IPv6 Multicast

Athanassios Liakopoulos (*aliako@grnet.gr*)

EFIPSANS IPv6 Training, Budapest, Hungary

Copy …Rights

IPv6 Deployment and Support

- *This slide set is the ownership of the 6DEPLOY project via its partners*
- *The Powerpoint version of this material may be reused and modified only with written authorization*
- *Using part of this material must mention 6DEPLOY courtesy*
- *PDF files are available from www.6deploy.org*
- *Looking for a contact ?*
  - *Mail to:* ***martin.potts@martel-consulting.ch***
  - *Or:* ***bernard.tuy@renater.fr***

Introduction

IPv6 Deployment and Support

- Multicast is inherent to the IPv6 protocol
- No broadcasts
  - Multicast used instead
- But some parts need to be configured
  - for building the multicast trees
  - for topology information (routing)

Outline

IPv6 Deployment and Support

- **Multicast Addressing**
- Multicast on the LAN (MLD)
- Intra-domain Multicast (PIM)
- Inter-domain Multicast
- Configuration examples

Generic Multicast Group Addresses (RFC 3513)

1111 1111	Flags	Scope	Group ID
8 bits	4 bits	4 bits	112 bits

- IPv6 multicast addresses are in the range of **FF00::/8**
- **Flag** field:
  - 000**T** values
    - T = 0, for permanent addresses defined by IANA
    - T = 1, for transient addresses
- **Scope** field: Allows limiting the scope of the multicasting
  - 0 - Reserved
  - 1 – Node-local
  - **2 – Link-local**
  - 3 – Subnet-local
  - 4 - Admin-local
  - 5 - Site-local
  - 8 - Organization-local
  - **E - Global (Internet)**

Generic Multicast Group Addresses - Examples

- IANA assigned addresses

  http://www.iana.org/assignments/ipv6-multicast-addresses

  - Flags = 0000
  - Scope
  - Group ID = 101 → NTP servers (variable scope)
    - FF01:0:0:0:0:0:0:101 : All NTP servers on the sender's host
    - FF02:0:0:0:0:0:0:101 : All NTP servers on the sender's link
    - FF05:0:0:0:0:0:0:101 : All NTP servers on the sender's site
    - FF0E:0:0:0:0:0:0:101 : All NTP servers on the Internet

IPv6 Multicast Address Assignments

IPv6 Deployment and Support

- Addresses for a given (fixed) scope
  - FF02::1 : All nodes of the link
  - FF02::2 : All routers of the link
  - FF05::1:3 : All DHCP servers of the site
  - FF02::D : All PIM routers of the link
  - ...
- Addresses for all scopes (variable scope)
  - FF0X :: 101 : Network Time Protocol (NTP)
  - FF0X :: FB : mDNSv6
  - ...

Group IDs

IPv6 Deployment and Support

- Should generally be 32 bits, although it is possible to use 112 bits when not using any of the specific addressing schemes
- When mapping IPv6 multicast groups to link-layer destination addresses, the last 32 bits are used. So one should try to always have different group IDs for different groups.
- RFC 3307 specifies a scheme with 32-bit IDs as follows
  - 0x00000001 - 0x3fffffff Groups assigned by IANA
  - 0x40000000 - 0x7fffffff Group IDs assigned by IANA
  - 0x80000000 - 0xffffffff Server/host allocation
- Note that IANA may assign group IDs, not just groups. This is something like scope relative addresses for IPv4.
- Whenever using unicast prefix based addresses, embedded-RP addresses or SSM, you should pick group IDs according to this scheme

Multicast Unicast Prefix-Based Addresses (RFC 3306)

IPv6 Deployment and Support

1111 1111	Flags	Scope	Resrv	Plen	Prefix	Group ID
8 bits	4 bits	4 bits	8 bits	8 bits	64 bits	32 bits

- Addresses are in the range of **FF30::/12**
- Flags = 00PT
  - P=0, for address not based on a unicast prefix
  - P=1, for address based on the unicast prefix
  - If P=1, then T=1 because address is not allocated by IANA
- Resrv : reserved, always 0
- Plen : prefix length
- Prefix : a unicast prefix

Multicast Unicast Prefix-Based Addresses - Example

IPv6 Deployment and Support

- GRNET-HQ prefix

  2001:648:2320::/48

- IPv6 Unicast Prefix-Based Address

  FF3E:30:2001:648:2320::cafe:babe

1111 1111	Flags	Scope	Resrv	Plen	Prefix	Group ID

SSM Addresses (RFC 4607)

IPv6 Deployment and Support

11111111	3	Scope	Resrv	0	0	Group ID
8 bits	4 bits	4 bits	8 bits	8 bits	64 bits	32 bits

- SSM addresses are a subset of unicast prefix-based addresses
- SSM addresses are in the range of **FF3x::/96**
- Flags = 00PT
  - T=1, for for transient addresses
  - P=1, for address based on the unicast prefix
- Plen : prefix length always 0
- Prefix : always 0

SSM Addresses - Example

IPv6 Deployment and Support

- SSM Address

FF3E:::cafe:babe

11111111	3	Scope	Resrv	0	0	Group ID

Multicast Embedded RP Addresses (RFC 3956)

IPv6 Deployment and Support

1111 1111	Flags	Scope	Resvd	RPadr	Plen	Prefix	Group ID
8 bits	4 bits	4 bits	4 bits	4 bits	8 bits	64 bits	32 bits

- Addresses are in the range of FF70::/12
- Flags = 0RPT
  - R=1, for address that embeds RP address
  - R=0, for address that do not embed RP address
  - If R=1, then (P,T)=(1,1).
- PRaddr : last four bits of RP address
- Plen : prefix length
- Prefix : a unicast prefix

Multicast Embeded RP Address - Example

IPv6 Deployment and Support

- GRNET-HQ address prefix

  2001:648:2320::/48

- GRNET-HQ RP address

  2001:648:2320::1/128

- Embedded RP Address

  FF7E:130:2001:648:2320::cafe:babe

11111111	Flags	Scope	Resvd	RPadr	Plen	Prefix	Group ID
8 bits	4 bits	4 bits	4 bits	4 bits	8 bits	64 bits	32 bits

Solicited-Node Multicast Addresses (RFC 4291)

IPv6 Deployment and Support

- Nodes build their own solicited multicast address using their unicast / anycast addresses
- Concatenate FF02::1:FF00:0000/104 with the 24-low-order bits of a unicast / anycast address
- Solicited multicast address are used …
  - … by Neighbor Discovery Protocol
  - … for Duplicate Address Detection management

Solicited-Node Multicast Addresses Example

IPv6 Deployment and Support

Concatenation of ff02::1:ffxx:xxxx with the last 24 bits of the IPv6 address

***IPv6 address*:**
2001:648:1a:4002:4421:21FF:FE24:87c1

***Sol. Mcast address:* FF02::1:FF24:87c1**

***Ethernet address:* 33-33-FF-24-87-c1**

Outline

IPv6 Deployment and Support

- Multicast Addressing
- **Multicast on the LAN (MLD)**
- Intra-domain Multicast (PIM)
- Inter-domain Multicast
- Configuration examples

Multicast Listener Discovery

- MLD is used among multicast-enabled routers and hosts to signal which groups (and sources) a host is interested in.
- MLDv1 (RFC 2710) supports only ASM, similar to IGMPv2.
- MLDv2 (RFC 3810) also supports SSM, similar to IGMPv3.
- MLD messages are sent in ICMPv6 packets

