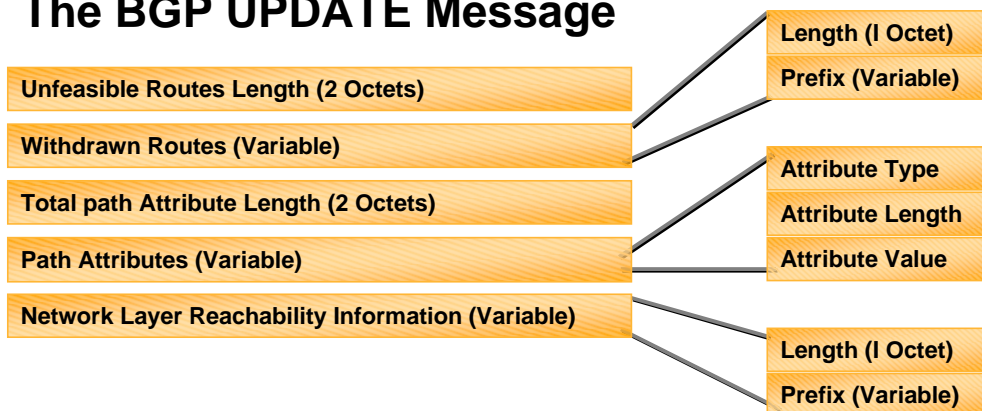
- **MBGP: Multiprotocol BGP**  
**(aka multicast BGP in multicast networks)**
  - **Defined in RFC 2283 (extensions to BGP)**
  - **Can carry different types of routes**
    - **Unicast**
    - **Multicast**
  - **Both routes carried in same BGP session**
  - **Does not propagate multicast state info**
  - **Same path selection and validation rules**
    - **AS-Path, LocalPref, MED, ...**

# MBGP Overview

- **New multiprotocol attributes**
  - MP\_REACH\_NLRI
  - MP\_UNREACH\_NLRI
- **MP\_REACH\_NLRI and MP\_UNREACH\_NLRI**
  - **Address Family Information (AFI) = 1 (IPv4)**
    - Sub-AFI = 1 (NLRI is used for unicast)
    - Sub-AFI = 2 (NLRI is used for multicast RPF check)
    - Sub-AFI = 3 (NLRI is used for both unicast and multicast RPF check)

# MBGP Overview

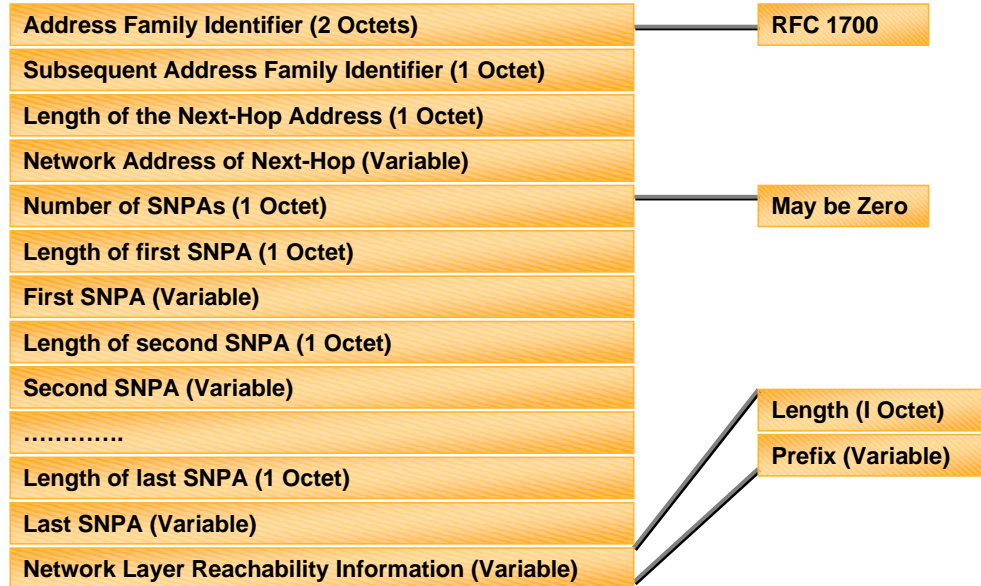
## The BGP UPDATE Message



- A BGP update is used to advertise a single feasible route to a peer, or to withdraw multiple unfeasible routes
- Each update message contains attributes, like origin, AS-Path, Next-Hop, ..... MP\_REACH\_NLRI

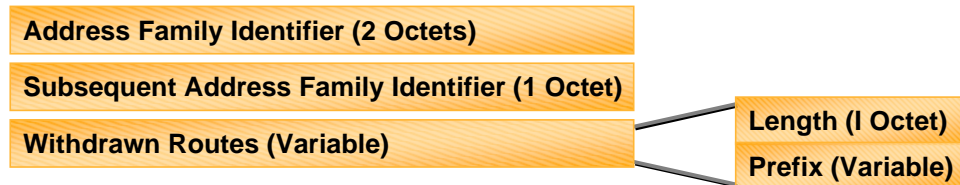
# MBGP Overview

## MP\_REACH\_NLRI Attribute



# MBGP Overview

## MP\_UNREACH\_NLRI Attribute



# MBGP Overview

- **Separate BGP tables maintained**
  - Unicast Routing Information Base (RIB)
  - Multicast Routing Information Base (MRIB)
- **Unicast RIB (U-RIB)**
  - Contains unicast prefixes for unicast forwarding
  - Populated with BGP unicast NLRI
    - AFI = 1, Sub-AFI = 1 or 3
- **Multicast RIB (M-RIB)**
  - Contains unicast prefixes for RPF checking
  - Populated with BGP multicast NLRI
    - AFI = 1, Sub-AFI = 2 or 3

# MBGP Overview

- **MBGP allows different unicast and multicast topologies and different policies**
  - Same IP address may have different signification
    - Unicast routing information
    - Multicast RPF information
  - For same IPv4 address two different NLRI with different next-hops
  - Can use existing or new BGP peering topology for multicast

# MBGP Overview

- **What is in the Cisco IOS® implementation?**
  - All the familiar BGP configuration knobs
  - Carries multicast routes in MP\_REACH\_NLRI
  - NLRI capability negotiation
  - Redistribution between MBGP and IGP (includes DVMRP)
  - Unicast to Multicast NLRI Translation
    - Used for older non-MBGP stub sites

## MBGP—Capability Negotiation

- **BGP routers establish BGP sessions through the OPEN message**
- **OPEN message contains optional parameters**
- **BGP session is terminated if OPEN parameters are not recognised**
- **New parameter: CAPABILITIES**
  - Multiprotocol extension
  - Multiple routes for same destination

## MBGP—Capability Negotiation

- **New keyword on neighbor command**

```
neighbor <foo> remote-as <asn> nlri multicast unicast
```

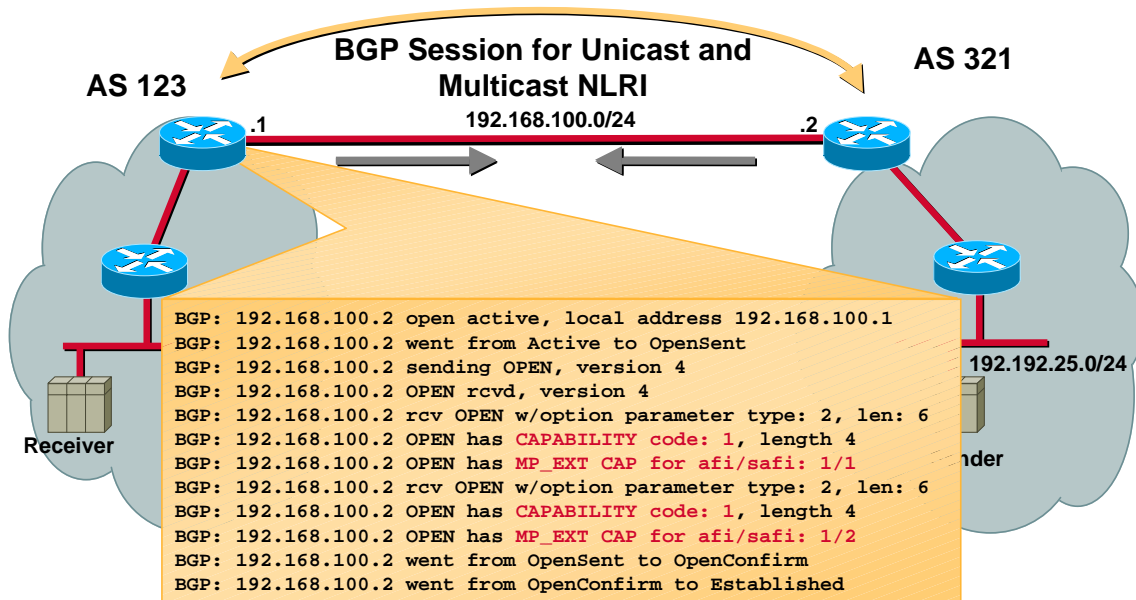
- **Configures router to negotiate either or both NLRI**
- **If neighbor configures both or subset, common NLRI is used in both directions**
- **If there is no match, notification is sent and peering doesn't come up**

## MBGP—Capability Negotiation

- **If neighbor doesn't include the capability parameters in open, Cisco backs off and reopens with no capability parameters**
- **Peering comes up in unicast-only mode**
- **Hidden command**

```
neighbor <foo> dont-capability-negotiate
```

# MBGP—Capability Negotiation



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

41

# MBGP NLRI Exchange

- **BGP/MBGP configuration allows you to:**
  - Define which NLRI type are exchanged (unicast, multicast, both)
  - Set NLRI type through route-maps (redistribution)
  - Define policies through standard BGP attributes (for unicast and/or multicast NLRI)
- **No redistribution allowed between MBGP and BGP tables**
  - NLRI type can be set with set nlri route-map command

2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

42

# MBGP NLRI Exchange

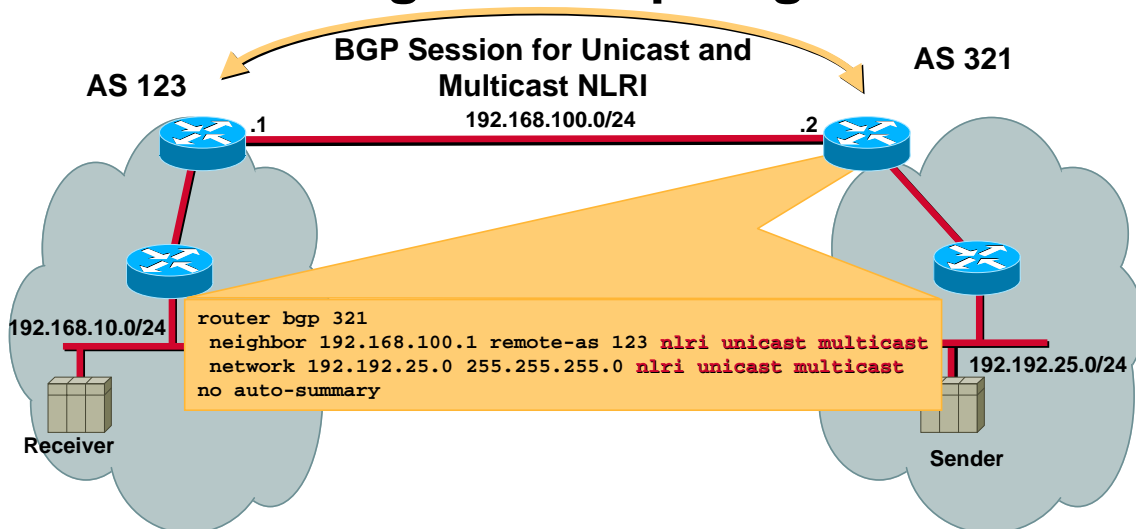
- MRIB is populated by:
  - Receiving AFI/SAFI 1/2 MP\_REACH\_NLRI from neighbors
  - Configured/stored locally by:

```
network <foo> <foo-mask> [nlri multicast unicast]  
redistribute <unicast> route-map <map>  
aggregate-address <foo> <foo-mask> [nlri multicast unicast]  
neighbor <foo> default-originate [nlri multicast unicast]
```

