PfR Enterprise WAN Application Control

Navigation

- Go to PfR home page
- Go to PfR Solution Guides home page

Contents

- 1 Version Used - PfR Version 2
- 2 Intelligent WAN (IWAN)
- 3 Enterprise Needs
  - 3.1 Presentation
  - 3.2 Traffic Classification and Overall Policies
- 4 PfR Network Topology Used
  - 4.1 Overview
  - 4.2 Transport and Overlay Backbones
  - 4.3 Routers and Servers used
  - 4.4 Traffic Generation
- 5 Dual DMVPN Setup
  - 5.1 DMVPN Phase Summary
  - 5.2 Front Door VRF
  - 5.3 DMVPN Configuration on the Hub
  - 5.4 DMVPN Configuration on the Spokes
  - 5.5 Routing on the Overlay Backbone
  - 5.6 Check Routing
- 6 Checking flows
- 7 PfR Configuration
  - 7.1 Presentation
  - 7.2 Provisioning

- 7.3 Enabling PfR Domain and Target Discovery
- 7.4 Learning Configuration
- 7.5 Policy Configuration
  - 7.5.1 Monitoring Modes
  - 7.5.2 Defining Advanced Policies per Group
  - 7.5.3 Configuration
- 8 Check Master Controllers
  - 8.1 Check Status
  - 8.2 Check Target Discovery and Peering
  - 8.3 Verify Traffic Class and Statistics
- 9 Check Border Routers (BRs)
  - 9.1 Active Probing
  - 9.2 Path Enforcement
- 10 More Information
- 11 Conclusion

Version Used - PfR Version 2

IOS 15.2(3)T and IOS-XE 3.6 releases are key milestones for PfR. Defaults have changed to better align with the customer deployments and the PfR Best practices.
Among the most important changes are:

- Mode route control is on by default
- Resolvers range and utilization have been removed. Load-balancing is used as the last resolver and cannot be disabled.
- Automatic learning is enabled by default
- Learning monitor-period is 1, and periodic-interval is 0 by default

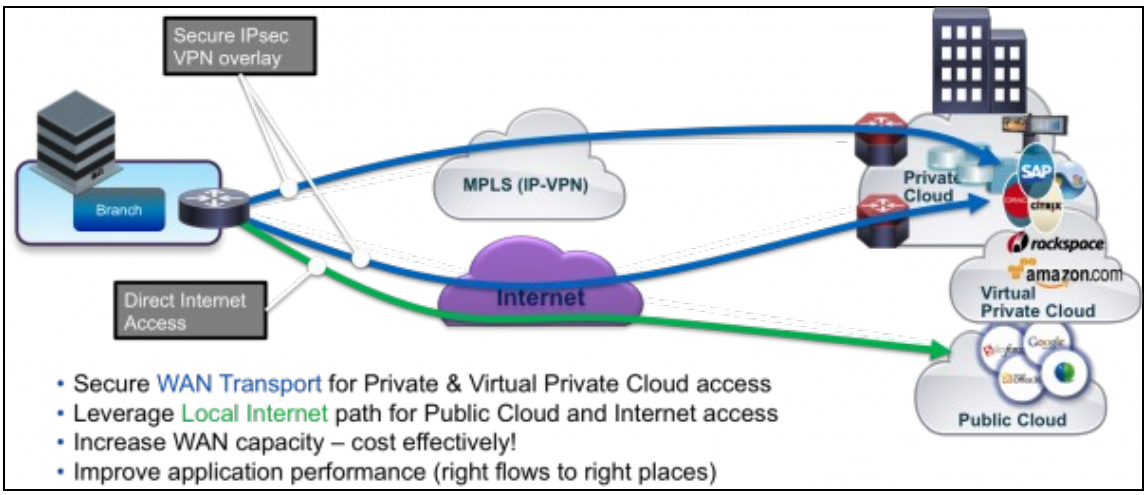
See PfR Simplification for more information.

This guide is based on IOS 15.3(3)M.

Intelligent WAN (IWAN)

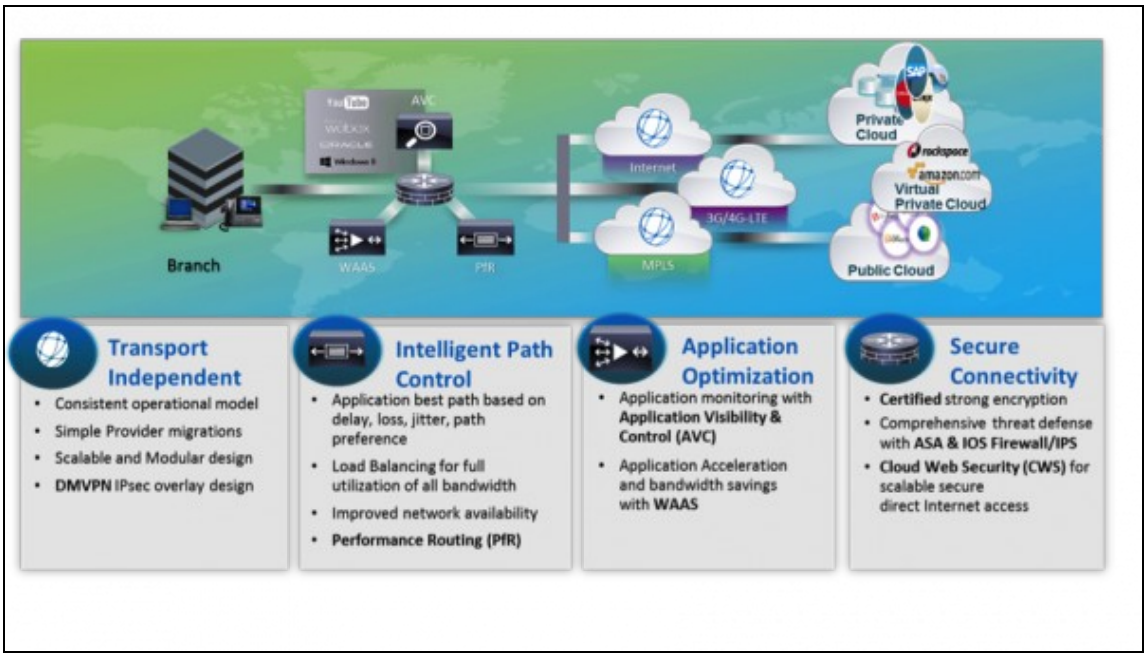
The Cisco Intelligent WAN (IWAN) is a system that enhances collaboration and cloud application performance while reducing the operating cost of the WAN. IWAN leverages low-cost high-bandwidth Internet services to increase bandwidth capacity without compromising performance, availability or security of collaboration or cloud based applications. Organizations can use IWAN to leverage the Internet as a WAN transport, as well as, for direct access to Public Cloud applications. (See Figure 1.)

Figure 1. Cisco IWAN works with both private and public clouds.



Cisco Intelligent WAN is based on four design components:

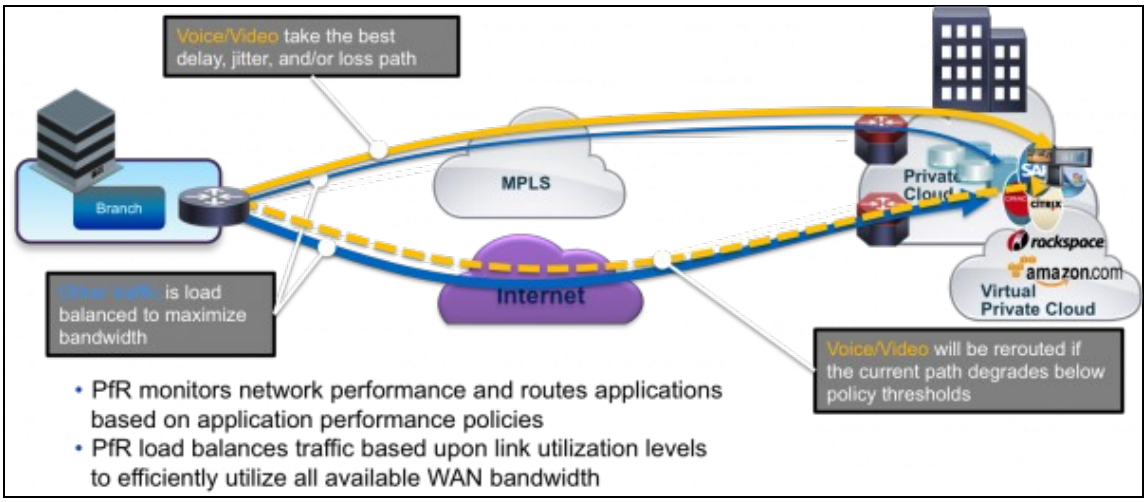
- Transport-Independent Design
- Intelligent Path Control
- Application Optimization
- Secure Internet Access



Transport-Independent Design simplifies the WAN design by using an IPsec VPN overlay over all WAN transport options including MPLS, Internet, and Cellular (3G/4G). The single VPN overlay reduces routing and security complexity, and provides flexibility in choosing providers and transport options. Cisco Dynamic Multipoint VPN (DMVPN) provides the IWAN IPsec overlay. Two or more WAN transport providers are recommended to increase network availability up to 99.999%.

Intelligent Path Control with Cisco Performance Routing (PfR) improves application delivery and WAN efficiency. PfR protects business applications from fluctuating WAN performance while intelligently load balancing traffic over all WAN paths. PfR monitors the network performance (delay, jitter, packet loss,?) to forward critical applications over the best performing paths based on the application policy. PfR?s advanced load balancing evenly distributes traffic to maintain equivalent link utilization levels ? even over links with differing bandwidth capacities. IWAN Intelligent Path Control is key to providing a business-class WAN over Internet transports. (See Figure 2.)

Figure 2. Cisco IWAN Intelligent Path Control.



Application Optimization is provided by Cisco Application Visibility and Control (AVC) and Cisco Wide Area Application Services (WAAS). With applications becoming increasingly ?opaque? (due to increased use of HTTP-based applications), static port classification of applications is no longer sufficient. AVC makes IWAN application-aware with deep packet inspection of traffic to identify and monitor application performance. With increased visibility into the applications on the network, better Quality of Service (QoS) policies can be enabled and fine-tuned to ensure that critical applications are properly prioritized across the network. Cisco WAAS provides application-specific acceleration capabilities that improve response times while reducing WAN bandwidth requirements.

Secure Internet Access offloads user traffic destined for Public Cloud or Internet out the local Internet service. This improves Public Cloud application performance while reducing traffic over the WAN. Cisco?s Cloud Web Security (CWS) service provides a cloud based web proxy to centrally manage and secure user traffic accessing the Internet.