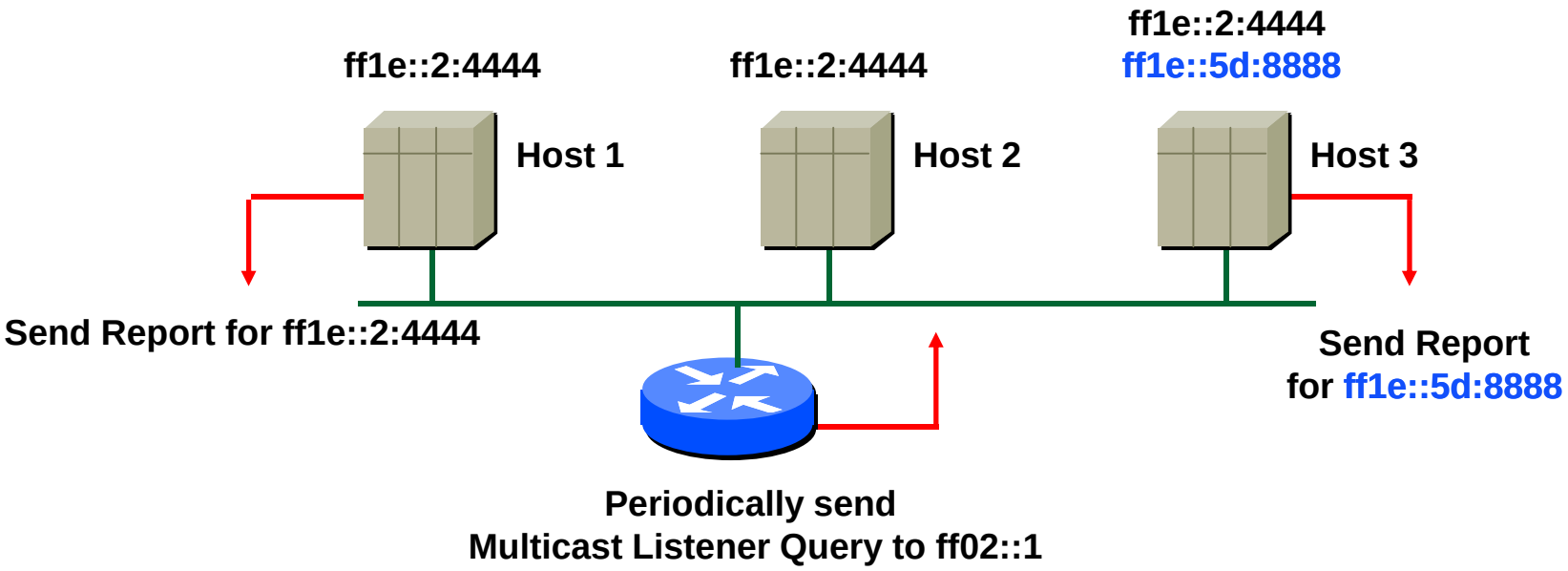
MLDv1 messages

IPv6 Deployment and Support

- Multicast listener Query (130)
  - Identify whether a group has listeners on a link
- Multicast Listener Report (131)
  - Response to a query
- Multicast Listener Done (132)
  - Indicate that a host stopped listening to a multicast address

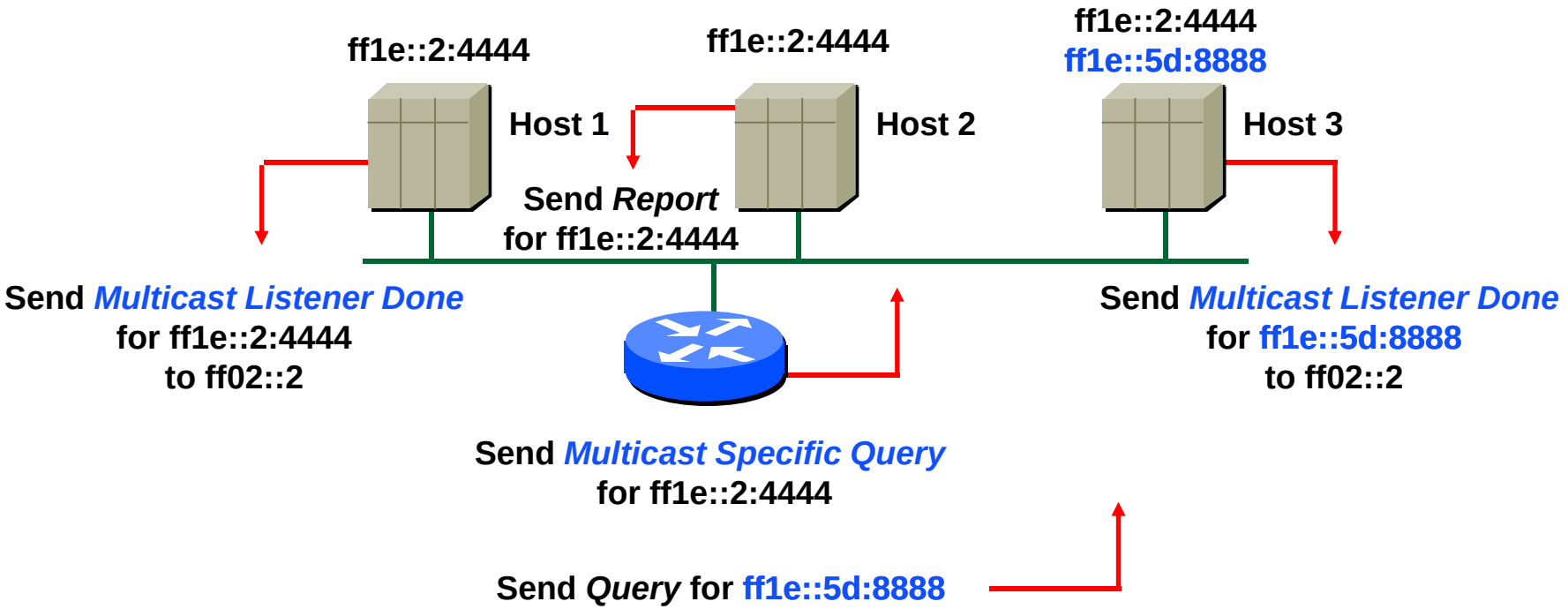
MLDv1 : Join a group

IPv6 Deployment and Support



MLDv1 : Leave a group

IPv6 Deployment and Support



MLDv2 (RFC 3810)

IPv6 Deployment and Support

- Management of groups & sources
  - ***INCLUDE*** : to receive packets from sources specified in the MLDv2 message
  - ***EXCLUDE*** : to receive packets from all sources except the ones specified in the MLDv2 message
- 2 types of messages
  - Multicast listener query messages
  - Multicast listener report messages
- Interoperable with MLDv1

Multicast and ethernet switches

- To avoid flooding multicast (like broadcast), switches may employ MLD (as well as IGMP and PIM) snooping
- Hosts should use MLD for link-local groups as well; considerable amount of state therefore kept on switch (see Cisco TCAM considerations)
  - ff02::1 unaffected
  - some switches may flood solicited-node or even all link-local traffic
- Problems arise when switches do not correctly understand MLD, when MLD not used or when suppression used wrong
  - common symptom: multicast forwarded for a while then stops
- IETF has snooping recommendations RFC but this is not a protocol, so implementation totally up to vendor
- Be very careful with snooping
  - especially for link-local traffic (it will eventually bite you)
  - disable when it does not make sense (backbone, metro)
- What about 802.11?
  - multicast = broadcast (at base rate of AP)
  - multicast->unicast vendor-specific tricks

Outline

IPv6 Deployment and Support

- Multicast Addressing
- Multicast on the LAN (MLD)
- **Intra-domain Multicast (PIM)**
- Inter-domain Multicast
- Current IPv6 Multicast Deployment
- Configuration examples

Protocol Independent Multicast (PIM)

IPv6 Deployment and Support

- PIM is *not* a routing protocol
  - Relies on RPF information from other routing protocols, such as MP-BGP
- PIM creates multicast trees between senders and receivers
- Only PIM Sparse Mode supported for IPv6
  - PIM-SM v2 (RFC 4601) should be used (SSM a subset of the spec)
  - see also bi-directional PIM (RFC 5015)
- Differences from IPv4:
  - link-local addresses used (HELLO option for listing all addresses)
  - tunnel interfaces used for PIM register

PIM Basics

IPv6 Deployment and Support

- **Shared Trees**
  - RP-based tree (RPT)
  - Source-based tree (SPT)
- **Rendez-Vous Points (RPs)**
  - Static configuration (and Anycast-RP)
  - Auto-RP and Bootstrap Router (BSR)
- **Designated Router (DR)**
  - **Join/Prune**: from receiver (hop-by-hop) towards
    - the RP (for *,G) – RPT
    - the source (S,G) – SPT
  - **Register**: from source unicast encapsulated to the RP

