MULTICAST ROUTING ALGORITHM

Protocol Independent Multicast Sparse Mode (PIM-SM)

PIYUSH HARSH
CONS, CISE University of Florida

INTERNET ROUTING SUMMARY

Unicast and Reverse Path Forwarding (RPF)	Any-Source Multicast (ASM)	Source-Specific Multicast (SSM)
	Inter domain	
Path vector {BGP (MBGP)}	Peer-RPF Flooding {MSDP}	No additional Protocol Needed
	Intra domain	
Link State {OSPF, ISIS(M-ISIS)} Distance Vector {RIP, DVMRP(Unicast)}	Sparse {PIM-SM} Dense {PIM-DM, DVMRP}	Sparse {PIM-SM (No-RP)} Dense {PIM-DM, DVMRP}

- PIM is protocol independent. What does this mean?
  - PIM uses the IP unicast forwarding table for performing RPF checks
  - IP Unicast forwarding table could be maintained using OSPF, RIP, DV, ISIS by the router (PIM does not care as long as there is some forwarding table)
  - Other prominent multicast routing protocol DVMRP maintains its own RPF table
  - Hence PIM is protocol independent

PIM Version 1

- Sent as IGMP message with IGMP version = 1 and IGMP type = 4
- Type of PIM Messages distinguished by IGMP Code
  - Code 0: Router-Query
  - Code 1: Register (PIM-SM only)
  - Code 2: Register-Stop (PIM-SM only)
  - Code 3: Join/Prune
  - Code 4: RP-Reachability (not used)
  - Code 5: Assert
  - Code 6: Graft (PIM-DM only)
  - Code 7: Graft-Ack (PIM-DM only)

PIM Version 2

- IP Protocol Number 103
- List of PIM Version 2 message types

  Type 0: PIM Hello Message
  Type 1: Register (PIM-SM only)
  Type 2: Register-Stop (PIM-SM only)
  Type 3: Join/Prune
  Type 4: Bootstrap
  Type 5: Assert
  Type 6: Graft (PIM-DM only)
  Type 7: Graft-Ack (PIM-DM only)
  Type 8: Candidate RP-Advertisement (PIM-SM only)
- Version 1 messages are subset of Version 2 messages
  - Router-Query (version 1) and Hello Messages (version 2) are same
  - RP-Reachability (version 1) message is not used

PIM Group to RP Mapping

All routers in domain must know and agree on active RP (Rendezvous Point) for each multicast group.

- Three mays to map the RP
  - Static group to RP mapping
  - Cisco Systems auto-RP (dynamic)
  - PIM bootstrap router (BSR) (dynamic)

Static Group to RP Mapping

- Least elaborate Method
- Each router in PIM domain must be manually configured with address of RP for each multicast group

Advantage: Simplicity

Disadvantages:

  - Requires reconfiguration on every router each time the address of RP changes
  - Fail over to backup RP requires additional configuration if primary RP becomes unreachable

CISCO Auto-RP

- Relies on dense mode operation to forward control messages
- Uses well known group addresses – 224.0.1.39 and 224.0.1.40
- All routers in PIM-SM Auto-RP enabled domain must be configured in sparse-dense mode
- Each router fits into one of three roles
  - Candidate RP
  - Mapping Agent
  - Discovery-only

CISCO Auto-RP (Contd ...)

- Candidate RP sends RP-Announce message detailing group ranges for which it intends to serve as RP to group 224.0.1.39 (CISCO-RP-ANNOUNCE)
- Routers configured as mapping agents join 224.0.1.39 group and listen to all RP-Announce messages
- Criteria for active RP selection by each mapping agents:
  - When multiple RPs announce same group prefix and mask, accept announcement from RP with highest IP add.
  - Reject group prefix if it is already covered by less-specific prefix advertised by same RP
  - Accept all other announcements
- Mapping agents announce RP for each group using RP-Mapping messages to 224.0.1.40 (CISCO-RP-DISCOVERY) after authoritatively selecting RP for group.

CISCO Auto-RP (contd ...)

