VRRP−Aware PIM with PIM NonDR Join Feature Configuration Example

Document ID: 118859

Contributed by Mohammed Muddasir Khan, Cisco TAC Engineer.
Mar 26, 2015

Contents

Introduction

This document describes how to configure a router in order to use Virtual Router Redundancy Protocol−aware (VRRP−aware) Protocol Independent Multicast (PIM).

Prerequisites

Requirements

Cisco recommends that you have knowledge of multicast and VRRP features.

Components Used

This document is not restricted to specific software and hardware versions.

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If your network is live, make sure that you understand the potential impact of any command.

Background Information

VRRP−aware PIM is supported in VRRP Version 3.10 Release (15.3(3)S). PIM has no inherent redundancy capability and its operation is completely independent of First Hop Redundancy Protocols (FHRP) such as VRRP. As a result, IP multicast traffic is not necessarily forwarded by the same router that is elected by VRRP.

There is a need to provide consistent IP multicast forwarding in redundant networks with Virtual Router Groups (VRGs) enabled. With PIM redundancy, Virtual Router Redundancy Service (VRRS) is leveraged and the Designated Router (DR) election and PIM Join/Prune processing decisions are made based on the VRRP states in the router. When you enable the ***PIM non−DR join*** feature, it allows the non−DR (NonDR) to create Multicast Route (mroute) states and pull traffic, but not forward the traffic. When a VRRP failover occurs, the new Master Router (MR) elected by the VRRP group takes over the First Hop Router (FHR) or Last Hop Router (LHR) DR responsibilities and begin to forward traffic.

New Interface Features

Cisco has introduced a new feature that is enabled with the ***ip pim non−dr−join*** CLI command. This new feature works independently of the VRRP−aware PIM feature, and it can be useful with other features, such as Bidirectional Forwarding Detection (BFD), in addition to VRRP. This CLI feature, once enabled, allows the NonDR to process Internet Group Management Protocol (IGMP) joins and function just as the DR, with these exceptions:

- The NonDR keeps the interfaces in the Outgoing Interface List (OIL), but it does not set the *F* flag (the *forward* flag in the Multicast Routing Information Base (MRIB)) so that traffic is not forwarded. When the NonDR becomes the DR, it sets the *F* flag and begins to forward traffic.

  Note: This logic works completely independently of VRRP groups states.
- If both the ***ip pim non−dr−join*** and ***ip pim redundancy <tag> vrrp dr−priority <value>*** features are enabled on an interface, traffic is pulled at all NonDRs as well, regardless of the VRRP state. PIM sets or clears the *F* flag on the interface based on the VRRP state, which allows fast convergence time upon VRRP switchover.

IP Pim Redundancy

The configuration that is described in this document makes use of the new interface CLI feature in order to bind PIM to a VRRS session via a tag (48 character string):

```
ip pim redundancy <tag> [vrrp ] dr−priority <value>

ipv6 pim redundancy <tag> [vrrp ] dr−priority <value>
```

PIM registers as a VRRS client and listens to the VRRP events notifications. In order to designate the VRRP MR as the PIM DR on a multi−access segment, increase the PIM DR priority in the *Hello* message that is sent from the physical IP address.

Once VRRP−aware PIM tracking is enabled on an interface, different behaviors might be observed, dependent upon whether the ***ip pim non−dr−join*** feature is enabled on the same interface:

- If the ***ip pim non−dr−join*** feature is enabled, the NonDRs process the IGMP reports and create mroute states as usual. Different from the default NonDR behavior, NonDRs add interfaces to the outbound interface list of the mroute entry, send PIM Join/Prune decisions upstream, and pull traffic just like the DR. However, NonDRs do not set the *F* flag on the interfaces in the MRIB, so traffic is

not forwarded from the interface. Instead, a new *b* flag (*blocked*) is set for the interface in the outbound interface list (OIL) of the MRIB, which indicates that forwarding is blocked on this interface (if in the VRRP Backup state). This allows fast convergence time upon switchover, at the cost of bandwidth.