- Note: Syntax changes in 12.07T/12.1 forward

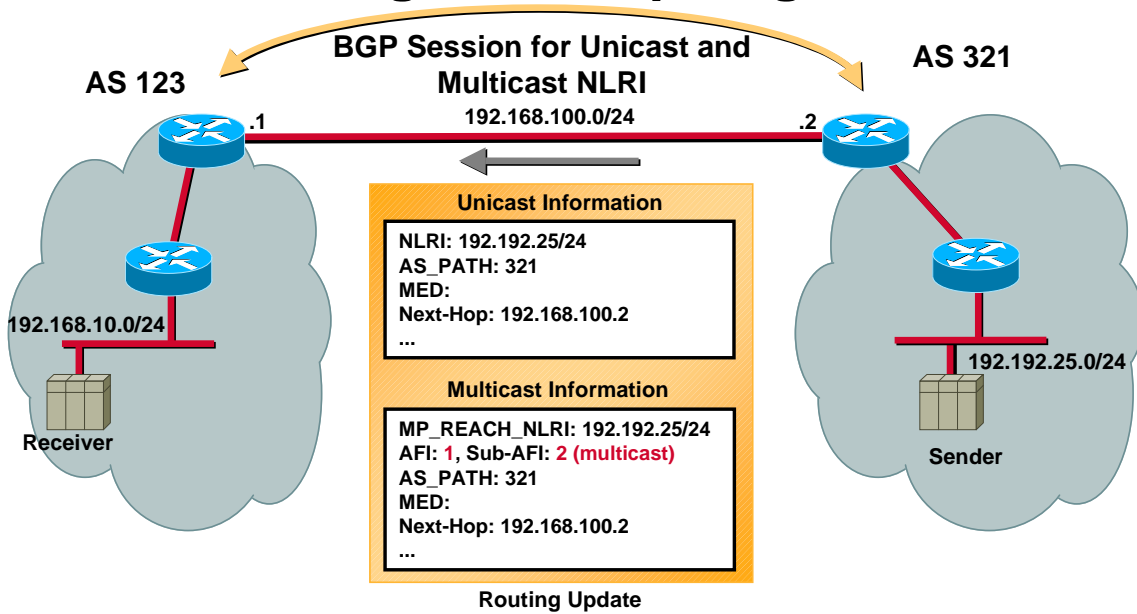
# MBGP — NLRI Information

## Congruent Topologies



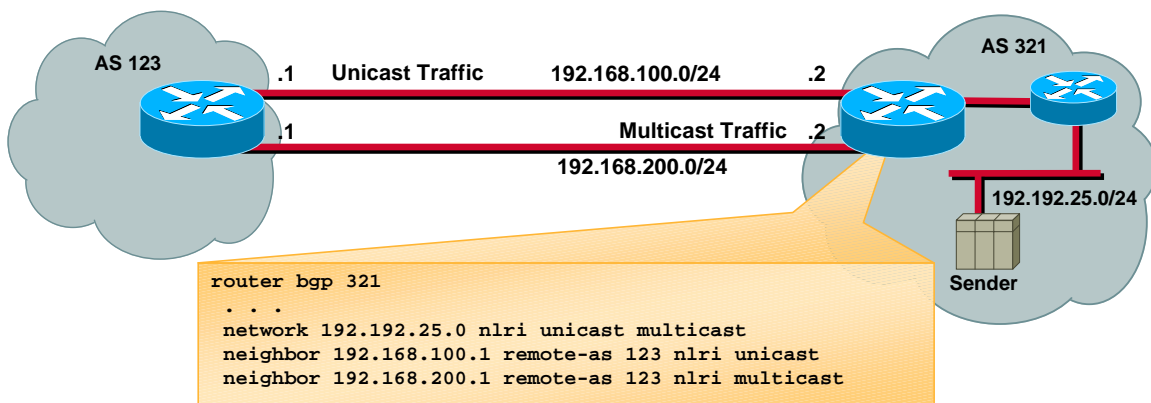
# MBGP — NLRI Information

## Congruent Topologies



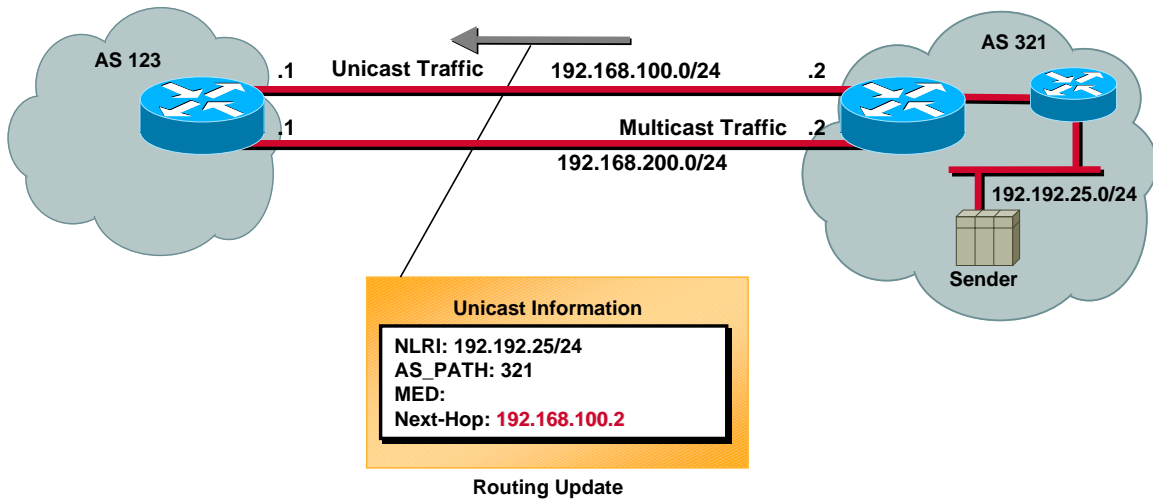
# MBGP — NLRI Information

## Incongruent Topologies



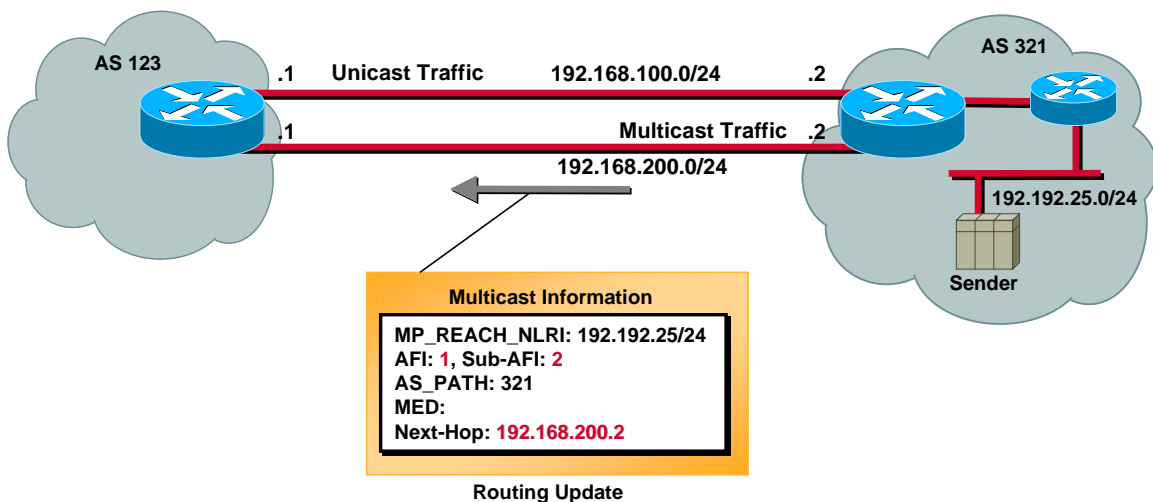
# MBGP — NLRI Information

## Incongruent Topologies



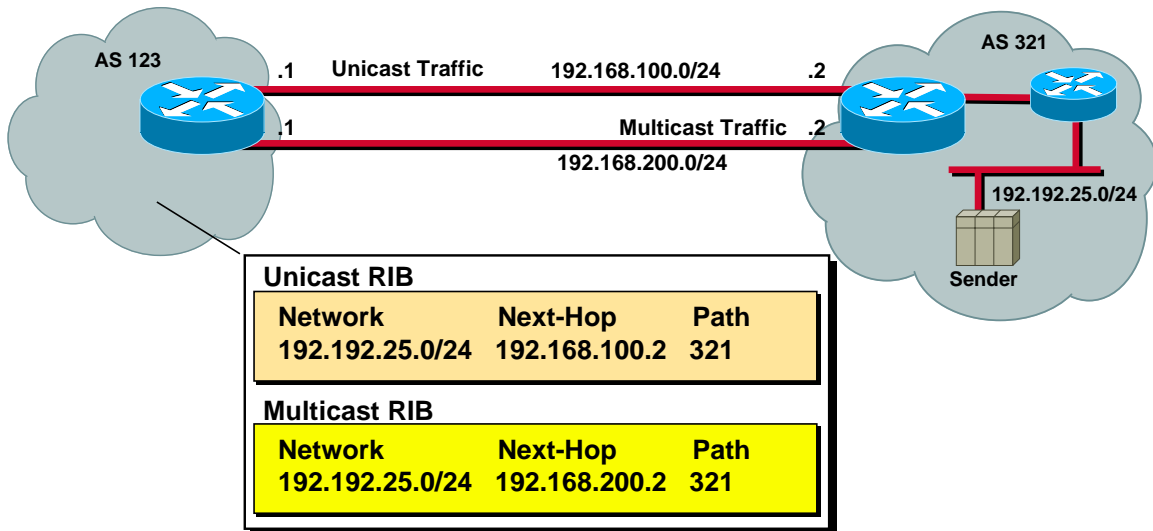
# MBGP — NLRI Information

## Incongruent Topologies

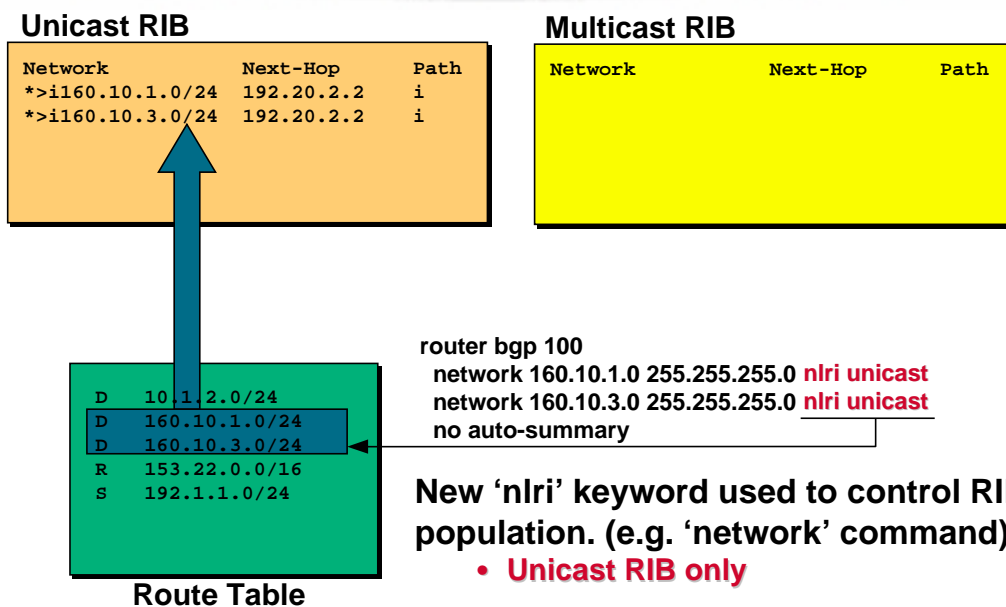


# MBGP — NLRI Information

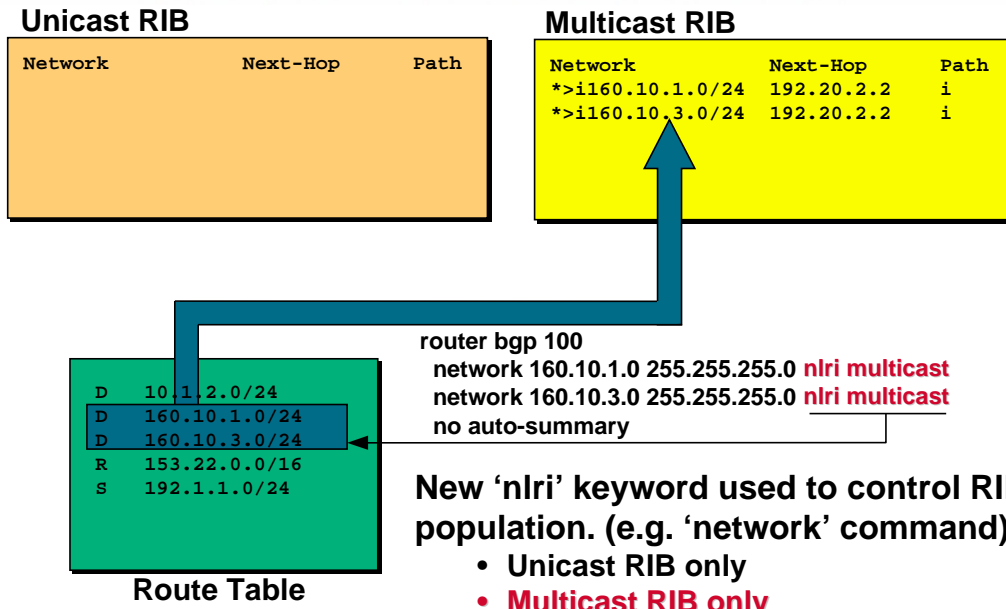
## Incongruent Topologies



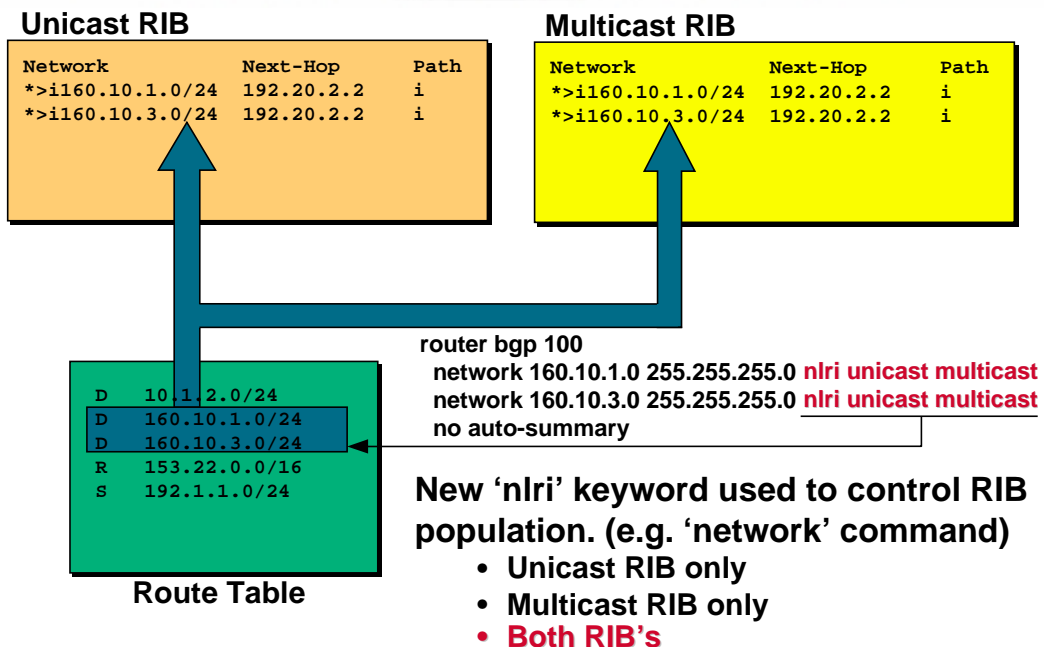
# MBGP — NLRI Information



# MBGP — NLRI Information



# MBGP — NLRI Information



# Populating the MRIB

- **Just to restate**

```
network <foo> <foo-mask> [nlri multicast unicast]  
redistribute <unicast> route-map <map>  
aggregate-address <foo> <foo-mask> [nlri multicast unicast]  
neighbor <foo> default-originate [nlri multicast unicast]
```

# DVMRP <-> MBGP Redistribution

- **DVMRP can be distributed as any other IGP**

```
router bgp <asn>  
redistribute dvmrp route-map <map>
```

- **You can do your typical set operations**
- **Used to connect legacy sites to the current native multicast Internet.**
- **Currently used at the AMES MIX to connect with the old DVMRP MBONE**

## Unicast-Multicast NLRI Translation

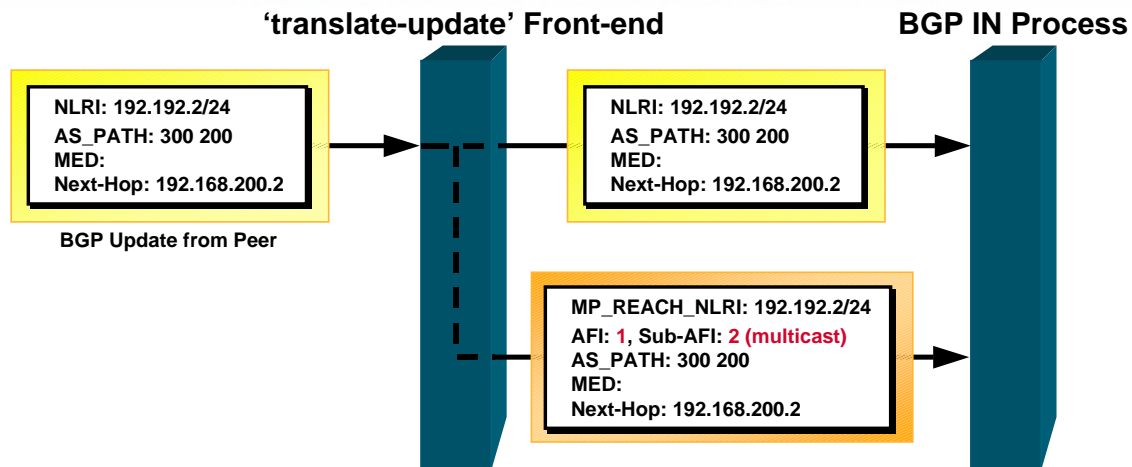
- **Strictly transitional—not recommended**
- **BGP stubs that don't have MBGP support need to get their routes into the MBGP backbone**
- **Stub gets external routes via existing BGP peering**
- **Use hidden command**

```
neighbor <foo> translate-update [nlri unicast multicast]
```

## Unicast-Multicast NLRI Translation

- **BGP Update received by translating router is translated into an MP\_REACH\_NLRI attribute**
  - **As if the neighbor sent AFI 1/SAFI 2 routes**

# Unicast-Multicast NLRI Translation



- Arriving Unicast update intercepted by 'translate-update' Front-end
- A translated Multicast update is created & passed to the IN Process
- Original Unicast update is passed on to the IN Process
- Both updates processed normally by the IN Process

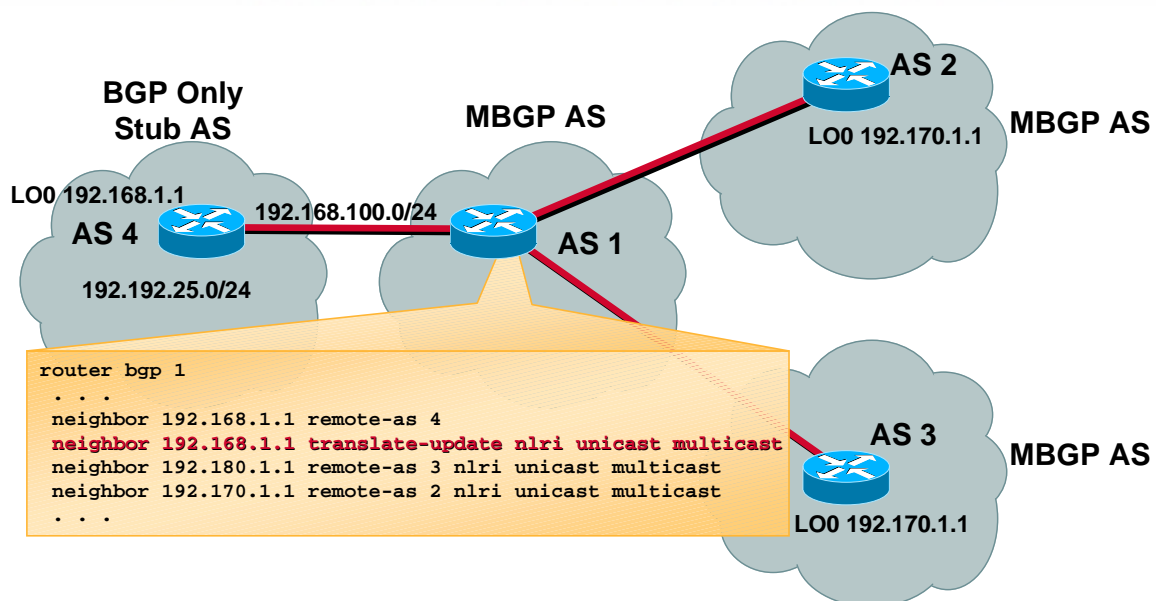
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

57

# Unicast-Multicast NLRI Translation



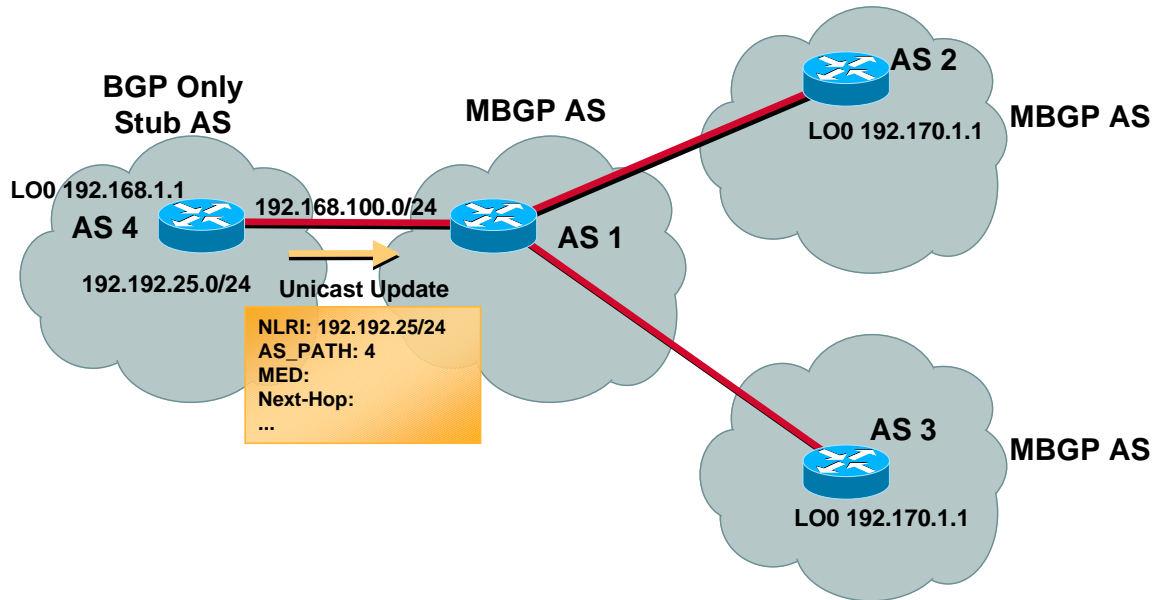
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