- Discovery-only routers join 224.0.1.40 and learn the RP for each group
- Candidate RP routers and mapping agents also join 224.0.1.40
- 224.0.1.39 and 224.0.1.40 must be treated as dense mode group other they would also need RP for control message delivery
- Inefficiencies of dense mode operation is not an issue here, why??
  - Control Traffic is limited
  - Group membership is stable

PIM Bootstrap

- Added in PIM version 2 as standardized way for dynamic group-to-RP mapping
- All routers in the domain MUST use PIM version 2 packet formats
- One or more routers must be configured to serve as candidate BSRs
  - By default: PIM router's BSR priority = 0 (router ineligible as candidate BSR)
  - At least one router in the domain must have BSR priority > 0
  - Candidate BSR router sends Bootstrap (type 4 PIM v 2 message) out all its interface
  - Neighboring routers process the Bootstrap messages and forward them to all interfaces except that from which the message was received
  - PIM routers drops Bootstrap message if the incoming message fails the RPF check for the originating BSR candidate's IP address
  - If a candidate BSR receives a Bootstrap message with BSR priority > its own, that router stops announcing itself as BSR candidate.

PIM Bootstrap (contd ...)

- Eventually there will be only one router that will send out the Bootstrap messages
- Bootstrap messages are sent periodically (default 60 seconds)
- The RP-set info is conveyed to the BSR by all candidate RPs in the domain
- Candidate RPs unicast RP advertisement messages to the BSR Address advertised in the Bootstrap messages.
- BSR does not select the RP for groups, it just collects all the candidate RPs information and sends it out in subsequent Bootstrap messages.
- The BSR performs the election of active RP of each group range for its own use
- Each router in the domain must independently run the active RP selection hash algorithm.

PIM Candidate RP-selection hash Algorithm

1. Find all RPs with the most specific group range covering group G
2. From subset in step 1 find all RPs with the highest priority value.
3. RPs that meet criteria in steps 1 and 2, compute the hash value based on G, the RP address and the hash mask included in the Bootstrap message.
4. The RP with the highest hash value is the elected RP for the group G.
5. In case of a tie, RP with the highest IP address is the active RP.

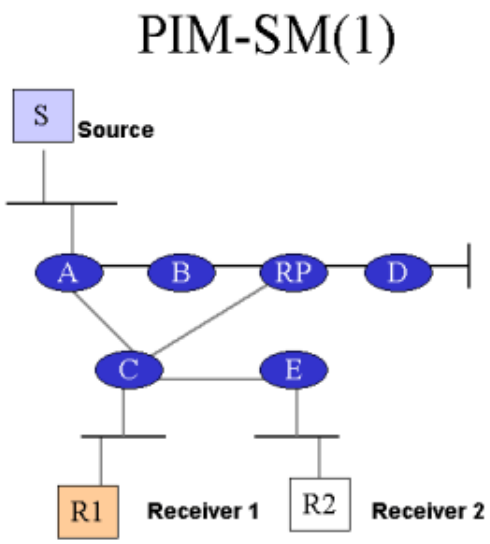
Formula for computing the hash value is:

(1103515245 * ((1103515245 * (G & M) + 12345) XOR C(i)) + 12345) mod 2 ^ 31

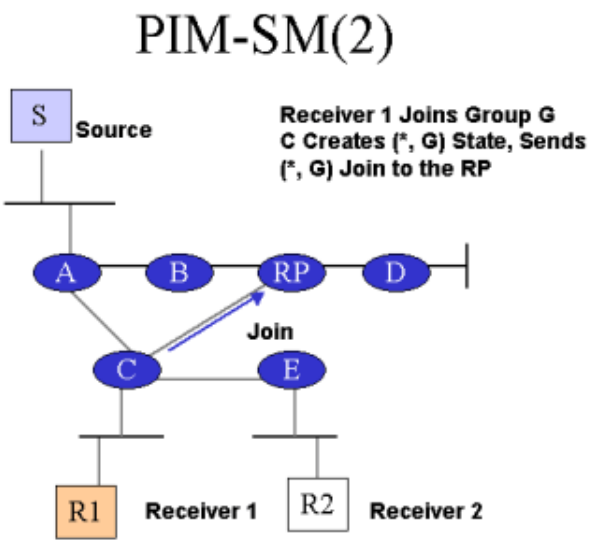
where C(i) is the RP address, M is the hash mask included in the Bootstrap message

- BSR addresses some deficiency of auto-RP with respect to robustness, load balancing and convergence.