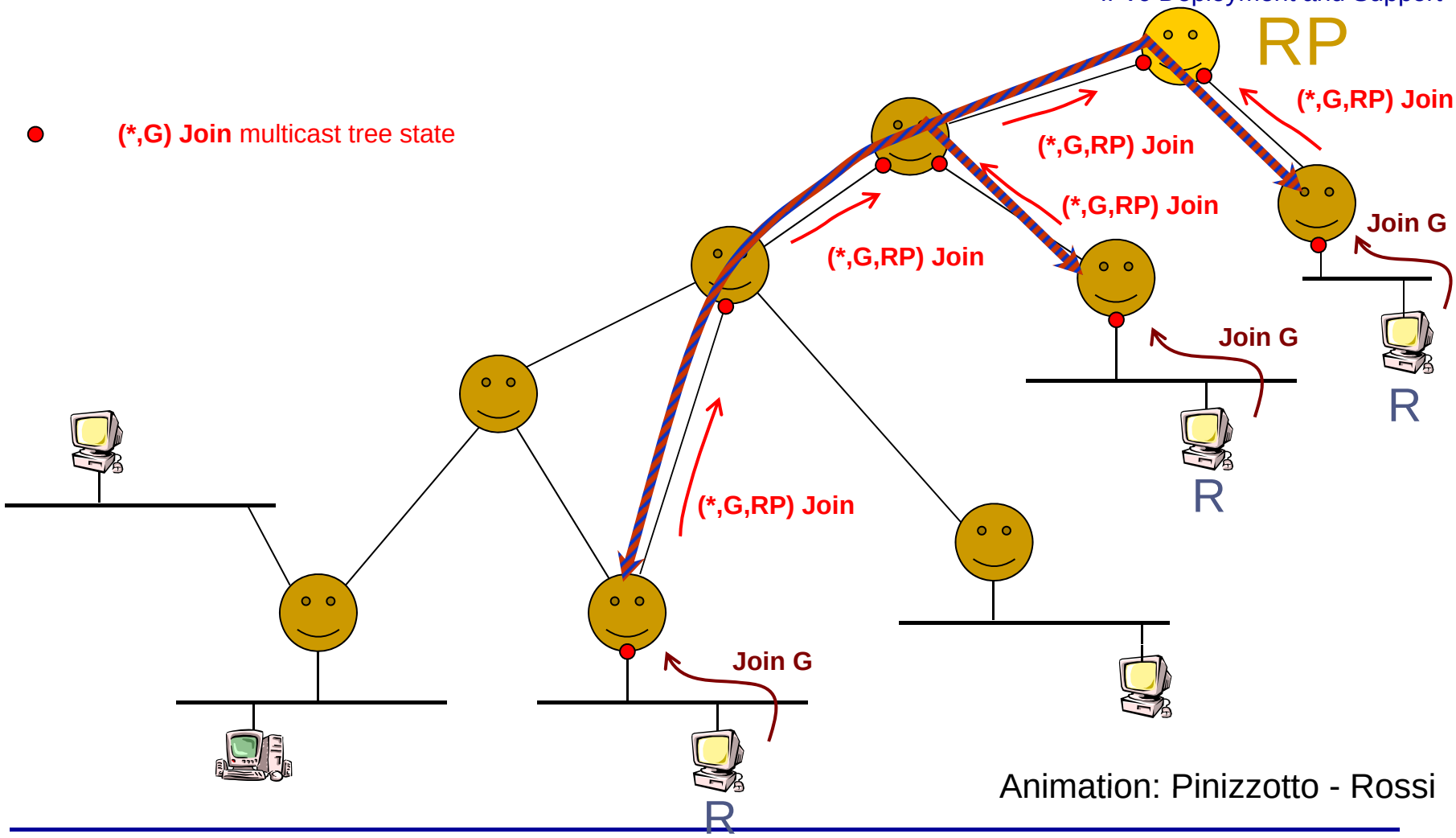
Reverse Path Forwarding

IPv6 Deployment and Support

- **PIM-SM uses RPF lookup to find where (interface/neighbor) to send join**
  - towards the RP (*,G join)
  - or towards the source (S,G join)
- **RPF also prevents loops: multicast not forwarded out the input interface**
- **RPF routing information provided by:**
  - **congruent sources:** EBGP, IGP, static routes etc, all routing protocols
  - **incongruent sources,** providing reachability information (NLRI) specifically for multicast RPF, such as MP-BGP