- If the ***ip pim non−dr−join*** feature is not enabled, then only the MR functions as the PIM DR and processes the PIM Join/Prune decisions while all of the backup routers ignore the IGMP Join and PIM Join/Prune requests. Upon switchover, the new MR sends the PIM Hello message with a virtual IP address. Hosts or downstream boxes are then triggered to re−send Join requests, so the new MR processes these requests and pulls the multicast traffic. This leads to a slower convergence time than the other approach, but it is more bandwidth−economic from a system point of view.

Since the only application setup of interest is the Last/First hop scenario, PIM is only allowed to track one VRRP group per interface. You cannot configure one interface in order to track multiple VRRP groups, which would create a situation where one interface is in the Master state for one VRRP group, and in the Backup state for another VRRP group.

Upon VRRP failover, the router that has become the new MR is elected as the new DR:

- If the ***ip pim non−dr−join*** feature is enabled, PIM walks all mroute entries, clears the *b* flag, and sets the *F* flag on the interfaces (since it is now the MR of the VRRP group). The previous MR clears the *F* flag and sets the *b* flag on the interfaces if it enters the Backup state.

- If the ***ip pim non−dr−join*** feature is not enabled, then the Hot Standby Router Protocol−aware (HSRP−aware) PIM logic is followed, the new MR sends the PIM Hello message with the new GenID in order to trigger the downstream boxes to re−send the PIM Join requests (or waits for the hosts to send the next periodic IGMP reports), recreates the mroute states, and pulls the traffic via the new DR.

- Traffic is now forwarded through the new MR (and PIM DR) to the LAN, and there is no operation required on the downstream routers at all upon failover.

Role of VRRP

VRRP specifies an election protocol that dynamically assigns responsibility for a virtual router that is represented by an IPv4/IPv6 address to one of the VRRP routers on a LAN (RFC5798). The VRRP router that controls the address(es) associated with a virtual router is called the Master, and it forwards the packets that are sent to a virtual Media Access Control (MAC) address.

When this new feature is implemented, VRRP is used in order to elect the VRRP MR. The VRRP MR performs the routing and forwarding for all of the traffic that is addressed to the VRRP group Virtual IP (VIP). This achieves three goals:

- It notifies VRRS about all VRRP server state change and updates.

- It allows all of the PIM Join/Prune requests to reach the VRRP group VIP, which minimizes changes and configurations at the downstream router side (they need to know the VIP only).

- It allows the PIM DR to run on the same gateway as the VRRP MR and maintain mroute states. Multicast traffic is forwarded trough the VRRP MR, and PIM can leverage VRRP redundancy in order to avoid potential duplicate traffic and cause failover to become enabled.

Role of PIM

PIM acts as a VRRS client, listens to state change and update notifications from the VRRS server (VRRP), and:

- Automatically adjusts the PIM DR priority based on the VRRP state.
- Receives state change notifications from VRRS for the tracked VRRP group(s) upon VRRP failover. In response, PIM manages the interface flags and ensures that traffic is forwarded through the VRRP MR.

Since the mroute states and traffic are available on both the Master and Backup routers, the switchover time is mostly decided by the redundancy infrastructure (VRRP and VRRS) as well as the setup scale (such as the number of mroute entries). Upon notification of state change, PIM immediately notifies the MRIB and the Multicast Forwarding Information Base (MFIB) to forward traffic through the VRRP MR.

Implementation Details

This section presents some important notes about the configuration that is described in this document.

Bind PIM with a VRRP Group

The PIM CLI command is introduced in order to enable PIM redundancy on an interface and bind it to a VRRS server group (VRRP group):

ip pim redundancy <tag> [vrrp] dr−priority <value>

ipv6 pim redundancy <tag> [vrrp] dr−priority <value>

When configured on an interface, PIM registers with the VRRS as a client and obtains a client ID that is assigned by the VRRS database. It also requests that VRRS send notifications to PIM for all of the events for the group that is identified by a ***<tag>***.

Note: The VRRS servers and clients bind with a name (48 character string), which is called a *Tag*. VRRS works via a registration and callback mechanism. Clients (such as PIM) that implement redundancy register with the VRRS.

