



# Cisco Support Community Expert series Webcast

## Layer 3 Multicast: Security and Best Practices

Luis Espejel CCIE #52804 & David Ramirez  
Telecommunications Manager & Network Engineer  
November 28<sup>th</sup> 2017



# News & Upcoming events



# Ask the Expert following the Webcast

Now through Friday December 8<sup>th</sup>

With Luis Espejel  
& Miguel Perez

<http://bit.ly/ATE-Muticast>



Luis Espejel



& Miguel Perez

# Cisco Support Community – Ask the Expert

Cisco AnyConnect implementation, troubleshooting & best practices

Aditya Ganjoo y  
Puneesh Chhabra

<http://bit.ly/ATE-AnyConnect>



Ask the Expert

Ask all your doubts about Cisco Anyconnect implementation, troubleshooting & best practices

Aditya & Puneesh

**NOV 20- DEC 8, 2017**

**Join the discussion!**

**SUPPORT COMMUNITY**

**-Event open only to Customers & Partners-**

The banner features a background image of a man in a grey hoodie sitting at a wooden bar counter in a cafe, looking at a tablet. A white mug is on the counter in front of him. The text is overlaid on this image.

-Event open only to Cisco  
Customers and Partners-

# Cisco Support Community – Webcast in Portuguese

## Hyperconvergence with Cisco HyperFlex

December 13<sup>th</sup>, 2017

14hrs Brasilia, UTC (-2hrs)

With  
Pedro Ivo Santos

<http://bit.ly/web-PT-Hyperflex>



The graphic features a central image of three people in a server room setting. A small inset in the top left shows a video feed of Pedro Ivo Santos with the text 'Webcast ao vivo' and 'Pedro Ivo Santos' below it. The bottom of the graphic has a blue banner with the title 'Hyperconvergence com Cisco HyperFlex', a note '-Este evento é exclusivo para os parceiros e clientes da Cisco-', and a 'Registre-se aqui!' button. The Cisco logo and 'COMUNIDADE DE SUPORTE' are also visible.

-Event open only to Cisco Partners & Customers-

# Cisco Support Community – Spanish Ask the Expert

## New Jabber release and best troubleshooting practices

Till December 15<sup>th</sup> , 2017

With  
Ricardo Rendon

<http://bit.ly/Jabber-troubleshoot>



Pregunte al Experto

Ricardo Rendon

Nuevo release de Jabber & mejores prácticas de Troubleshooting

NOV 27 - DIC 15 -Evento Privado- ¡Haz Preguntas!

COMUNIDAD DE SOPORTE CISCO

-Event open only to Cisco  
Customers and Partners-

# Become an event Top Contributor!

Participate in Live  
Interactive Technical  
Events and much more

<http://bit.ly/Event-Top-Contributors>



Welcome to Cisco Support Community. We would love to have your [feedback](#).  
For an introduction to the new site, [click here](#). If you'd prefer to explore, try our [test area](#) to get started. And see [here for current known issues](#).

## Events Top Contributors



This program recognizes Cisco experts in the Cisco Support Community (CSC) that host technical events (Webcasts, Ask the Experts, Tech Talks, and Facebook Forums.) With this program, Cisco recognizes the positive, valuable influence that our top Cisco experts exert on the communities. To learn more, please visit our [FAQs](#).

2014 2013



## Cisco Designated VIPs

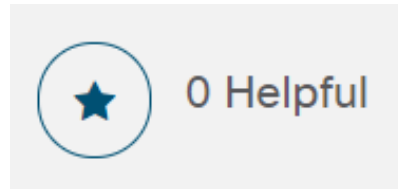


The Cisco Designated VIP program recognizes the top external individual contributors in Cisco's online communities, including the Cisco Support Community (CSC), Cisco Learning Network (CLN) and the Cisco Developers Network (CDN). Cisco Designated VIPs are recognized by their peers for their expertise and tireless contributions, and their abundant participation is vital to community success. With this program, Cisco formally recognizes the positive, valuable influence our top individual members exert on the communities overall. [FAQs](#)

# Rate content at the Cisco Support Community

Help us to recognize the quality content in the community

Rate documents,  
Videos & blogs!



Encourage and acknowledge people who  
generously share their  
time and expertise





# Cisco Support community Experts



Luis Espejel  
Telecommunications Manager  
CCIE R&S #52804



David Ramirez  
Network Engineer

# Question Manager



Miguel Perez  
Technical Support Engineer

Thank You For  
Joining Us Today!



Download Today's Presentation  
[http://bit.ly/webcast-slides\\_multicast](http://bit.ly/webcast-slides_multicast)

# Submit Your Questions Now!

Use the **Q&A** panel to submit your questions and the panel of experts will respond.