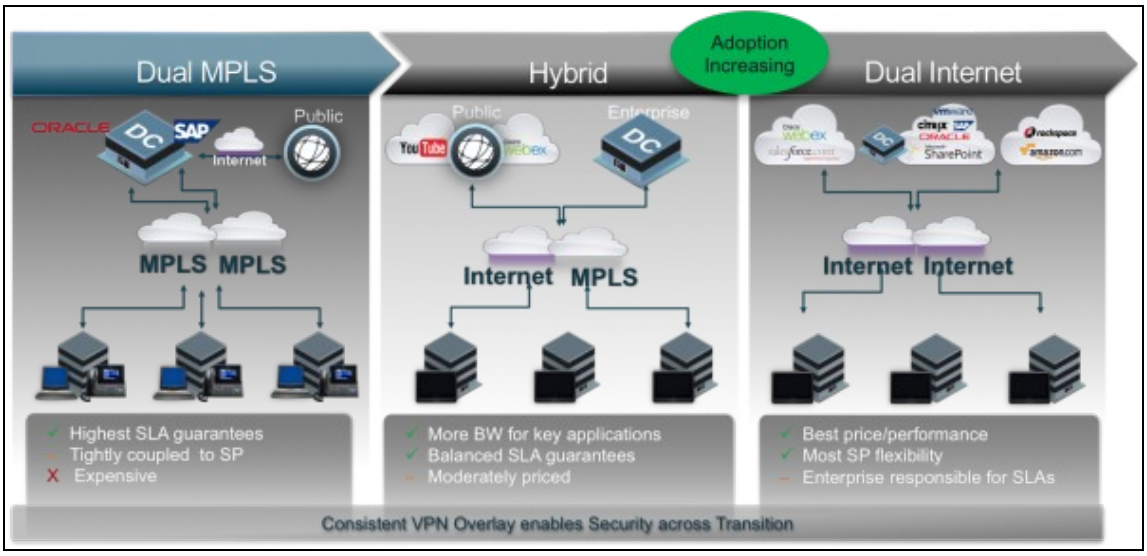
For More Information Read about IWAN at www.cisco.com/go/iwan or contact your local Cisco account representative.

Enterprise Needs

Presentation

This solution is for an Enterprise which is dual-attached to two Service Providers. Multiple scenarios are covered like:

- Dual MPLS-VPN service
- or a primary MPLS-VPN and a secondary DMVPN over the Public Internet.
- or dual DMVPN over the Public Internet



PfR being a key pillar of the IWAN initiative, this guide will focus on a design with dual DMVPN clouds.

In this guide we will cover several concepts including:

- Building the tunnels with DMVPN
- Target Discovery to simplify the active mode configuration
- Advanced learning: defining groups of traffic classes, e.g. VOICE_VIDEO, CRITICAL and BEST_EFFORT using a PfR feature called 'learn-list'
- Measurement modes: passive for BEST_EFFORT and active for VOICE_VIDEO and CRITICAL
- Defining specific policies for each group.
- Path Preference, also called Link-groups

Traffic Classification and Overall Policies

In this guide, the enterprise traffic is divided into 3 main application groups (Service Class) that have their own policies. This is just an example, one can adjust based on his requirements but that's a good starting point.

Voice and Video traffic - VOICE_VIDEO Service Class

- Voice and video traffic is running between the central site and the branches. This traffic is already marked with DSCP EF and the transport is RTP. To capture video we also match DSCP AF41 and CS4 which is used by Tandberg endpoints.
- We would like to track delay, jitter and loss.
- We want to have deterministic path selection and therefore use the primary SP called SP1.
- If the primary WAN experiences brownouts or blackout, we want PfR to failover this traffic to the secondary path.

Critical Application - CRITICAL Service Class

- There is a highly important application running between branches and the central site network and running over TCP.
- This traffic is marked with DSCP AF31.
- This application is interactive, so it is sensitive to delay. In addition, it does not respond well to packet loss. The servers being used for this important application are also responsible for other applications that are not only less important, but that have different requirements of the network.
- We want to have deterministic path selection and therefore use the primary SP SP1.
- If the primary WAN experiences brownouts or blackout, we want PfR to failover this traffic to the secondary path.

Best Effort Applications - BEST_EFFORT Service Class

- The rest of the traffic is just HTTP or email and should be load balanced over all exit interfaces, SP1 and SP2.

Qos is already in place with classification/marking procedures. Therefore packets entering on the Border Router are already classified and marked directly on the access switch connecting the station, IP Phone or multimedia terminal. That means PfR can use the DSCP field as an efficient way for the traffic profiling phase. This guide will use the DSCP values as the classification criteria. But you could also use a combination of destination prefixes and DSCP or even NBAR (see [[PfR:Solutions:NBAR | PfR NBAR Based Application Control] for more information). If the traffic is not marked when entering the BR, then a good option is to classify on ingress, mark the DSCP and then use the DSCP as the classification criteria within PfR. Classification can be done with the usual use of access-lists.

PfR Network Topology Used

Overview

The IWAN independent transport solution requires a DMVPN dual-cloud design, each with a single hub router. The DMVPN routers use tunnel interfaces that support IP unicast as well as IP multicast and broadcast traffic, including the use of dynamic routing protocols. After the initial spoke-to-hub tunnel is

active, it is possible to create dynamic spoke-to-spoke tunnels when site-to-site IP traffic flows require it.

DMVPN requires the use of ISAKMP keepalive for Dead Peer Detection (DPD), which is essential to facilitate fast reconvergence and for spoke registration to function properly in case a DMVPN hub is reloaded. This design enables a spoke to detect that an encryption peer has failed and that the ISAKMP session with that peer is stale, which then allows a new one to be created.

Transport and Overlay Backbones

Addressing Plan:

- 172.16.0.0/16 - Transport
- 10.0.0.0/16 - Overlay - DMVPN1 (10.0.100.0/24) and DMVPN2 (10.0.200.0/24)
- 10.1.0.0/16 - Branch inside prefixes
- 10.2.0.0/16 - Branch loopbacks
- 10.10.0.0/16 - Central Sites

Central site (10.10.0.0/16):

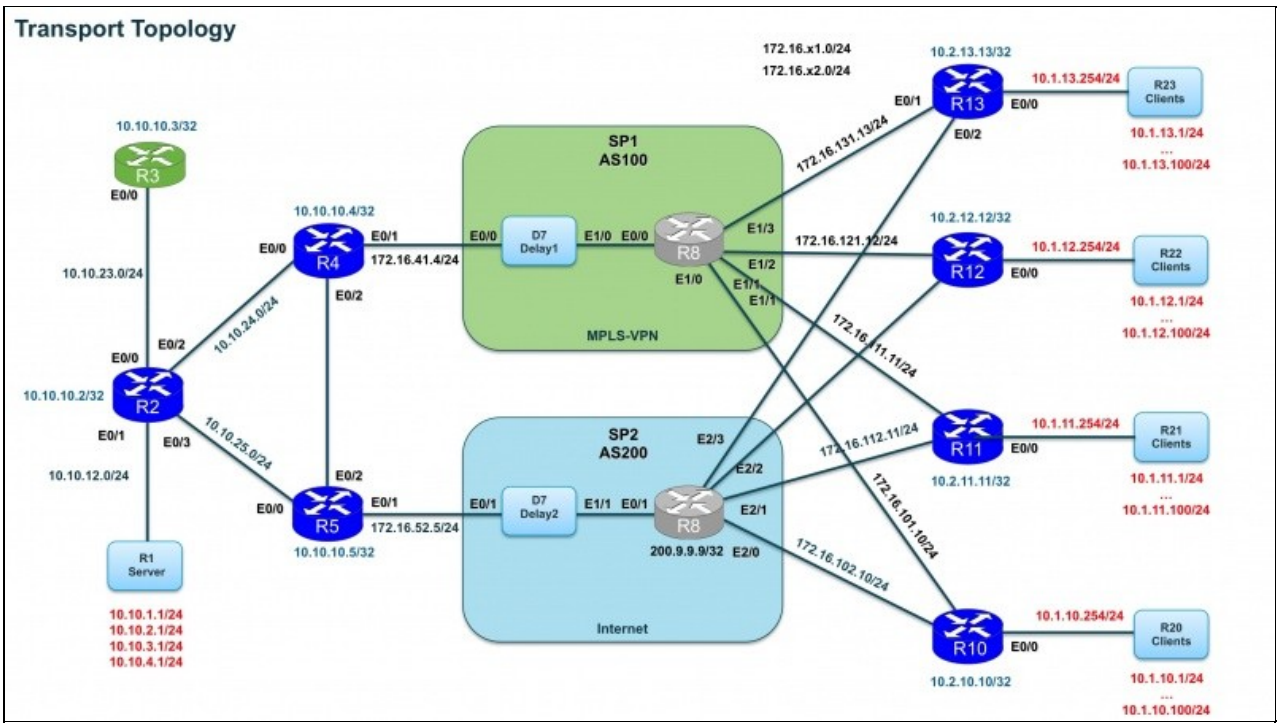
- A dedicated Master Controller (MC) R3 and two Border Routers (BRs) R4 and R5
- R2 is a campus core router or L3 switch and used as an IP SLA responder shadow router.
- Server S1: HTTP and Mail server, voice peer.

Branch Sites:

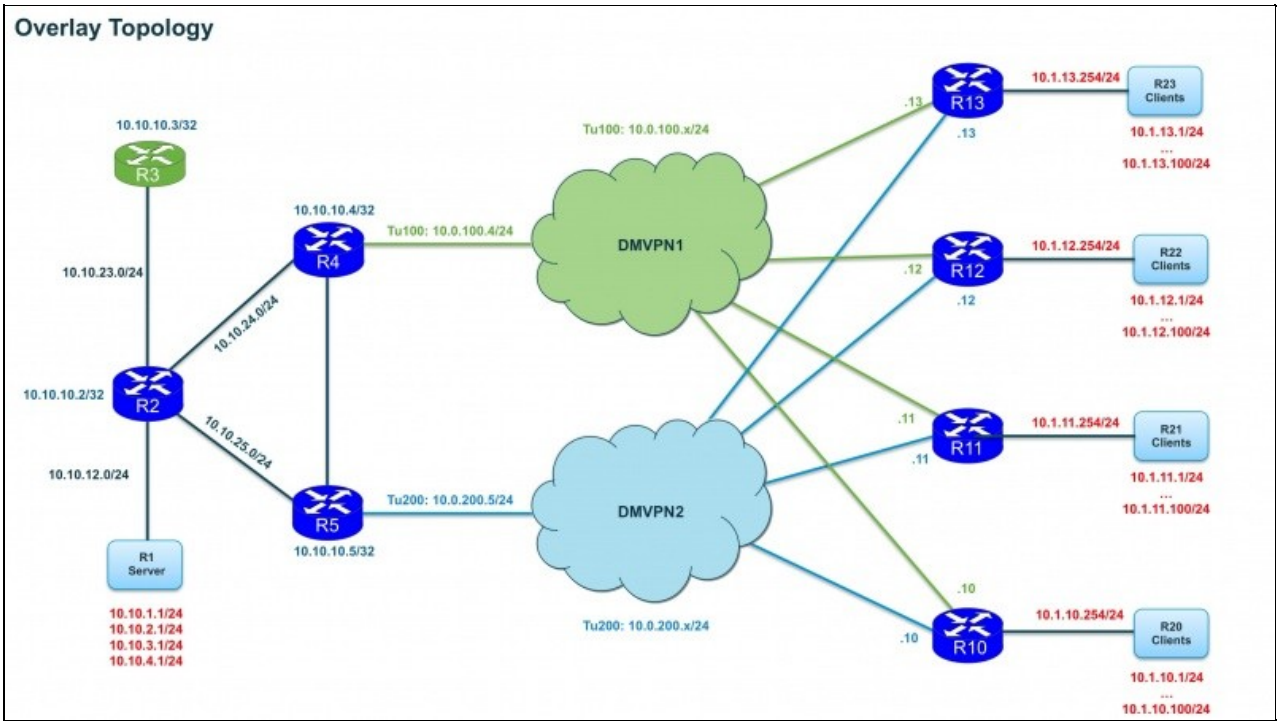
- site R10 (10.1.10.0/24)
- site R11 (10.1.11.0/24)
- site R12 (10.1.12.0/24)
- site R13 (10.1.13.0/24)

Two SP Clouds (SP1 and SP2) with a delay generator (D7).