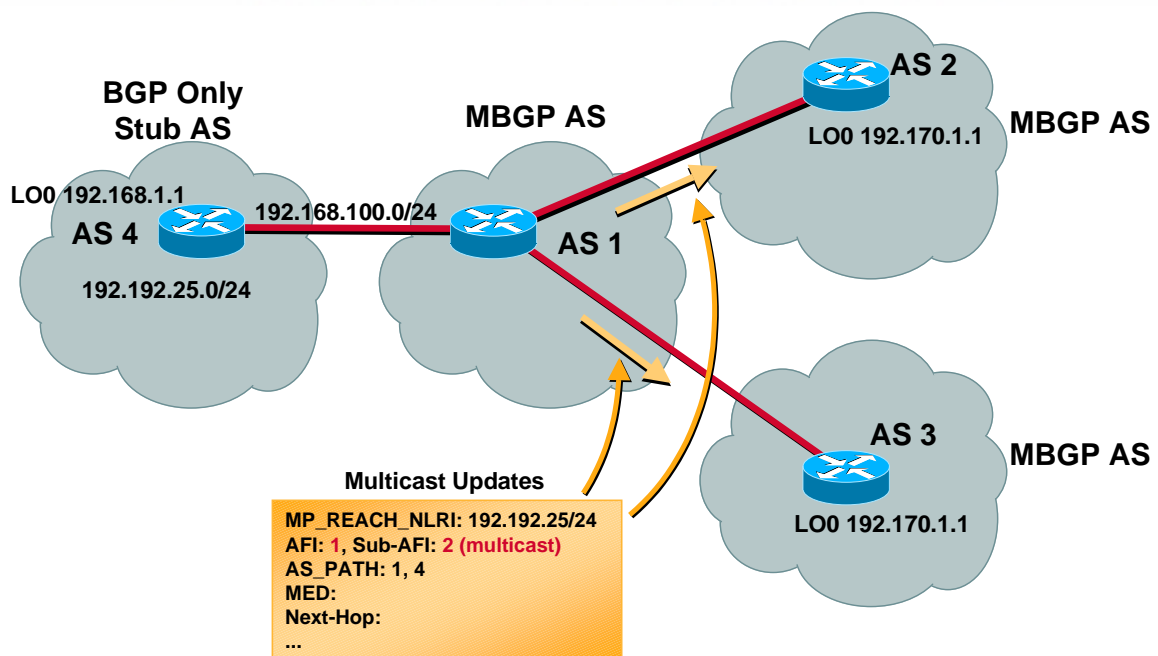
cisco.com

58

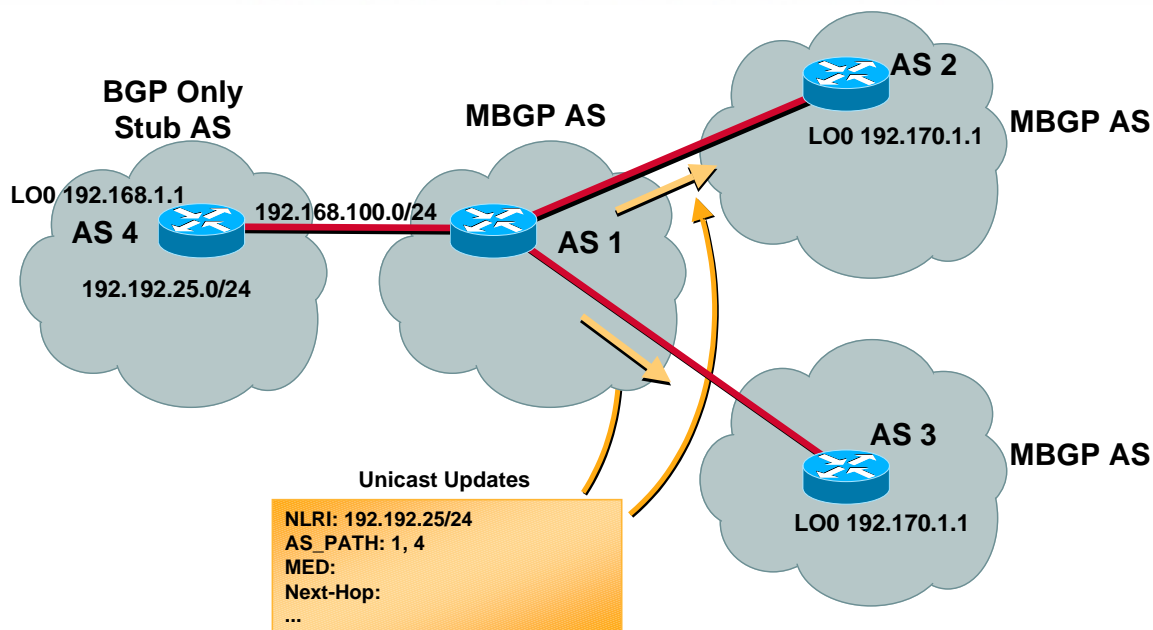
# Unicast-Multicast NLRI Translation



# Unicast-Multicast NLRI Translation



# Unicast-Multicast NLRI Translation



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

61

## MBGP—Summary

- **Solves part of inter-domain problem**
  - Can exchange multicast routing information
  - Uses standard BGP configuration knobs
  - Permits separate unicast and multicast topologies if desired
- **Still must use PIM to:**
  - Build distribution trees
  - Actually forward multicast traffic
  - PIM-SM recommended
    - But there's still a problem using PIM-SM here... (more on that later)

2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

62

## ISP Requirements to Deploy Now

- **Want an explicit join protocol for efficiency**
  - ✓ PIM-SM (forwarding)
- **Use existing (unicast) operation model**
  - ✓ MBGP (routing)
- **Need interdomain source discovery**
  - Separate PIM domain RP's need to share source information
  - Hmmm

## ISP Requirements to Deploy Now

- **Will not share RP with competitors**
  - Firm requirement
  - Third-party resource dependency
    - “If my customers are multicasting on group G whose RP is in my competitor's network and that RP goes down, my customers lose connectivity.”
- **Want flexibility re: RP placement**
  - May need to place RP(s) someplace other than a single interconnect point

# ISP Requirements to Deploy Now

- **Must interconnect PIM-SM domains**
  - **Inter-domain rendezvous mechanism?**
    - Requires dynamic DNS (or something similar)
    - Still results in third-party RP problem
  - **Interconnect using shared trees**
    - That's BGMP! Can't wait
  - **Interconnect using source trees**
    - Need a way to discover all multicast sources
      - Hmm. Interesting idea!
- **Solution: MSDP**
  - **Multicast Source Discovery Protocol**

# MSDP—Multicast Source Discovery Protocol

- **MSDP concepts**
- **MSDP design points**
- **MSDP example**
- **Cisco MSDP implementation**
- **MSDP configuration**
- **MSDP application—Anycast RP**

# MSDP Concept

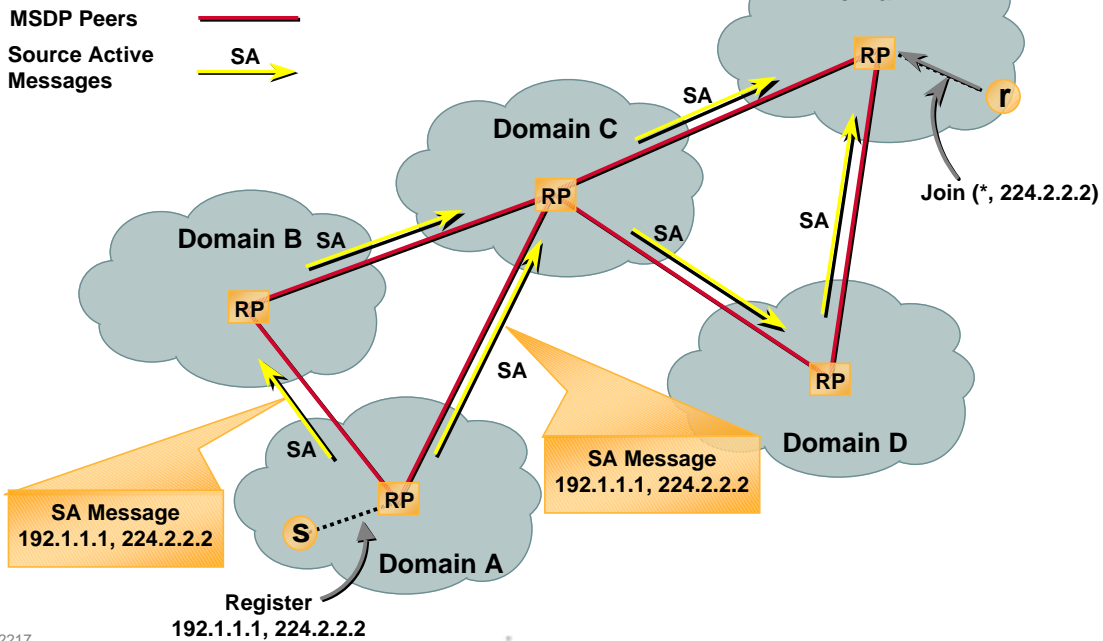
- **Simple but elegant**
  - Utilize inter-domain source trees
  - Reduces problem to locating active sources
  - RP or receiver last-hop can join inter-domain source tree

# MSDP Concepts

- **Works with PIM-SM only**
  - RP's knows about all sources in a domain
    - Sources cause a "PIM Register" to the RP
    - Can tell RP's in other domains of its sources
      - Via MSDP SA (Source Active) messages
  - RP's know about receivers in a domain
    - Receivers cause a "(\*, G) Join" to the RP
    - RP can join the source tree in the peer domain
      - Via normal PIM (S, G) joins

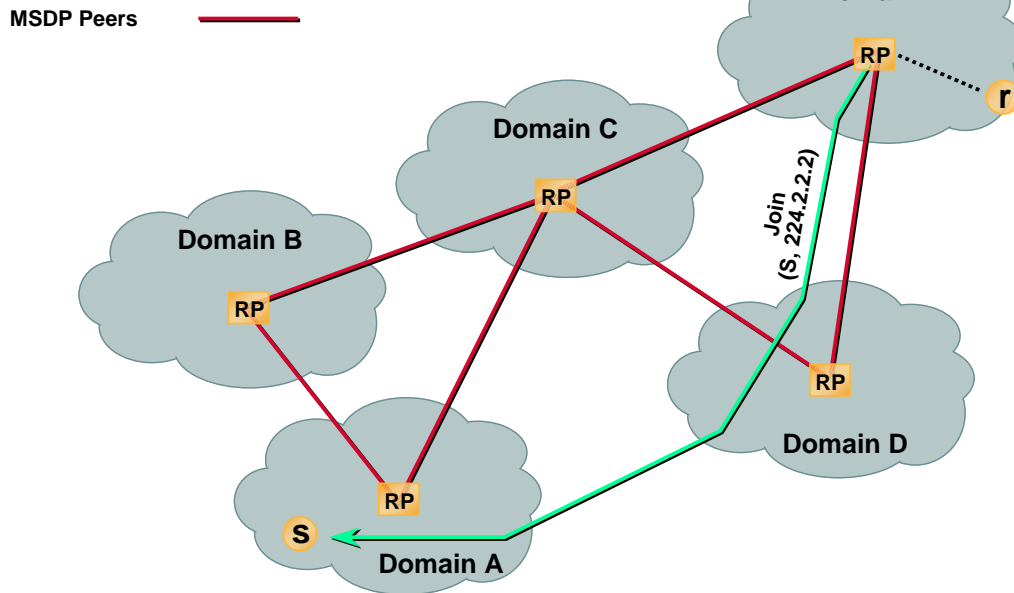
# MSDP Overview

## MSDP Example



# MSDP Overview

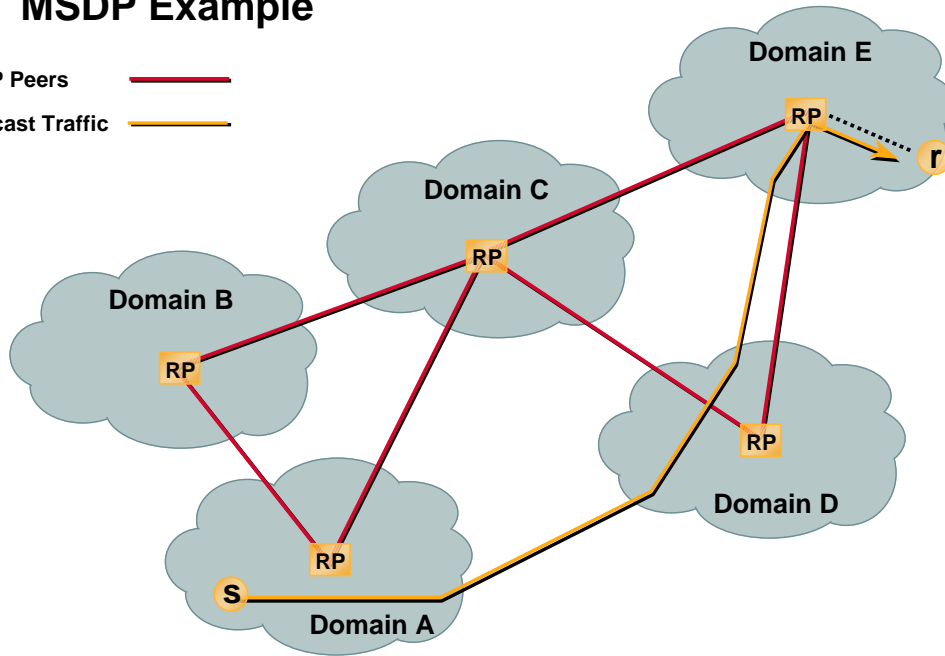
## MSDP Example



# MSDP Overview

## MSDP Example

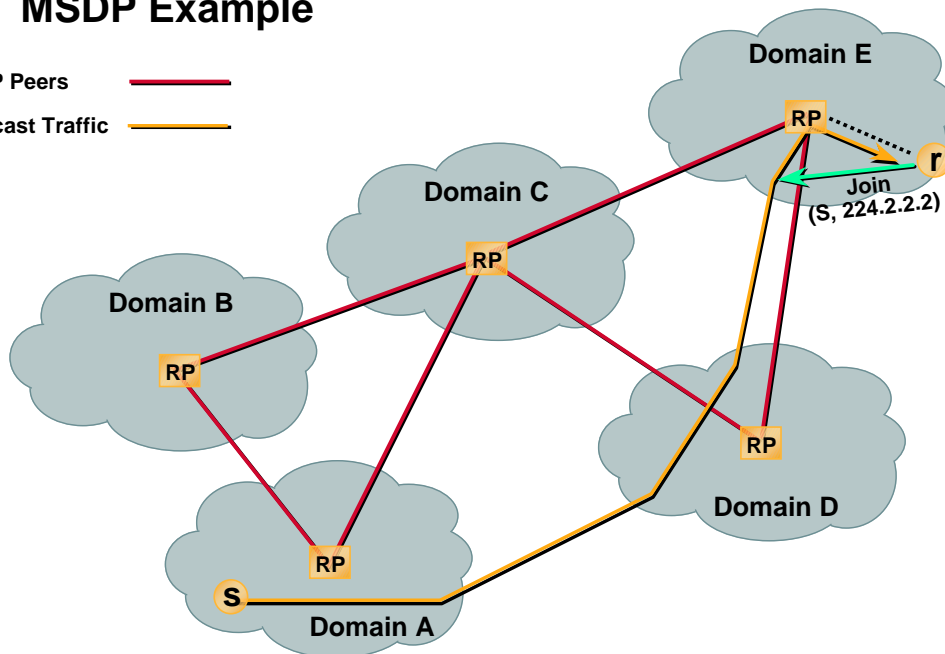
MSDP Peers   
Multicast Traffic 



# MSDP Overview

## MSDP Example

MSDP Peers   
Multicast Traffic 



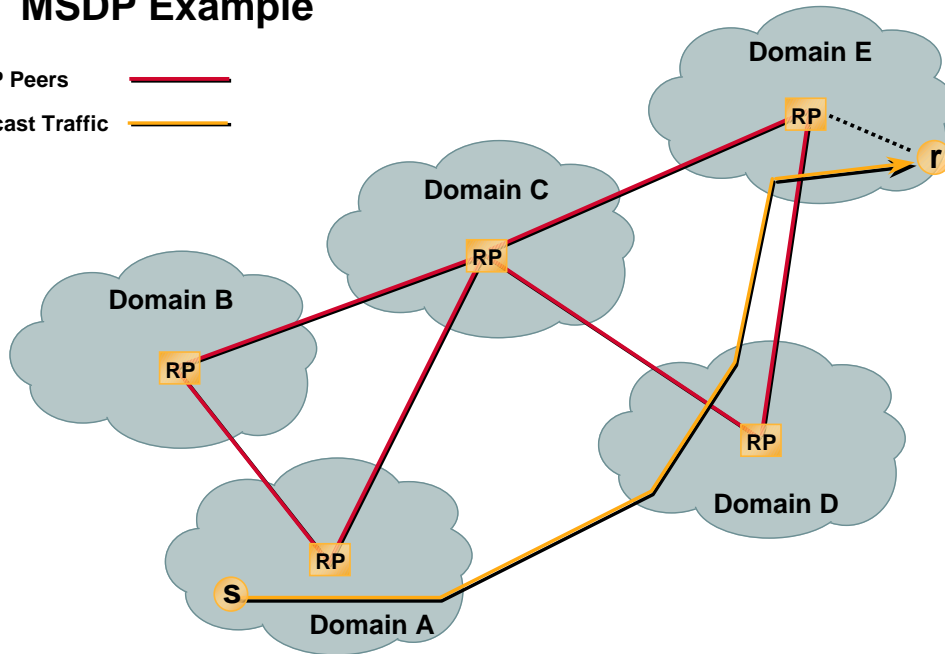
# MSDP Overview

## MSDP Example

MSDP Peers



Multicast Traffic



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

73

# MSDP Design Points

- **MSDP peers talk via TCP connections**
- **Source Active (SA) messages**
  - **Peer-RPF forwarded to prevent loops**
    - RPF check on AS-PATH back to the peer RP
      - other rules from draft apply
    - If successful, flood SA message to other peers
    - Stub sites can be configured to accept all SA messages (similar to static default for routing)
      - Since they have only one exit (e.g., default peer)
  - **MSDP speaker may cache SA messages**
    - Reduces join latency

2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

74

# MSDP Peers

- **MSDP establishes a neighbor relationship between MSDP peers**
  - Peers connect using TCP port 639
  - Peers send keepalives every 60 secs (fixed)
  - Peer connection reset after 75 seconds if no MSDP packets or keepalives are received
- **MSDP peers must run BGP!**
  - May be an MBGP peer, a BGP peer or both
  - Exception: BGP is unnecessary when peering with only a single MSDP peer.