They will be answered eventually



Please take a moment to complete the survey at the end of the webcast

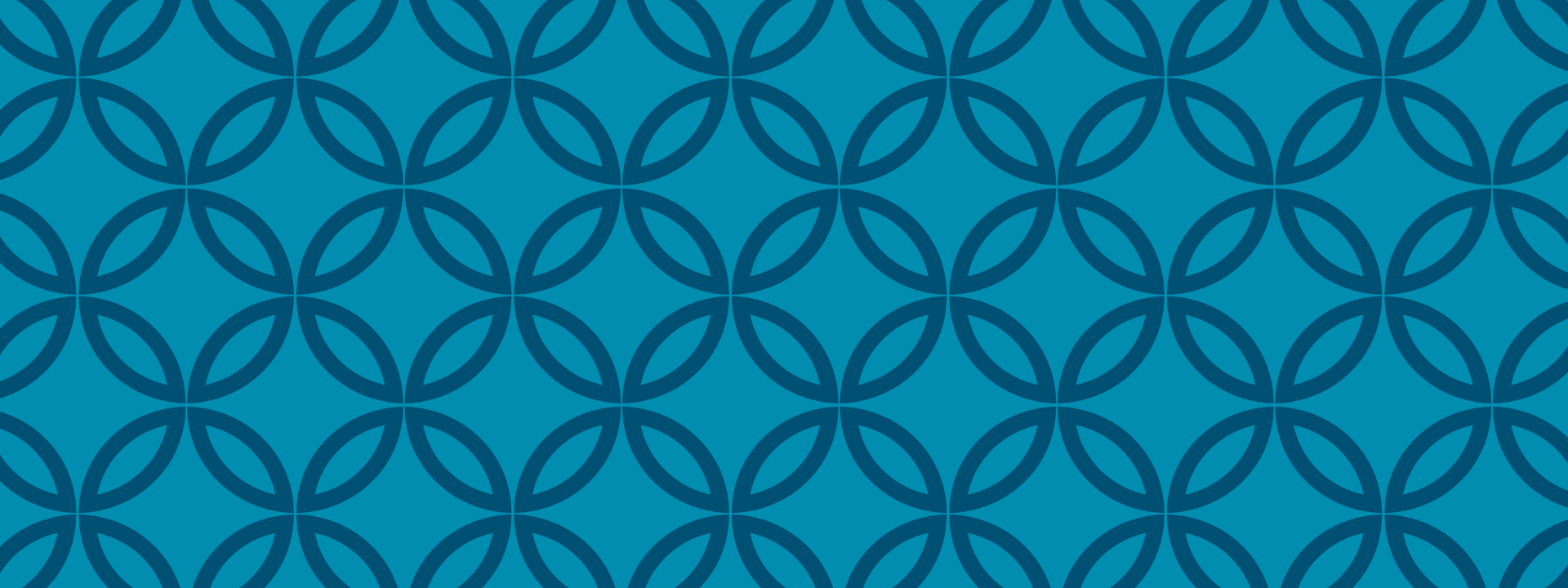


# Cisco Support Community Expert series Webcast

## Layer 3 Multicast: Security and Best Practices

Luis Espejel CCIE #52804 & David Ramirez  
Telecommunications Manager & Network Engineer  
November 28<sup>th</sup> 2017





# IP Multicast Security

Luis Espejel  
David Ramirez

# Polling Question 1

Do you use multicast in your network?

- A. Yes
- B. No

# Multicast – what problem solves?

- Sending traffic for a group of devices in a better way than using broadcast
- Bandwidth conservation
- One to many applications: TV, Radio, Distance learning, Application Updates, Data Distribution (weather, news, NTP, IoT sensors)
- Many to Many: Audio and Video Communication, Data distribution and synchronization, group chat, Financial, Polling, Multiplayer games.
- Many to one: Data Collection, video surveillance, Polling, IoT sensors.



# Multicast – PIM Basic packets

## Multicast:

- PIM hello
- PIM join/prune
- PIM DF-elect (PIM Bidir)
- PIM Assert
- PIM Bootstrap

## Unicast:

- Source Register Packets
- Register Stop packet
- BSR Candidate-RP advertisement

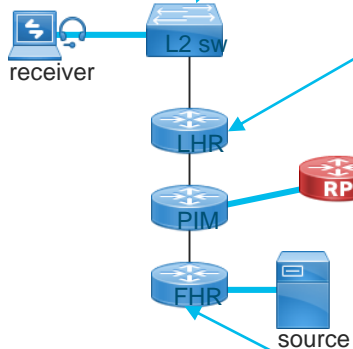
# Multicast – Basic functioning

```
IGMPSN: Received IGMPv2 Report for group 239.1.2.3 received on Vlan 10, port Et0/0
IGMPSN: group: Received IGMPv2 report for group 239.1.2.3 from Client 10.2.3.1 received on Vlan
10, port Et0/0
IGMPSN: Add v2 group 239.1.2.3 member port Et0/0, on Vlan 10
IGMPSN: group: Added port Et0/0 to group 239.1.2.3
IGMPSN: group: Forwarding 239.1.2.3 report to router ports
```

```
*Nov 18 23:04:03.850: MRT(0): Create (*,239.1.2.3), RPF (unknown, 0.0.0.0, 0/0)
*Nov 18 23:04:03.850: MRT(0): WAVL Insert interface: Vlan10 in (*,239.1.2.3) Successful
*Nov 18 23:04:03.850: MRT(0): set min mtu for(0.0.0.0, 239.1.2.3) 0->1500
*Nov 18 23:04:03.850: MRT(0): Set the C-flag for (*, 239.1.2.3)
*Nov 18 23:04:03.850: MRT(0): Add Vlan10/239.1.2.3 to the olist of (*, 239.1.2.3), Forward state -
MAC not built
*Nov 18 23:04:03.850: MRT(0): Update Vlan10/239.1.2.3 in the olist of (*, 239.1.2.3), Forward state -
MAC not built
*Nov 18 23:04:03.850: PIM(0): Building Triggered (*,G) Join / (S,G,RP-bit) Prune message for
239.1.2.3
```

```
RP#
PIM(0): Received v2 Join/Prune on Ethernet0/2 from 10.4.7.4, to us
PIM(0): Join-list: (*, 239.1.2.3), RPT-bit set, WC-bit set, S-bit set
PIM(0): Check RP 10.0.0.4 into the (*, 239.1.2.3) entry
PIM(0): Adding register decap tunnel (Tunnel1) as accepting interface of (*, 239.1.2.3).
PIM(0): Add Ethernet0/2/10.4.7.4 to (*, 239.1.2.3), Forward state, by PIM *G Join
PIM(0): Prune-list: (10.5.6.6/32, 239.1.2.3) RPT-bit set
```

```
FHR#
PIM(0): Adding register encap tunnel (Tunnel0) as forwarding interface of (10.5.6.6, 239.1.2.3).
PIM(0): Received v2 Join/Prune on Ethernet0/1 from 10.4.5.4, to us
PIM(0): Join-list: (10.5.6.6/32, 239.1.2.3), S-bit set
PIM(0): Add Ethernet0/1/10.4.5.4 to (10.5.6.6, 239.1.2.3), Forward state, by PIM SG Join
PIM(0): Building Periodic (*,G) Join / (S,G,RP-bit) Prune message for 239.1.2.3
```



# Multicast – Basic functioning

## RP Selection:

- Static RP: We define an IP address for an RP in the network
- AutoRP: dynamic RP propagation using announce and discovery messages (224.0.1.39 and 224.0.40) to announce and RP to mapping agent that distributes RP information in the network; Cisco proprietary
- BSR: Dynamic RP discovery using PIM messages (224.0.0.13), standard protocol.
- Anycast: using the “same” ip address in several devices to improve scalability.
- MSDP: State communication between RP devices.