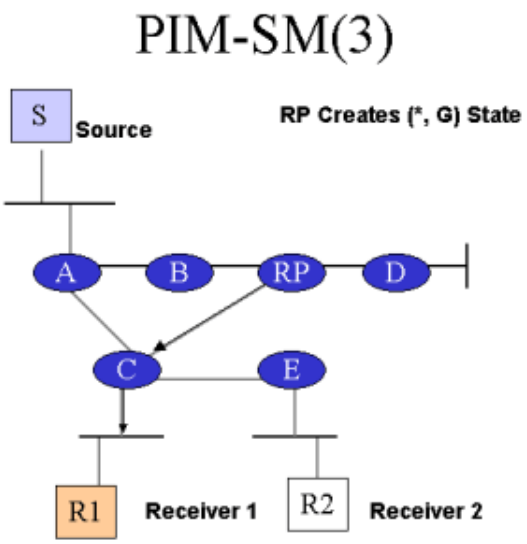
PIM in Action



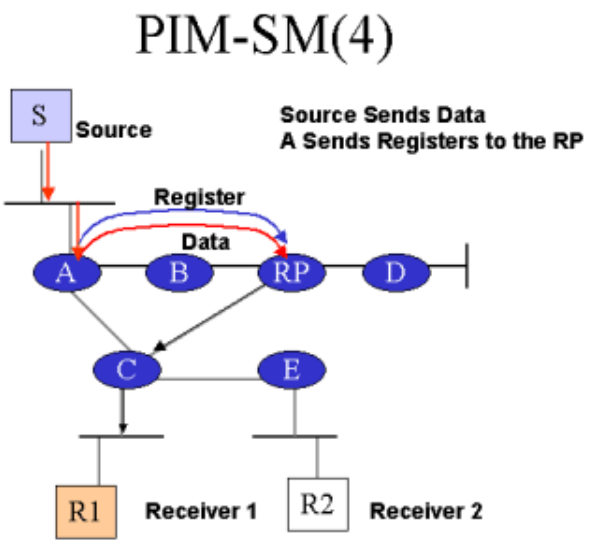
PIM in Action (contd ...)



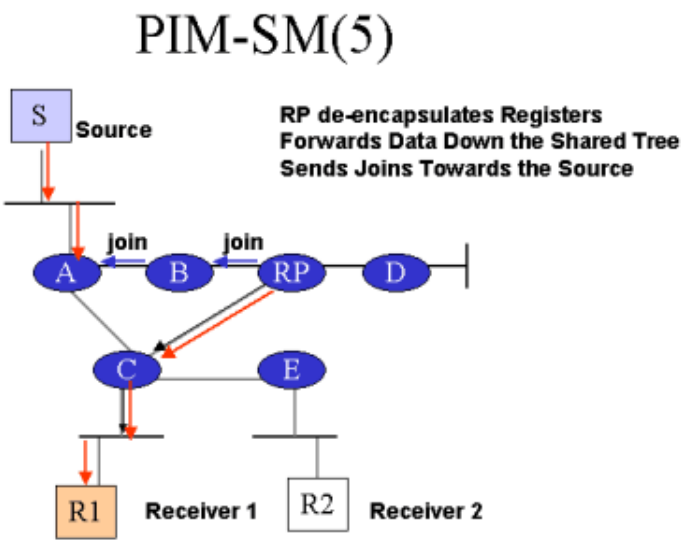
PIM in Action (contd ...)



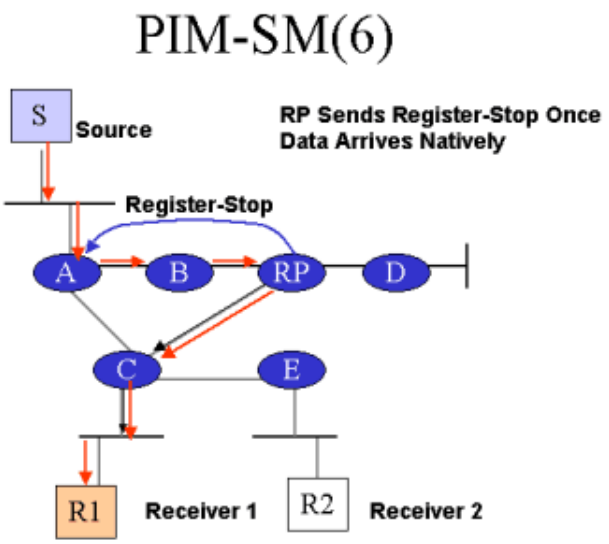
PIM in Action (contd ...)