The Transport Topology is based on two Service Providers, SP1 being considered as the primary one with known SLAs. Therefore SP1 is the preferred path for voice/video and critical applications:



The overlay topology is based on two DMVPN clouds:



This overlay design with two DMVPN clouds can accommodate any kind of transports. The primary path can connect to an MPLS-VPN or even to the Public Internet. The configuration of PfR (and QoS) will remain

the same even if the transport design changes.

Routers and Servers used

Servers:

- S1 is a server on the central site.
- S20, S21, S22 and S23 are clients on their branch office

Routers:

- R3 is the MC for the central site
- R4 and R5 are the BRs for the central site
- R10, R11, R12 and R13 are MC/BR for their branch offices
- R2 is just a router or L3 switch in the central site. It's the campus core backbone.

Delay Generator

- D7 is a delay generator

Traffic Generation

Traffic generation between R1 and R11/R12:

- Voice traffic on RTP ? DSCP EF between spokes and the hub but also between spokes
- Business applications running on port TCP 7000 ? DSCP AF31 between spokes and hub
- Best effort application ? DSCP 0 between spokes and hub

Traffic details:

BEST EFFORT

http

- dest-ip: 10.10.1.1
- src-ip: 10.1.10.1 to 10.1.10.100 by 0.0.0.1 (branch10)
- src-ip: 10.1.10.1 to 10.1.11.100 by 0.0.0.1 (branch11)
- src-ip: 10.1.10.1 to 10.1.12.100 by 0.0.0.1 (branch12)
- src-ip: 10.1.10.1 to 10.1.13.100 by 0.0.0.1 (branch13)

- TCP port 80
- DSCP 0

smtp

- dest-ip: 10.10.2.1
- src-ip: 10.1.10.2 to 10.1.10.100 by 0.0.0.1 (branch10)
- src-ip: 10.1.10.2 to 10.1.11.100 by 0.0.0.1 (branch11)
- src-ip: 10.1.10.2 to 10.1.12.100 by 0.0.0.1 (branch12)
- src-ip: 10.1.10.2 to 10.1.13.100 by 0.0.0.1 (branch13)

- TCP port 25
- DSCP 0

BUSINESS

http

- dest-ip: 10.10.3.1
- src-ip: 10.1.10.3 to 10.1.10.100 by 0.0.0.1 (branch10)
- src-ip: 10.1.10.3 to 10.1.11.100 by 0.0.0.1 (branch11)
- src-ip: 10.1.10.3 to 10.1.12.100 by 0.0.0.1 (branch12)
- src-ip: 10.1.10.3 to 10.1.13.100 by 0.0.0.1 (branch13)

- TCP port 7000
- DSCP = AF31 (26, 0x1A)
- TOS = 0x68, 104

VOICE

rtp

- 10.10.1.1 (hub) <----> 10.1.10.200 (branch10)
- 10.10.1.1 (hub) <----> 10.1.11.200 (branch11)
- 10.10.1.1 (hub) <----> 10.1.12.200 (branch12)
- 10.10.1.1 (hub) <----> 10.1.13.200 (branch13)
- 10.10.10.200 (branch10) <----> 10.1.11.200 (branch11)

- dest-port 20000
- src-port 10000
- DSCP = EF (46, 0x2E)
- TOS = 0xB8, 184

Dual DMVPN Setup

DMVPN Phase Summary

DMVPN has multiple phases that are summarized below:

Phase 1 - 12.2(13)T	Phase 2 – 12.3(4)T (Phase 1+)	Phase 3 – 12.4(6)T
• Hub and spoke functionality • p-pGRE interface on spokes, mGRE on hubs • Simplified and smaller configuration on hubs • Support dynamically addressed CPEs (NAT) • Support for routing protocols and multicast • Spokes don't need full routing table – can summarize on hubs	• Spoke to spoke functionality • mGRE interface on spokes • Direct spoke to spoke data traffic reduces load on hubs • Hubs must interconnect in daisy-chain • Spoke must have full routing table – no summarization • Spoke-spoke tunnel triggered by spoke itself • Routing protocol limitations	• More network designs and greater scaling • Same Spoke to Hub ratio • No hub daisy-chain • Spokes don't need full routing table – can summarize • Spoke-spoke tunnel triggered by hubs • Remove routing protocol limitations • NHRP routes/next-hops in RIB (15.2(1)T)

DMVPN Phase 2 has no summarization on the hub:

- Each spoke has the next-hop (spoke address) for each spoke destination prefix.
- PfR has all information to enforce the path with dynamic PBR and the correct next-hop information

DMVPN phase3 allows route summarization:

- When parent route lookup is performed, only the route to the hub is available
- NHRP dynamically installs shortcut tunnel and hence populates RIB/CEF.
- PfR still has the hub next-hop information and is currently unaware of the next-hop change.

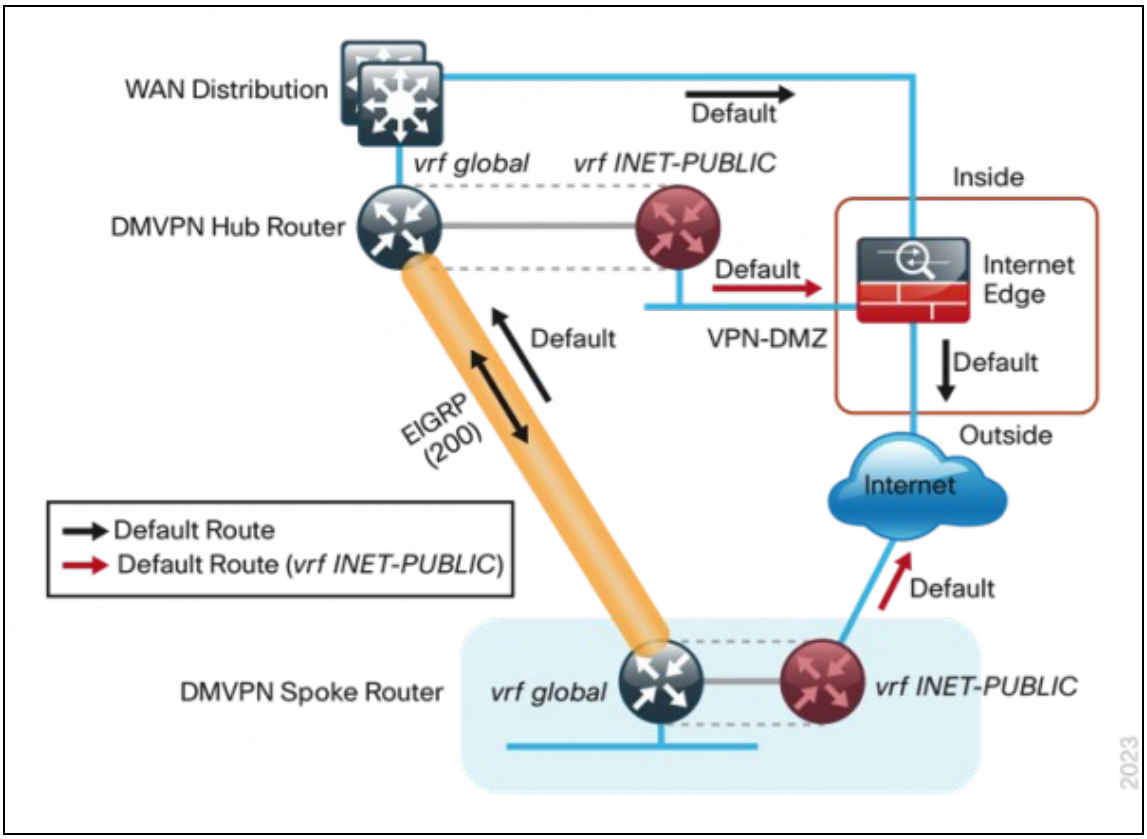
PfR currently supports DMVPN Phase 2 only.

Front Door VRF

Virtual Route Forwarding (VRF) is a technology that allows multiple instances of a routing table to co-exist within the same router at the same time. Because the routing instances are independent, you can use the same or overlapping IP Addresses without conflicting with each other. The simplest form of VRF implementation is VRF Lite. In this implementation, each router within the network participates in the virtual routing environment on a peer-by-peer basis. VRF Lite configurations are only locally significant. The global VRF corresponds to the traditional routing table, and additional VRFs are given names and route descriptors (RDs). Certain features on the router are VRF aware, including static routing and routing protocols, interface forwarding and IPSec tunneling.

The IP routing policy used in this design for the WAN remote sites does not allow direct Internet access for web browsing or other uses; any remote-site hosts that access the Internet must do so via the Internet edge at the primary site. The end hosts require a default route for all Internet destinations; however, this route must force traffic across the primary or secondary WAN transport DMVPN tunnels. This requirement conflicts with the more general VPN spoke router requirement for an Internet-facing default route to bring up the VPN tunnel.

The multiple default route conflict is solved through the use of Front VRFs on the router. This is used in conjunction with DMVPN to permit the use of multiple default routes for both the DMVPN hub routers and DMVPN spoke routers. This combination of features is referred to as front-door (FVRF), because the VRF faces the Internet and the router internal interfaces and the mGRE tunnel all remain in the global VRF.



Note:

- PfR is not VRF-aware
- Tunnels IP addresses are still in the global routing table

The DMVPN hub requires a connection to the Internet, and the DMVPN hub is usually connected through a Firewall using a DMZ interface specifically created and configured for a VPN termination router. This is not represented here.

