Building Scalable Cisco Internetworks (BSCI)

IP Multicast Routing

http://www.INE.com

What Is Multicast?

- Unicast
  - One to One
- Broadcast
  - One to All
- Multicast
  - One to Many

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Unicast Transmission Review

- Layer 3
  - Build routing table based on dynamic protocols or static configuration
  - Use destination IP address in packet to find outgoing interface
- Layer 2
  - Build CAM table through flooding of traffic
  - Use destination MAC address in frame to find outgoing port

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Broadcast Transmission Review

- Layer 3
  - Do not route broadcasts between interfaces
- Layer 2
  - No broadcast CAM table
  - Flood frames everywhere except port received on

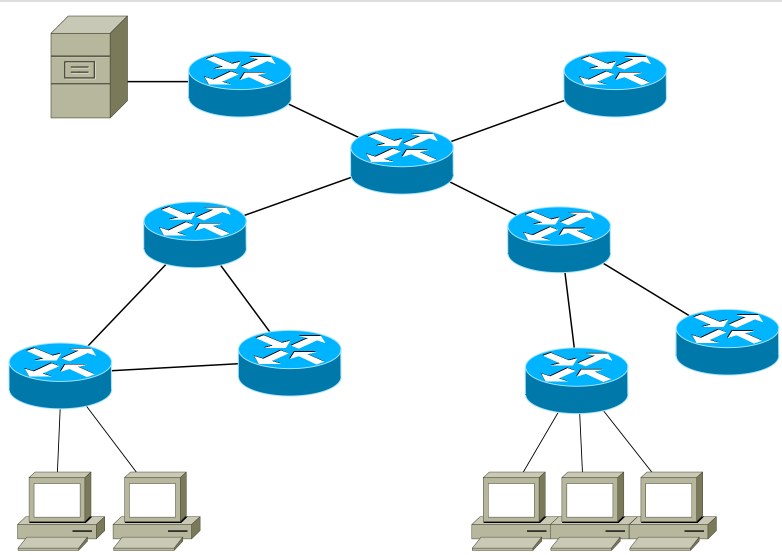
Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Why Use Multicast?

- Unicast
  - When transmitting similar feeds to multiple hosts, bandwidth is wasted
    - e.g. IPTV
- Broadcast
  - Can't be forwarded between router interfaces
- Multicast
  - Generate only one layer 3 packet
  - Let the layer 2 process do replication if needed

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Unicast vs. Broadcast vs. Multicast



Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Multicast Transmission Issues

- Layer 3
  - How do I prevent traffic from looping?
    - Build multicast tree based on reverse path away from source
      - i.e. IGP already assumed to be loop free
    - Use destination IP address in packet to find outgoing interface(s)
      - i.e. Multicast routing table
- Layer 2
  - Multicast is destination only… how do I build CAM?
    - Without help, treat like broadcast
    - With help, treat like unicast
      - Static CAM entries
      - CGMP
      - IGMP Snooping

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

IPv4 Multicast Layer 3 Addressing

- Uses "Class D" address range
- 224.0.0.0/4
  - Binary 11100000 - 11101111
  - Decimal 224.0.0.0 – 239.255.255.255
- Special uses…
  - 224.0.0.0/8
    - Link Local (TTL = 1)
  - 232.0.0.0/8
    - Source Specific Multicast (SSM)
  - 239.0.0.0/8
    - Administratively Scoped (private addressing)
- Traffic always sent ***to*** group address, never ***from***

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Common Multicast Addresses

- 224.0.0.1
  - All multicast hosts
- 224.0.0.2
  - All multicast routers
- 224.0.0.5
  - All OSPF routers
- 224.0.0.6
  - OSPF DR/BDR
- 224.0.0.9
  - RIPv2
- 224.0.0.10
  - EIGRP
- 224.0.0.13
  - PIM
- 224.0.0.22
  - IGMPv3

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

IPv4 Multicast Layer 2 Addressing

- Ethernet MAC address is 48 bits total
- Multicast MAC has…
  - 25 most significant bits fixed
    - Starts with 0100.5e
    - 25^{th} most significant bit always zero
  - 23 least significant bits derived from layer 3 multicast address, 9 most significant discarded
- Implies overlap of layer 3 to layer 2 mappings
  - (32 bit IPv4 address) – (4 most significant bits in common) – (23 least significant bits unique) = 5 bits of overlap
  - i.e. 2^5 or 32 addresses possible per MAC

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Layer 3 to Layer 2 Conversion Example