Enter one of these commands into the CLI in order to enable the NonDR Join feature:

ip pim non−dr−join

ipv6 pim non−dr−join

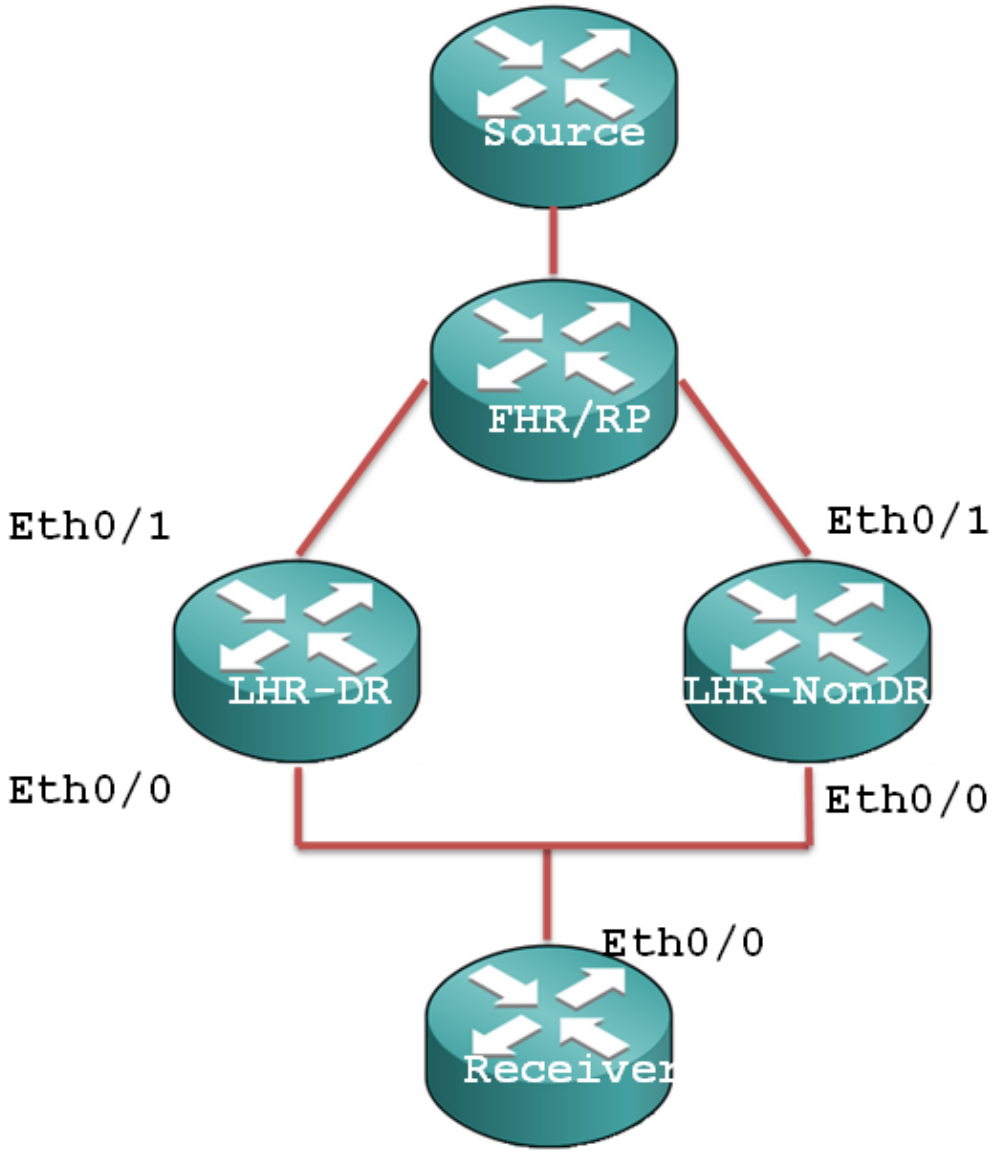
Track Multiple VRRP Groups on an Interface

As the target application scenario is only First/Last hop setup, the most common setup is where all LHR interfaces on the LAN track the same VRRP group. Therefore, PIM is only allowed to track one VRRP group per interface, even if you can enable VRRS to track multiple tags per interface.