The Front Door VRF implementation requires the following steps:

- Creating the VRF

- Assigning the external interface to the FVRF
- Defining a default route in the FVRF to allow the creation of the DMVPN tunnel

Front Door VRF Configuration on R4:

```
!
ip vrf INET1
 rd 65512:1
!

interface Ethernet0/1
 description --ISP1--
 bandwidth 1000
 ip vrf forwarding INET1
 ip address 172.16.41.4 255.255.255.0
!
ip route vrf INET1 0.0.0.0 0.0.0.0 172.16.41.8
!
```

Front Door VRF Configuration on R5:

```
!
ip vrf INET2
 rd 65512:2
!
interface Ethernet0/1
 description --ISP2--
 bandwidth 1000
 ip vrf forwarding INET2
 ip address 172.16.52.5 255.255.255.0
!
!
ip route vrf INET2 0.0.0.0 0.0.0.0 172.16.52.8
!
```

The DMVPN spoke routers at the WAN remote sites connect to the Internet directly through a router interface without a separate firewall. This connection is secured in two ways. Because the Internet interface is in a separate VRF, no traffic can access the global VRF except traffic sourced through the DMVPN tunnel. This design provides implicit security. Additionally, an IP access list permits only the traffic required for an encrypted tunnel, as well as DHCP and various ICMP protocols for troubleshooting. The IP access list must permit the protocols specified in the following configuration sample. The access list is applied inbound on the WAN interface, so filtering is done on traffic destined to the router.

```
interface Ethernet0/1
 ip access-group ACL-INET-PUBLIC in
!
interface Ethernet0/2
```

```
 ip access-group ACL-INET-PUBLIC in
!
ip access-list extended ACL-INET-PUBLIC
 permit udp any any eq non500-isakmp         ! IPsec via NAT-T
 permit udp any any eq isakmp                ! ISAKMP (UDP 500)
 permit esp any any                          ! IPSEC
 permit udp any any eq bootpc                ! DHCP
```

The additional protocols listed in the following table may assist in troubleshooting, but are not explicitly required to allow DMVPN to function properly.

```
ip access-list extended ACL-INET-PUBLIC
 permit icmp any any echo                     ! Allow remote pings
 permit icmp any any echo-reply               ! Allow ping replies (from our requests)
 permit icmp any any ttl-exceeded             ! Allow traceroute replies (from our requests)
 permit icmp any any port-unreachable         ! Allow traceroute replies (from our requests)
 permit udp any any gt 1023 ttl eq 1          ! Allow remote traceroute
```

Front Door VRF Configuration on R10 which is dual homed:

```
!
ip vrf INET1
 rd 65512:1
!
ip vrf INET2
 rd 65512:2
!
!
interface Ethernet0/1
 description --ISP1--
 ip vrf forwarding INET1
 ip address 172.16.101.10 255.255.255.0
!
interface Ethernet0/2
 description --ISP2--
 ip vrf forwarding INET2
 ip address 172.16.102.10 255.255.255.0
!
ip route vrf INET1 0.0.0.0 0.0.0.0 172.16.101.8
ip route vrf INET2 0.0.0.0 0.0.0.0 172.16.102.8
```

DMVPN Configuration on the Hub

PfR runs over Dual DMVPN clouds. On the hub site R4 is the hub for DMVPN1 and R5 is the hub for DMVPN2. Each spoke as 2 tunnels, one per DMVPN cloud.

R4 DMVPN configuration:

```
!
! ---------------------------------------------------------------------------
! 1. Configure the crypto keyring
! The crypto keyring defines a pre-shared key (or password) valid for IP sources
! reachable within a particular VRF.
! This key is a wildcard pre-shared key if it applies to any IP source.
! A wildcard key is configured using the 0.0.0.0 0.0.0.0 network/mask combination.
!
crypto keyring DMVPN-KEYRING1 vrf INET1
  pre-shared-key address 0.0.0.0 0.0.0.0 key cisco123
!
!
!
! ---------------------------------------------------------------------------
! 2. Configure the ISAKMP policy
! The ISAKMP policy for DMVPN uses the following:
!   - Advanced Encryption Standard (AES)
!   - Authentication by pre-shared key
!
crypto isakmp policy 10
 encr aes
 authentication pre-share
!
crypto isakmp invalid-spi-recovery
!
!
! ---------------------------------------------------------------------------
! 3. Configure the ISAKMP Profile
! The ISAKMP profile creates an association between an identity address, a VRF, and a crypto keyri
! A wildcard address within a VRF is referenced with 0.0.0.0.
!
crypto isakmp profile ISAKMP-INET1
   keyring DMVPN-KEYRING1
   match identity address 0.0.0.0 INET1
!
!
! ---------------------------------------------------------------------------
! 4. Define the IPSec transform set
! A transform set is an acceptable combination of security protocols, algorithms,
! and other settings to apply to IPsec-protected traffic.
! Peers agree to use a particular transform set when protecting a particular data flow.
!
crypto ipsec transform-set AES256/SHA/TRANSPORT esp-aes 256 esp-sha-hmac
 mode transport
!
!
! ---------------------------------------------------------------------------
! 5. Create the IPSec profile
! The IPsec profile creates an association between an ISAKMP profile and an IPsec transform-set.
!
crypto ipsec profile DMVPN-PROFILE1
 set transform-set AES256/SHA/TRANSPORT
 set isakmp-profile ISAKMP-INET1
!
!
!
! ---------------------------------------------------------------------------
! 6. Configure the mGRE tunnel
! DMVPN uses multipoint GRE (mGRE) tunnels.
! This type of tunnel requires a source interface only.
!
!   - Use the same source interface that you use to connect to the Internet.
!   - Set the tunnel vrf command to the VRF defined previously for FVRF.
!   - Configure basic interface settings
```

```
!      The IP MTU should be configured to 1400
!      The ip tcp adjust-mss should be configured to 1360.
!      There is a 40 byte difference which corresponds to the combined IP and TCP header length.
!   - Configure NHRP
!   - Set NHRP holdtime to 600
!   - Set Delay to 1000
!   - Apply the IPSec profile to the tunnel
!
interface Tunnel100
 bandwidth 1000
 ip address 10.0.100.4 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp authentication cisco
 ip nhrp map multicast dynamic
 ip nhrp network-id 1
 ip nhrp holdtime 600
 ip tcp adjust-mss 1360
 load-interval 30
 delay 1000
 tunnel source Ethernet0/1
 tunnel mode gre multipoint
 tunnel key 100
 tunnel vrf INET1
 tunnel protection ipsec profile DMVPN-PROFILE1
!
```

R5 DMVPN configuration:

```
!
crypto keyring DMVPN-KEYRING2 vrf INET2
  pre-shared-key address 0.0.0.0 0.0.0.0 key cisco123
!
!
crypto isakmp policy 10
 encr aes
 authentication pre-share
!
crypto isakmp invalid-spi-recovery
!
crypto isakmp profile ISAKMP-INET2
   keyring DMVPN-KEYRING2
   match identity address 0.0.0.0 INET2
!
!
crypto ipsec transform-set AES256/SHA/TRANSPORT esp-aes 256 esp-sha-hmac
 mode transport
!
crypto ipsec profile DMVPN-PROFILE2
 set transform-set AES256/SHA/TRANSPORT
 set isakmp-profile ISAKMP-INET2
!
!
interface Tunnel200
 bandwidth 1000
 ip address 10.0.200.5 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp authentication cisco
 ip nhrp map multicast dynamic
 ip nhrp network-id 2
```

```
 ip nhrp holdtime 600
 ip tcp adjust-mss 1360
 load-interval 30
 delay 1000
 tunnel source Ethernet0/1
 tunnel mode gre multipoint
 tunnel key 200
 tunnel vrf INET2
 tunnel protection ipsec profile DMVPN-PROFILE2
!
```

DMVPN Configuration on the Spokes

R10 Spoke Configuration with 2 DMVPN tunnels:

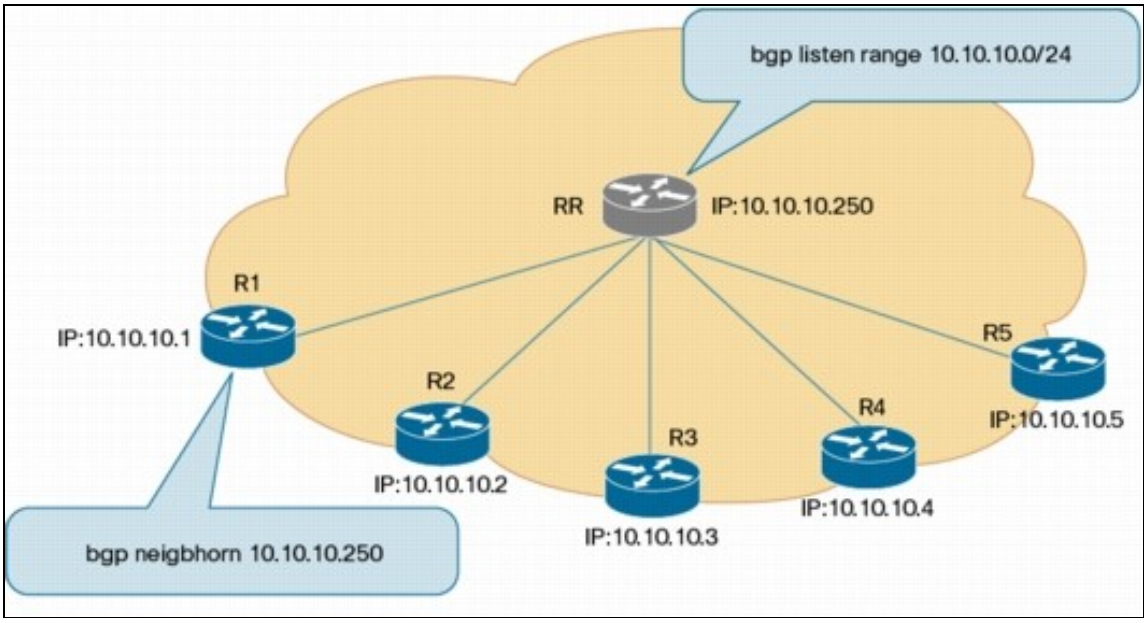
```
!
! ---------------------------------------------------------------------------
! 1. Configure the crypto keyring
! The crypto keyring defines a pre-shared key (or password) valid for IP sources
! reachable within a particular VRF.
! This key is a wildcard pre-shared key if it applies to any IP source.
! A wildcard key is configured using the 0.0.0.0 0.0.0.0 network/mask combination.
!
crypto keyring DMVPN-KEYRING1 vrf INET1
  pre-shared-key address 0.0.0.0 0.0.0.0 key cisco123
!
crypto keyring DMVPN-KEYRING2 vrf INET2
  pre-shared-key address 0.0.0.0 0.0.0.0 key cisco123
!
!
! ---------------------------------------------------------------------------
! 2. Configure the ISAKMP policy
! The ISAKMP policy for DMVPN uses the following:
!   - Advanced Encryption Standard (AES)
!   - Authentication by pre-shared key
!
crypto isakmp policy 10
 encr aes
 authentication pre-share
!
crypto isakmp invalid-spi-recovery
!
!
! ---------------------------------------------------------------------------
! Enable DPD
!  - with keepalive intervals sent at 30-second intervals
!  - with a 5-second retry interval,
! which is considered to be a reasonable setting to detect a failed hub.
!
crypto isakmp keepalive 30 5
!
!
! ---------------------------------------------------------------------------
! 3. Configure the ISAKMP Profiles
! The ISAKMP profile creates an association between an identity address, a VRF, and a crypto keyri
! A wildcard address within a VRF is referenced with 0.0.0.0.
!
crypto isakmp profile ISAKMP-INET1
   keyring DMVPN-KEYRING1
   match identity address 0.0.0.0 INET1
```

```
!
crypto isakmp profile ISAKMP-INET2
   keyring DMVPN-KEYRING2
   match identity address 0.0.0.0 INET2
!
!
! ---------------------------------------------------------------------
! 4. Define the IPSec transform set
! A transform set is an acceptable combination of security protocols, algorithms,
! and other settings to apply to IPsec-protected traffic.
! Peers agree to use a particular transform set when protecting a particular data flow.
!
crypto ipsec transform-set AES256/SHA/TRANSPORT esp-aes 256 esp-sha-hmac
 mode transport
!
!
! ---------------------------------------------------------------------
! 5. Create the IPSec profiles
! The IPsec profile creates an association between an ISAKMP profile and an IPsec transform-set.
!
crypto ipsec profile DMVPN-PROFILE1
 set transform-set AES256/SHA/TRANSPORT
 set isakmp-profile ISAKMP-INET1
!
crypto ipsec profile DMVPN-PROFILE2
 set transform-set AES256/SHA/TRANSPORT
 set isakmp-profile ISAKMP-INET2
!
!
! ---------------------------------------------------------------------
! 6. Configure the mGRE tunnel on DMVPN1
! DMVPN uses multipoint GRE (mGRE) tunnels.
! This type of tunnel requires a source interface only.
!
!   - Use the same source interface that you use to connect to the Internet.
!   - Set the tunnel vrf command to the VRF defined previously for FVRF.
!   - Configure basic interface settings
!      The IP MTU should be configured to 1400
!      The ip tcp adjust-mss should be configured to 1360.
!      There is a 40 byte difference which corresponds to the combined IP and TCP header length.
!   - Configure NHRP
!   - Set NHRP holdtime to 600
!   - Set Delay to 1000
!   - Apply the IPSec profile to the tunnel
!
!
interface Tunnel100
 ip address 10.0.100.10 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp authentication cisco
 ip nhrp map 10.0.100.4 172.16.41.4
 ip nhrp map multicast 172.16.41.4
 ip nhrp network-id 1
 ip nhrp holdtime 600
 ip nhrp nhs 10.0.100.4
 ip nhrp registration timeout 60
 ip tcp adjust-mss 1360
 delay 1000
 tunnel source Ethernet0/1
 tunnel mode gre multipoint
 tunnel key 100
 tunnel vrf INET1
 tunnel protection ipsec profile DMVPN-PROFILE1
```

```
!
!
! ---------------------------------------------------------------------------
! 7. Configure the mGRE tunnel on DMVPN2
!
interface Tunnel200
 ip address 10.0.200.10 255.255.255.0
 no ip redirects
 ip mtu 1400
 ip nhrp authentication cisco
 ip nhrp map 10.0.200.5 172.16.52.5
 ip nhrp map multicast 172.16.52.5
 ip nhrp network-id 2
 ip nhrp holdtime 600
 ip nhrp nhs 10.0.200.5
 ip nhrp registration timeout 60
 ip tcp adjust-mss 1360
 delay 1000
 tunnel source Ethernet0/2
 tunnel mode gre multipoint
 tunnel key 200
 tunnel vrf INET2
 tunnel protection ipsec profile DMVPN-PROFILE2
!
```