- 224.0.0.1
  - **11100000.00000000.00000000.00000001**
  - **12345678 90123456 78901234 56789012**
  - Result - 0100.5e00.0001
- 230.220.18.5
  - **11100110.11011100.00010010.00000101**
  - **12345678 90123456 78901234 56789012**
  - Result - 0100.5e5c.1205

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Multicast Group Membership

- End hosts only receive traffic for multicast groups they have subscribed to
  - e.g. tune to a channel in IPTV
- How does the host tell the network what it wants?
  - Internet Group Management Protocol (IGMP)

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

IGMP

- Internet Group Management Protocol
- Host to Router multicast protocol
- Used to “join” a multicast feed
- IP protocol number 2
- Three versions
  - IGMPv1
    - RFC 1112
  - IGMPv2
    - RFC 2236
  - IGMPv3
    - RFC 3376

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

IGMPv1

- When host wants to join multicast group, it sends an IGMP “Membership Report” to the group address
  - e.g. to join 224.1.2.3 send IGMP to 224.1.2.3
- Multicast routers listen for IP protocol 2 (IGMP) and keep track of the members
- Router periodically asks if hosts still want the feed
  - i.e. “Query” sent to 224.0.0.1 (all hosts)
- If no responses, router removes the group
  - i.e. implicit leave

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

IGMPv2

- Two main enhancements
  - Group specific queries
    - In IGMPv1 router asks everyone (224.0.0.1) if they still want the feed
    - In IGMPv2 router can ask just that group
  - Explicit leave message
    - In IGMPv1 the group times out if no one responds to the router's query message
    - In IGMPv2 the host can tell the router it's leaving
      - i.e. IGMPv2 "Leave Group" message
      - Router immediately responds with group specific query
- Misc enhancements
  - Querier election
    - Which router asks if the are multiple on the link?
  - Query-interval & response time configurable

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

IGMPv3

- IGMPv1/v2 only allows hosts to join based on destination address
  - Called "star comma G" join
  - Denoted as (*,G)
  - There could be multiple servers sending to same destination address
- IGMPv3 allows host to join based on source and destination address
  - Called "S comma G" join
  - Denoted as (S,G)
  - Used for Source Specific Multicast (SSM)

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Building the Multicast CAM Table

- Without help, layer 2 Ethernet switches must treat multicast as broadcast
  - Defeats the purpose of group membership at the LAN segment level
- Possible solutions
  - Static CAM entries
    - Works, but too much administrative overhead
  - CGMP
    - Cisco Group Management Protocol (deprecated)
    - Have the router tell the switch who joined
  - IGMP Snooping
    - Have the switch listen for join/leave messages

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

IGMP Snooping

- When a host wants to join/leave the group, it sends IGMP membership report / explicit leave
- If the switch listens for IP protocol 2, it knows who joined / left the group
  - Host A on port 1 joined group 224.1.2.3
  - Add MAC for 224.1.2.3 on port 1
- Requires implementation in hardware to avoid excessive delay
  - All frames must be checked at layer 3 now

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Building the Multicast Routing Table

- Last hop router & switch now know what traffic the host(s) want
- How do we route the traffic loop free?
- Two possibilities
  - Run a separate protocol to advertise loop free multicast topology
    - e.g. DVMRP & MOSPF
  - Use our already loop-free IGP topology
    - i.e. Protocol Independent Multicast (PIM)

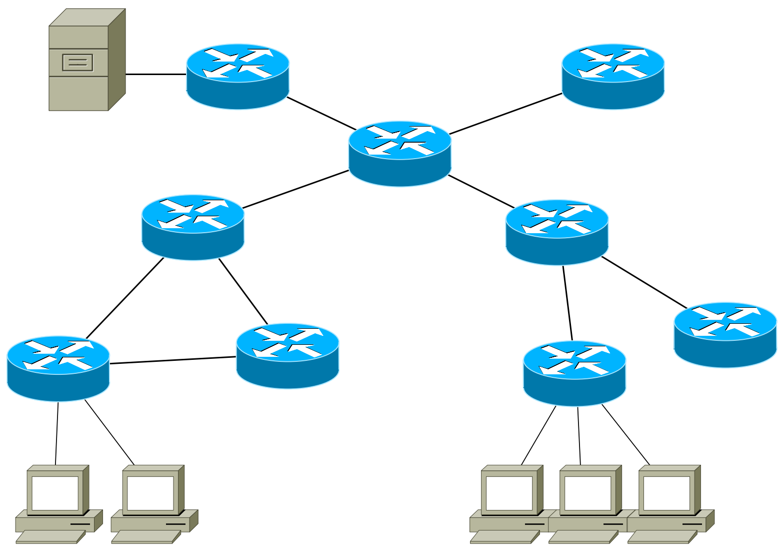
Copyright © 2009 Internetwork Expert, Inc
www.INE.com