# MSDP SA Messages

- **MSDP SA Message Contents**
  - **One or more messages (in TLV format)**
    - Keepalives
    - Source Active (SA) Messages
    - Source Active Request (SA-Req) Messages
    - Source Active Response (SA-Resp) Message
  - **Source Active (SA) Messages**
    - Used to advertise active Sources in a domain
    - Can also carry initial multicast packet from source
      - Hack for Bursty Sources (ala SDR)
    - SA Message Contents:
      - IP Address of Originator (usually an RP)
      - Number of (S, G)'s pairs being advertised
      - List of active (S, G)'s in the domain
      - Encapsulated Multicast packet

# MSDP SA Messages

- **MSDP Message Contents (cont.)**
  - **SA Request (SA-Req) Messages**
    - Used to request a list of active sources for a group
      - Sent to an MSDP SA Cache Server
      - Reduces Join Latency to active sources
    - SA Request Messages contain:
      - Requested Group Address
  - **SA Response (SA-Resp) Messages**
    - Sent in response to an SA Request message
    - SA Response Messages contain:
      - IP Address of Originator (usually an RP)
      - Number of (S, G)'s pairs being advertised
      - List of active (S, G)'s in the domain
  - **Keepalive messages**
    - Used to keep MSDP peer connection up

# Receiving SA Messages

- **Skip RPF Check and process SA if:**
  - Sending MSDP peer = only MSDP peer  
(i.e. the 'default-peer' or only 'msdp-peer' configured.)
  - Sending MSDP peer = Mesh-Group peer
- **RPF Check the received SA message.**
  - **Lookup best BGP path to RP in SA message**
    - Search MBGP RIB first then BGP RIB
      - If path to RP not found, RPF Check Fails; ignore SA message
  - **Sending MSDP Peer = BGP peer?**
    - **Yes: Is best path to RP via this BGP peer?**
      - If yes, RPF Check Succeeds; process SA message
    - **No: Is next AS in best path to RP = AS of MSDP peer?**
      - If yes, RPF Check Succeeds; process SA message

# Receiving SA Messages

- Detailed RPF Check rules
  - Case 1: Sending MSDP Peer = iBGP peer
    - Is best path to RP via this BGP peer?
      - Translation: Is the address of the iBGP peer that advertised the best path to the RP = address of the sending MSDP peer?
  - Case 2: Sending MSDP Peer = eBGP peer
    - Is best path to RP via this BGP peer?
      - Translation: Is the AS of eBGP peer = next AS in the best path to the RP?
  - Case 3: Sending MSDP Peer != BGP peer
    - Is the next AS in best path to RP = AS of the sending MSDP peer?
      - Translation: None needed.

2217  
1198\_05\_2000\_c2

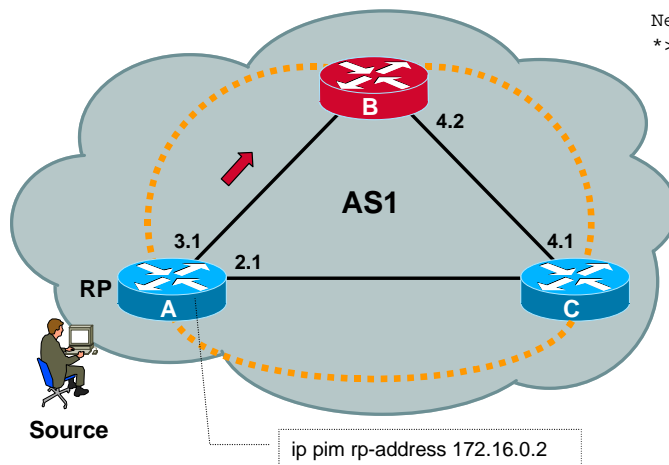
© 2000, Cisco Systems, Inc.

cisco.com

79

# RPF Check Example

RPF rule when MSDP == internal (m)BGP peer



BGP Table router B

```
Network      Next Hop      Path
*> 172.16.0.2/32 172.16.3.1    i
```

Who is the iBGP peer advertising this route, in our example 172.16.3.1

MSDP Peers router B

MSDP Peer	AS	State
172.16.3.1	1	UP
172.16.4.1	1	UP

Is the MSDP == BGP peer

**RPF Success!**

➔ MSDP SA

⋯ MSDP/iMBGP mesh-peering

2217  
1198\_05\_2000\_c2

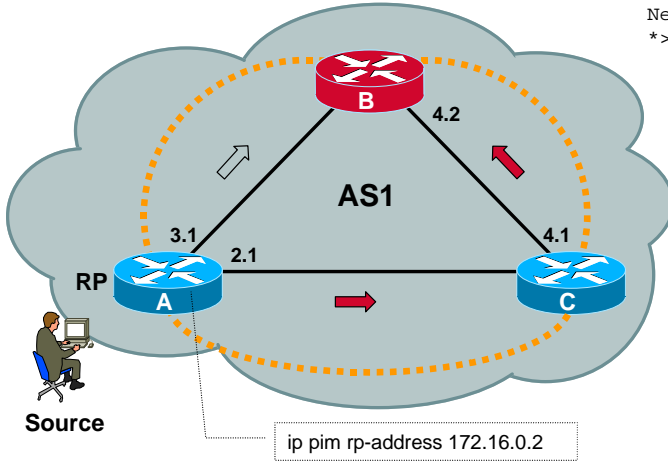
© 2000, Cisco Systems, Inc.

cisco.com

80

# RPF Check Example

RPF rule when MSDP == internal (m)BGP peer



BGP Table router B

Network	Next Hop	Path
*> 172.16.0.2/32	172.16.3.1	i

Who is the iBGP peer advertising this route, In our example 172.16.3.1

MSDP Peers router B

MSDP Peer	AS	State
172.16.3.1	1	UP
172.16.4.1	1	UP

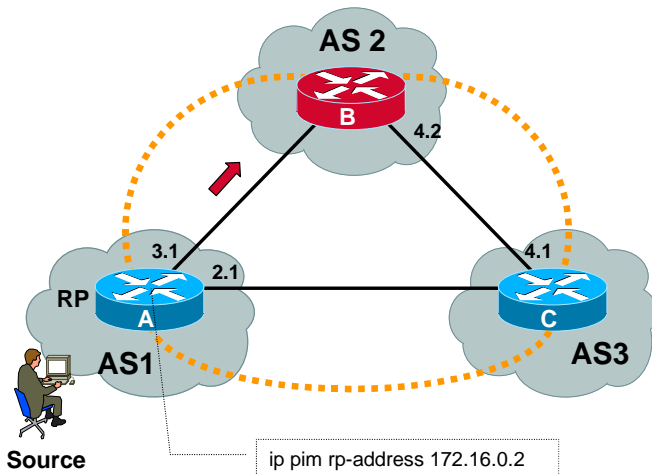
Is the MSDP == BGP peer

## RPF Failure!

- ➔ MSDP SA
- - - MSDP/iMBGP mesh-peering

# RPF Check Example

RPF rule when MSDP == external (m)BGP peer



MSDP Peers router B

MSDP Peer	AS	State
172.16.3.1	1	UP
172.16.4.1	3	UP

BGP Neighbours router B

Neighbor	AS	Up/Down
172.16.3.1	1	06:06:08
172.16.4.1	3	06:09:06

BGP Table

Network	Next Hop	Path
*> 172.16.0.2/32	172.16.3.1	1 i
172.16.0.2/32	172.16.4.1	3 1 i

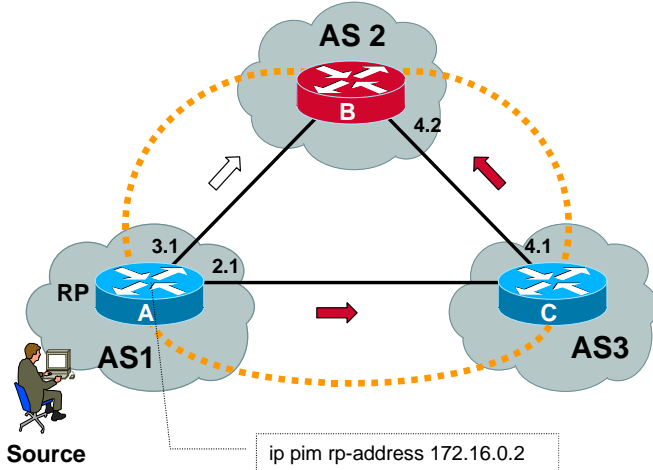
Is the MSDP AS == Last AS in RP route

## RPF Success!

- ➔ MSDP SA
- - - MSDP/eMBGP mesh-peering

# RPF Check Example

RPF rule when MSDP == external (m)BGP peer



Source

ip pim rp-address 172.16.0.2

→ MSDP SA

--- MSDP/eMBGP mesh-peering

Who is the BGP peer advertising this route

MSDP Peers router B

MSDP Peer	AS	State
172.16.3.1	1	UP
172.16.4.1	3	UP

BGP Neighbours router B

Neighbor	AS	Up/Down
172.16.3.1	1	06:06:08
172.16.4.1	3	06:09:06

BGP Table

Network	Next Hop	Path
*> 172.16.0.2/32	172.16.3.1	1 i
172.16.0.2/32	172.16.4.1	3 1 i

Is the MSDP AS == Last AS in RP route

**RPF Failure!**

2217  
1198\_05\_2000\_c2

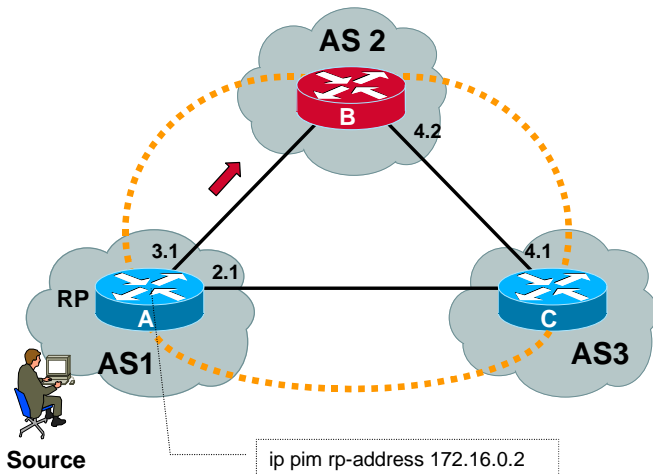
© 2000, Cisco Systems, Inc.

cisco.com

83

# RPF Check Example

RPF rule when MSDP != (m)BGP peer



Source

ip pim rp-address 172.16.0.2

→ MSDP SA

--- MSDP/eMBGP mesh-peering

MSDP Peers router B

MSDP Peer	AS	State
172.16.3.1	1	UP
172.16.4.1	3	UP

BGP Table router B

Network	Next Hop	Path
*> 172.16.0.2/32	172.16.3.1	1 i
172.16.0.2/32	172.16.4.1	3 1 I
*> 172.16.4.0/24	172.16.4.1	3 i
*> 172.16.3.0/24	172.16.3.1	1 i

Is the MSDP AS == Last AS in RP route

**RPF Success!**

2217  
1198\_05\_2000\_c2

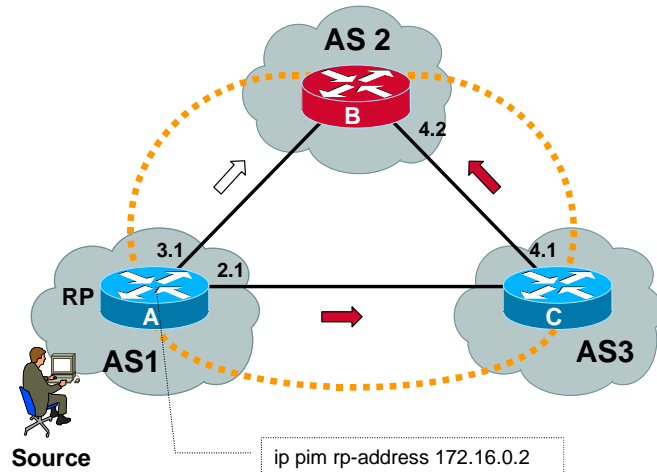
© 2000, Cisco Systems, Inc.

cisco.com

84

# RPF Check Example

RPF rule when MSDP != (m)BGP peer



→ MSDP SA  
 - - - - - MSDP/eMBGP mesh-peering

MSDP Peers router B

MSDP Peer	AS	State
172.16.3.1	1	UP
172.16.4.1	3	UP

BGP Table router B

Network	Next Hop	Path
*> 172.16.0.2/32	172.16.3.1	1 i
172.16.0.2/32	172.16.4.1	3 1 I
*> 172.16.4.0/24	172.16.4.1	3 i
*> 172.16.3.0/24	172.16.3.1	1 i

Is the MSDP AS == Last AS in RP route

**RPF Failure!**

2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

85

# Cisco MSDP Implementation

- Cisco implementation current with ID:
  - draft-ietf-msdp-spec-02.txt
- Multiple peer support
  - Peer with BGP, MBGP, or static peers
- SA caching (off by default)
- Sending and receiving SA-requests
- Sending and receiving SA-responses

2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

86

# Cisco MSDP Implementation

- SA input and output filtering
- SA-request input filtering
- Default peer support
  - So a tail site can MSDP with a backbone provider without requiring the two to BGP peer
- Triggered join support when creating an (S,G) learned by MSDP
- Mesh groups
  - Reduces RPF-flooding of SA messages between fully meshed MSDP peers

# MSDP Configuration

- **Configure peers**  
`ip msdp peer <ip-address> [connect-source <i/f>]`
- **Configure default peer**  
`ip msdp default-peer <ip-address> [prefix-list acl]`
- **SA caching**  
`ip msdp cache-sa-state [list <acl>]`
- **Mesh groups**  
`ip msdp mesh-group <name> <ip-address>`

# MSDP Configuration

- **Filtering**
  - Can filter SA in/out, groups, with acls or route-maps
- **TTL Scoping**

```
ip msdp ttl-threshold <ip-address> <ttl>
```
- **For more configuration commands see:**  
<ftp://ftpeng.cisco.com/ipmulticast/msdp-commands>

# ISP Requirements to Deploy Now

- **Want an explicit join protocol for efficiency**
  - ✓ PIM-SM (forwarding)
- **Use existing (unicast) operation model**
  - ✓ MBGP (routing)
- **Need interdomain source discovery**
  - ✓ MSDP (source discovery)

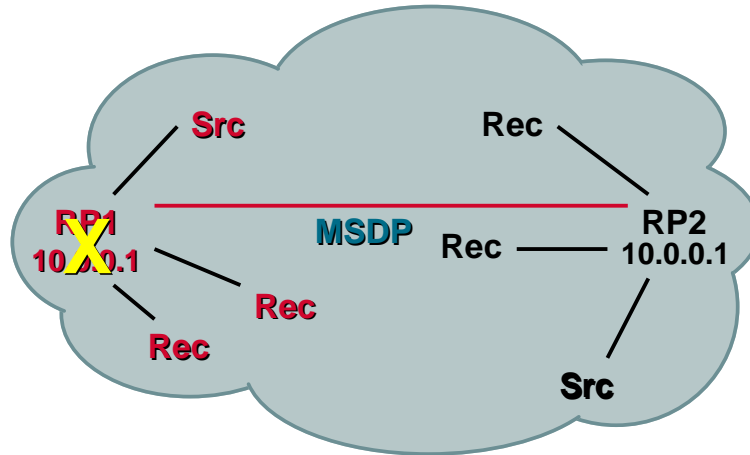
## MSDP Application—Anycast RP

- **draft-ietf-mboned-anycast-rp-05.txt**
- **Within a domain, deploy more than one RP for the same group range**
- **Give each RP the same IP address assignment**
- **Sources and receivers use closest RP**
- **May be used intra-domain (enterprise) to provide redundancy and RP load sharing**

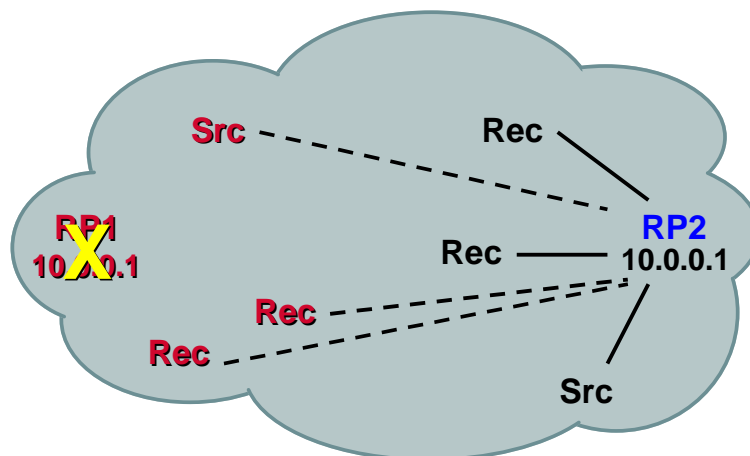
## MSDP Application—Anycast RP

- **Sources from one RP are known to other RPs using MSDP**
- **When an RP goes down, sources and receivers are taken to new RP via unicast routing**
  - **Fast convergence**

# Anycast RP—Convergence



# Anycast RP—Convergence



# Midterm Exam

## Which protocol...

- ...is used to propagate forwarding state?  
**PIM-SM - Protocol Independent Multicast-Sparse Mode**
- ...maintains routing information for interdomain RPF checking?  
**MBGP - Multiprotocol Border Gateway Protocol**
- ...exchanges active source information between RPs?  
**MSDP - Multicast Source Discovery Protocol**

## Putting it all together..

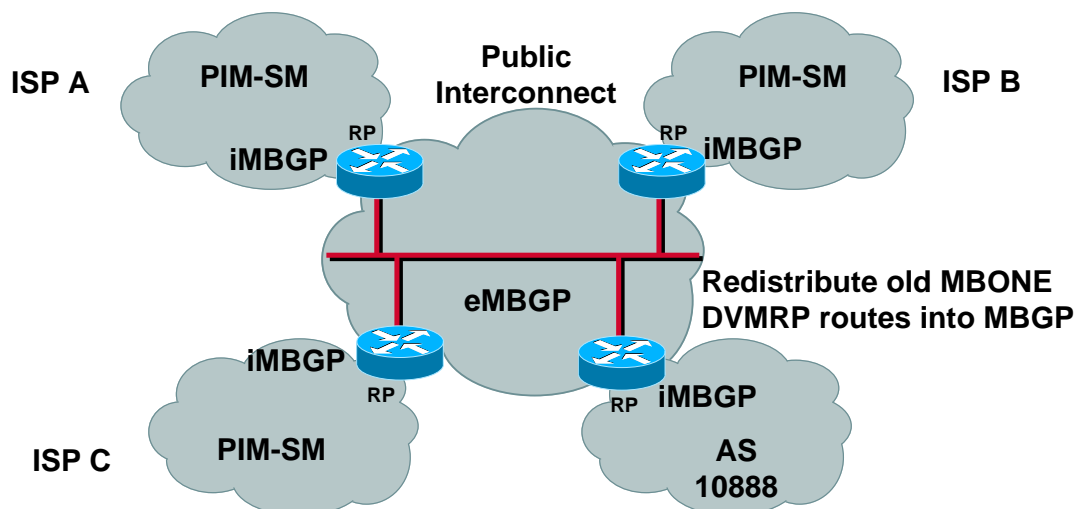
- **PIM-SM, MBGP, MSDP**
- **Public multicast peering (MIX)**
- **Transit topology examples**
- **Address allocation**
  - GLOP draft-ietf-mboned-static-allocation-00.txt