Note: By default, the feature is disabled.

Configure

Network Diagram



Enable the PIM Redundancy Feature

Note: There is only one CLI command that you can use in order to enable PIM redundancy. You can use the current ***show*** and ***debug*** commands for PIM and HSRP.

Enter one of these commands into the CLI in order to enable the PIM redundancy feature and specify the PIM DR priority for each VRRP group:

```
[no] ip pim redundancy <tag> [hsrp | vrrp] dr-priority <value>
```

```
[no] ipv6 pim redundancy <tag> [hsrp | vrrp] dr-priority <value>
```

Enter one of these commands into the CLI in order to enable PIM DR functionalities (except forwarding on NonDRs):

```
[no] ip pim non-dr-join
```

```
[no] ipv6 pim non-dr-join
```

LHR Configurations

Use this configuration for LHR DRs:

```
interface Ethernet0/0
ip address 10.10.10.1 255.255.255.0
ip pim redundancy VRRP vrrp dr-priority 150
ip pim non-dr-join
ip pim sparse-mode
vrrp 1 address-family ipv4
  vrrs leader VRRP
  priority 120
  track 1 decrement 30
  address 10.10.10.5 primary
exit-vrrp   !   track 1 interface Ethernet0/1 line-protocol
```

Use this configuration for LHR NonDRs:

```
interface Ethernet0/0
ip address 10.10.10.2 255.255.255.0
ip pim redundancy VRRP vrrp dr-priority 150
ip pim non-dr-join
ip pim sparse-mode
vrrp 1 address-family ipv4
  address 10.10.10.5 primary
exit-vrrp
```

Enter the ***show vrrp brief*** command in order to view the LHR configuration:

```
LHR-DR#show vrrp brief
Interface         Grp A-F Pri Time Own Pre State   Master addr/Group addr
Et0/0               1 IPv4 120     0 N   Y MASTER 10.10.10.1(local) 10.10.10.5
LHR-DR#

LHR-NonDR#show vrrp brief
Interface         Grp A-F Pri Time Own Pre State   Master addr/Group addr
Et0/0               1 IPv4 100 3609 N   Y BACKUP 10.10.10.1 10.10.10.5
LHR-NonDR#
```

Verify

Use the information that is described in this section in order to verify that your configuration works properly.

Verify VRRS Database Information

Enter the ***show vrrs server VRRP*** command into the CLI in order to verify that the VRRS database is populated per the previous configuration:

```
LHR-DR#show vrrs server VRRP

Server Name: vrrpEthernet0/0v41
  Address Family: IPv4
  Interface: Ethernet0/0
  State: ACTIVE
  vMAC: 0000.5E00.0101
  vIP Address: 10.10.10.5
  Tags Connected:
    Tag Name VRRP
LHR-DR#
```

```
LHR-NonDR#show vrrs server VRRP

Server Name: vrrpEthernet0/0v41
  Address Family: IPv4
  Interface: Ethernet0/0
  State: BACKUP
  vMAC: 0000.5E00.0101
  vIP Address: 10.10.10.5
  Tags Connected:
LHR-NonDR#
```

Verify Interface Information

Enter one of these commands in order to verify that the interfaces are correctly programmed for the ***non-dr-join*** feature and that the NonDR has the tree built with a blocked flag:

```
LHR-DR#show ip pim int e0/0 det | i Non|DR
   PIM DR: 10.10.10.1 (this system)
   PIM Non-DR-Join: TRUE

LHR-NonDR#show ip pim int e0/0 det | i Non|DR
   PIM DR: 10.10.10.1
   PIM Non-DR-Join: TRUE
LHR-NonDR#
```