PIM Loop Prevention

- Logic is…
  - Assume my IGP is already loop free
    - e.g. EIGRP, OSPF, static, etc.
  - When a multicast packet is received, compare the source address against the unicast routing table
    - Did the packet come in the interface I would use to get back to the source?
    - If so I can assume it's loop free and can forward the packet
    - Did the packet come in an interface I would ***not*** use to get back to the source?
    - If so I ***can not*** assume it's loop free, ***I must drop the packet***
- Called Reverse Path Forwarding lookup (RPF)
- PIM is considered "Independent" because it doesn't care what IGP is used to build the RPF tree

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

RPF Lookup Example



Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Finding the Servers

- I now know…
  - How to prevent loops when traffic comes in
  - What hosts on the LAN want traffic
- How do we find the servers?
- Two options…
  - Flood their traffic everywhere
    - i.e. PIM Dense Mode
  - Ask for the traffic before flooding
    - i.e. PIM Sparse Mode

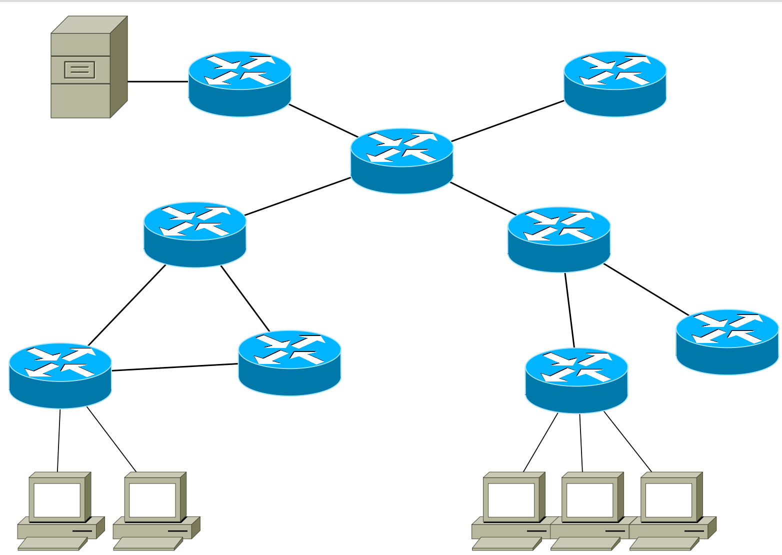
Copyright © 2009 Internetwork Expert, Inc
www.INE.com

PIM Dense Mode

- Uses "periodic flood and prune" behavior to send server traffic everywhere
  - Flooding
    - When a multicast packet comes in, send it out all other PIM enabled interfaces
  - Pruning
    - If someone tells me they *don't* want the traffic, I'll stop sending
  - Periodic
    - After a while flood it again just in case
- AKA "implicit join"
  - All traffic unless you say you don't want it
- All dense trees are considered Source Based Trees (SBTs) or Shortest Path Trees (SPTs)
  - RPF failure automatically prunes off looped paths
  - Efficient routing, but inefficient bandwidth utilization