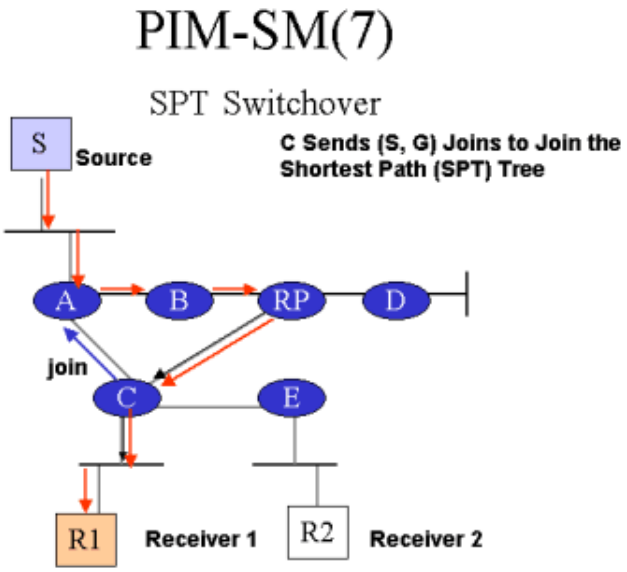
PIM in Action (contd ...)



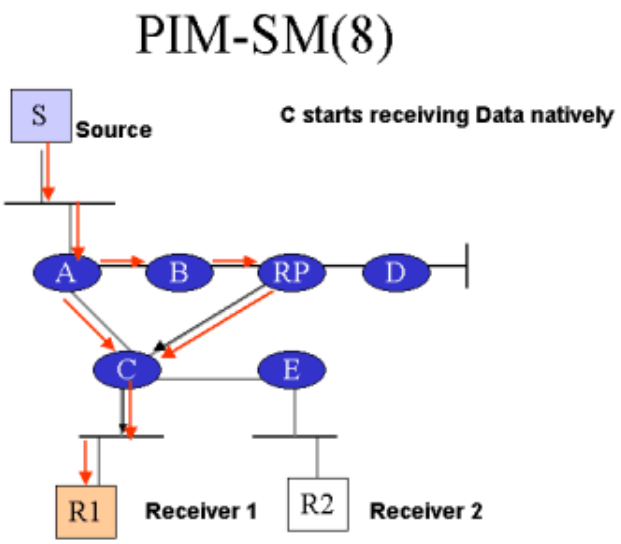
PIM in Action (contd ...)



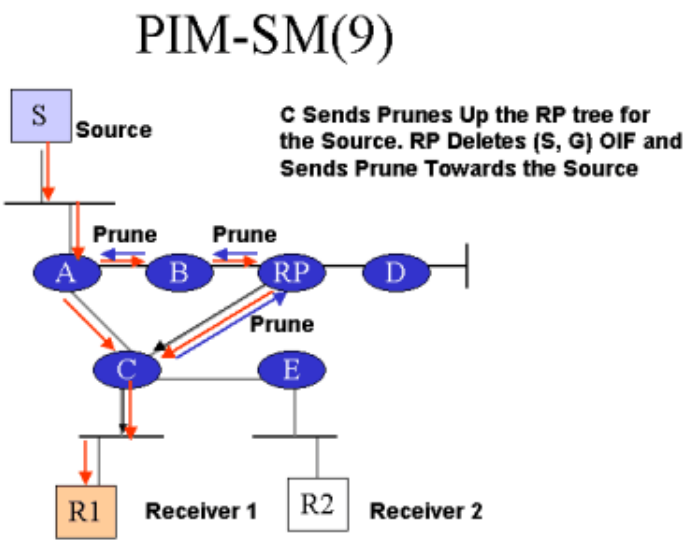
PIM in Action (contd ...)



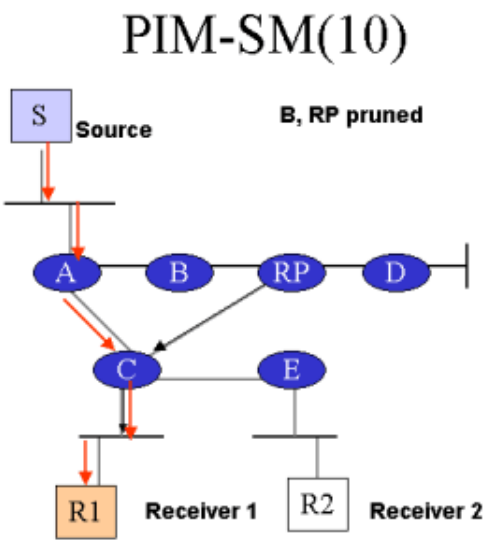
PIM in Action (contd ...)