Enter the ***show ip mroute sparse*** command into the LHR-NonDR CLI in order to view the new *Blocked* field:

```
LHR-NonDR#show ip mroute sparse
(*, 239.1.1.1), 01:26:15/stopped, RP 192.168.1.254, flags: SJC
Incoming interface: Ethernet0/1, RPF nbr 192.168.2.2
Outgoing interface list:
   Ethernet0/0, Forward/Sparse, 00:00:16/00:02:43 Blocked

(192.168.7.2, 239.1.1.1), 00:11:56/00:02:50, flags: T
Incoming interface: Ethernet0/1, RPF nbr 192.168.2.2
Outgoing interface list:
   Ethernet0/0, Forward/Sparse, 00:00:16/00:02:43 Blocked
```

Enter the ***show mrib route*** command into the CLI of the LHR-NonDR in order to verify that the MRIB route does NOT have *F* flag set:

```
LHR-NonDR#show ip mrib route 239.1.1.1 | b \(
(*,239.1.1.1) RPF nbr: 192.168.2.2 Flags: C
Ethernet0/1 Flags: A NS

(192.168.7.2,239.1.1.1) RPF nbr: 192.168.2.2 Flags:
Ethernet0/1 Flags: A
```

As desired, the MRIB route does have the *F* flag set on the LHR-DR:

```
LHR-DR#show ip mrib route 239.1.1.1 | b \(
(*,239.1.1.1) RPF nbr: 192.168.3.2 Flags: C
Ethernet0/0 Flags: F NS
Ethernet0/1 Flags: A NS

(192.168.7.2,239.1.1.1) RPF nbr: 192.168.3.2 Flags:
Ethernet0/1 Flags: A
Ethernet0/0 Flags: F NS
```

Enter the ***conf t*** command into the CLI of the LHR-DR in order to trigger a VRRP state change via Ethernet0/1 shutdown:

```
LHR−DR#conf t
Enter configuration commands, one per line. End with CNTL/Z.
LHR−DR(config)#int e0/1
LHR−DR(config−if)#shutdown
LHR−DR(config−if)#end
```

When you observe the outputs from the LHR−NonDR, you can see that the VRRP status has changed (which is informed to VRRS) and that PIM takes the notification from VRRS and changes the DR role accordingly:

```
LHR−NonDR#show ip pim int e0/0 det | i DR
   PIM DR: 10.10.10.2 (this system)
   PIM Non−DR−Join: TRUE
LHR−NonDR#

LHR−NonDR# show vrrp brief
Interface         Grp A−F Pri Time Own Pre State   Master addr/Group addr
Et0/0               1 IPv4 100     0 N   Y MASTER 10.10.10.2(local) 10.10.10.5

LHR−NonDR# show vrrs server VRRP

Server Name: vrrpEthernet0/0v41
Address Family: IPv4
Interface: Ethernet0/0
State: ACTIVE
vMAC: 0000.5E00.0101
vIP Address: 10.10.10.5
Tags Connected:
```

As expected, the *F* flag is set and the NonDR begins to forward the Multicast traffic without the need to build a fresh multicast tree:

```
LHR−NonDR# show ip mrib route 239.1.1.1 | b \(
(*,239.1.1.1) RPF nbr: 192.168.2.2 Flags: C
Ethernet0/0 Flags: F NS
Ethernet0/1 Flags: A NS

(192.168.7.2,239.1.1.1) RPF nbr: 192.168.2.2 Flags:
Ethernet0/0 Flags: F NS
Ethernet0/1 Flags: A
```

Troubleshoot

Two packets were lost in the transaction from the previous section. You can verify this on the source router:

```
Source#ping 239.1.1.1 rep 1000
Type escape sequence to abort.
Sending 1000, 100−byte ICMP Echos to 239.1.1.1, timeout is 2 seconds:
Reply to request 0 from 10.10.10.3, 2 ms
Reply to request 1 from 10.10.10.3, 2 ms
Reply to request 2 from 10.10.10.3, 1 ms..
Reply to request 5 from 10.10.10.3, 1 ms
```

Deployments that run on a High Availability (HA) Multicast design require a standby tree formation on NonDRs and can benefit from the ***non−dr−join*** feature. This feature pulls the multicast traffic but does not forward it until it is elected as the DR.

Related Information

- ***VRRPv3 Protocol Support***
- ***Technical Support & Documentation − Cisco Systems***