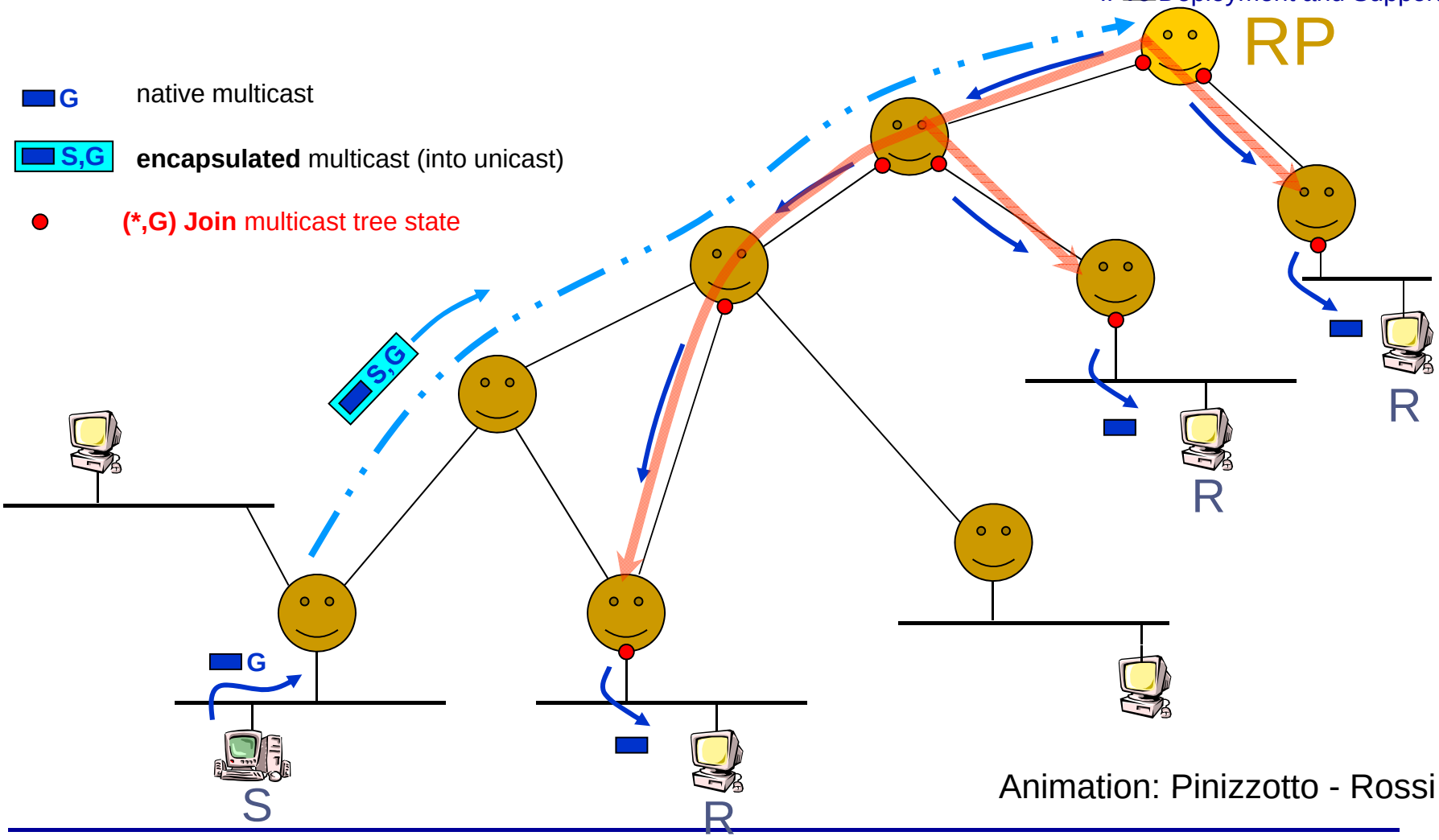
PIM-SM Example

IPv6 Deployment and Support



Animation: Pinizzotto - Rossi

PIM-SM Example

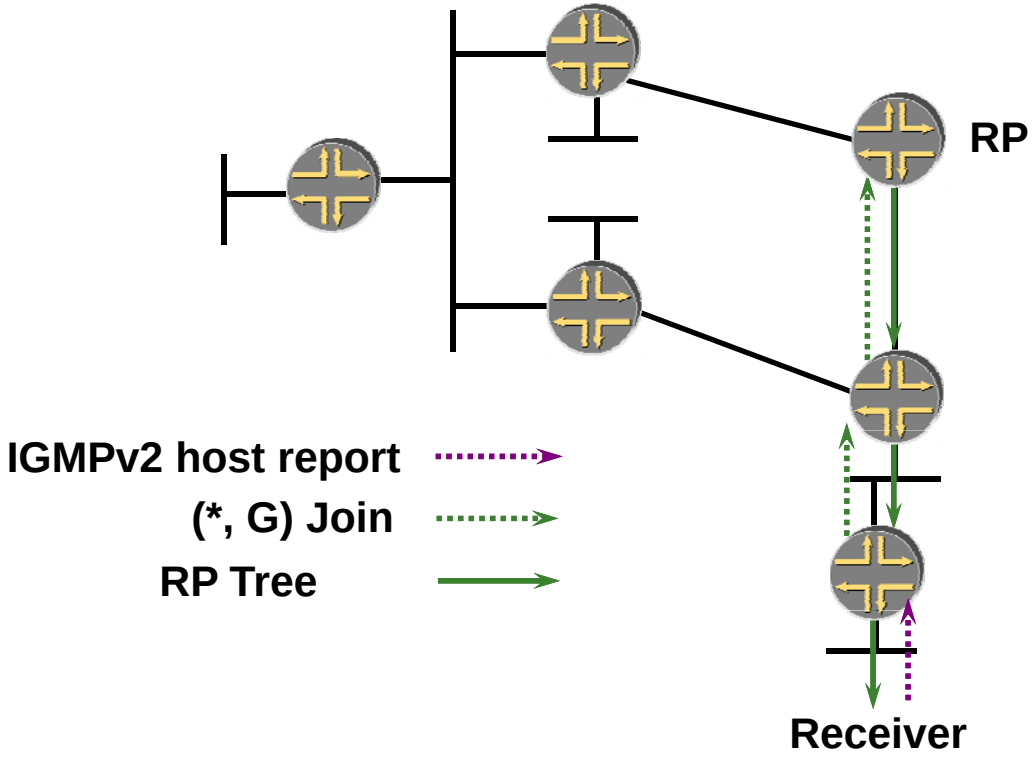


Animation: Pinizzotto - Rossi

Athanassios Liakopoulos
(aliako@grnet.gr)

ASM RP Tree Join

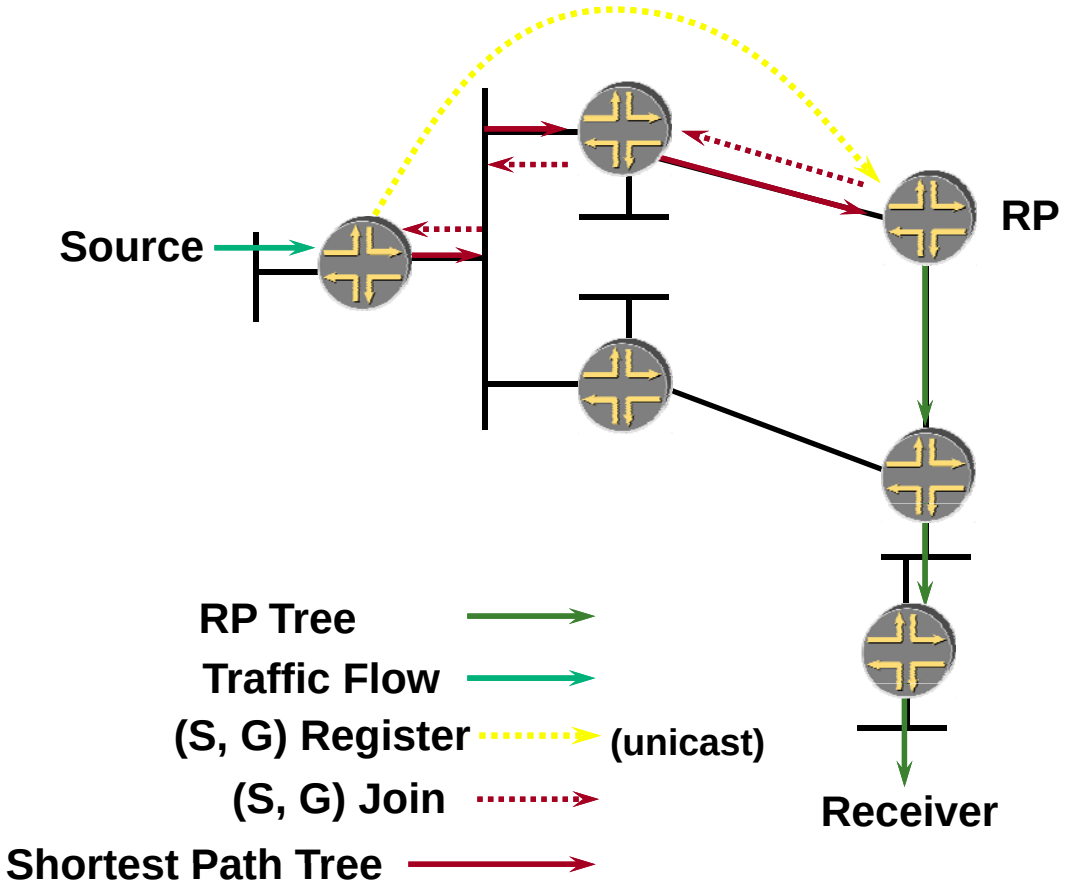
IPv6 Deployment and Support



Receiver announces desire to join group G with IGMPv2 host report – (*,G).

DR sends PIM (*,G) Join toward the RP; subsequent routers do likewise.

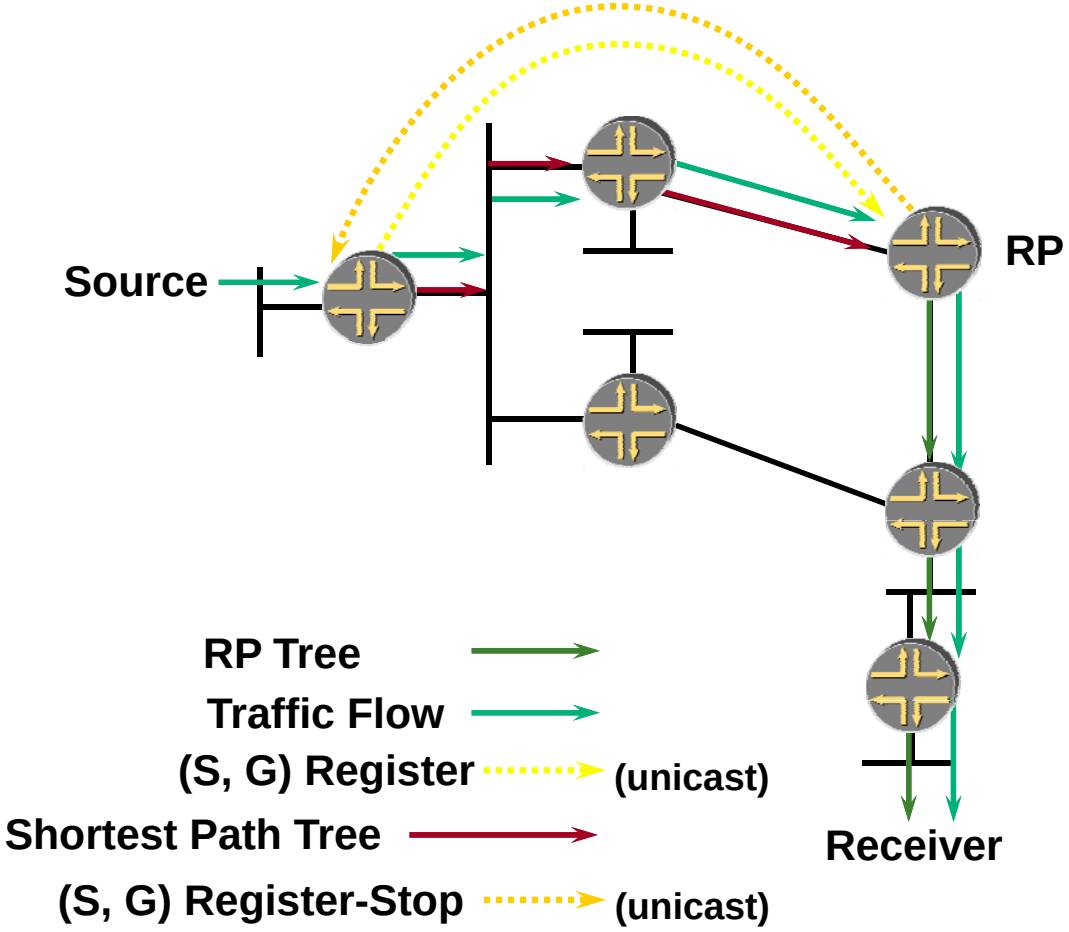
ASM Sender Registration



Active source triggers DR to send (S,G) Register message to RP.