Routing on the Overlay Backbone

R4 and R5 are iBGP peers and implement a feature called Dynamic Neighbors.
BGP dynamic neighbor support allows BGP peering to a group of remote neighbors that are defined by a range of IP addresses. Each range can be configured as a subnet IP address.



With this feature R4 and R5 just listen to incoming BGP connections. This avoids the manual configuration of all remote sites neighbors. In this design, there is no mutual redistribution, BGP is only redistributed into OSPF.

R4 Hub Configuration:

```
router ospf 100
 router-id 10.10.10.4
 redistribute bgp 10 metric 100 subnets
!
router bgp 10
 bgp router-id 10.10.10.4
 bgp cluster-id 10.10.10.4
 bgp log-neighbor-changes
 bgp listen range 10.0.100.0/24 peer-group SPOKES-1
 neighbor SPOKES-1 peer-group
 neighbor SPOKES-1 remote-as 10
 neighbor SPOKES-1 update-source Loopback0
 neighbor SPOKES-1 timers 5 25
 neighbor 10.10.10.5 remote-as 10
 neighbor 10.10.10.5 update-source Loopback0
 !
 address-family ipv4
  bgp redistribute-internal
  network 10.0.100.0 mask 255.255.255.0
  network 10.10.10.4 mask 255.255.255.255
  aggregate-address 10.10.0.0 255.255.0.0 summary-only
  neighbor SPOKES-1 activate
  neighbor SPOKES-1 route-reflector-client
  neighbor 10.10.10.5 activate
  distance bgp 20 21 21
 exit-address-family
!
!
```

Notes:

- All spokes are iBGP peers
- R4 listens from incoming connections from range 10.0.100.0/24
- R4 summarizes hub prefix to 10.10.0.0/16
- R4 does NOT summarize spoke prefixes 10.1.0.0/16

R5 Hub Configuration:

```
!
router ospf 100
 router-id 10.10.10.5
 redistribute bgp 10 metric 100 subnets
!
router bgp 10
 bgp router-id 10.10.10.5
 bgp cluster-id 10.10.10.5
 bgp log-neighbor-changes
 bgp listen range 10.0.200.0/24 peer-group SPOKES-2
 neighbor SPOKES-2 peer-group
 neighbor SPOKES-2 remote-as 10
 neighbor SPOKES-2 update-source Loopback0
 neighbor SPOKES-2 timers 5 25
 neighbor 10.10.10.4 remote-as 10
 neighbor 10.10.10.4 update-source Loopback0
 !
```

```
 address-family ipv4
  bgp redistribute-internal
  network 10.0.200.0 mask 255.255.255.0
  network 10.10.10.5 mask 255.255.255.255
  aggregate-address 10.10.0.0 255.255.0.0 summary-only
  neighbor SPOKES-2 activate
  neighbor SPOKES-2 route-reflector-client
  neighbor 10.10.10.4 activate
  distance bgp 20 21 21
 exit-address-family
!
```

Notes:

- All spokes are iBGP peers
- R5 listens from incoming connections from range 10.0.200.0/24
- R5 summarizes hub prefix to 10.10.0.0/16
- R5 does NOT summarize spoke prefixes 10.1.0.0/8

R9 Spoke Configuration:

```
!
router bgp 10
 bgp router-id 10.2.10.10
 bgp log-neighbor-changes
 neighbor HUBS-1 peer-group
 neighbor HUBS-1 remote-as 10
 neighbor HUBS-1 update-source Tunnel100
 neighbor HUBS-1 timers 5 25
 neighbor HUBS-2 peer-group
 neighbor HUBS-2 remote-as 10
 neighbor HUBS-2 update-source Tunnel200
 neighbor HUBS-2 timers 5 25
 neighbor 10.0.100.4 peer-group HUBS-1
 neighbor 10.0.200.5 peer-group HUBS-2
 !
 address-family ipv4
  network 10.1.10.0 mask 255.255.255.0
  network 10.2.10.10 mask 255.255.255.255
  neighbor 10.0.100.4 activate
  neighbor 10.0.200.5 activate
 exit-address-family
!
!
```

Notes:

- All spokes are BGP Route Reflector clients

Check Routing

PfR being not active, the parent routes are BGP based on R4 and R5 as seen below. There is no preference applied to a particular DMVPN.

On R4 (hub) - Route to destination prefixes BRANCH10 (10.1.10.0/24):

```
R4#sh ip bgp 10.1.10.0
BGP routing table entry for 10.1.10.0/24, version 4
Paths: (2 available, best #2, table default)
  Advertised to update-groups:
     1          2
  Refresh Epoch 1
  Local
    10.0.200.10 (metric 11) from 10.10.10.5 (10.10.10.5)
      Origin IGP, metric 0, localpref 100, valid, internal
      Originator: 10.2.10.10, Cluster list: 10.10.10.5
      rx pathid: 0, tx pathid: 0
  Refresh Epoch 1
  Local, (Received from a RR-client)
    10.0.100.10 from *10.0.100.10 (10.2.10.10)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
R4#
```

What to check:

- Best route is directly DMVPN1

On R5 (hub) - Route to destination prefixes BRANCH2 (10.1.10.0/24):

```
R5#sh ip bgp 10.1.10.0
BGP routing table entry for 10.1.10.0/24, version 3
Paths: (2 available, best #2, table default)
  Advertised to update-groups:
     1          2
  Refresh Epoch 1
  Local
    10.0.100.10 (metric 11) from 10.10.10.4 (10.10.10.4)
      Origin IGP, metric 0, localpref 100, valid, internal
      Originator: 10.2.10.10, Cluster list: 10.10.10.4
      rx pathid: 0, tx pathid: 0
  Refresh Epoch 1
  Local, (Received from a RR-client)
    10.0.200.10 from *10.0.200.10 (10.2.10.10)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
R5#
```

What to check:

- Best route is directly DMVPN2

On R10 (spoke) - Routes to the hub:

```
R10#sh ip bgp 10.10.0.0
BGP routing table entry for 10.10.0.0/16, version 11
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  Local, (aggregated by 10 10.10.10.5)
    10.0.200.5 from 10.0.200.5 (10.10.10.5)
      Origin IGP, metric 0, localpref 100, valid, internal, atomic-aggregate
      rx pathid: 0, tx pathid: 0
  Refresh Epoch 1
  Local, (aggregated by 10 10.10.10.4)
    10.0.100.4 from 10.0.100.4 (10.10.10.4)
      Origin IGP, metric 0, localpref 100, valid, internal, atomic-aggregate, best
      rx pathid: 0, tx pathid: 0x0
R10#
```

What to check:

- Hub subnets 10.10.0.0/16 advertized through DMVPN1 (10.0.100.4) and DMVPN2 (10.0.200.5)
- Best route is DMVPN1

On R10 (spoke) - Routes to the other spoke BRANCH11 (R11):

```
R10#sh ip bgp 10.1.11.0
BGP routing table entry for 10.1.11.0/24, version 9
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.0.200.11 from 10.0.200.5 (10.10.10.5)
      Origin IGP, metric 0, localpref 100, valid, internal
      Originator: 10.2.11.11, Cluster list: 10.10.10.5
      rx pathid: 0, tx pathid: 0
  Refresh Epoch 1
  Local
    10.0.100.11 from 10.0.100.4 (10.10.10.4)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Originator: 10.2.11.11, Cluster list: 10.10.10.4
      rx pathid: 0, tx pathid: 0x0
R10#
```

What to check:

- R10 route advertized through R4 (DMVPN1, 10.0.100.4) - best route
- R10 route advertized through R5 (DMVPN2, 10.0.200.5.

Checking flows

This section is optional. There is absolutely **no need** to configure Flexible NetFlow (FNF) for PfR to run. Flexible Netflow (FNF) is used here to check the active flows between the central site and the branch offices. This also shows that FNF can be used as a troubleshooting tool. But keep in mind that Performance Routing version2 (PfRv2) makes use of **Traditional NetFlow (TNF)** and learn the traffic going from an internal to an external interface based on the TNF cache.