# Multicast – Basic functioning

- **LHR:**

```
Auto-RP(0): Received RP-discovery packet of length 48, from 1.1.1.1, RP_cnt 1, ht 181
(0): pim_add_prm:: 239.1.2.3/255.255.255.255, rp=10.0.0.4, repl = 1, ver =3, is_neg =0, bidir = 0, crp =
0
Auto-RP(0): Update
prm_rp->bidir_mode = 0 vs bidir = 0 (239.1.2.3/32, RP:10.0.0.4), PIMv2 v1
```

- **RP:**

```
Auto-RP(0): Build RP-Announce for 10.0.0.4, PIMv2/v1, ttl 5, ht 181
Auto-RP(0): Build announce entry for (239.1.2.3/32)
Auto-RP(0): Send RP-Announce packet of length 48 on Ethernet0/2
Auto-RP(0): Send RP-Announce packet of length 48 on Loopback0(*)
Auto-RP(0): Received RP-announce packet of length 54, from 10.0.0.4, RP_cnt 1, ht 181
pim_add_prm:: 239.1.2.3/255.255.255.255, rp=10.0.0.4, repl = 0, ver =3, is_neg =0, bidir = 0, crp = 0
Auto-RP(0): Update
prm_rp->bidir_mode = 0 vs bidir = 0 (239.1.2.3/32, RP:10.0.0.4), PIMv2 v1
Auto-RP(0): Build RP-Announce for 1.1.1.1, PIMv2/v1, ttl 5, ht 181
Auto-RP(0): Build announce entry for (224.0.1.39/32)
Auto-RP(0): Build announce entry for (224.0.1.40/32)
Auto-RP(0): Send RP-Announce packet of length 54 on Ethernet0/2
Auto-RP(0): Send RP-Announce packet of length 54 on Loopback1(*)
Auto-RP(0): Received RP-announce packet of length 54, from 1.1.1.1, RP_cnt 1, ht 181
```

# Multicast – Attack protection

## **Possible attacks from hosts:**

- AutoRP or BSR messages
- PIM hellos (changes for DR)
- PIM Joins
- Register or Register-stop messages (sending fake traffic)

## **Hosts should never run PIM**

## **Considerations:**

- Passive interfaces
- CoPP
- ACL
- Good routing design (set boundaries, stub devices, passive interfaces)
  
- `ip pim rp-address override` (when static RP)
- Autorp control
- Inbound ACL to filter IGMP queries or PIM messages\*

# Multicast – PIM Neighbor Control

```
•ip pim multicast-routing
•ip access-list 1 permit 10.0.0.2
•ip access-list 1 deny any
•interface ethernet0/0
• ip address 10.0.0.1 255.255.255.0
• ip pim sparse-mode
• ip pim neighbor-filter 1
```

```
•ip pim multicast-routing
•ip access-list 1 permit 10.0.0.1
•ip access-list 1 deny any
•interface ethernet0/0
• ip address 10.0.0.2 255.255.255.0
• ip pim sparse-mode
• ip pim neighbor-filter 1
```

## •MSDP Control:

```
•ip msdp peer 10.0.0.1
•ip msdp password peer 10.0.0.2 $ecret
```



# Multicast – AutoRp Control

```
•ip access-list 1 permit 10.0.0.4
•ip access-list 1 permit 10.0.0.3
•ip access-list 2 permit 224.0.0.0 15.255.255.255
•ip access-list 3 deny 224.0.1.39
•ip access-list 3 deny 224.0.1.40

•ip pim rp-announce-filter rp-list 1 group-list 2

•Interface ethernet0/0
•ip multicast boundary 3
•ip multicast boundary filter-autorp
•ip multicast boundary out (prevents to include interface in the outgoing list)
•ip multicast boundary in (interface can't create an (S,G)

•Auto-RP(0): Received RP-announce packet of length 48, from 1.1.1.1, RP_cnt 1, ht 181
•Auto-RP(0): Filtered 224.0.1.39/32 for RP 1.1.1.1
•Auto-RP(0): Filtered 224.0.1.40/32 for RP 1.1.1.1
•Auto-RP(0): Build RP-Announce for 10.0.0.4, PIMv2/v1, ttl 5, ht 181
•Auto-RP(0): Build announce entry for (239.1.2.3/32)
•Auto-RP(0): Send RP-Announce packet of length 48 on Ethernet0/2
•Auto-RP(0): Send RP-Announce packet of length 48 on Loopback0(*)
•Auto-RP(0): Received RP-announce packet of length 48, from 10.0.0.4, RP_cnt 1, ht 181
•pim_add_prm:: 239.1.2.3/255.255.255.255, rp=10.0.0.4, repl = 0, ver =3, is_neg =0, bidir = 0, crp = 0
•Auto-RP(0): Update
```

# Multicast – BSR Control

```
•interface ethernet0/0  
•ip pim bsr-border
```

- **BSR is PIM based, and it's forwarded hop by hop.**
- **Not possible to use an ACL.**
- **Blocks all BSR messages**



# Multicast – Access control

Any Source or Source Specific Multicast:

Groups to receive in an interface: `ip igmp access-group`

Usable Groups: `ip multicast group-range`

Control Plane scoping: `ip multicast boundary [acl|ext-acl]`

Packet level filtering: `ip access-group [in|out]`

Known Sources: `ip pim accept-registry`

`mroutes limit:`  
`[warning]` `ip multicast route-limit [limit]`

`mroutes limit per interface:` `ip multicast limit [rpf|out|connected]`

# References

The Multicast Security Tool Kit: <https://www.cisco.com/c/en/us/about/security-center/multicast-toolkit.html>

Configuring a Rendezvous Point :

[https://www.cisco.com/c/en/us/td/docs/ios/solutions\\_docs/ip\\_multicast/White\\_papers/rps.html](https://www.cisco.com/c/en/us/td/docs/ios/solutions_docs/ip_multicast/White_papers/rps.html)

IP Multicast Best Practices for Enterprise Customers: [https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/multicast-enterprise/whitepaper\\_c11-474791.html](https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/multicast-enterprise/whitepaper_c11-474791.html)

IP Multicast: PIM Configuration Guide: [https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipmulti\\_pim/configuration/imc-pim-15-mt-book.html](https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipmulti_pim/configuration/imc-pim-15-mt-book.html)

IP Multicast, Volume I: Cisco IP Multicast Networking: <http://www.ciscopress.com/store/ip-multicast-volume-i-cisco-ip-multicast-networking-9781587144592>

IP Multicast: IGMP Configuration Guide: [https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipmulti\\_igmp/configuration/xs-16/imc-igmp-xe-16-book/imc-customizing-igmp.html](https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipmulti_igmp/configuration/xs-16/imc-igmp-xe-16-book/imc-customizing-igmp.html)