RP sends (S,G) Join to source.

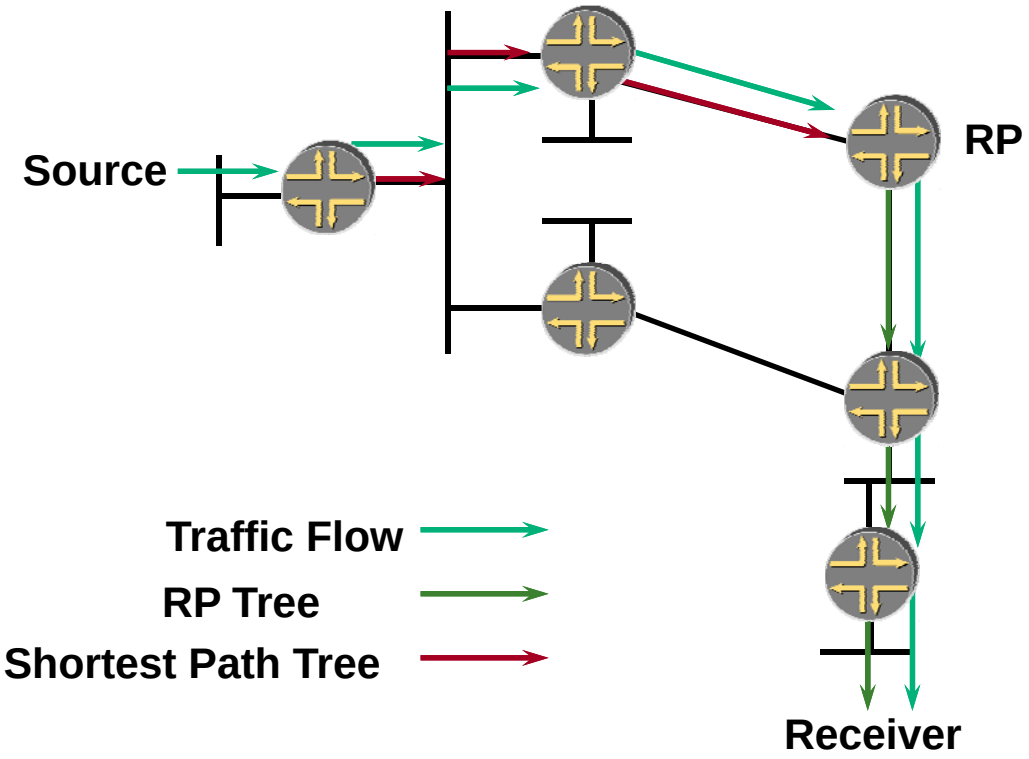
ASM Sender Registration



(S, G) traffic begins arriving at the RP via the SPT.

RP sends a Register-Stop back to the DR to stop the Register process.

ASM Sender Registration

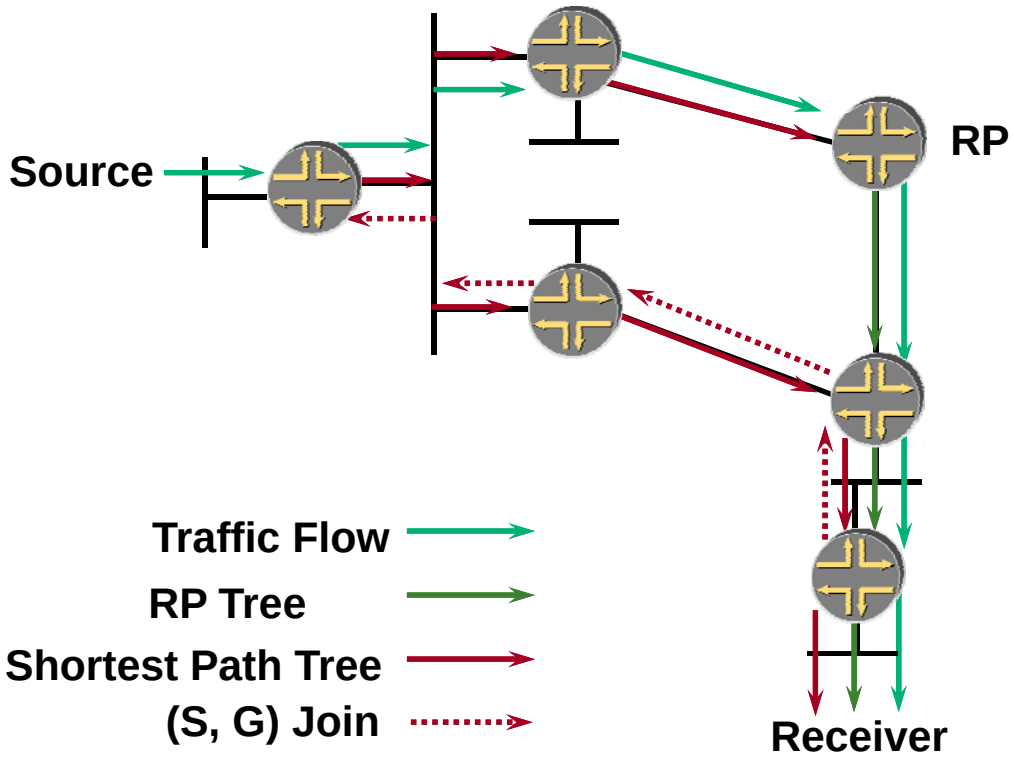


Source traffic flows natively along SPT to RP.

From RP, traffic flows down the RPT to the receiver.

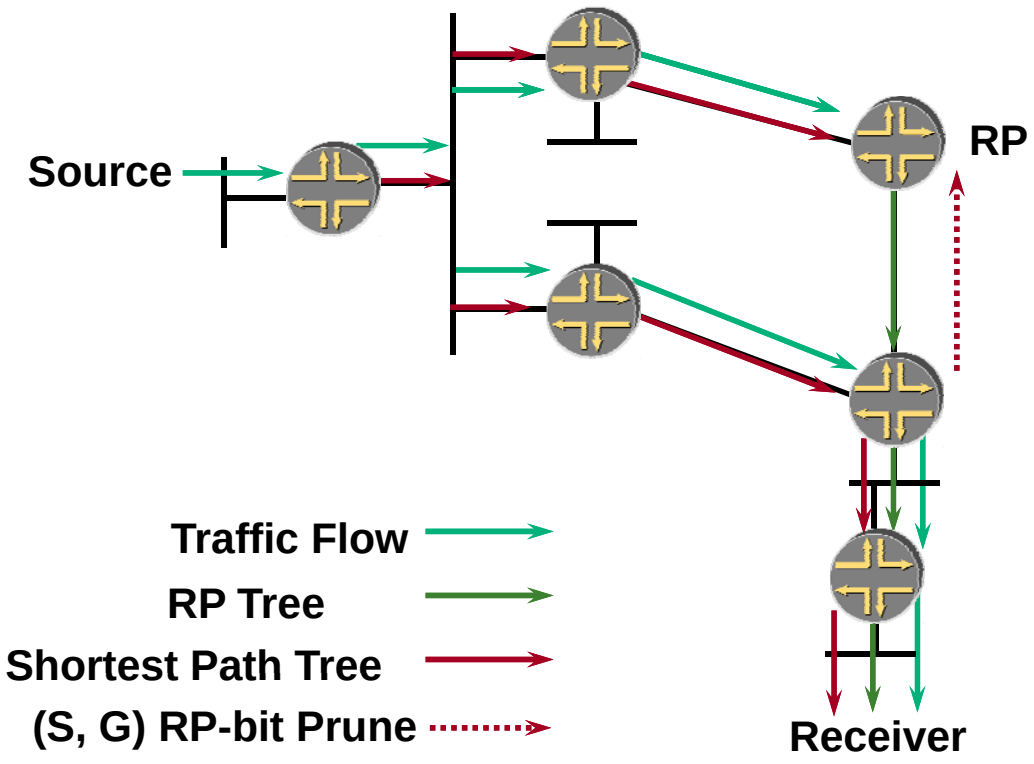
ASM SPT Cutover

IPv6 Deployment and Support



Receiver's DR joins the SPT.