# ISP Requirements at the MIX

- **Current solution: MBGP + PIM-SM + MSDP**
  - **Environment**
    - ISPs run iMBGP and PIM-SM (internally)
    - ISPs multicast peer at a public interconnect
  - **Deployment**
    - Border routers run eMBGP
    - The interfaces on interconnect run PIM-SM
    - RPs' MSDP peering is fully meshed
    - All peers set a common distance for eMBGP

# ISP Requirements at the MIX

## Peering Solution: MBGP + PIM-SM + MSDP



# Multicast Transit Design Objectives

- **PIM Border Constraints**
  - Confine registers within domain
  - Confine local groups
  - Confine RP announcements
  - Control SA advertisements via MSDP
- **Border RPF check**
  - RPF check against unicast routes to multicast sources
- **MSDP RPF check**
  - RPF check toward RP in received SAs

## Recommended MSDP SA Filter

<ftp://ftp-eng.cisco.com/ipmulticast/msdp-sa-filter.txt>

```
! domain-local applications
access-list 111 deny ip any host 224.0.2.2 !
access-list 111 deny ip any host 224.0.1.3 ! Rwhod
access-list 111 deny ip any host 224.0.1.24 ! Microsoft-ds
access-list 111 deny ip any host 224.0.1.22 ! SVRLOC
access-list 111 deny ip any host 224.0.1.2 ! SGI-Dogfight
access-list 111 deny ip any host 224.0.1.35 ! SVRLOC-DA
access-list 111 deny ip any host 224.0.1.60 ! hp-device-disc
! auto-rp groups
access-list 111 deny ip any host 224.0.1.39
access-list 111 deny ip any host 224.0.1.40
! scoped groups
access-list 111 deny ip any 239.0.0.0 0.255.255.255
! loopback, private addresses (RFC 1918)
access-list 111 deny ip 10.0.0.0 0.255.255.255 any
access-list 111 deny ip 127.0.0.0 0.255.255.255 any
access-list 111 deny ip 172.16.0.0 0.15.255.255 any
access-list 111 deny ip 192.168.0.0 0.0.255.255 any
access-list 111 permit ip any any
```

# Multicast Boundary Filter

## Minimum Recommended

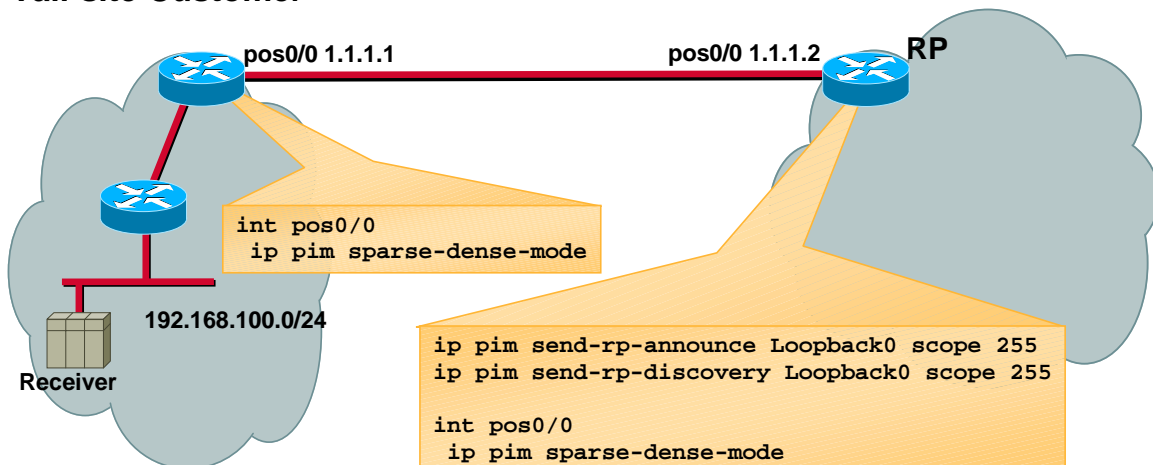
```
!deny auto-rp groups
access-list 1 deny 224.0.1.40
access-list 1 deny 224.0.1.39
!deny scoped groups
access-list 1 deny 239.0.0.0 0.255.255.255
!permit the rest of 224/4
access-list 1 permit 224.0.0.0 15.255.255.255
```

# Single-Homed, ISP RP, Non-MBGP

## PIM Border Constraints

### Tail-site Customer

### Transit AS109

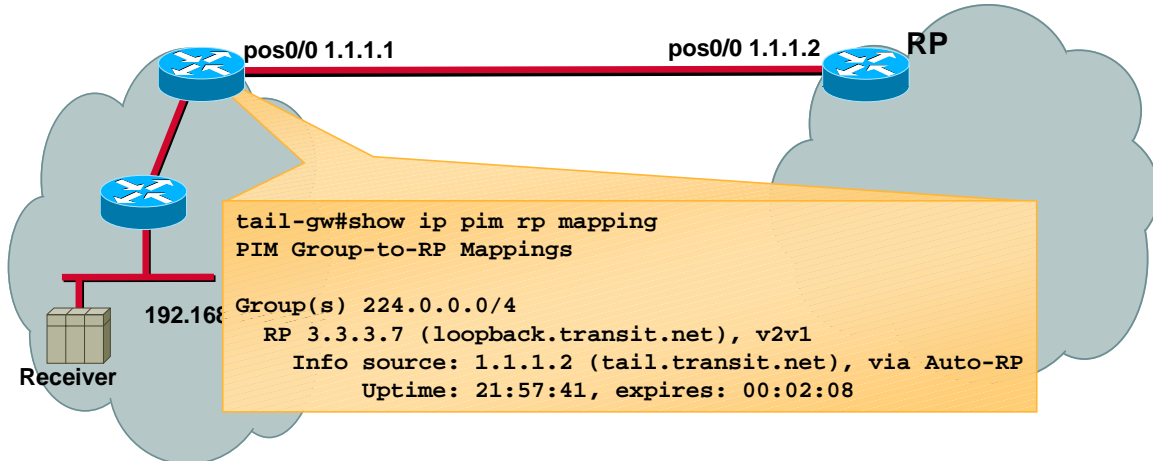


# Single-Homed, ISP RP, Non-MBGP

## Checking PIM Border (RP mapping)

Tail-site Customer

Transit AS109



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

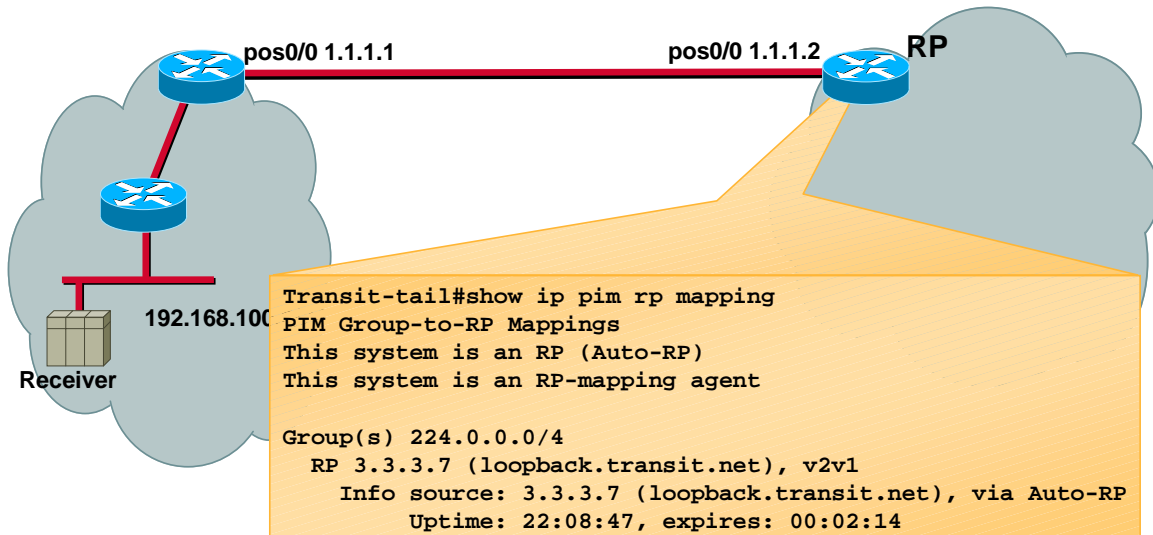
103

# Single-Homed, ISP RP, Non-MBGP

## Checking PIM Border (RP mapping)

Tail-site Customer

Transit AS109



2217  
1198\_05\_2000\_c2

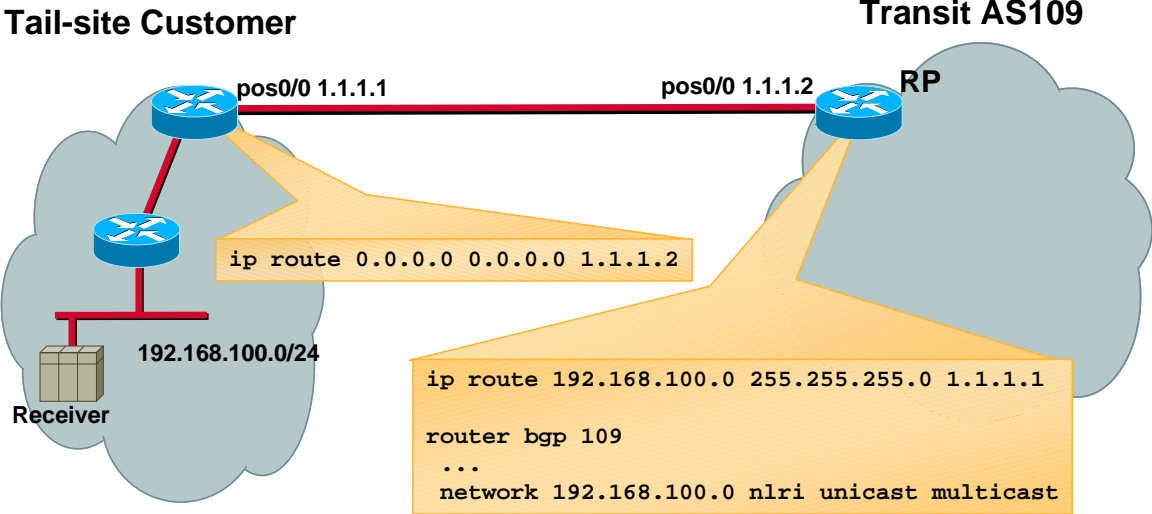
© 2000, Cisco Systems, Inc.

cisco.com

104

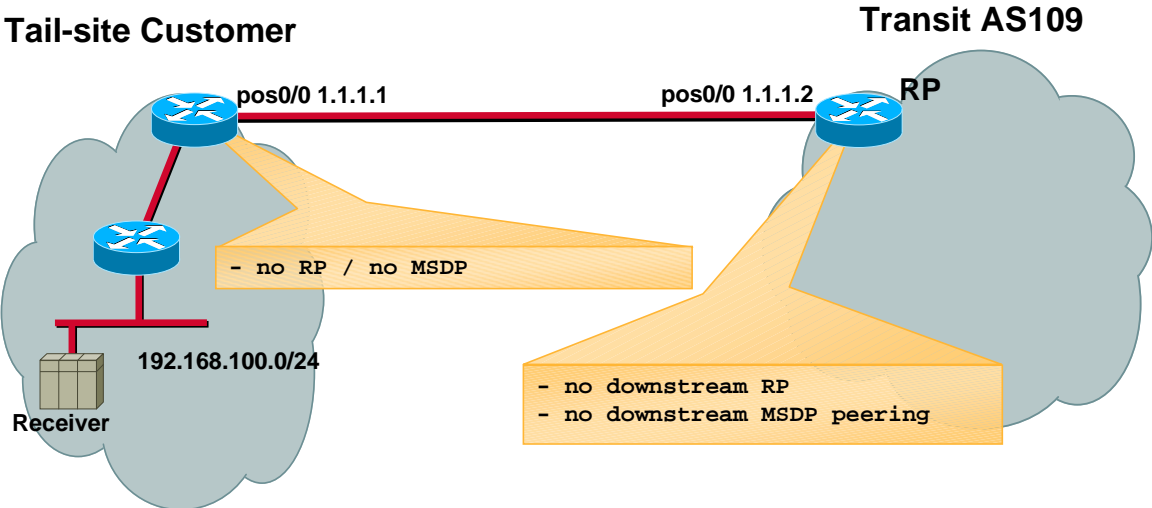
# Single-Homed, ISP RP, Non-MBGP

## Border RPF Check



# Single-Homed, ISP RP, Non-MBGP

## MSDP RPF Check

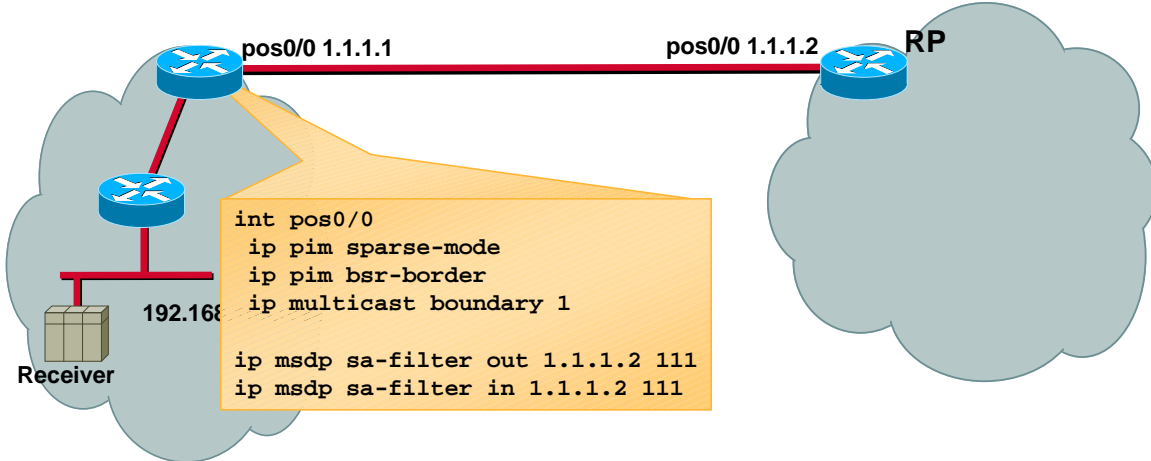


# Single-Homed, Customer RP, Non-MBGP

## PIM Border Constraints

Tail-site Customer

Transit AS109



Note: Access-list 111 = Recommended SA Filter

2217  
1198\_05\_2000\_c2 © 2000, Cisco Systems, Inc.

cisco.com

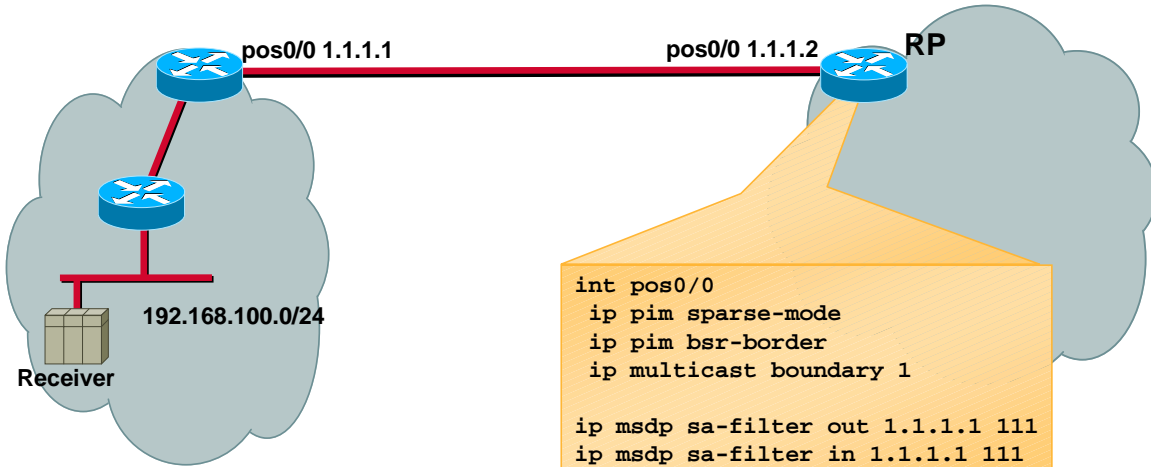
107

# Single-Homed, Customer RP, Non-MBGP

## PIM Border Constraints

Tail-site Customer

Transit AS109



Note: Access-list 111 = Recommended SA Filter

2217  
1198\_05\_2000\_c2 © 2000, Cisco Systems, Inc.