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

PIM Dense Mode Example



Copyright © 2009 Internetwork Expert, Inc
www.INE.com

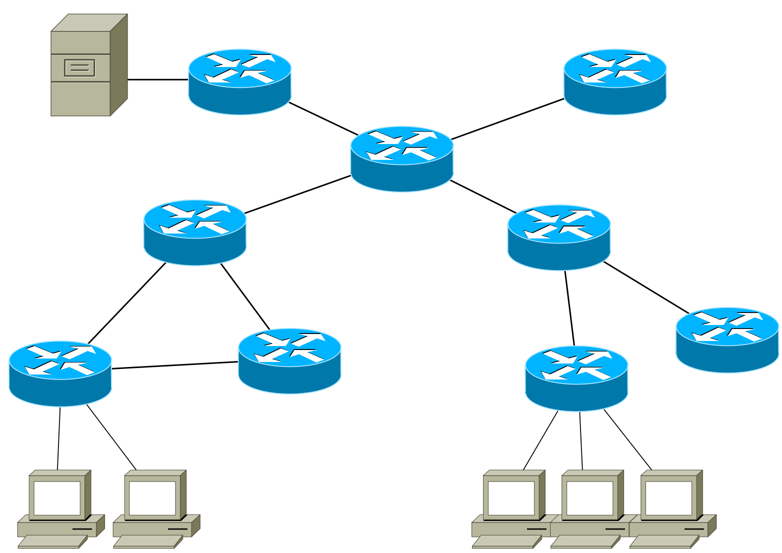
Copyright © 2009 Internetwork Expert

PIM Sparse Mode

- Uses “explicit join” model
  - No traffic unless you ask for it
- Who should we ask?
  - Rendezvous Point (RP)
- RP learns about the sources
  - PIM Register process
- RP learns about the clients
  - PIM Join process
- RP ties the trees together
  - Shared tree through the RP
  - Can do SPT switchover afterwards

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

PIM Sparse Mode Example



Copyright © 2009 Internetwork Expert, Inc
www.INE.com

RP Assignment

- Manually
  - **`ip pim rp-address`**
- Automatically
  - Auto-RP
    - Cisco proprietary
  - Bootstrap Router (BSR)
    - Standard per PIMv2

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

How RP Discovers Source

- Source S1 sends traffic to group G1
- PIM DR on LAN segment hears (S1,G1) traffic and sends Unicast “Register” message to RP
  - **`debug ip pim`** on DR and RP shows this
- RP now knows that (S1,G1) is sending and replies to DR with “Register Stop”
  - I know the source exists, stop sending me the traffic
- DR will periodically refresh Register message
  - Called null register

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

How RP Discovers Destination

- Destination sends IGMP join for (*,G1) onto LAN
- PIM DR on LAN segment hears IGMP Join for (*,G1) and sends PIM Join for (*,G1) up reverse path tree towards RP
  - **`debug ip pim`** on DR, RP, and in transit path
- PIM routers in transit path install (*,G1) entry in multicast routing table
- RP now knows that a host wants (*,G1)

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

How the Shared Tree is Merged

- RP now knows source (S1,G1) and destination (*,G1)
- PIM (*,G1) Join is forwarded up reverse path tree towards S1's DR
- PIM routers in transit path install (*,G1) entry in multicast routing table
- S1's DR receives (*,G1) join and adds incoming interface of PIM join to OIL for (S1,G1)
- (S1,G1) traffic now flows from source to destination through RP

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Shortest Path Switchover

- Once destination’s DR receives (S1,G1) feed it may choose to switch to a Shortest Path Tree (SPT) by sending a new (*,G1) PIM Join towards S1 instead of towards RP

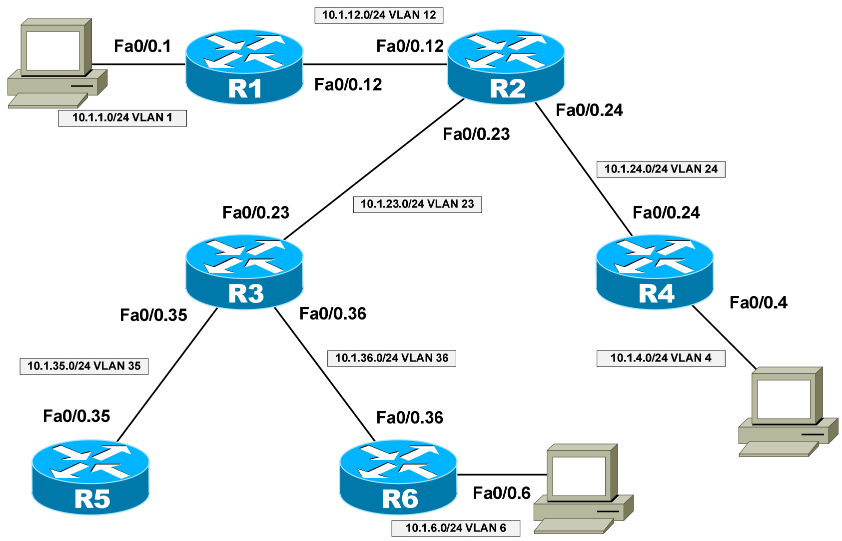
Copyright © 2009 Internetwork Expert, Inc
www.INE.com

PIM Sparse Dense Mode

- Combination of both modes
- Groups with an RP run as sparse
- Groups without an RP run as dense
- If RP goes down, fail open to dense mode

Copyright © 2009 Internetwork Expert, Inc
www.INE.com

PIM Examples



Copyright © 2009 Internetwork Expert, Inc
www.INE.com

Multicast Q&A

Copyright © 2009 Internetwork Expert, Inc
www.INE.com