Create the flow record:

```
!*******************************
! FLOW RECORD
! What data do I want to meter?
! First define the flow record that you want.
! ?match? is used to identify key fields, ie, those fields used to define what a flow is
! ?collect? is used to define what fields we want to monitor
!
flow record RECORD-STATS
 match ipv4 dscp
 match ipv4 protocol
 match ipv4 source address
 match ipv4 destination address
 match transport source-port
 match transport destination-port
 match interface input
 match flow direction
 collect routing next-hop address ipv4
 collect counter bytes
!
```

Create the flow monitor

```
!*******************************
! FLOW MONITOR
! Creates a new NetFlow cache
! Attach the flow record
! Exporter is attached to the cache
! Potential sampling configuration
!
flow monitor MONITOR-STATS
 cache timeout inactive 60
 cache timeout active 60
 cache timeout update 1
 record RECORD-STATS
!
```

Then apply the flow monitor on the interface:

```
!
interface Ethernet0/2
 description -- TO BORDER ROUTERS --
 ip flow monitor MONITOR-STATS input
 ip flow monitor MONITOR-STATS output
!
```

You can now check the NetFlow cache on R2 and check that we have critical applications with DSCP 0x1A (AF31) and voice flows with DSCP 0x2E (EF) (this is just and extract of the cache):

```
R2#sh flow monitor MONITOR-STATS cache format table
  Cache type:                               Normal
```

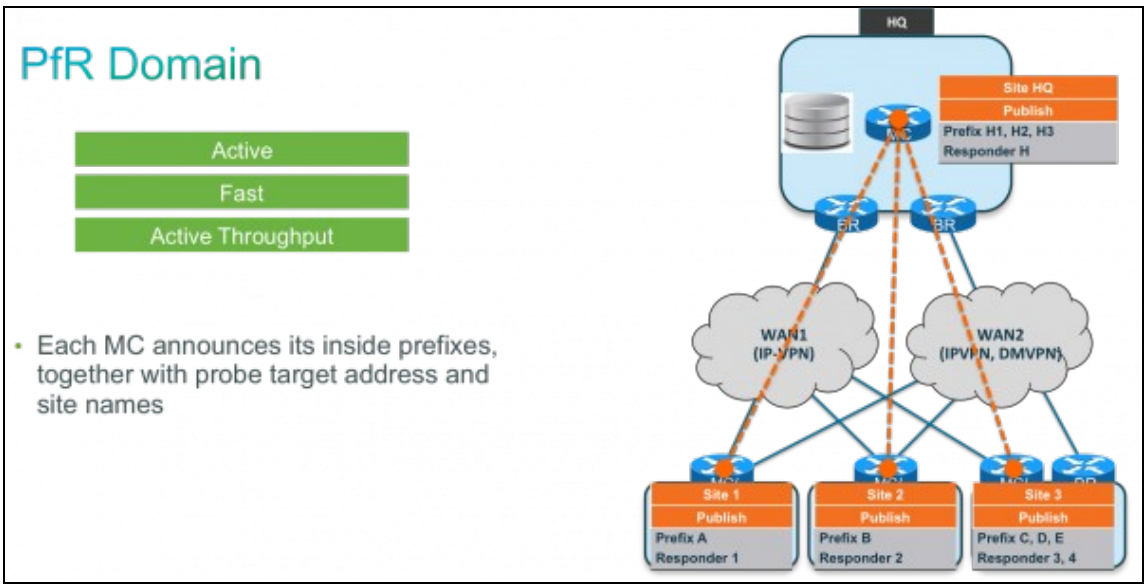
```
  Cache size:                                 4096
  Current entries:                             859
  High Watermark:                              882

  Flows added:                                5958
  Flows aged:                                 5099
    - Active timeout      (    60 secs)       5099
    - Inactive timeout    (    60 secs)          0
    - Event aged                                 0
    - Watermark aged                             0
    - Emergency aged                             0

IPV4 SRC ADDR    IPV4 DST ADDR    TRNS SRC PORT  TRNS DST PORT  INTF INPUT            FLOW DIRN  I
===============  ===============  =============  =============  ====================  =========  =

10.10.1.1        10.1.10.200              20000          10000  Et0/1                 Output     0
10.10.1.1        10.1.11.200              10000          20000  Et0/1                 Output     0
10.10.1.1        10.1.13.200              10000          20000  Et0/1                 Output     0
10.10.1.1        10.1.12.200              10000          20000  Et0/1                 Output     0
10.1.12.19       10.10.3.1                 7094           7000  Et0/2                 Input      0
10.1.12.23       10.10.2.1                 1100             25  Et0/2                 Input      0
10.0.100.10      10.10.10.2               60324           5000  Et0/2                 Input      0
10.0.200.10      10.10.10.2               53253           5000  Et0/3                 Input      0
10.1.12.69       10.10.3.1                 7053           7000  Et0/2                 Input      0
10.10.3.1        10.1.12.69                7000           7053  Et0/1                 Output     0
10.1.12.34       10.10.1.1                 2016             80  Et0/3                 Input      0
10.10.1.1        10.1.12.34                  80           2016  Et0/1                 Output     0
10.1.10.28       10.10.1.1                 2044             80  Et0/3                 Input      0
10.10.1.1        10.1.10.28                  80           2044  Et0/1                 Output     0

{SNIP]
```

PfR Configuration

Presentation

In this guide we want to protect voice/video and critical applications against soft errors, brownouts and blackouts. We want to be able to track jitter, delay and loss but we also want to have a fast failover. Therefore performance measurement will be in active mode and will use jitter probes. Jitter probe configuration for all the remote sites is a painful process and Cisco Performance Routing is now using a new feature called Target Discovery to help simplifying the configuration and deployment in such cases.

A peering session is configured between the Master Controller on the central site and the Master Controllers on the remote sites. The MCs which participate in this peering will exchange the list of inside prefixes and IP SLA probe responder addresses that belong to the sites, along with the site name (configurable).



Provisioning

Central Site:

Basic configuration to establish session between MC and BR on the central site:

```
! This is the basic configuration required to establish session between MC and BR.
! With this basic configuration, learning is enabled, route control is in place and
! load-balancing happens on all configured external interfaces.
!
! It includes
! ? Key-chain configuration for authentication.
!
! ? Specification of BR?s IP Address and internal/external interface on the BR.
!  There is a direct interface between the BRs, it should be configured as internal
!
! - Specification of the Carrier name (link-group). Link-group is used to color the exit interface
!   We will define an administrative policy for voice/video and critical application to choose SP1
!   has the primary path and SP2 as the fallback path.
!
! - Load balancing range set to 30%. This is conservative and helps to cut down on the churn resul
!   from movement of traffic-classes across exits.
!
! - Disabling auto-tunnels as BRs are L2 adjacent.
! - Forcing the use of PBR
! - Specification of the probe packet number. Used to decrease the CPU utilization.
!
! Notes on the new defaults:
! - Automatic learning is enabled
! - mode route control is enabled
! - Maximum transmit utilization is 90%
! - Load-balancing is now used as the last resolver and cannot be disabled.
!
!
pfr master
 max-range-utilization percent 30
 !
 border 10.10.10.4 key-chain pfr
  interface Ethernet0/0 internal
```

```
  interface Ethernet0/2 internal
  interface Tunnel100 external
   link-group SP1
 !
 border 10.10.10.5 key-chain pfr
  interface Ethernet0/0 internal
  interface Ethernet0/2 internal
  interface Tunnel200 external
   link-group SP2
 !
 mode route protocol pbr
 no mode auto-tunnels
 periodic 90
 probe packets 20
 !
!
```

Enable NetFlow version9 Export (optional)

```
! Flow Exporter Definition
! Where do I want my data sent?
!
flow exporter MYEXPORTER
 destination 10.151.1.131
 source Loopback0
 transport udp 9991
 option interface-table timeout 300
 option sampler-table timeout 300
 option application-table timeout 300
!
!
! Add NetFlow export to PfR
!
pfr master
 exporter MYEXPORTER
!
```

Enable logging

```
! Following configuration is to enable logging.
! This will print PfR related syslog messages on the console.
!
pfr master
 logging
!
```

Branch Sites:

Basic configuration to establish session between MC and BR on the branch sites:

```
!
```

```
! BRANCH R10
!
key chain pfr
 key 0
  key-string cisco
!
pfr master
 logging
 !
 border 10.2.10.10 key-chain pfr
  interface Ethernet0/0 internal
  interface Tunnel100 external
   link-group SP1
  interface Tunnel200 external
   link-group SP2
 !
!
pfr border
 logging
 local Loopback0
 master 10.2.10.10 key-chain pfr
!
```

Similar configuration applies to R11, R12 and R13.

Enabling PfR Domain and Target Discovery

The Performance Routing (PfR) Target Discovery feature introduces a scalable solution for managing the performance of video, voice and critical applications across large Enterprise branch networks by automating the identification and configuration of IP SLA responders. To optimize media applications using voice and video traffic, PfR uses jitter, loss, and delay measurements. The IP SLA udp-jitter probe provides these measurements but requires an IP SLA responder. The initial PfR solution was to manually configure the probes for all remote sites. PfR Target Discovery uses the PfR Domain Peering framework to advertize site prefixes and to identify and advertize IP SLA responder IP addresses.

Enabling Target Discovery on R3:

```
!
! Target Discovery is used in a hub and spoke model where R3 is the hub and R9/R10 are the spokes.
! R3 is listening to incoming connections from the spokes.
! We want to define the IP address of the shadow router used as a responder on the hub. We also wa
!
pfr master
 mc-peer domain 65000 head-end Loopback0
 target-discovery responder-list RESPONDER_PREFIX inside-prefixes HQ_PREFIX
!
ip prefix-list HQ_PREFIX seq 5 permit 10.10.1.0/24
ip prefix-list HQ_PREFIX seq 10 permit 10.10.2.0/24
ip prefix-list HQ_PREFIX seq 15 permit 10.10.3.0/24
ip prefix-list HQ_PREFIX seq 20 permit 10.10.4.0/24
ip prefix-list RESPONDER_PREFIX seq 5 permit 10.10.10.2/32
!
```

Enabling Target Discovery on R10:

```
!
! Target Discovery is used in a hub and spoke model where R3 is the hub and R10/R11/R12/R13 are th
! R10 is configured with the IP address of the hub (R3)
! To simplify the configuration on the branch offices, we don't want to manually define the target
! PfR will automatically generate what's needed. Bu default the responder address is the IP addres
!
pfr master
 mc-peer domain 65000 10.10.10.3 Loopback0
 target-discovery
!
```

A similar configuration applies to R11, R12 and R13.

Learning Configuration

Central Site

We would like PfR to learn and optimize traffic going to the branch offices. More specifically, of the streams that are headed to that branch, we would like PfR to optimize certain mission critical traffic based on certain policy parameters and the rest of the traffic based on a different set of parameters.

We also want to apply policies not on a global basis but for specific application groups. The choice here is to define application groups by filtering on DSCP values.

We assume that traffic is already marked when it enters the BRs on all sites. We can therefore define learning based on DSCP values while filtering based on the branch prefix. The filtering (based on the prefix) will make sure that PfR only learns traffic destined to the remote branch offices, while the different access-lists (based on DSCP values) help categorize the interesting traffic into different buckets like critical, best effort, etc., which can then be associated with different optimization policies.