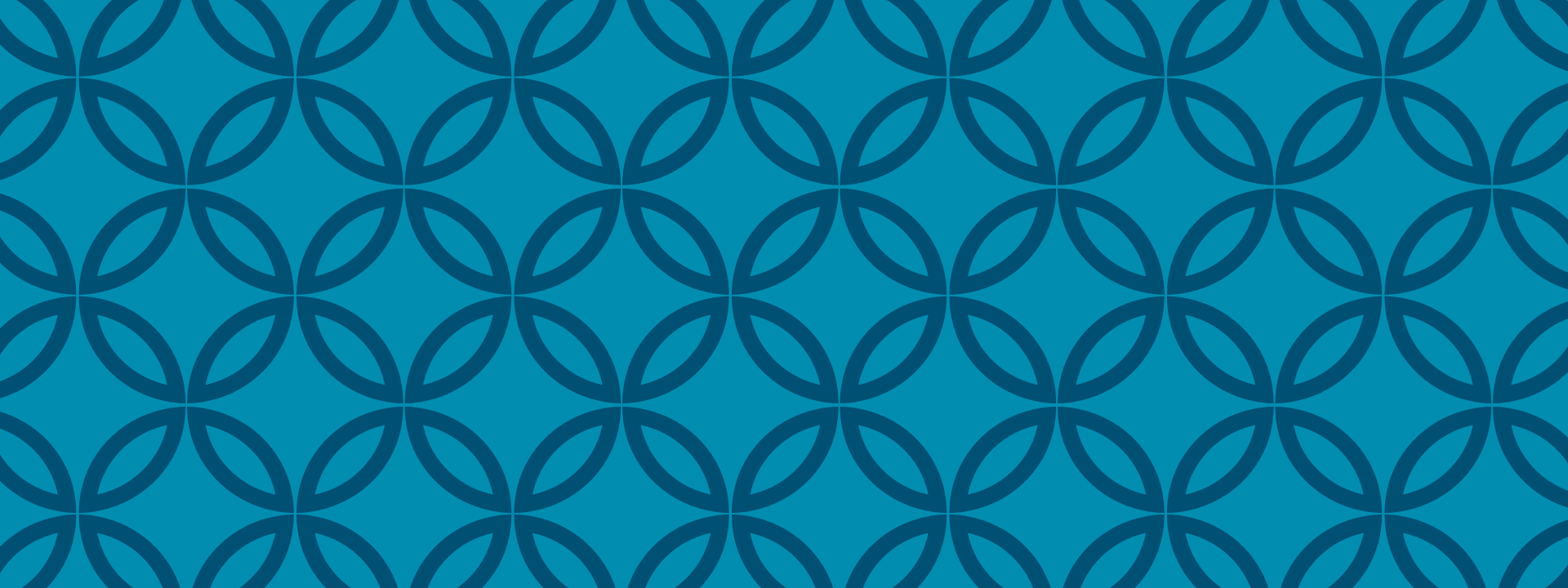
IP Multicast Troubleshooting Guide: <https://www.cisco.com/c/en/us/support/docs/ip/ip-multicast/16450-mcastguide0.html>

Basic Multicast Troubleshooting Tools: <https://www.cisco.com/c/en/us/support/docs/ip/ip-multicast/13726-57.html>

## Polling Question 2

What is the device used to distribute the multicast routes to devices in the network?

- A. Rendezvous Point
- B. Replicate Point
- C. Roundabout Point



# IP Multicast Best Practices |

# Using Point-to-Point Links in the Core

- A collapsed backbone should be avoided in the core of the network. A common layer 2 segment between routers introduces a number of unnecessary complexities and inefficiencies as described below
  - **a. Triggered events on link failure**
- When a router or a link fails in a P2P environment the carrier signal is dropped and creates a triggered event that will cause immediate IGP convergence, which will be followed by IP Multicast convergence.
- In a switched environment, a router can fail and it will not be detected until several hello messages are missing at a layer 3 protocol level. This will increase the convergence time.
- Using BFD may be able to minimize the effect on convergence time.

# Using Point-to-Point Links in the Core (Cont)

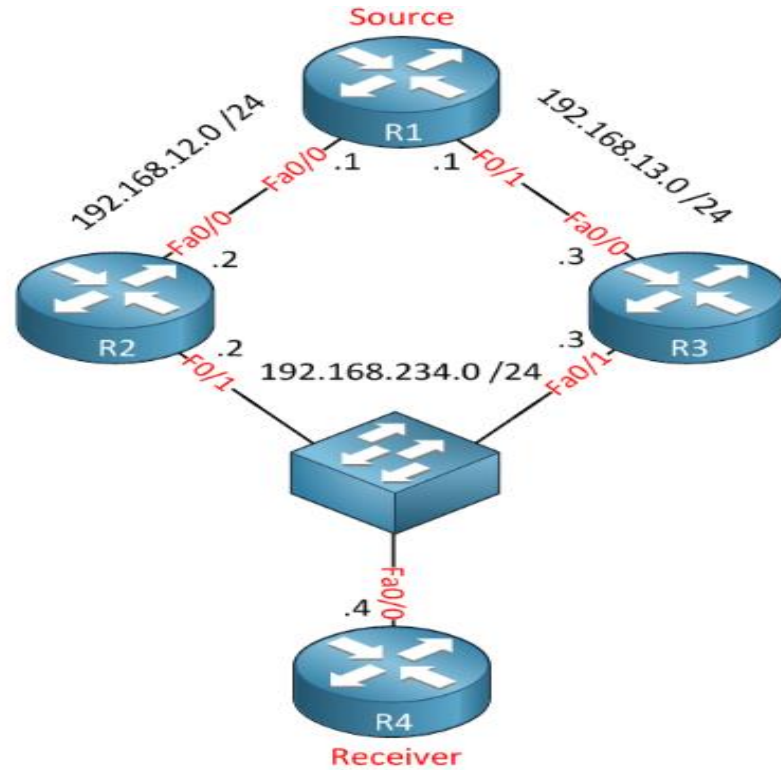
- **b. Avoid situations which require PIM snooping**

- In networks where a Layer 2 switch interconnects several routers, the switch floods IP Multicast packets to all multicast router ports by default, even if there are no multicast receivers downstream. In these environments, PIM snooping should be used to constrain the multicast to the interested routers.
- With PIM snooping enabled, the switch restricts multicast packets for each IP multicast group to only those multicast router ports that have downstream receivers joined to that group. When you enable PIM snooping, the switch learns which multicast router ports need to receive the multicast traffic within a specific VLAN by listening to the PIM hello messages, PIM join and prune messages, and bidirectional PIM designated forwarder-election messages.
- Point-to-point interfaces will avoid the additional complexity that requires PIM snooping.

- **c. Assert issues**

- The PIM Assert mechanism prevents duplicate traffic from flowing into a multi-access network. Assert messages are sent by the two forwarding routers and the router with the best metric will win the assert. There are several known cases in which assert can cause temporary routing loops and sub optimal behavior.
- Point-to-point interfaces will avoid assert issues with IP Multicast.

# PIM ASSERT



# PIM ASSERT

- when R1 starts streaming multicast traffic towards R2 and R3 they will both forward multicast packets to R4 resulting in duplicate traffic. To stop this PIM will **elect one PIM forwarder** for this segment. PIM doesn't have any routing information itself but relies on other routing protocols that are configured, it will use this information to select the best forwarding path with the PIM assert mechanism.
- Note: Don't confuse the PIM forwarder with the [PIM DR \(Designated Router\)](#). those are two different things!