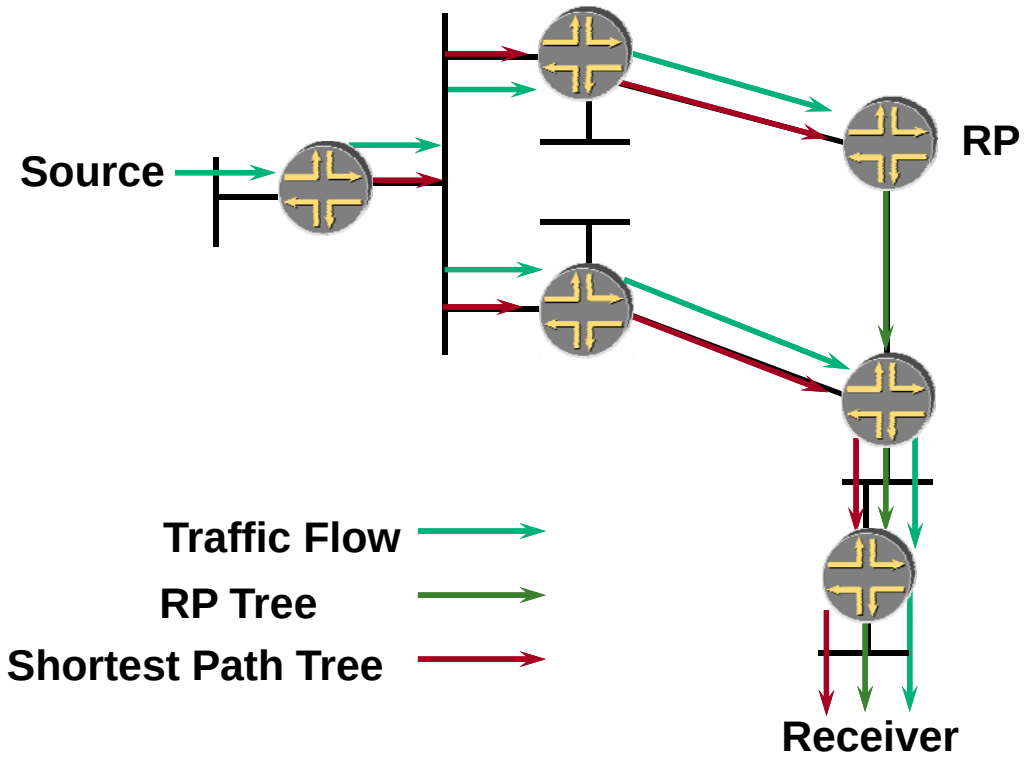
ASM SPT Cutover



Traffic begins flowing down the new branch of the SPT.

Additional (S, G) state is created along the RPT to prune off (S, G) traffic.

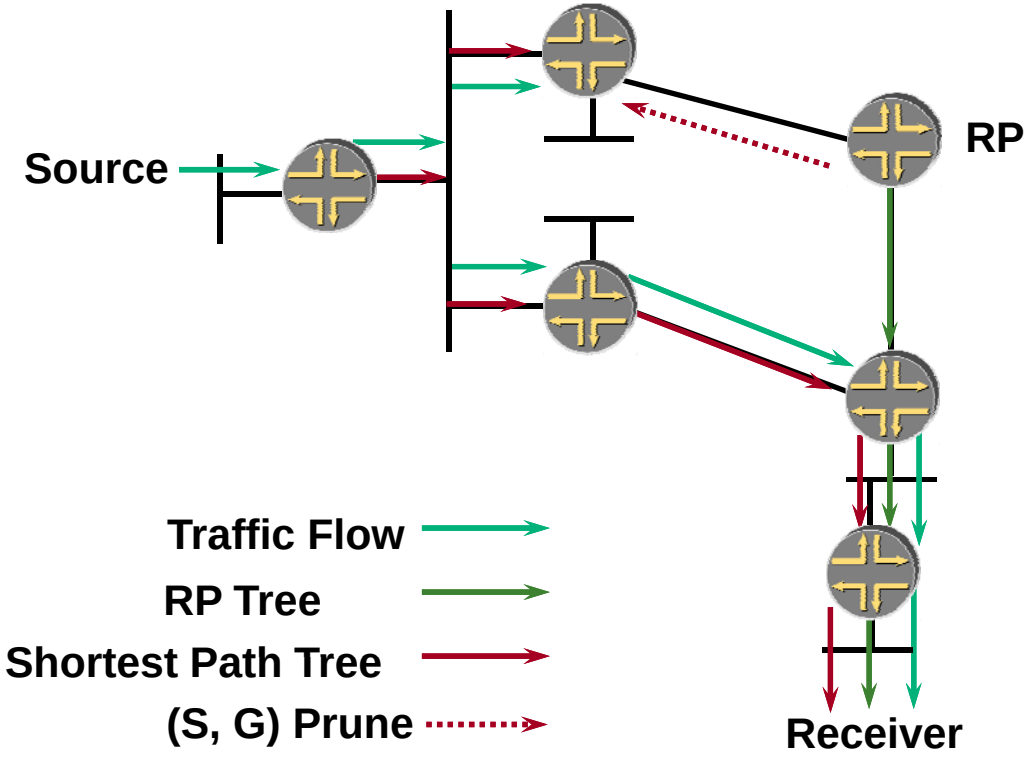
ASM SPT Cutover



(S,G) traffic flow is now pruned off of this branch of the RPT and is flowing to the receiver via the SPT.

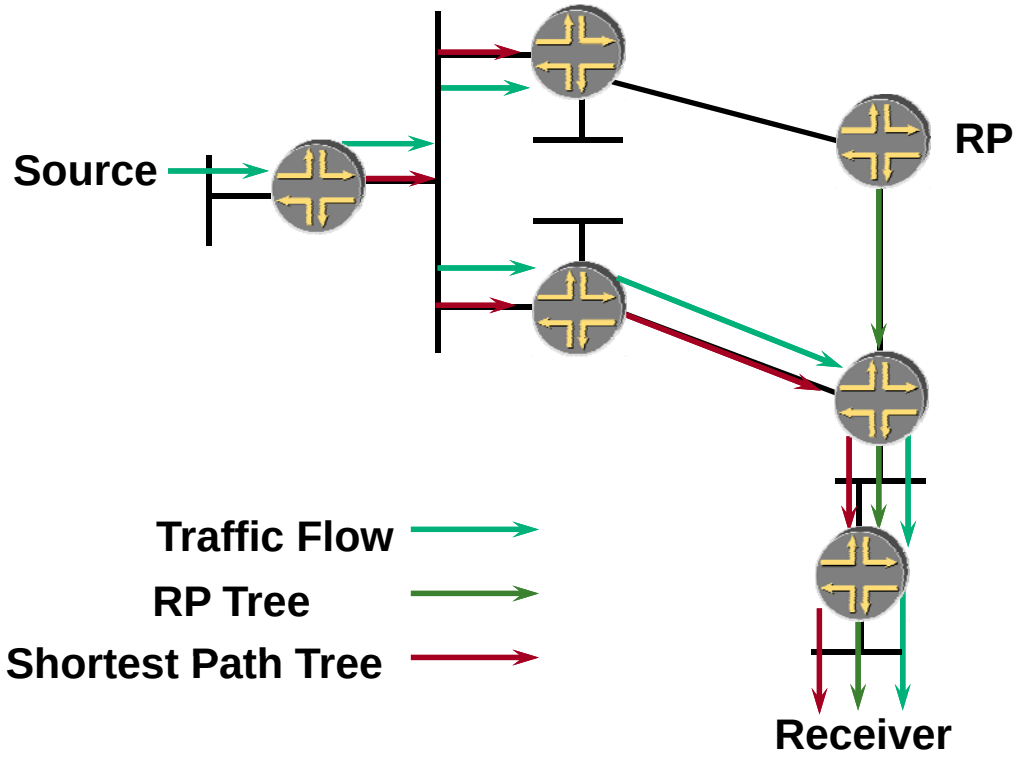
Traffic for other sources may still be flowing down the RPT.