Define learning parameters:

```
! By default now:
!
! - Automatic learning is enabled
!
! - Continuous learn cycle, each 1 minute duration
!   periodic-interval = 0 (forever)
!   monitor-period = 1 (minutes)
!
! - traffic-class sorting based on ?throughput? at the end of each learning cycle.
!
! - Anything traffic that doesn?t match the learn list will be learned under global learn and will
!
pfr master
 !
 mc-peer domain 65000 head-end Loopback0
 target-discovery responder-list RESPONDER_PREFIX inside-prefixes HQ_PREFIX
 logging
 !
 learn
  ! - Define 3 groups of applications
```

```
  ! - disable global learn.
  ! - Sort the traffic-class based on ?throughput? at the end of each learning cycle.
  !
  throughput
  !
  ! For large scale deployment you may want to define the periodic interval from 0 (infinite) to 1
  ! This allows the MC to have time to learn the new Traffic Classes and avoid having possible ove
  ! In this solution guide we will keep the default configuration
  !
  ! periodic-interval 1
  !
  ! We do not want to use global learning.
  ! Learn-list has been optimized for learning and high number of TCs.
  ! We also want to use this for all Traffic Classes including the best effort ones.
  ! Therefore we disable global learning.
  !
  traffic-class filter access-list DENY_GLOBAL_LEARN_LIST
  !
  !-------------------------------------------------------------------
  ! Service Class for Voice and Video traffic
  ! Control the max number of new TCs per learning period with 'count'
  ! Control the max total number of TCs in this learn-list with 'max'
  ! Define the Aggregation Mask
  !
  list seq 10 refname LEARN_VOICE_VIDEO
   traffic-class access-list VOICE_VIDEO filter BRANCH_PREFIX
   aggregation-type prefix-length 24
   count 2000 max 10000
   throughput
  !
  !-------------------------------------------------------------------
  ! Service Class for Business applications
  ! Control the max number of new TCs per learning period with 'count'
  ! Control the max total number of TCs in this learn-list with 'max'
  ! Define the Aggregation Mask
  !
  list seq 20 refname LEARN_CRITICAL
   traffic-class access-list CRITICAL filter BRANCH_PREFIX
   aggregation-type prefix-length 24
   count 2000 max 10000
   throughput
  !
  !-------------------------------------------------------------------
  ! Service Class for Best Effort applications
  ! Control the max number of new TCs per learning period with 'count'
  ! Control the max total number of TCs in this learn-list with 'max'
  ! Define the Aggregation Mask
  !
  list seq 30 refname LEARN_BEST_EFFORT
   traffic-class access-list BEST_EFFORT filter BRANCH_PREFIX
   aggregation-type prefix-length 24
   count 2000 max 10000
   throughput
 !
!
!-------------------------------------------------------------------
! ACL to deny global learning
!
ip access-list extended DENY_GLOBAL_LEARN_LIST
 deny   ip any any
!
!-------------------------------------------------------------------
! Voice and Video traffic classified based on DSCP EF, AF41 and CS4 (Tandberg)
!
```

```
ip access-list extended VOICE_VIDEO
 permit ip any any dscp ef
 permit ip any any dscp af41
 permit ip any any dscp cs4
!
!-----------------------------------------------------------------------
! Business application classified based on DSCP AF31
!
ip access-list extended CRITICAL
 permit ip any any dscp af31
!
!-----------------------------------------------------------------------
! Everything else is best effort
!
ip access-list extended BEST_EFFORT
 permit ip any any dscp default
!
!
!-----------------------------------------------------------------------
! Filter to track traffic going to the branch only.
!
ip prefix-list BRANCH_PREFIX seq 10 permit 10.1.0.0/16 ge 24
!
!
```

Important notes:

- A Traffic Class is an aggregation of individual flows based on destination prefix, DSCP and application name. Defining the appropriate aggregation mask length is key for TD operation to work properly and generates jitter probes to the remote sites.
- Aggregation mask should match the mask length advertized by Target Discovery (TD).
- In this case all branch sites advertized a /24 prefix. Therefore the aggregation mask is set to 24 (this is the default and won't appear in the final configuration.

Branch Offices

A very similar configuration applies to branches. Here is the configuration for R10:

```
!
pfr master
 mc-peer domain 65000 10.10.10.3 Loopback0
 target-discovery
 logging
 !
 border 10.2.10.10 key-chain pfr
  interface Tunnel100 external
   link-group SP1
  interface Tunnel200 external
   link-group SP2
  interface Ethernet0/0 internal
 !
 learn
  traffic-class filter access-list DENY_GLOBAL_LEARN_LIST
  list seq 10 refname LEARN_VOICE_VIDEO
   traffic-class access-list VOICE_VIDEO filter HQ_PREFIX
   count 2000 max 10000
   throughput
  list seq 20 refname LEARN_CRITICAL
   traffic-class access-list CRITICAL filter HQ_PREFIX
```

```
    count 2000 max 10000
    throughput
   list seq 30 refname LEARN_BEST_EFFORT
    traffic-class access-list BEST_EFFORT filter HQ_PREFIX
    count 2000 max 10000
    throughput
  !
!
ip access-list extended DENY_GLOBAL_LEARN_LIST
 deny   ip any any
!
ip access-list extended VOICE_VIDEO
 permit ip any any dscp ef
 permit ip any any dscp af41
 permit ip any any dscp cs4
!
ip access-list extended CRITICAL
 permit ip any any dscp af31
!
ip access-list extended BEST_EFFORT
 permit ip any any dscp default
!
!
ip prefix-list HQ_PREFIX seq 10 permit 10.10.0.0/16 ge 24
!
```

Policy Configuration

Monitoring Modes

We have 3 groups of applications that require different monitoring modes. See Monitoring Modes for more information.

For best effort applications, running passive mode is enough because traffic is mainly based on TCP. For voice, traffic is based on UDP (RTP) and we may want to get additional metrics like delay, jitter or even MOS. So we have to go with an active mode. For critical application we may also want to run in active mode.

To acquire jitter metrics, the router will need to inject instrumented synthetic traffic and derive the required information from there. With TCP traffic, we cannot make passive jitter measurements---only latency, loss and throughput. PfR makes use of the IOS IPSLA feature to generate probes and collect information about the state of the network.

In the case of jitter measurements, IPSLA requires that the probe destination needs to be an IPSLA responder. The IPSLA responder function is available on IOS devices as well as Cisco Telepresence units. The IPSLA responder is able to make accurate latency time stamping and has the vital ability of factoring out the processing time taken within the responder between reception of the probe and the generation of the probe reply.

Realistically, having access to a IPSLA responder along the path of a data traffic is usually only within the scope of private IP networks.

Note that if a traffic-class is monitored under active mode (which is completely active measurement based), passive measurements are disabled for that traffic-class. Notably, the throughput information is unavailable.

PfR offers an hybrid monitoring mode called ?active throughput? that will trigger OOP based on the active measurements, but throughput information is also collected. The CLI for this command is ?mode monitor active throughput? and is used for the critical applications in this lab.

Defining Advanced Policies per Group

PfR-maps are somewhat similar to route-maps in functionality in that a pfr-map will ?match? traffic and then apply (or in other words ?set?) a form of policy to that traffic. Now that the network is able to identify that traffic (via the ACLs and the learn-lists) we can configure the pfr-map entries. So far, PfR was optimizing all traffic based on the inherited global policies, which emphasized load sharing. Also, measurements were done using the default monitor mode of ?both? (meaning a combination of ?passive? and ?active?) for all traffic.

VIDEO group:

- Our voice/video traffic is matching DSCP EF marked flows
- Latency, jitter and loss would be more important. Another possibility (rather than specifying delay, jitter and loss individually) is that we could have represented our performance target in terms of Mean Opinion Score (MOS), which is a composite metric based on loss, jitter, latency and the specific codec being used. Specifying the target in terms of MOS is much simpler and the reference being used is voice quality (for example, MOS 4 is toll quality voice service).
- Measurement mode should be ?fast? to be able to track delay and loss quickly and more accurately. So, jitter probe is needed and an explicit jitter configuration per branch would be required. Because we use Target Discovery, there is no need for jitter probe configuration anymore.

CRITICAL group:

- Our critical traffic class is matching DSCP AF31 marked flows.
- Latency and/or loss would be more important.
- Measurement mode should be ?active throughput? to be able to track delay and loss as well as getting the bandwidth per traffic class. So, jitter probe is needed and an explicit jitter configuration per branch would be required. Because we use Target Discovery, there is no need for jitter probe configuration anymore.

BE group:

- Best effort traffic on the other hand can still be optimized for load-balancing.
- Measurement mode can be passive.

Configuration

Policy configuration for controlling applications:

```
!-------------------------------------------------------------------------------
! Policies for Voice and Video
!
! ? match command is to specify that this policy should be applied
!   to all the traffic-classes learned under list LEARN_VOICE_VIDEO
!
! - monitor mode is set to fast. This means probe all external interfaces
!   all the time. When Out-of-Policy condition is detected on the current
```

```
!   exit results on alternate exit is available for quick decision. In other
!   modes alternate exits are probed only when current link is determined to
!   be OOP. The fast mode helps in switching the path quickly when the
!   problem is detected.
!
! - Re-evaluate exit every 90 sec (periodic 90)
!
! ? delay threshold is configured as 150 msec. The delay measured by PfR is Round-Trip-Time.
! - loss threshold is configured as 5%
! - jitter threshold is configured as 30 ms
!
! ? Probe frequency is set to 8 seconds and can be changed to a lesser value
!   if the application being controlled is critical.
!
! - Probes are automatically configured and generated by Target Discovery
!
!-------------------------------------------------------------------------------
!
pfr-map MYMAP 10
 match PfR learn list LEARN_VOICE_VIDEO
 set periodic 90
 set delay threshold 150
 set loss threshold 50000
 set jitter threshold 30
 set mode monitor fast
 set resolve delay priority 1 variance 5
 set resolve loss priority 2 variance 5
 set resolve jitter priority 3 variance 5
 set probe frequency 8
 set link-group SP1 fallback SP2
!
!
!-------------------------------------------------------------------------------
! Policies for Critical Data
!
! ? match command is to specify that this policy should be applied
!   to all the traffic-classes learned under list LEARN_CRITICAL
!
! - monitor mode is set to fast. This means probe all external interfaces
!   all the time. When Out-of-Policy condition is detected on the current
!   exit results on alternate exit is available for quick decision. In other
!   modes alternate exits are probed only when current link is determined to
!   be OOP. The fast mode helps in switching the path quickly when the
!   problem is detected.
!
! - Re-evaluate exit every 90 sec (periodic 90)
!
! ? delay threshold is configured as 200 msec. The delay measured by PfR is Round-Trip-Time.
!
! ? Probe frequency is set to 8 seconds and can be changed to a lesser value
!   if the application being controlled is critical.
!
! - Probes are automatically configured and generated by Target Discovery
!
!-------------------------------------------------------------------------------
!
pfr-map MYMAP 20
 match pfr learn list LEARN_CRITICAL
 set periodic 90
 set delay threshold 200
 set loss threshold 50000
 set mode monitor fast
 set resolve delay priority 1 variance 20
 set resolve loss priority 5 variance 10
```

```
 set probe frequency 8
 set link-group SP1 fallback SP2
!
!
!-------------------------------------------------------------------------------
! Policies for Best Effort
!
! ? match command is to specify that this policy should be applied
!   to all the traffic-classes learned under list LEARN_BEST_EFFORT
!
! - monitor mode is set to both. This is the default mode. Monitoring is passive and
!    echo probes are used to help finding a new exit when a TC is OOP
!
! - Default policies used: link utilization first then load-balancing
!
!-------------------------------------------------------------------------------
!
pfr-map MYMAP 30
 match pfr learn list LEARN_BEST_EFFORT
 set periodic 90
 set mode monitor both
!
```