# PIM Assert

- When R2 and R3 both forward multicast packets to the 192.168.234.0 /24 segment they will see each others multicast traffic, this will trigger the PIM assert mechanism. We will elect a PIM forwarder based on the following rules:
- The router with the **lowest administrative distance** to the source of the multicast stream will be the elected PIM forwarder. This only happens if you are using two routing protocols or if you used a static route pointing to the source.
- If the AD is equal we will compare the **unicast routing metric** towards the source.
- If the AD and metric are both the same we will elect the PIM forwarded based on the **highest IP address**.
- The elected PIM forwarder will keep forwarding traffic to the receiver while the loser will prune its interface.

# Tuning at Access Layer Edge

- a. Loop Free Layer 2
- Limit VLANs to a single closet whenever possible. There should be no STP loops - all interfaces should be in forwarding state-no interfaces in blocked state.
- There are many reasons why STP/RSTP convergence should be avoided for the most deterministic and highly available network topology. In general, when you avoid STP/RSTP, convergence can be predictable, bounded, and reliably tuned. Additionally, it should be noted that in soft failure conditions, where keepalives (BPDU or routing protocol hellos) are lost, L2 environments fail open, forwarding traffic with unknown destinations on all ports and causing potential broadcast storms; while L3 environments fail closed, dropping routing neighbor relationships, breaking connectivity, and isolating the soft failed devices.

# Tuning at Access Layer Edge

- Spanning-tree is a layer 2 protocol designed to prevent broadcast storms caused by loops in the layer 2 topology. There are a couple of potential interactions between spanning-tree and multicast that should be kept in mind.
- In many switches, when a switch receives spanning-tree Topology Change Notice (TCN), IGMP snooping will flood multicast traffic to all ports. This is based on the assumption that multicast traffic is critical and until the topology has converged multicast should be forwarded everywhere. If this is not desired or is causing a problem “TCN flooding” can usually be disabled.
- When a switch receives spanning-tree Topology Change Notice (TCN) it puts all its ports in listening and then learning mode during which they do not forward packets. Depending on the spanning-tree protocol used, this can be from 10-50 seconds. This may not be a huge issue for traditional TCP traffic, but for a real time UDP service it is a major disruption. On at least the streaming source ports on the switch, an immediate forwarding setting like Cisco’s PortFast or Juniper’s Edge, should be set to avoid service interruption.

# Tuning at Access Layer Edge

- **b. If STP is absolutely required, use Rapid PVST+**
- Older applications that require L2 connectivity between Data Center or L2 switches need to be updated and/or replaced. Example: Old Tibco middleware versions required the use of a L2 broadcast for a heartbeat. It has been a decade since that middleware version has been updated to use a L3 IP Multicast heartbeat.
- If you are compelled by application requirements to depend on STP to resolve convergence events, use Rapid PVST+. Rapid PVST+ is far superior to 802.1d and even PVST+ (802.1d plus Cisco enhancements) from a convergence perspective.

# Tuning at Access Layer Edge

- **Hardcode all interface settings**
- Hardcode duplex, speed and trunking capability on router and switch interfaces and then turn off auto-negotiation. This tuning can save seconds during re-convergence when restoring a failed link or node. Unused VLANs should be manually pruned from trunked interfaces to avoid broadcast propagation.
- Finally, VTP transparent mode should be used because the need for a shared common VLAN database is reduced.

# Tuning at Access Layer Edge

- **Traffic Storm Control**
- A traffic storm occurs when packets flood the LAN, creating excessive traffic and degrading network performance. The traffic storm control feature prevents LAN ports from being disrupted by a broadcast, multicast, or unicast traffic storm on physical interfaces.
- Traffic storm control (also called traffic suppression) monitors incoming traffic levels over a 1-second traffic storm control interval, and during the interval it compares the traffic level with the traffic storm control level that you configure. The traffic storm control level is a percentage of the total available bandwidth of the port. Each port has a single traffic storm control level that is used for all types of traffic (broadcast, multicast, and unicast).
- Traffic storm control monitors the level of each traffic type for which you enable traffic storm control in 1-second traffic storm control intervals.

# Tuning at Access Layer Edge

- **Traffic Storm Control**
- Because hardware traffic storm control uses a bandwidth-based method to measure traffic, the most significant implementation factor is setting the percentage of total available bandwidth that can be used by controlled traffic. Because packets do not arrive at uniform intervals, the 1-second interval during which controlled traffic activity is measured can affect the behaviour of traffic storm control.
- **Note:** When the storm control threshold for multicast traffic is reached, all multicast traffic except control traffic, such as bridge protocol data unit (BDPU) and Cisco Discovery Protocol (CDP) frames, are blocked. However, the switch does not differentiate between routing updates, such as OSPF, and regular multicast data traffic, so both types of traffic are blocked.

# Tuning at Access Layer Edge

- Storm Control Level Configurations

- Because of hardware limitations and the method by which packets of different sizes are counted, the level percentage is an approximation. Depending on the sizes of the frames making up the incoming traffic, the actual enforced level might differ from the configured level by several percentage points.
- All routers in a VLAN see copies of all broadcast traffic. To avoid high RP CPU utilization caused by a high volume of broadcast traffic, the threshold typically is set to a very low value; for example, less than 1 percent on a Gigabit Ethernet port (which would still represent 10Mb of broadcast traffic)
- The peak broadcast traffic that is acceptable for an access port is 5 percent. The peak multicast traffic that is acceptable for an access port is 30 percent.
- `interface <<interface_name>>`
- `storm-control broadcast level 5.00`
- `storm-control multicast level 30.00`



# IGP Tuning

- IP Multicast traffic will converge after unicast routing converges. Therefore it is important to minimize convergence on the edge by tuning IGP timers.
- **a. EIGRP**
  - Set hello and dead timers to 1 and 3:
    - interface GigabitEthernet1/0
    - ip hello-interval eigrp 100 1
    - ip hold-time eigrp 100 3
  - **b. OSPF**
    - Tune OSPF Fast hello, dead-interval, SPF and LSA throttle timers.
    - The example below sets the dead interval to 1 second and the hello interval to 250 ms.
      - interface GigabitEthernet1/0
      - ip ospf dead-interval minimal hello-multiplier 4

# IGP Tuning

- The OSPF Link-State Advertisement (LSA) Throttling feature provides a dynamic mechanism to slow down link-state advertisement (LSA) updates in OSPF during times of network instability. It also allows faster Open Shortest Path First (OSPF) convergence by providing LSA rate limiting in milliseconds.
- The SPF and LSA throttle timers should be tuned to these recommended settings.