ASM SPT Cutover



(S, G) traffic flow is no longer needed by the RP, so it prunes the flow of (S, G) traffic.

ASM SPT Cutover



(S, G) Traffic flow is now only flowing to the receiver via a single branch of the SPT.

As long as the source remains active, its DR sends Null-Register messages to the RP, enabling the RP to maintain a list of all active sources.

Outline

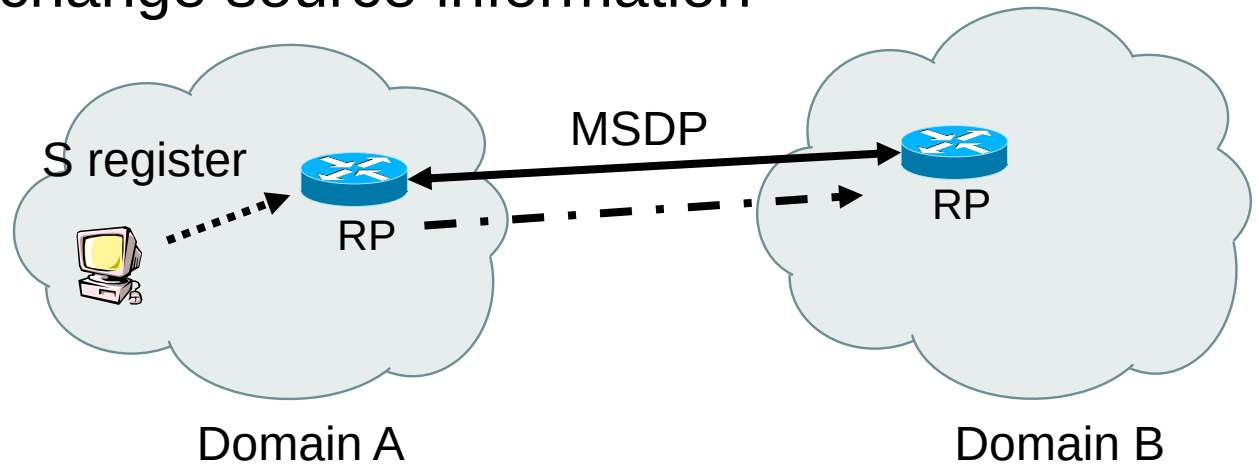
IPv6 Deployment and Support

- Multicast Addressing
- Multicast Listener Discovery
- Protocol Independent Multicast
- **Inter-domain Multicast**
- Configuration examples

Inter-domain Multicast

IPv6 Deployment and Support

- MSDP not supported for IPv6
  - In IPv4, each domain has typically one or more RPs. RPs in different domains use MSDP to learn about sources in remote domains
  - Also used between Anycast-RPs in same domain to exchange source information



Inter-domain Multicast

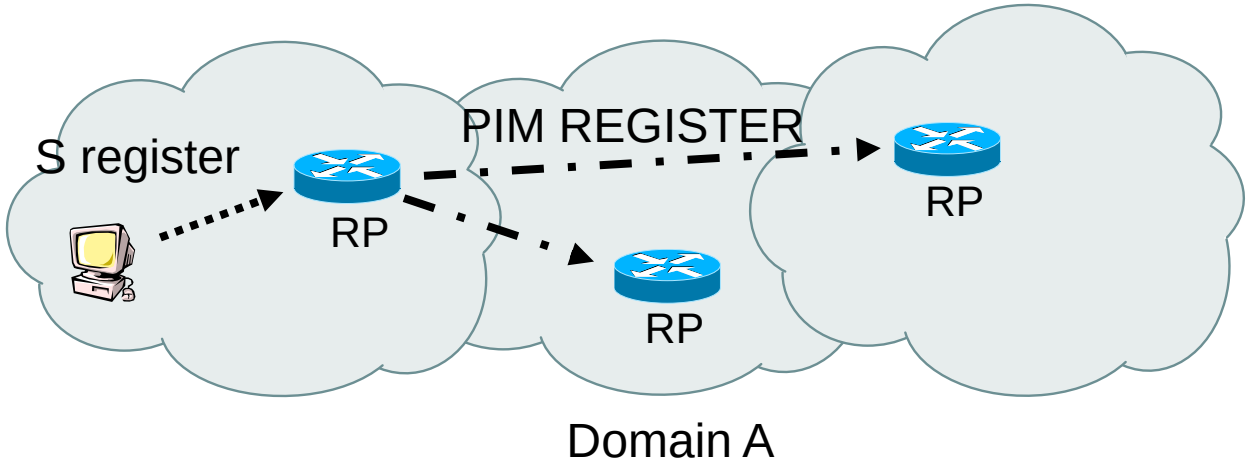
IPv6 Deployment and Support

- Without MSDP, there is no inter-domain source (and RP) discovery mechanism defined for IPv6
- For a given global group there can be only one single RP on the Internet
  - not scalable
- Only SSM? MSDP replacement? The end of inter-domain ASM?
- Embedded-RP (RFC 3956)
  - RP address is embedded in the group – RP discovery
  - PIM routers only need to support E-RP, no configuration necessary (except on RP)

  However:
  - the RP address is again the SPoF
  - and E-RP is open to inter-domain abuse of the RP resource
- Work in progress…

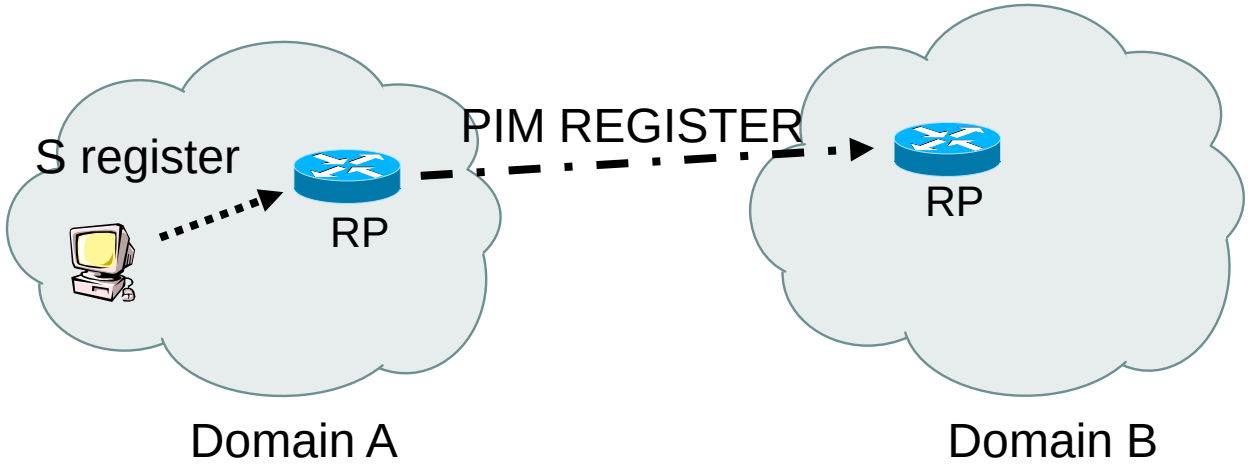
PIM Anycast-RP (RFC 4610) for Embedded-RP

- Each RP forwards PIM REGISTER received from a DR to other RPs in a domain (RP-set)
- Substitutes MSDP mesh-groups (among Anycast-RPs) with PIM REGISTER message forwarding



PIM Anycast-RP (RFC 4610) for Embedded-RP

- Consider extending this mechanism across domains with separate RP-sets
  - of course also sharing the anycast E-RP address
- Maybe half a solution for the inter-domain ASM problem but… could end up recreating MSDP