cisco.com

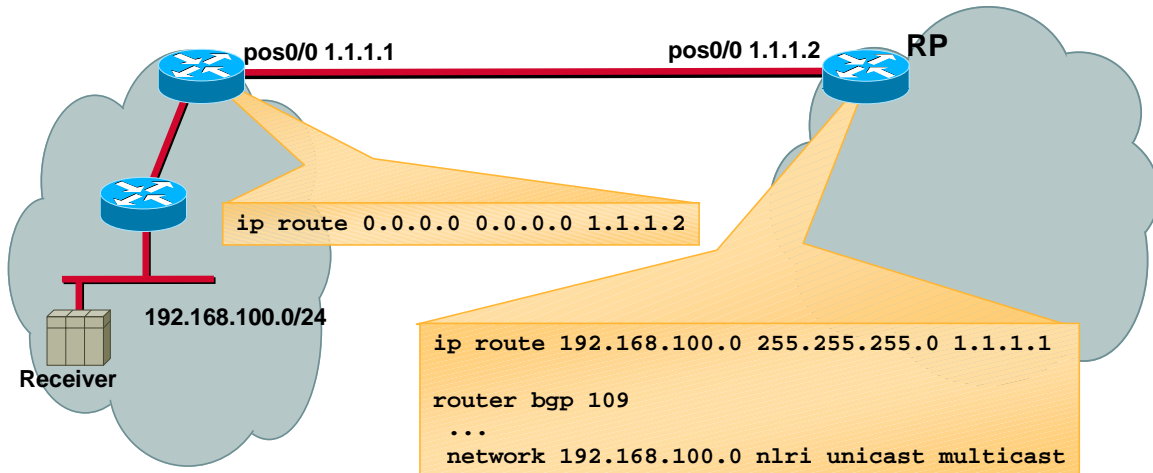
108

# Single-Homed, Customer RP, Non-MBGP

## Border RPF Check

Tail-site Customer

Transit AS109



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

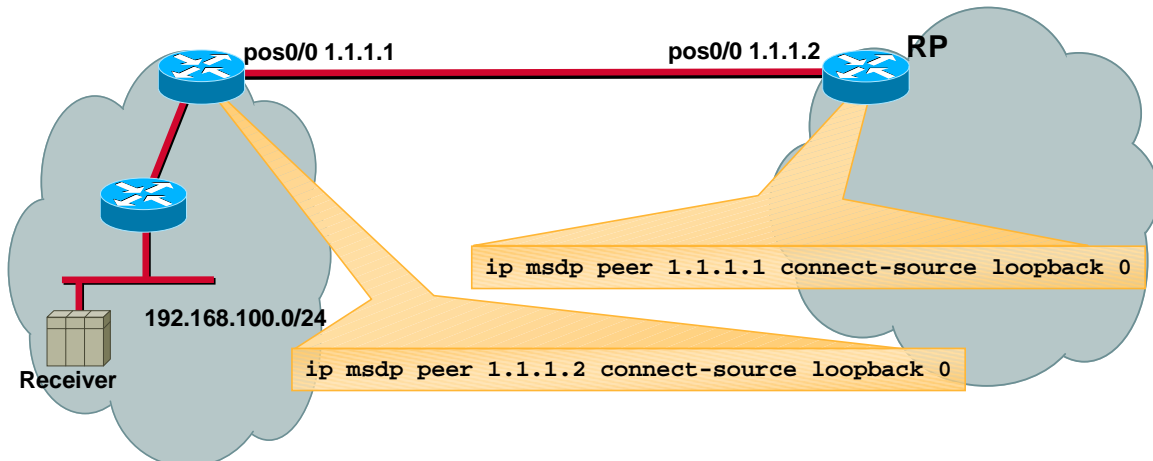
109

# Single-Homed, Customer RP, Non-MBGP

## MSDP RPF Check

Tail-site Customer

Transit AS109



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

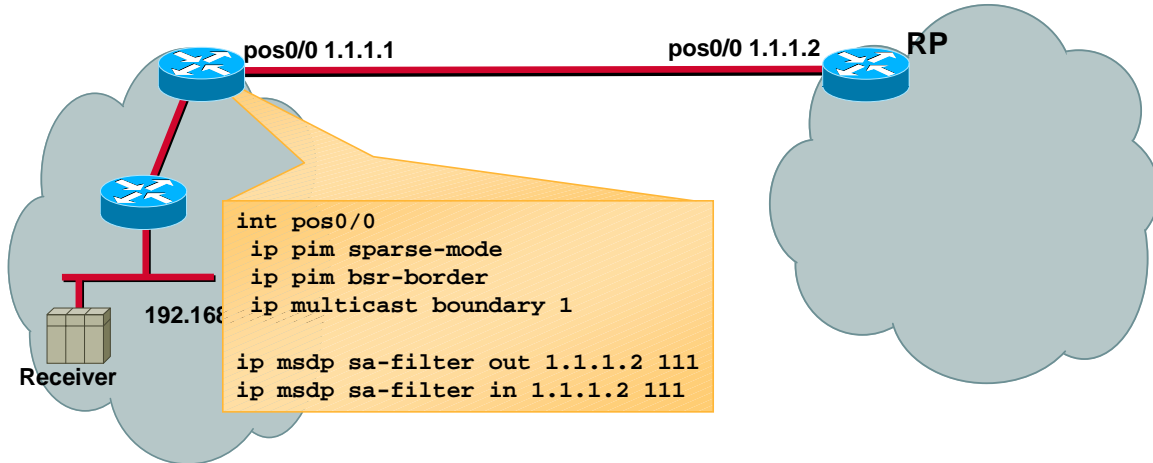
110

# Single-Homed, Customer RP, MBGP

## PIM Border Constraints

Tail-site Customer

Transit AS109



**Note: Access-list 111 = Recommended SA Filter**

2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

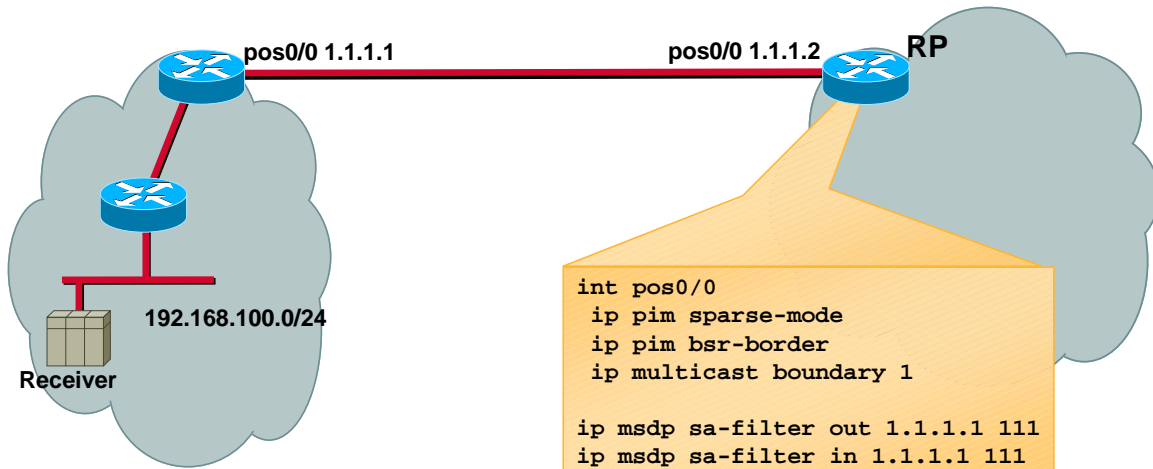
111

# Single-Homed, Customer RP, MBGP

## PIM Border Constraints

Tail-site Customer

Transit AS109



**Note: Access-list 111 = Recommended SA Filter**

2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

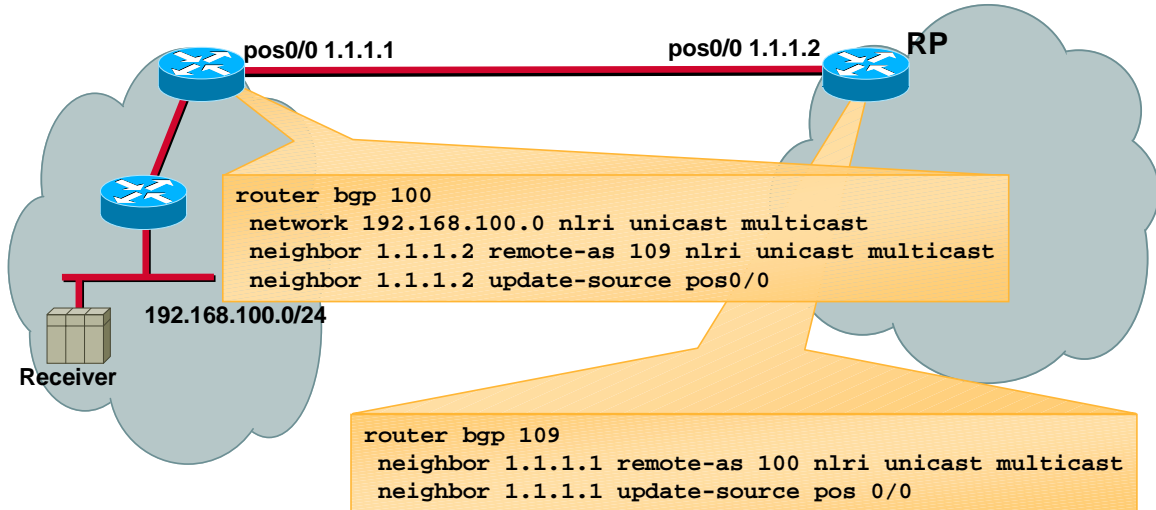
112

# Single-Homed, Customer RP, MBGP

## Border RPF Check

Tail-site Customer

Transit AS109



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

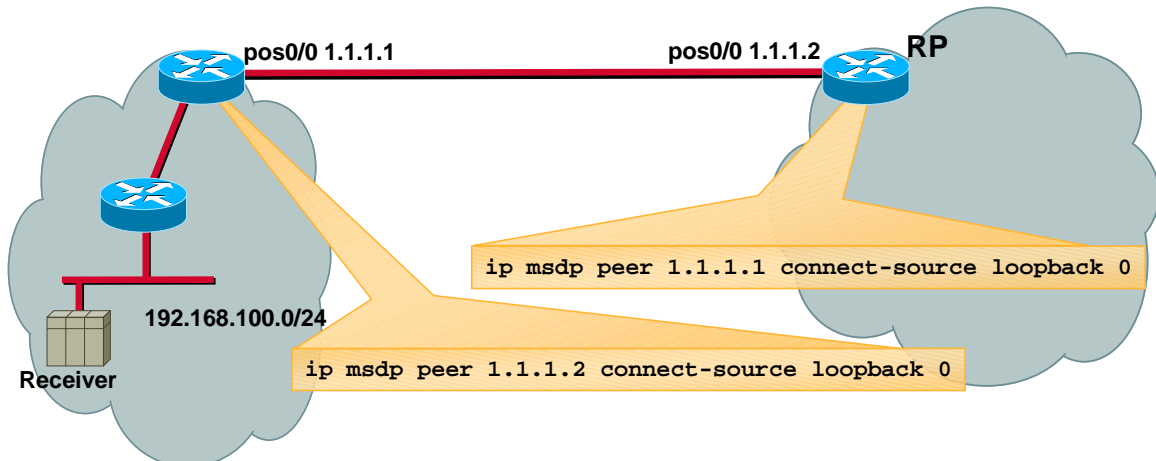
113

# Single-Homed, Customer RP, MBGP

## MSDP RPF Check

Tail-site Customer

Transit AS109



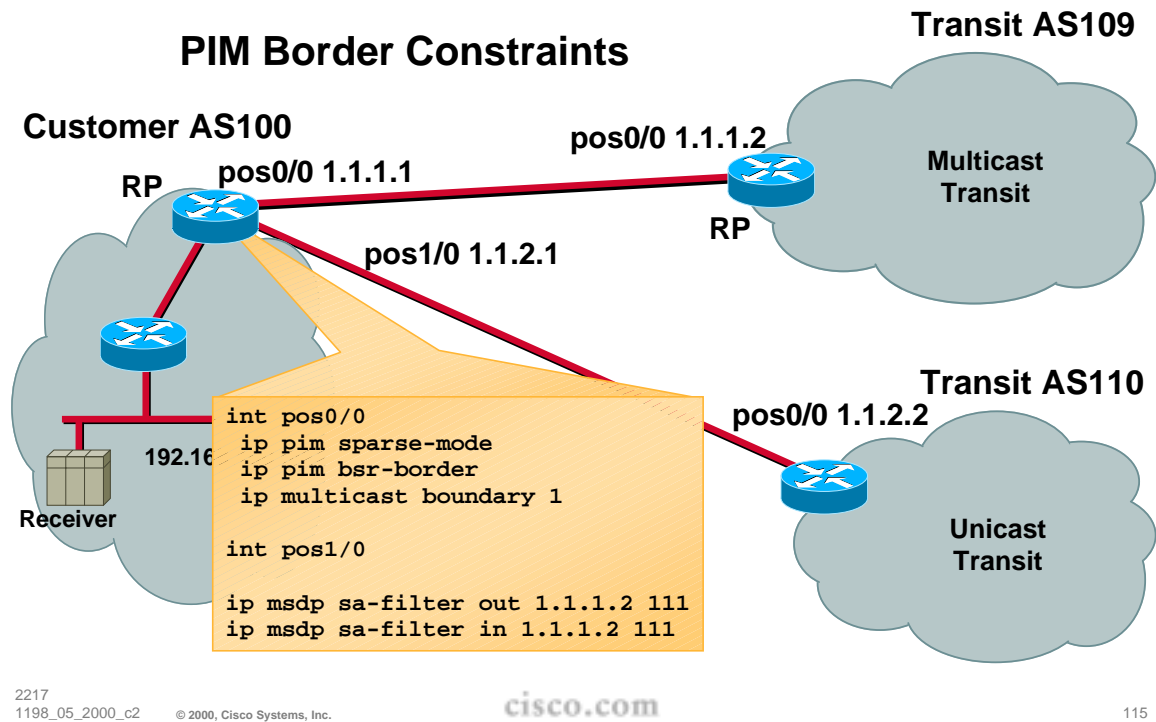
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

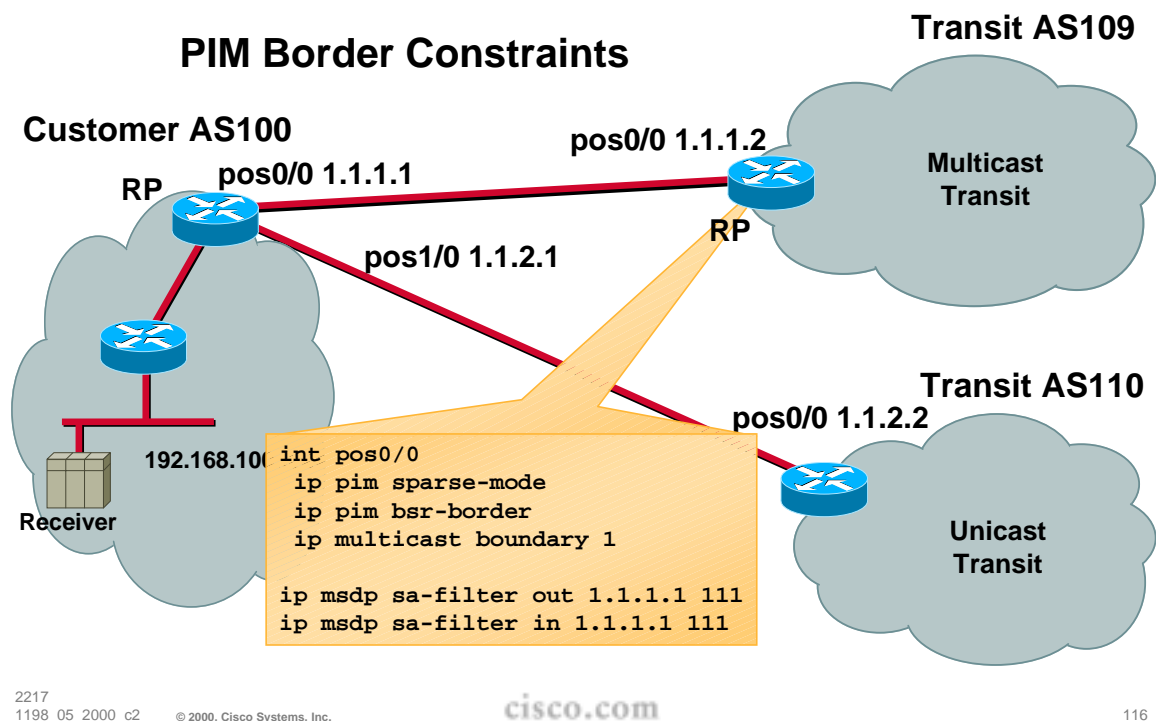
cisco.com

114

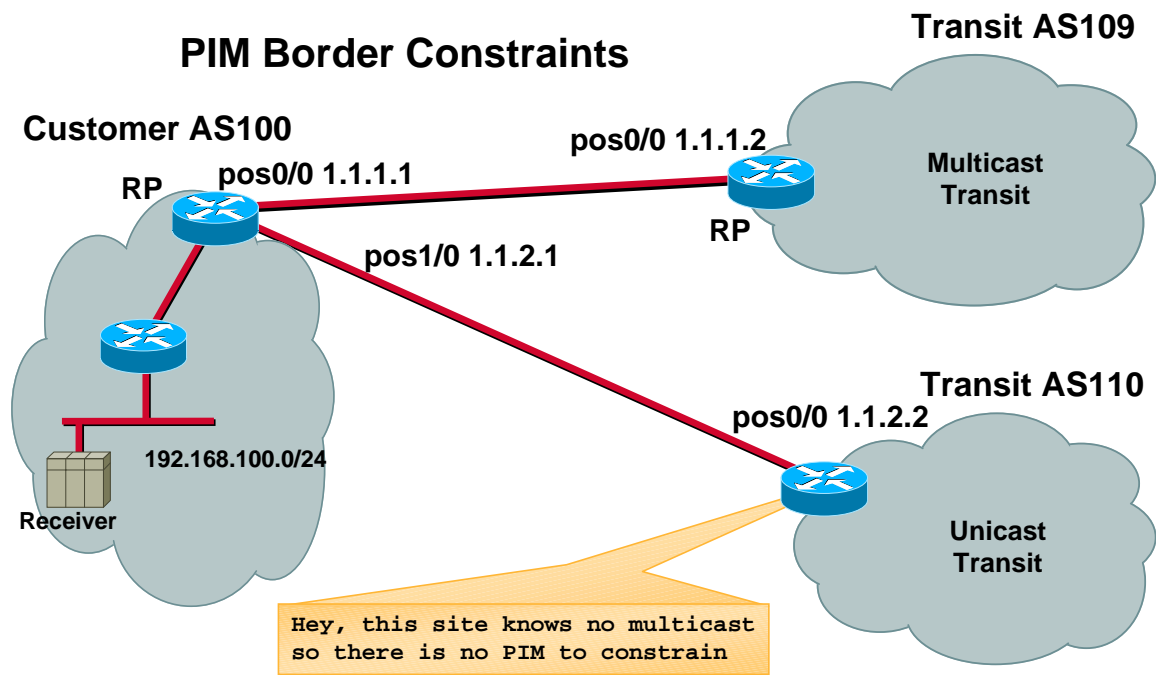
# Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



# Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



# Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



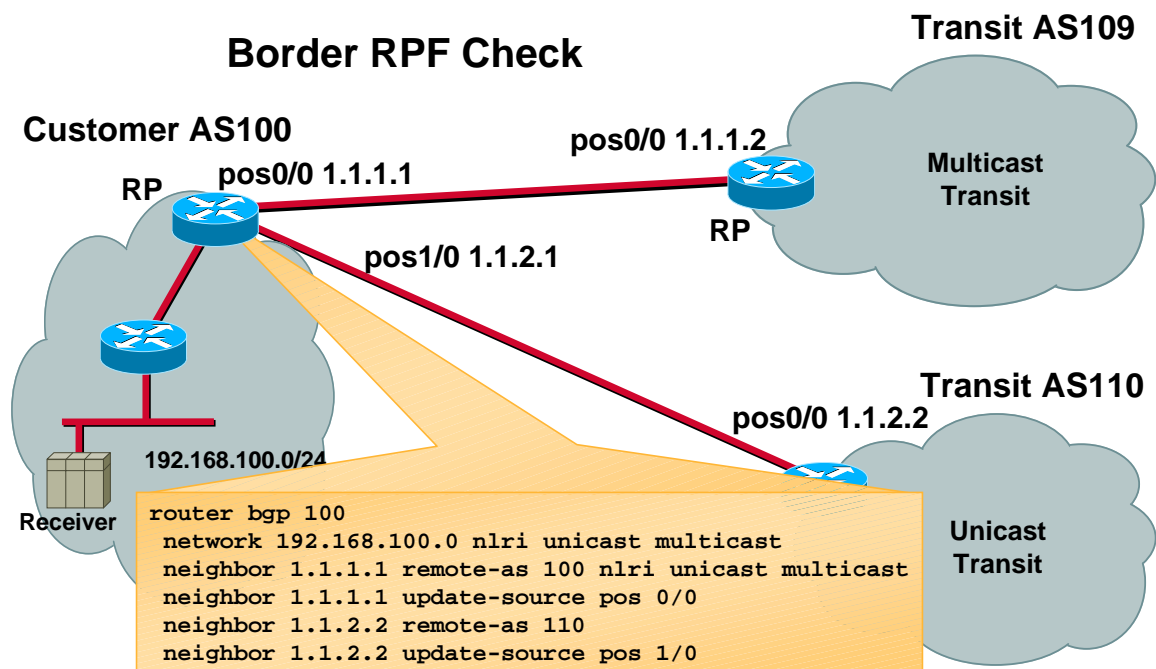
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

117

# Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



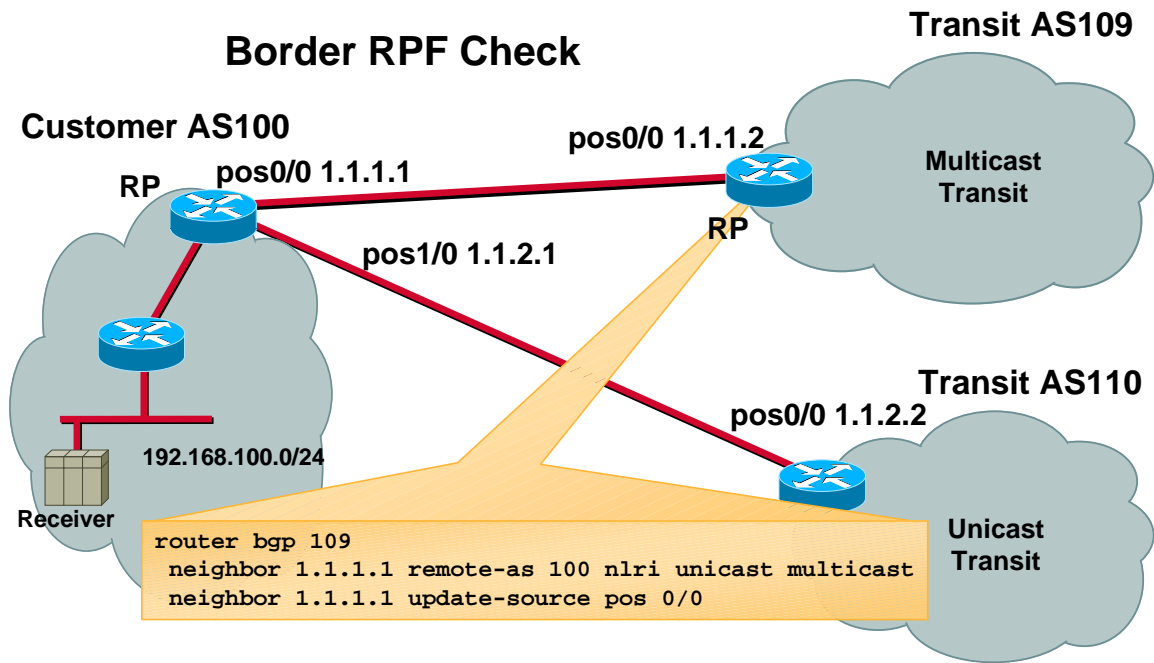
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

118

# Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



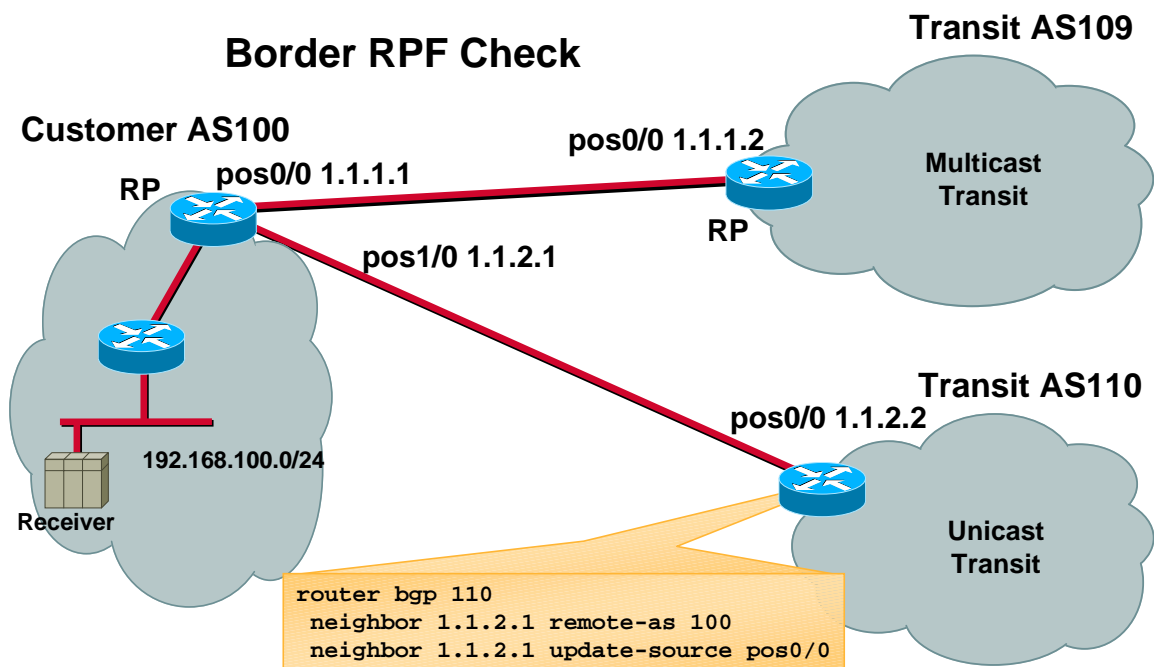
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

119

# Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



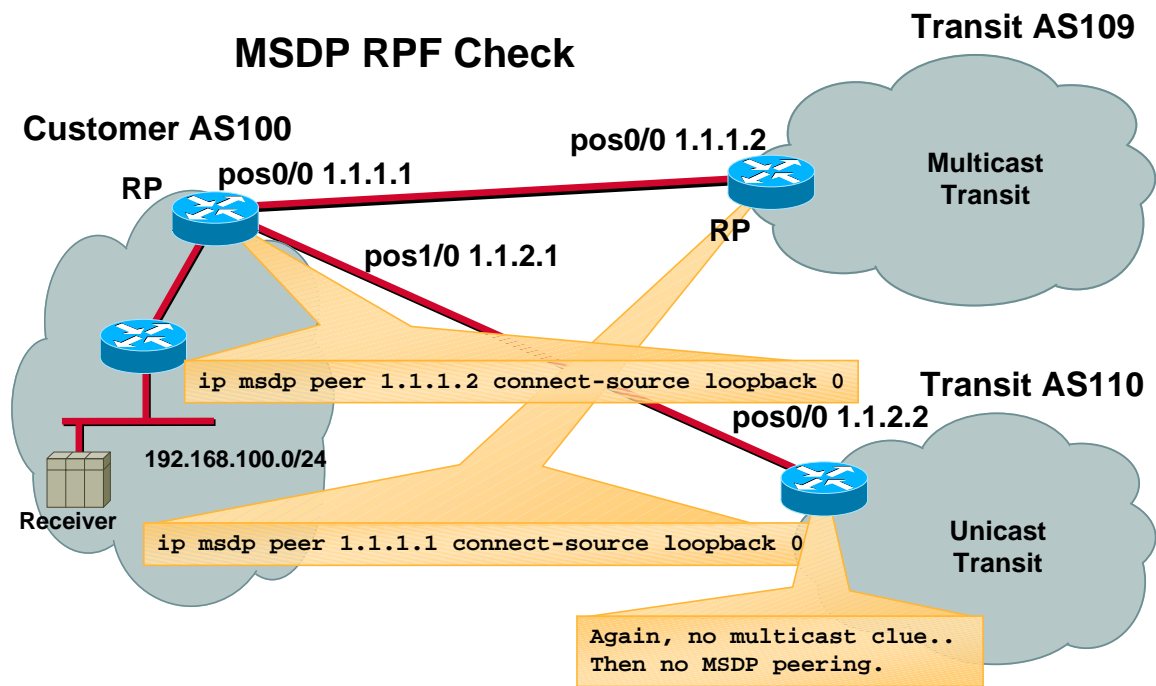
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

120

# Dual-Homed, Customer RP, MBGP Incongruent Multicast—Unicast



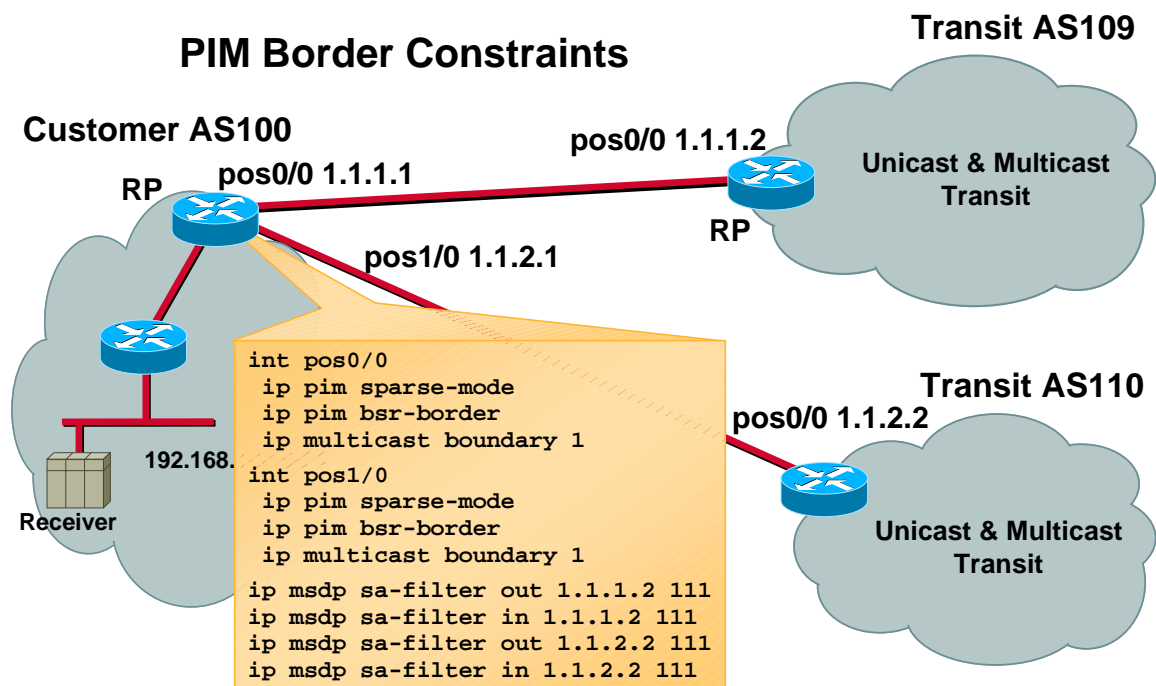
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

121

# Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast



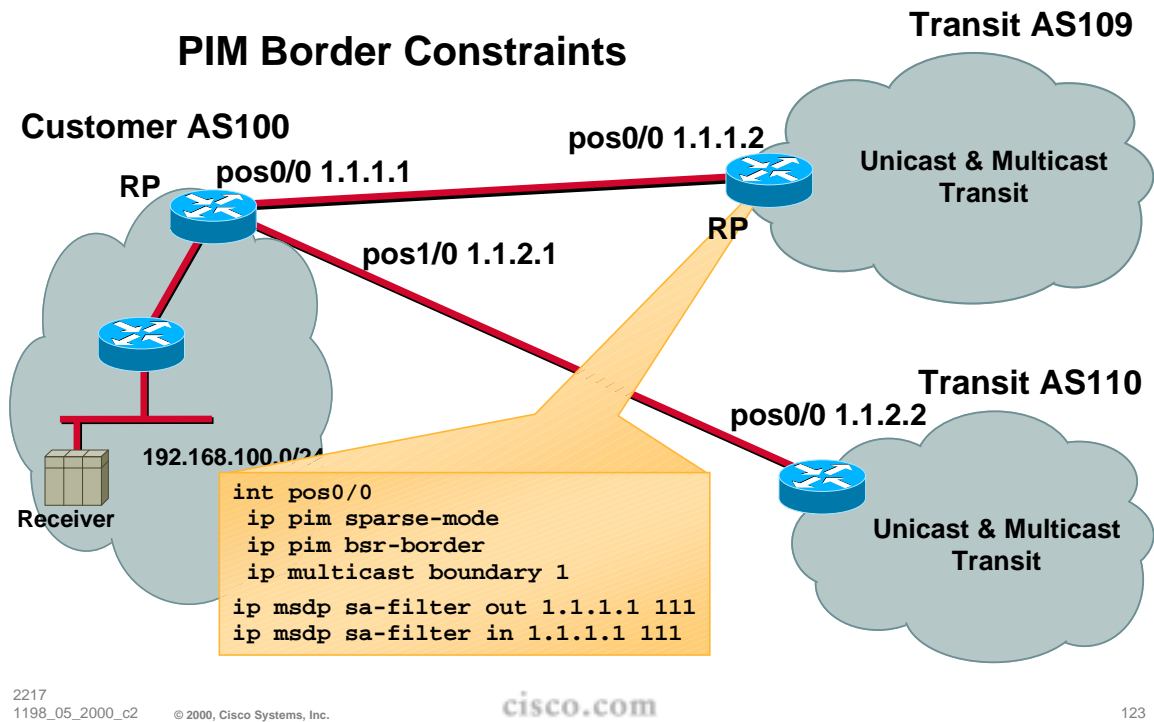
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

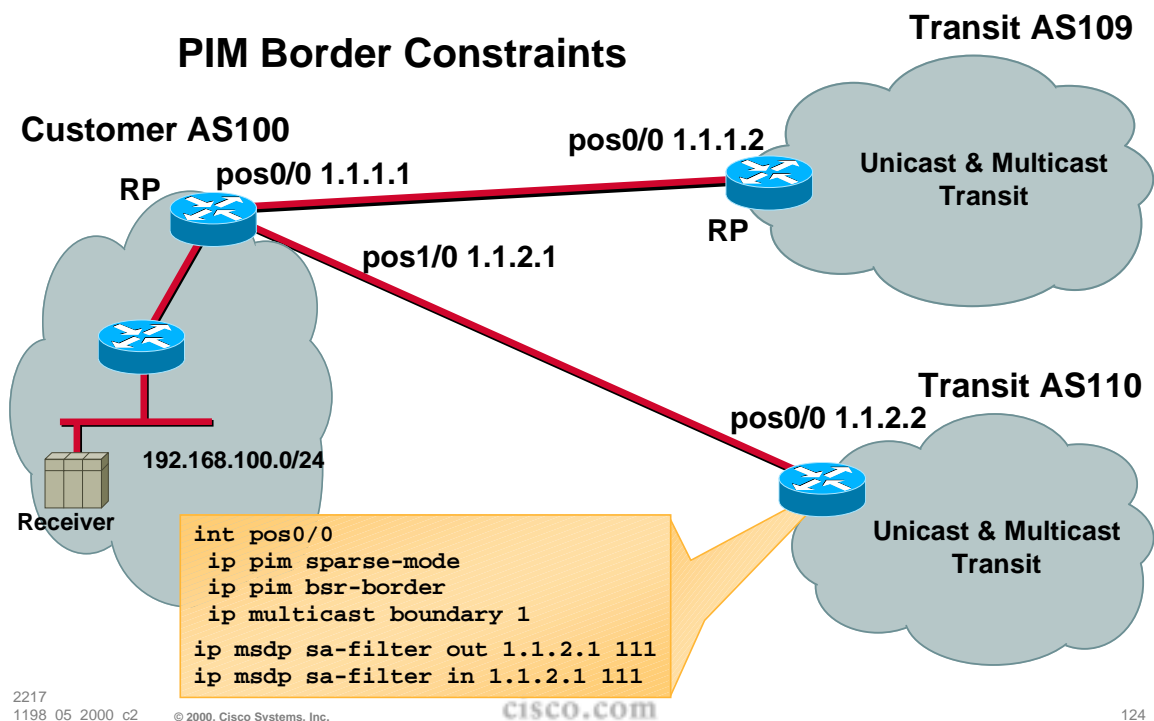
cisco.com

122

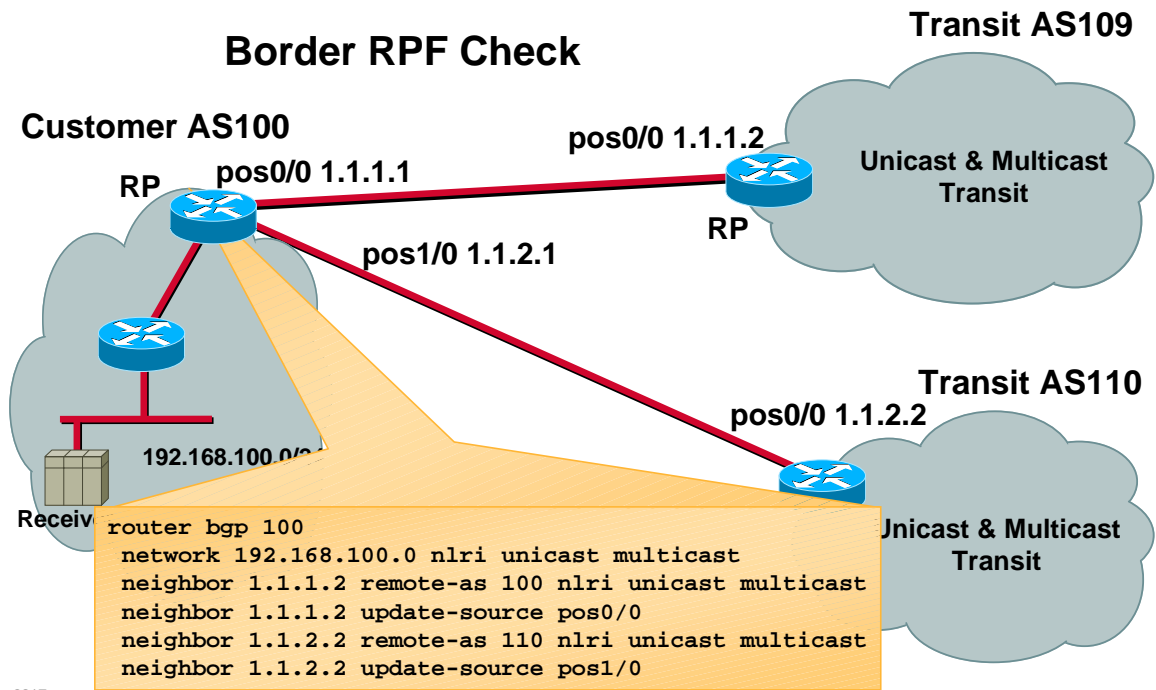
# Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast



# Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast



# Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast



2217

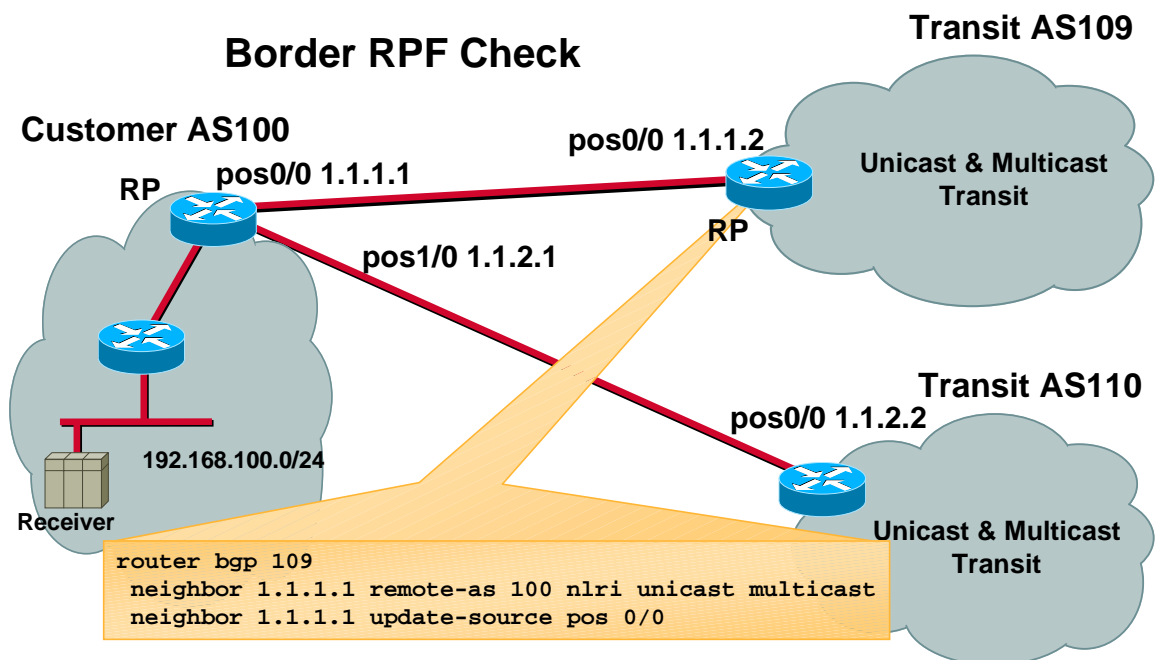
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

125

# Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast



2217

1198\_05\_2000\_c2

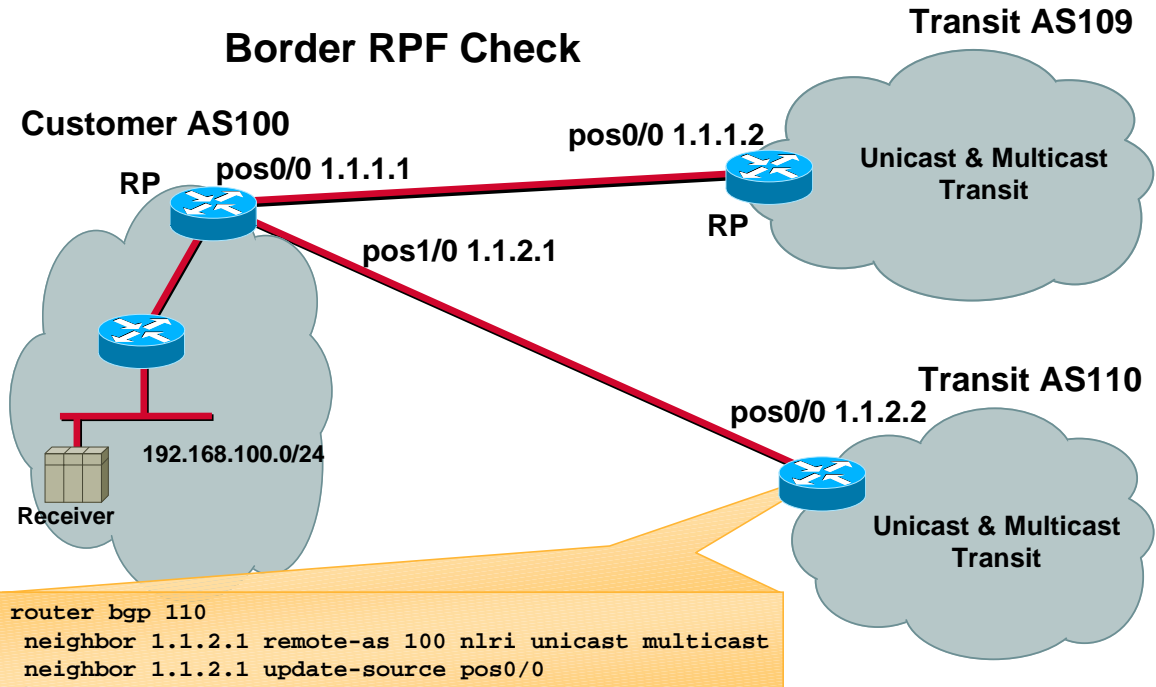
© 2000, Cisco Systems, Inc.

cisco.com

126

# Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

## Border RPF Check



2217  
1198\_05\_2000\_c2

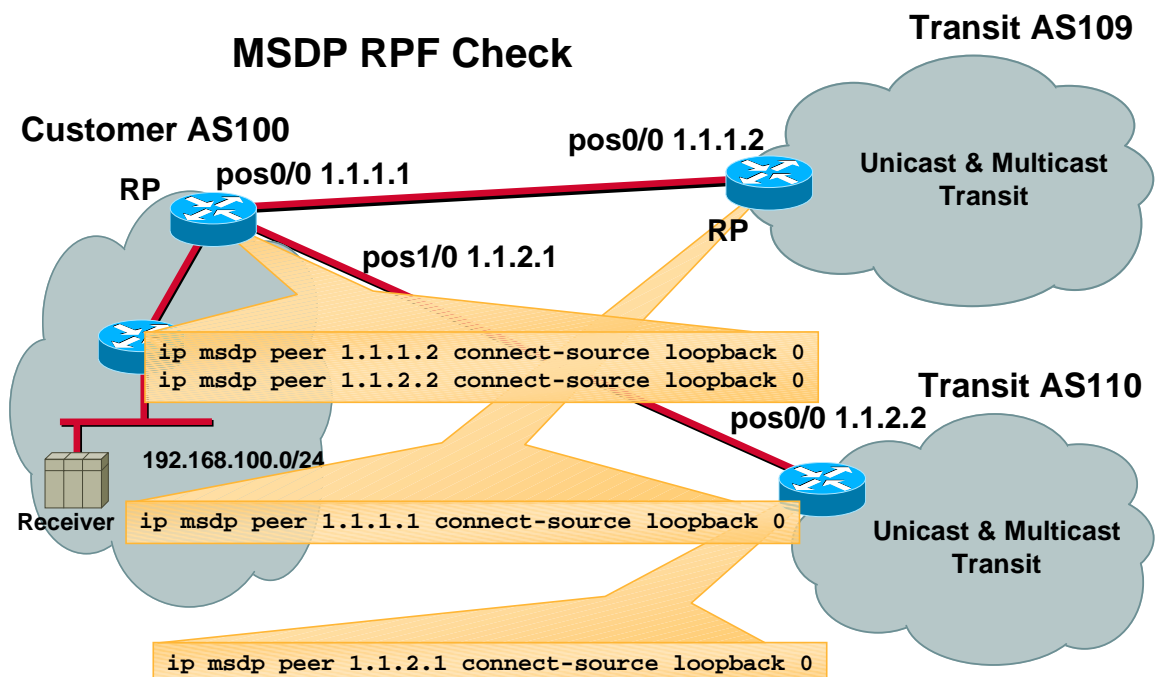
© 2000, Cisco Systems, Inc.

cisco.com

127

# Dual-Homed, Customer RP, MBGP Congruent Multicast—Unicast

## MSDP RPF Check



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

128

# GLOP—Static Allocation of 233/8

- **Temporary allocation of 233/8**
  - rfc2770
- **Statically assigned by mapping AS number into middle octets.**
  - <http://gigapop.uoregon.edu/glop/index.html>
- **Provides each AS with /24 addresses to use while waiting another solution**

## Agenda

- **PIM-SM review**
- **Inter-Domain Multicast**
  - MBGP
  - MSDP
  - Topology Examples
- **SSM (Source Specific Multicast)**

# Source Specific Multicast

- Simplify solution for well-known sources, particularly in cases where there is a single source sending to a given group.
  - Allow immediate use of shortest forwarding path to a specific source, without need to create shared tree.
  - Eliminate dependence on MSDP for finding sources.
  - Simplify address allocation for global, single source groups when combined with elimination of shared trees (232/8).

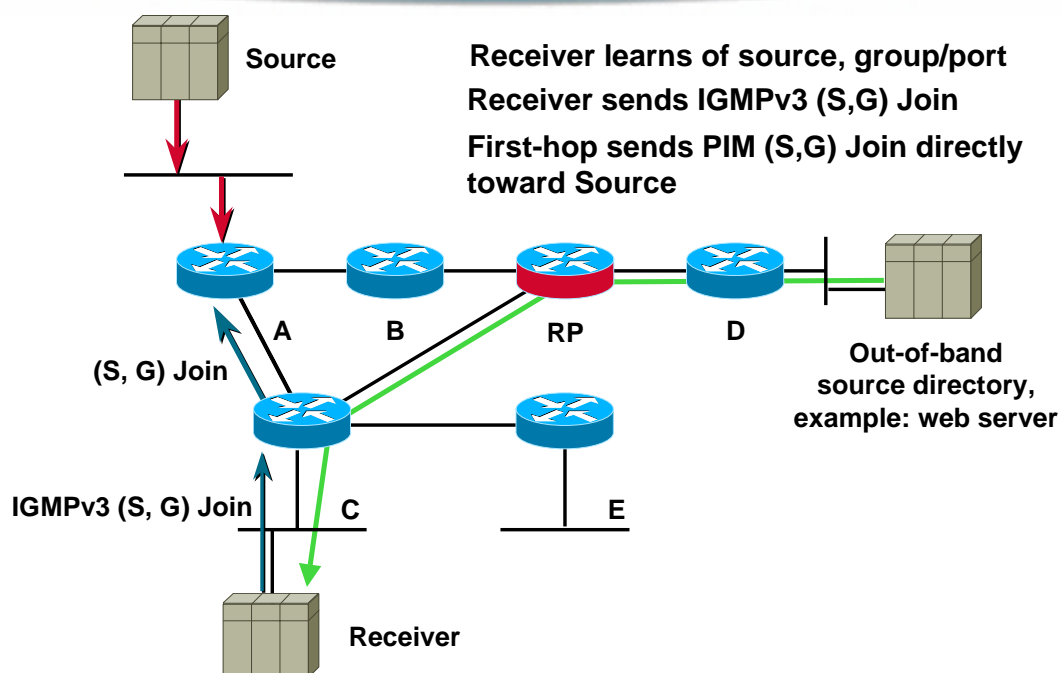
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

131

# PIM Source Specific Mode



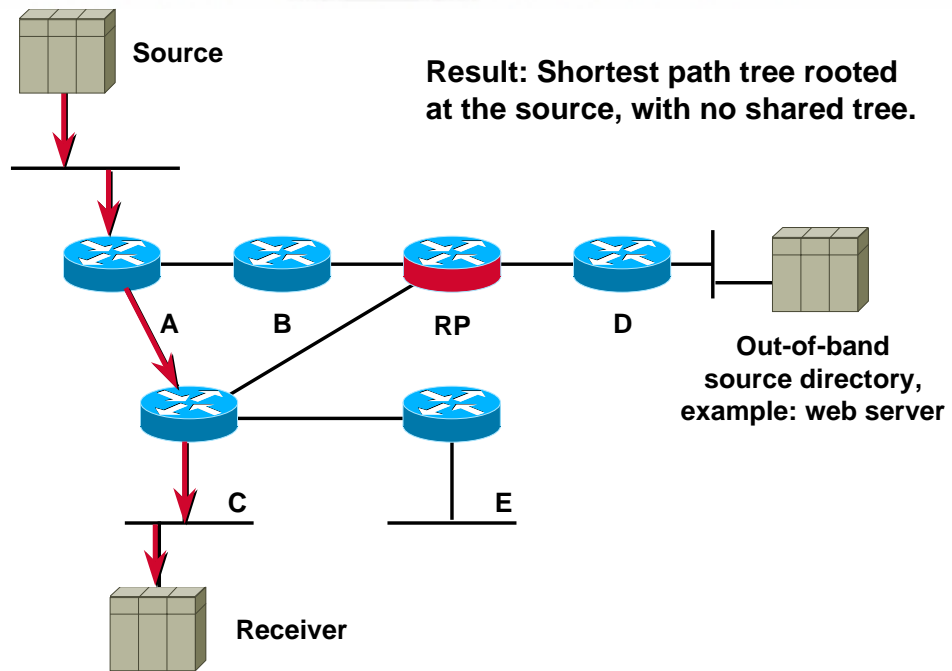
2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

132

# PIM Source Specific Mode



2217  
1198\_05\_2000\_c2

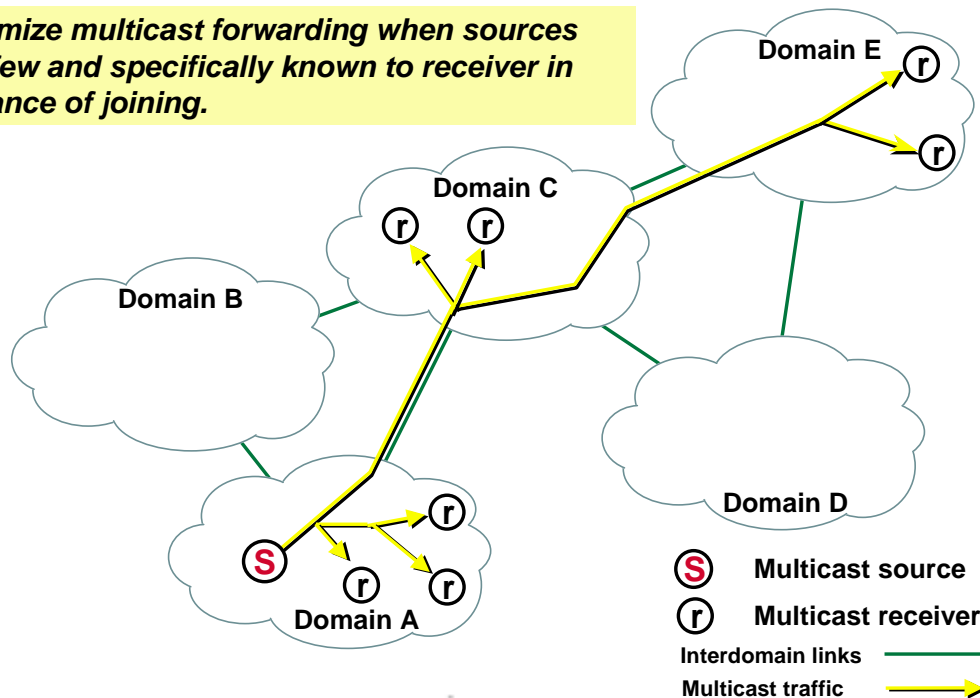
© 2000, Cisco Systems, Inc.

cisco.com

133

# Objective of SSM Multicast

*Optimize multicast forwarding when sources are few and specifically known to receiver in advance of joining.*



2217  
1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

cisco.com

134

## Where Is SSM?

- **Framework**  
draft-holbrook-ssm-00.txt
- **BCP proposal**  
draft-shepherd-232-00.txt
- **Supported in IOS 12**
- **Waiting for IGMPv3 on hosts**

## Documentation and Contact Info

- **EFT/Beta Site Web Page:**  
<ftp://ftpeng.cisco.com/ipmulticast.html>
- **TAC Support Mailing List:**  
[tac@cisco.com](mailto:tac@cisco.com)
- **Customer Support Mailing List:**  
[cs-ipmulticast@cisco.com](mailto:cs-ipmulticast@cisco.com)

# If All Else Fails—RTFB<sup>1</sup>



<sup>1</sup> Read this fine book

2217

1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

[cisco.com](http://cisco.com)

137

## Advanced IP Multicast Routing

### Session 2217

2217

1198\_05\_2000\_c2

© 2000, Cisco Systems, Inc.

[cisco.com](http://cisco.com)

138



**Please Complete Your  
Evaluation Form**

**Session 2217**

2217  
1198\_05\_2000\_c2 © 2000, Cisco Systems, Inc. [cisco.com](http://cisco.com) 139

**CISCO SYSTEMS**



**EMPOWERING THE  
INTERNET GENERATION<sup>SM</sup>**