<b>spf-start</b>	<b>10 ms</b>
<b>msecspf-hold</b>	<b>100 to 500 ms</b>
<b>msecspf-max-wait</b>	<b>5 seconds</b>
<b>lsa-start</b>	<b>10 ms</b>
<b>mseclsa-hold</b>	<b>100 to 500 ms</b>
<b>mseclsa-max-wait</b>	<b>5 seconds</b>
<b>lsa arrival</b>	<b>80 ms (less than lsa-hold of 100 ms)</b>

# IGP Tuning

- This is an example on setting those timers:

```
router ospf 100
```

```
timers throttle spf 10 100 5000
```

```
timers throttle lsa all 10 100 5000
```

```
timers lsa arrival 80
```

- All these timers must be set consistently on both sides of the link.

# IGMP Snooping

- IGMP snooping is an IP Multicast constraining mechanism that runs on a Layer 2 LAN switch. Without IGMP snooping enabled, all multicast traffic will be forwarded to all hosts connected to the switch. IGMP snooping will insure that only hosts that are interested in the data stream will receive it.
- Every Cisco switch supports IGMP snooping. IGMP snooping should always be enabled if you are running IP Multicast. Some platform and switch software combinations may not have IGMP snooping enabled by default. Make sure IGMP snooping is enabled before running any multicast streams.
- There are some situations in which network administrators would like to run multicast in a contained environment and not have it forwarded to the rest of the network. In those cases, PIM is not enabled on the routers and there is no IGMP querier elected.

# IGMP Snooping

- In order for IGMP Snooping to operate correctly there needs to be an IGMP Querier sending out periodic IGMP Queries, so that the receivers will respond and send out IGMP Membership reports. These reports control which switchports will receive the multicast traffic for a particular group.
- If PIM is not enabled on at least one router in the switch environment then one router or switch needs to be configured as the IGMP querier. This is accomplished with this interface command:
  - **ip igmp snooping querier**
- An alternative would be to configure PIM on the interface facing the switch environment. In this case, the igmp querier will not have to be explicitly configured.

# Choosing the Right Multicast Groups

- There are some basic rules that must be followed for selecting which IP Multicast address range to use.
- **a. Do not use x.0.0.x or x.128.0.x group addresses**
- Multicast addresses in the 224.0.0.x range are considered link local multicast addresses. They are used for protocol discovery and are flooded to every port. For example, OSPF uses 224.0.0.5 and 224.0.0.6 for neighbor and DR discovery.
- These addresses are reserved and will not be constrained by IGMP snooping. Do not use these addresses for an application.
- Further, since there is a 32:1 overlap of IP Multicast addresses to Ethernet MAC addresses, any multicast address in the [224-239].0.0.x and [224-239].128.0.x ranges should NOT be considered.
- **b. Use 239/8 addresses for internal applications**
- RFC 2365 describes the use of administratively scoped IP Multicast addresses. This address range should be used for all internal applications. The concept is similar to the use of RFC 1918 addresses for unicast.

# Choosing the Right Multicast Groups

- **c. Use 233 GLOP addresses for interdomain applications**
- RFC 3180 describes the use of GLOP addresses that can be used based on an AS number. Exchanges should be encouraged to use these addresses for interdomain multicast data streams.  
<https://tools.ietf.org/html/rfc3180>
- **d. Use PIM-SSM and 232/8 for interdomain one to many applications**
- RFC 4608 describes the use of the 232/8 address range for PIM-SSM interdomain applications. Exchanges and FSPs are encouraged to use PIM-SSM and the 232/8 address range for one-to-many unidirectional multicast data delivery.
- **e. Petition IANA for a 224 to use externally**
- The last choice for external addresses is to petition IANA for a 224 address range to use for your interdomain application. This should be considered a last resort for content providers such as stock exchanges that need to insure there will not be an address collision globally with any provider or customer. This address space is extremely limited but many of the largest exchanges have successfully been assigned 224 address ranges.

# PIM Query-Interval Tuning

- The 'ip pim query-interval' controls the interval that a PIM hello packet is transmitted out each pim enabled interface.
- The PIM hello packets are used to discover PIM neighbors and to determine the Designated Router (DR) on each network segment. The default interval for the PIM hello packets to be sent is 30 seconds. A PIM neighbor is considered down after 3 consecutive missed messages. Therefore, it could take 90 seconds for the DR to failover. If you lower the query interval to 1 second, then the DR failover time is reduced to 3 seconds.
- The goal is not to set the query-interval too low so that there is unnecessary flapping. Cisco generally recommends a 1 second query-interval, which would give you a 3 second failover at the receiver edge. Some customers may choose to use the sub-second option. Cisco does not recommend an interval less than 500 ms. Due to queue lengths and processing delays on the switch platforms, lower intervals have been known to cause problems



# PIM Query-Interval Tuning

- In summary:
  - Turn down the pim query-interval on the receiver edge to reduce DR failover time
  - This only needs to be done when there are redundant edge routers and receivers
  - A general recommendation is a query interval of 1 second and no less than 500ms. This should be used with care as the number of interfaces increase

# Register Rate Limits

- When a new source starts transmitting in a PIM Sparse Mode network, the packets will be encapsulated and sent using unicast to the Rendezvous Point (RP). This process can be taxing on the CPU of the Designated Router (DR) and the RP if the source is running at a high data rate and/or there are many new sources starting at the same time. This scenario can potentially occur immediately after a network failover.
- In order to protect both the edge routers and the RP, it is recommended to set the 'ip pim register-rate limit' to a relatively low value. Normally, there is no limit to the number of packets that will be encapsulated and sent to the RP.
- Use this command to limit the number of register messages that the Designated Router (DR) will allow for each (S, G) entry. Enabling this command will limit the load on the DR and the Rendezvous Point (RP) at the expense of dropping register messages that exceed the set limit. Receivers may experience data packet loss in the first seconds in which register messages are sent from bursty sources.
- The number to limit the register packets will depend on the number of potential sources registering at the same time and their data rate. A typical setting in a PIM Sparse Mode (PIM-SM) network is between 4 and 10 messages per second.
- **ip pim register-rate-limit rate**

# Static RP vs. AutoRP Listener

- The main tradeoff between using static RP configuration and AutoRP is administrative overhead.
- **a. Static RP**
  - An RP could be statically defined with as little as 1 line on each router. If the network does not have many different RPs defined and/or they don't change very often this could be an attractive option.
  - The override option can be used with the rp-address configuration for additional security. This option will cause the router to ignore any AutoRP or BSR announcements that conflict with the statically defined RP.
  - Sample config:
    - ip pim rp-address 1.1.1.1 1
    - access-list 1 permit 239.254.1.0 0.0.0.15