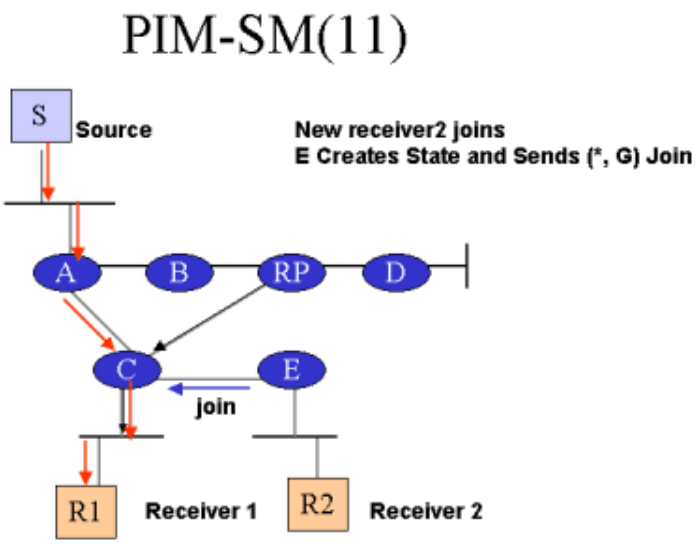
PIM in Action (contd ...)



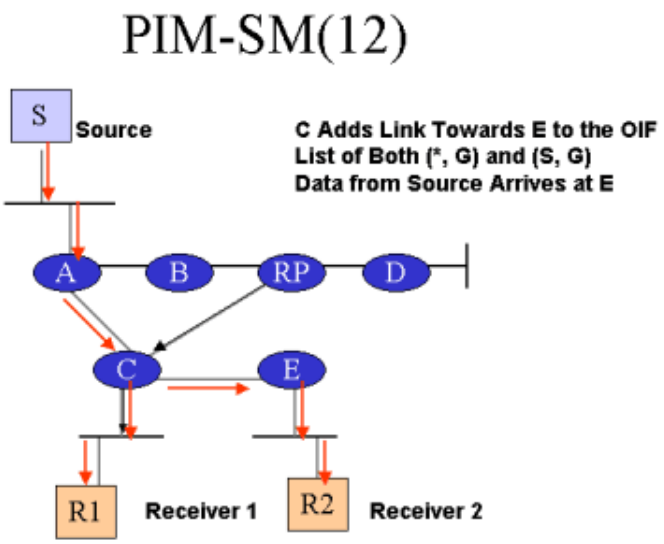
PIM in Action (contd ...)



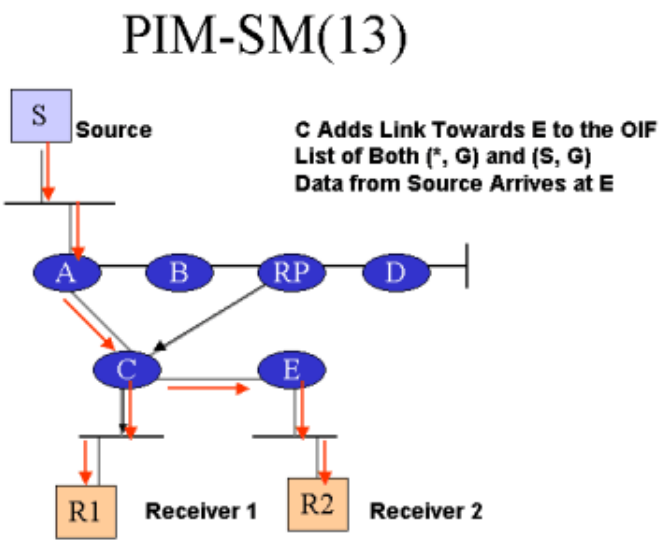
PIM in Action (contd ...)



PIM in Action (contd ...)



PIM in Action (contd ...)



References and Credits

- Interdomain Multicast Routing

  Practical Juniper Networks and Cisco Systems Solutions
- http://www.soi.wide.ad.jp/class/99007/slides/25/
- Developing IP Multicast Networks: Volume 1, Cisco Press

THANK YOU!