Outline

IPv6 Deployment and Support

- Multicast Addressing
- Multicast Listener Discovery
- Protocol Independent Multicast
- Inter-domain Multicast
- **Configuration examples**

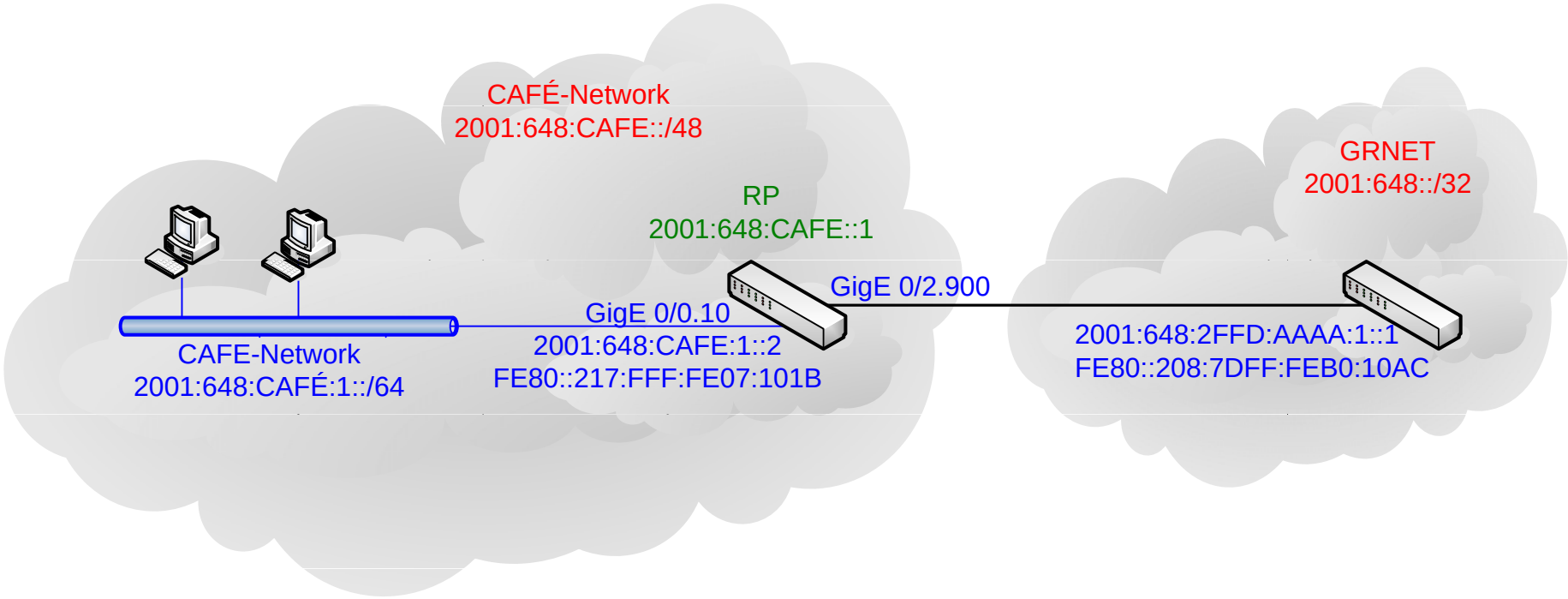
Necessary (?) IPv6 Multicast Features

IPv6 Deployment and Support

- Multicast Listener Discovery (MLD) Protocol, Ver.1&2
- PIM Sparse Mode (PIM-SM)
- PIM Source-Specific Multicast (PIM-SSM)
- PIM Embedded RP Support
- Address Family Support for Multiprotocol BGP
- Static Multicast Routing (mroute) for IPv6
- Scope Boundaries
- IPv6 Static RPs
- PIM-Bidir
- Bootstrap Router (BSR)

Topology

IPv6 Deployment and Support



Enable IPv6 multicast routing

IPv6 Deployment and Support

```
!
ipv6 multicast-routing
!
```

Verify MLD activation

IPv6 Deployment and Support

```
gw1#sh ipv6 mld interface gigabitEthernet 0/0.10
GigabitEthernet0/0.10 is up, line protocol is up
  Internet address is FE80::217:FFF:FE07:101B/10
  MLD is enabled on interface
  Current MLD version is 2
  MLD query interval is 125 seconds
  MLD querier timeout is 255 seconds
  MLD max query response time is 10 seconds
  Last member query response interval is 1 seconds
  MLD activity: 1454 joins, 1436 leaves
  MLD querying router is FE80::217:FFF:FE07:101B (this system)
```

Verify PIM neighbors

IPv6 Deployment and Support

```
gw1#sh ipv6 pim neighbor
Neighbor Address          Interface          Uptime    Expires DR pri Bidir

FE80::208:7DFF:FEB0:10AC  Gi0/2.900          3w3d      00:01:35 1      B

gw1#
```

MLD Access Control

IPv6 Deployment and Support

```
!
interface GigabitEthernet0/0
 ipv6 mld access-group mld-access-ctrl
!
ipv6 access-list mld-access-ctrl
 deny ipv6 …
 permit ipv6 …
…
!
```

Exchange IPv6 Multicast Routes


```
router bgp 65057
 neighbor 2001:648:2FFD:AAAA:1::1 remote-as 5408
 !
 address-family ipv6 multicast
 neighbor 2001:648:2FFD:AAAA:1::1 activate
 network 2001:648:CAFE::/48
 exit-address-family
!
```

Exchange IPv6 Multicast Routes


```
gw1#show bgp ipv6 multicast summary
BGP router identifier 195.251.29.207, local AS number 65057
BGP table version is 3182, main routing table version 3182
59 network entries using 8791 bytes of memory
59 path entries using 4484 bytes of memory
1704/49 BGP path/bestpath attribute entries using 211296 bytes of memory
1367 BGP AS-PATH entries using 39804 bytes of memory
295 BGP community entries using 14982 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 279357 total bytes of memory
BGP activity 57034/50250 prefixes, 63492/56708 paths, scan interval 60 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
2001:648:2FFD:AAAA:1::1
                4  5408  217451   93229     3182    0    0 3w0d           59
gw1#
```

Static RP


```
ipv6 pim rp-address 2001:660:3007:300:1:: m6bone-
  global-rp
!
ipv6 access-list m6bone-global-rp
 permit ipv6 any FF0E::/16
 permit ipv6 any FF1E::/16
 permit ipv6 any FF3E::/16
!
```

Static RP

IPv6 Deployment and Support

```
gw1#sh ipv6 pim tunnel
…
Tunnel2*
  Type  : PIM Encap
  RP    : 2001:660:3007:300:1::
  Source: 2001:648:CAFE:1::2
…

gw1#show int tunnel 2
Tunnel2 is up, line protocol is up
  Hardware is Tunnel
  MTU 1514 bytes, BW 9 Kbit, DLY 500000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation TUNNEL, loopback not set
  Keepalive not set
  Tunnel source 2001:648:CAFE:1::2 (GigabitEthernet0/0.10), destination
   2001:660:3007:300:1::
  Tunnel protocol/transport PIM/IPv6
  Tunnel TTL 255
  Tunnel is transmit only
…
```

Define Embedded RP

IPv6 Deployment and Support

```
ipv6 pim rp-address 2001:648:CAFE::1 grnethq-
  embedded-rp
!
ipv6 access-list grnethq-embedded-rp
 permit ipv6 any FF73:130:2001:648:CAFE::/96
…
!
```

This piece of configuration is needed only at the RP router!

```
gw1#sh ipv6 pim group-map FF73:130:2001:648:CAFE::/96

FF73:130:2001:648:2320::/96*
    SM, RP: 2001:648:CAFE::1
    RPF: Tu4,2001:648:CAFE::1 (us)
    Info source: Static
    Uptime: 2w6d, Groups: 0

gw1#
```

BSR domain boundaries

IPv6 Deployment and Support

```
!
interface GigabitEthernet0/2.900
 ipv6 pim bsr border
 ipv6 multicast boundary scope 8
!
```

Possible scope boundary values:

Scope 5: Universities (sites)

Scope 8: Universities

Scope A: NRENs

Scope B: SEEREN?


deploy
Questions?