# Static RP vs. AutoRP Listener

- b. AutoRP with AutoRP Listener
- Previously, sparse-dense mode was required on the interfaces to run AutoRP. Today, sparse-mode is configured on the interfaces and the autorp listener option is configured globally.
- On every router:
  - ip pim autorp listener
  - interface GigabitEthernet3/40
    - ip address 126.1.3.11 255.255.255.0
    - ip pim sparse-mode

# Static RP vs. AutoRP Listener

- On the RP routers:

```
ip pim send-rp-announce Loopback0 scope 16 group-list 7
```

```
ip pim send-rp-discovery Loopback1 scope 16
```

```
access-list 7 permit 239.254.2.0 0.0.0.255
```

```
interface Loopback0
```

```
ip address 126.0.4.1 255.255.255.255
```

```
ip pim sparse-mode
```

```
interface Loopback1
```

```
ip address 126.0.1.15 255.255.255.255
```

```
ip pim sparse-mode
```

- This example is advertising the Anycast RP address of 126.0.4.1 and the AutoRP announcement messages are being sent with a source address from Loopback 1.
- It is recommended that with either Static RP or AutoRP Listener you also have RP redundancy with Anycast RP or the Phantom RP.

# Anycast RP & Phantom RP

- **Anycast RP for PIM-SM**

- The RP is a critical function for PIM-SM and PIM-Bidir deployments.

RP redundancy is always recommended. The best form of redundancy for PIM-SM is Anycast RP which is described in the document:

- Anycast RP:

[http://www.cisco.com/en/US/docs/ios/solutions\\_docs/ip\\_multicast/White\\_papers/anycast.html](http://www.cisco.com/en/US/docs/ios/solutions_docs/ip_multicast/White_papers/anycast.html)

- **Phantom RP for PIM-Bidir**

- The RP is a critical function for PIM-SM and PIM-Bidir deployments.

RP redundancy is always recommended. The best form of redundancy for PIM-Bidir is the Phantom RP which is described in the document:

- Bidirectional PIM Deployment Guide : <http://www.cisco.com/en/US/prod/collateral>

# MSDP Timers

- In PIM-SM deployments that use MSDP, there may be situations in which faster convergence of the Source Active (SA) messages is desired. A typical scenario is when the MSDP session is reset and new sources start up during the time the session is being reestablished. Potentially it may take as long as one minute for the new traffic stream to start forwarding.
- For these situations, you may want to consider adjusting the MSDP timers down to as low as 5 seconds:

```
ip msdp keepalive <peer-name-or-address> 5 15
```

```
ip msdp timer 5
```

- **Note:** The source information in the SA Cache will remain active for as long as 6 minutes. Modifying these times will only apply to new sources that start up during the time that the MSDP session is down. As with any timer settings, there is a tradeoff between higher CPU utilization and network convergence.

# Intermittent Sources

- A common issue with market data applications is servers that send data to a multicast group and then go silent for more than 3.5 minutes. These intermittent sources may cause thrashing of state on the network and can introduce packet loss during the window of time when soft state exists, and then hardware shortcuts are being created.
- There are a few scenarios in which the outage can be more severe. One case would be if the source starts sending again right around the 3.5 minute mark. At that point, state has started to time out in some of the routers along the data path and there may be inconsistent state in the network. This could create a situation in which data from the source would be dropped for as long as a minute until state clears out and then is created again on the intermediate routers.



# Intermittent Sources

These are the best solutions to deal with intermittent sources.

## a. PIM-Bidir or PIM-SSM

- The first and best solution for intermittent sources is to use PIM-Bidir for many-to-many applications and PIM-SSM for one-to-many applications.
- Both of these optimizations of the PIM protocol do not have any data driven events in creating forwarding state. That means that as long as the receivers are subscribed to the streams, the network will have the forwarding state created in the hardware switching path.
- Intermittent sources are not an issue with PIM-Bidir and PIM-SSM.

## b. Null packets

- In PIM-SM environments, a common method to make sure forwarding state is created is to send a burst of null packets to the multicast group before the actual data stream. The application needs to efficiently ignore these null data packets to make sure it doesn't affect performance. The sources would only need to send the burst of packets if they have been silent for more than 3 minutes. A good practice would be to send the burst if the source was silent for more than a minute.
- Many financial applications send out an initial burst of traffic in the morning and then all well behaved sources will not have a problem.

# Intermittent Sources

## c. Periodic keepalives or heartbeats

An alternative approach for PIM-SM environments is for sources to send periodic heartbeat messages to the multicast groups. This is a similar approach to the null packets, but the packets can be sent on a regular timer so that the forwarding state will never expire. A typical timer for the heartbeat message is 60 seconds.

## d. S,G expiry timer

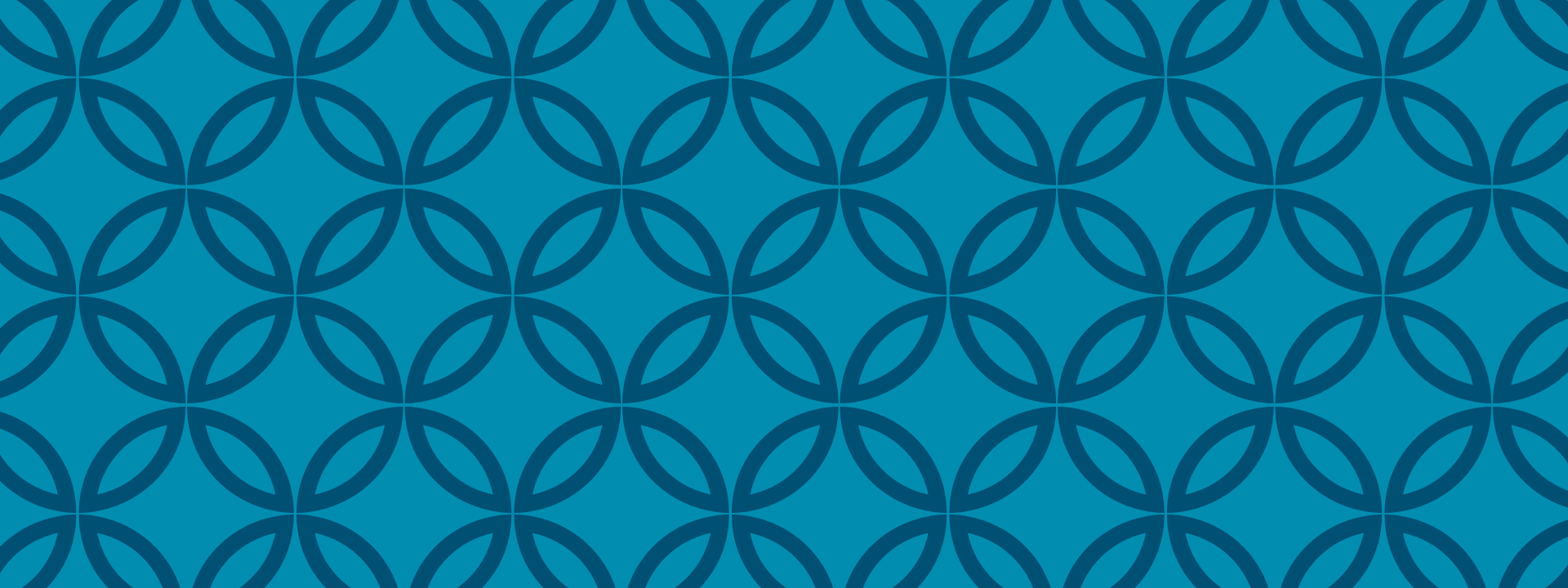
Finally, Cisco has made a modification to the operation of the S, G expiry timer in IOS. There is now a CLI knob to allow the state for a S, G to stay alive for hours without any traffic being sent. This fix was in response to a customer request in a PIM-SM environment to maintain the state and not fall back to \*, G forwarding. The command is "ip pim sparse sg-expiry-timer" and is documented in the command reference:

- [http://www.cisco.com/en/US/docs/ios/ipmulti/command/reference/imc\\_04.html#wp1018443](http://www.cisco.com/en/US/docs/ios/ipmulti/command/reference/imc_04.html#wp1018443)
- This approach should be considered a workaround until PIM-Bidir or PIM-SSM is deployed or the app is fixed.

## Polling Question 3

What applications does multicast has?

- A. Voice
- B. Video
- C. Data distribution
- D. Data monitoring
- E. All of the above



# Demonstration

Luis Espejel  
David Ramirez

# References

- Guidelines for Enterprise IP Multicast Address Allocation:  
[http://www.cisco.com/en/US/tech/tk828/technologies\\_white\\_paper09186a00802d4643.shtml](http://www.cisco.com/en/US/tech/tk828/technologies_white_paper09186a00802d4643.shtml)
- IP Multicast Best Practices for Enterprise Customers
- [https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/multicast-enterprise/whitepaper\\_c11-474791.html](https://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/multicast-enterprise/whitepaper_c11-474791.html)
- IP Multicast: Multicast Optimization Configuration Guide
- [https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipmulti\\_optim/configuration/12-4t/imc-optim-12-4t-book/imc\\_pim\\_sparse.html](https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipmulti_optim/configuration/12-4t/imc-optim-12-4t-book/imc_pim_sparse.html)

Submit Your  
Questions Now!



Use the Q&A panel to submit your  
questions, our expert will respond

# Ask the Expert following the Webcast

Now through Friday November 3<sup>rd</sup>

With Luis Espejel  
& Miguel Perez

<http://bit.ly/ATE-Muticast>



Luis Espejel



& Miguel Perez

# Collaborate within our Social Media



## Twitter

- @Cisco\_Support
- <http://bit.ly/csc-twitter>

## Facebook

- Cisco Support Community
- <http://bit.ly/csc-facebook>

Learn About Upcoming Events



# Lo invitamos a nuestros próximos eventos en Redes Sociales

## Google+

- Description, details
- <http://bit.ly/csc-googleplus>



## App

- Cisco Technical Support



## LinkedIn

- CSC-Cisco-Support-Community
- <http://bit.ly/csc-linked-in>



## YouTube

- Ciscosupportchannel
- <http://bit.ly/csc-youtube>



# Cisco has support communities in other languages!

If you speak Spanish, Portuguese, Japanese, Russian or Chinese we invite you to participate & collaborate



[Comunidad de Soporte  
De Cisco](#)  
Spanish

[Comunidade de  
Suporte de Cisco](#)  
Portuguese

[思科服务支持社区](#)  
Chinese

[Сообщество  
Технической Поддержки  
Cisco](#)  
Russian

[ツスコサポートコミュ  
ニティ](#)  
Japanese



# More IT Training Videos and Technical Seminars on the Cisco Learning Network

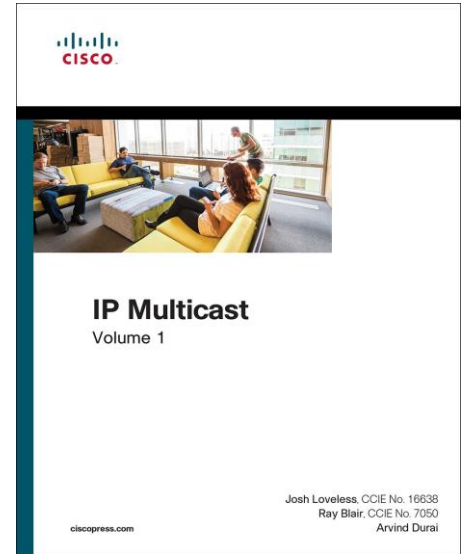
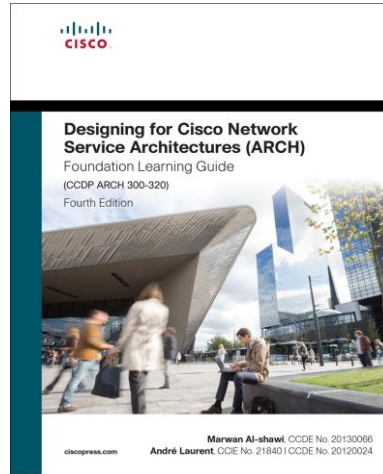
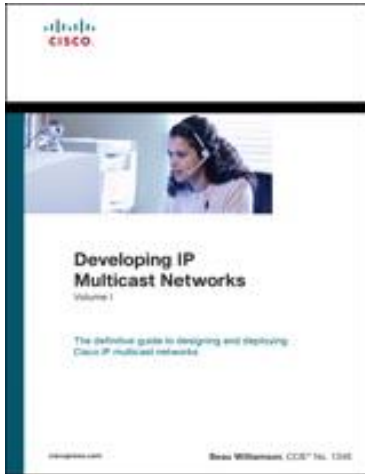
View Upcoming Sessions Schedule

<https://cisco.com/go/techseminars>

# Thank you for participating, you earned a discount!

Redeem your 35% discount offer by entering code: CSC when checking out.

<http://bit.ly/CSC-CiscoPress-2017>





Thank you for Your  
Time!

Please take a moment to complete the  
survey



*Thanks For Joining today!*