Now that the policies are defined using pfr-map, we need to apply them on the PfR MC for them to take effect. This is done through the ?policy-rules? command under ?pfr master? global mode:

```
!
pfr master
 max-range-utilization percent 30
 policy-rules MYMAP
!
```

Check Master Controllers

Check Status

Verify the Border Routers, verify the parameters used (default and configured) and check the learn-list.

```
R3#sh pfr master
OER state: ENABLED and ACTIVE
  Conn Status: SUCCESS, PORT: 3949
  Version: 3.3
  Number of Border routers: 2
  Number of Exits: 2
  Number of monitored prefixes: 16 (max 5000)
  Max prefixes: total 5000 learn 2500
  Prefix count: total 16, learn 12, cfg 0
  PBR Requirements met
  Nbar Status: Inactive

Border           Status                UP/DOWN             AuthFail  Version  DOWN Reason
10.10.10.5       ACTIVE                UP       1w2d          0  3.3
10.10.10.4       ACTIVE                UP       1w2d          0  3.3
```

```
Global Settings:
  max-range-utilization percent 30 recv 0
  rsvp post-dial-delay 0 signaling-retries 1
  mode route metric bgp local-pref 5000
  mode route metric static tag 5000
  mode route protocol pbr
  trace probe delay 1000
  logging
  exit holddown time 60 secs, time remaining 0

Default Policy Settings:
  backoff 90 900 90
  delay relative 50
  holddown 90
  periodic 90
  probe frequency 56
  number of jitter probe packets 100
  mode route control
  mode monitor both
  loss relative 10
  jitter threshold 20
  mos threshold 3.60 percent 30
  unreachable relative 50
  trigger-log percentage 30

Learn Settings:
  current state : STARTED
  time remaining in current state : 118 seconds
  throughput
  no delay
  no inside bgp
  traffic-class filter access-list DENY_GLOBAL_LEARN_LIST
  monitor-period 1
  periodic-interval 0
  aggregation-type prefix-length 24
  prefixes 100 appls 100
  expire after time 720

  Learn-List seq 10 refname LEARN_VOICE_VIDEO
    Configuration:
     Traffic-Class Access-list: VOICE_VIDEO
     Filter: BRANCH_PREFIX
     Aggregation-type: prefix-length 24
     Learn type: throughput
     Session count: 2000 Max count: 10000
     Policies assigned: 10
     Status: ACTIVE
    Stats:
     Traffic-Class Count: 4
  Learn-List seq 20 refname LEARN_CRITICAL
    Configuration:
     Traffic-Class Access-list: CRITICAL
     Filter: BRANCH_PREFIX
     Aggregation-type: prefix-length 24
     Learn type: throughput
     Session count: 2000 Max count: 10000
     Policies assigned: 20
     Status: ACTIVE
    Stats:
     Traffic-Class Count: 4
  Learn-List seq 30 refname LEARN_BEST_EFFORT
    Configuration:
     Traffic-Class Access-list: BEST_EFFORT
```

```
     Filter: BRANCH_PREFIX
     Aggregation-type: prefix-length 24
     Learn type: throughput
     Session count: 2000 Max count: 10000
     Policies assigned: 30
     Status: ACTIVE
    Stats:
     Traffic-Class Count: 4
R3#
```

What to check:

- Both Border Routers are up and running
- Check to make sure that PBR Requirements are met
- Number of automatically learned applications ? 4 in this case for Voice, 4 for Critical and 4 for best effort. We have 4 branch offices and aggregation mask is per branch.
- All default policy settings are displayed
- Learn is started (current state : STARTED)
- Learn-list status is ACTIVE and learn-type is throughput
- Make sure that all the configured learn lists show up
- For each learn-list, check the Traffic Class access-list (if configured), prefix-list (if configured) and the policy number associated (which is the pfr-map it refers to).
- The Traffic Class count is also displayed which allows you to check whether the learning process worked well.

Check the policies:

```
R3#sh pfr master policy
Default Policy Settings:
  backoff 90 900 90
  delay relative 50
  holddown 90
  periodic 90
  probe frequency 56
  number of jitter probe packets 100
  mode route control
  mode monitor both
  loss relative 10
  jitter threshold 20
  mos threshold 3.60 percent 30
  unreachable relative 50
  trigger-log percentage 30
oer-map MYMAP 10
  sequence no. 8444249301975040, provider id 1, provider priority 30
    host priority 0, policy priority 10, Session id 0
  match oer learn list LEARN_VOICE_VIDEO
  backoff 90 900 90
 *delay threshold 150
  holddown 90
 *periodic 90
 *probe frequency 8
  number of jitter probe packets 100
  mode route control
 *mode monitor fast
 *loss threshold 50000
 *jitter threshold 30
```

```
  mos threshold 3.60 percent 30
  unreachable relative 50
  next-hop not set
  forwarding interface not set
  trigger-log percentage 30
 *resolve delay priority 1 variance 5
 *resolve loss priority 2 variance 5
 *resolve jitter priority 3 variance 5
 *link-group SP1 fallback SP2
oer-map MYMAP 20
  sequence no. 8444249302630400, provider id 1, provider priority 30
    host priority 0, policy priority 20, Session id 0
  match oer learn list LEARN_CRITICAL
  backoff 90 900 90
 *delay threshold 200
  holddown 90
 *periodic 90
 *probe frequency 8
  number of jitter probe packets 100
  mode route control
 *mode monitor fast
 *loss threshold 50000
  jitter threshold 20
  mos threshold 3.60 percent 30
  unreachable relative 50
  next-hop not set
  forwarding interface not set
  trigger-log percentage 30
 *resolve delay priority 1 variance 20
 *resolve loss priority 5 variance 10
 *link-group SP1 fallback SP2
oer-map MYMAP 30
  sequence no. 8444249303285760, provider id 1, provider priority 30
    host priority 0, policy priority 30, Session id 0
  match oer learn list LEARN_BEST_EFFORT
  backoff 90 900 90
  delay relative 50
  holddown 90
 *periodic 90
  probe frequency 56
  number of jitter probe packets 100
  mode route control
 *mode monitor both
  loss relative 10
  jitter threshold 20
  mos threshold 3.60 percent 30
  unreachable relative 50
  next-hop not set
  forwarding interface not set
  trigger-log percentage 30

* Overrides Default Policy Setting
R3#
```

Notes:

- Default policies that will be applied are somewhat similar to the concept of default-class in QoS. All Traffic that does not match learn-list falls under global policies. Load balancing is enabled by default for this kind of traffic. In this solution guide we have disable global learning.

- Then note the additional section under ?pfr-map MYMAP 10?, ?pfr-map MYMAP 20? and ?pfr-map MYMAP 30? (just like route-maps, we can have multiple stanzas of match critera and policies).
- Note how the pfr-map inherits the policy settings from default for parameters that were not explicitly configured under the pfr-map entry.

Check Target Discovery and Peering

On the central site:

```
R3#sh pfr master target-discovery
PfR Target-Discovery Services
 Mode: Static  Domain: 65000
 Responder list: RESPONDER_PREFIX  Inside-prefixes list: HQ_PREFIX
 SvcRtg: client-handle: 1  sub-handle: 1  pub-seq: 1

PfR Target-Discovery Database (local)

 Local-ID: 10.10.10.3        Desc: R3
   Target-list: 10.10.10.2
   Prefix-list: 10.10.4.0/24, 10.10.3.0/24, 10.10.2.0/24, 10.10.1.0/24

PfR Target-Discovery Database (remote)

 MC-peer: 10.2.10.10         Desc: R10
   Target-list: 10.1.10.254
   Prefix-list: 10.1.10.0/24

 MC-peer: 10.2.11.11         Desc: R11
   Target-list: 10.1.11.254
   Prefix-list: 10.1.11.0/24

 MC-peer: 10.2.12.12         Desc: R12
   Target-list: 10.1.12.254
   Prefix-list: 10.1.12.0/24

 MC-peer: 10.2.13.13         Desc: R13
   Target-list: 10.1.13.254
   Prefix-list: 10.1.13.0/24

R3#
```

What to Check:

- You get the list of all remote sites
- You get the list of destination prefixes per remote site
- You also get the probe target addresses per remote site (there could be multiple responders within a given site)

You can display the state of active probes on the MC.
Remember that IP SLA probes are generated by the BRs (R4 and R5), which report back the results to the MC.

```
R3#sh pfr master active-probes target-discovery
PfR Master Controller active-probes (TD)
Border  = Border Roter running this probe
MC-Peer = Remote MC associated with this target
Type    = Probe Type
Target  = Target Address
TPort   = Target Port
N - Not applicable

Destination Site Peer Addresses:

MC-Peer           Targets
10.2.10.10        10.1.10.254
10.2.11.11        10.1.11.254
10.2.12.12        10.1.12.254
10.2.13.13        10.1.13.254

The following Probes are running:

Border          Idx  State     MC-Peer            Type     Target           TPort
10.10.10.5       8   TD-Actv   10.2.10.10         jitter   10.1.10.254      5000
10.10.10.4       8   TD-Actv   10.2.10.10         jitter   10.1.10.254      5000
10.10.10.5       8   TD-Actv   10.2.10.10         jitter   10.1.10.254      5000
10.10.10.4       8   TD-Actv   10.2.10.10         jitter   10.1.10.254      5000
10.10.10.4       8   TD-Actv   10.2.11.11         jitter   10.1.11.254      5000
10.10.10.5       8   TD-Actv   10.2.11.11         jitter   10.1.11.254      5000
10.10.10.5       8   TD-Actv   10.2.11.11         jitter   10.1.11.254      5000
10.10.10.4       8   TD-Actv   10.2.11.11         jitter   10.1.11.254      5000
10.10.10.5       8   TD-Actv   10.2.12.12         jitter   10.1.12.254      5000
10.10.10.4       8   TD-Actv   10.2.12.12         jitter   10.1.12.254      5000
10.10.10.5       8   TD-Actv   10.2.12.12         jitter   10.1.12.254      5000
10.10.10.4       8   TD-Actv   10.2.12.12         jitter   10.1.12.254      5000
10.10.10.4       8   TD-Actv   10.2.13.13         jitter   10.1.13.254      5000
10.10.10.5       8   TD-Actv   10.2.13.13         jitter   10.1.13.254      5000
10.10.10.5       8   TD-Actv   10.2.13.13         jitter   10.1.13.254      5000
10.10.10.4       8   TD-Actv   10.2.13.13         jitter   10.1.13.254      5000


R3#
```

Check Target Discovery on the branch site R10:

```
R10#sh pfr master target-discovery
PfR Target-Discovery Services
 Mode: Dynamic  Domain: 65000
 SvcRtg: client-handle: 1  sub-handle: 1  pub-seq: 1

PfR Target-Discovery Database (local)

 Local-ID: 10.2.10.10        Desc: R10
   Target-list: 10.1.10.254
   Prefix-list: 10.1.10.0/24

PfR Target-Discovery Database (remote)

 MC-peer: 10.10.10.3         Desc: R3
   Target-list: 10.10.10.2
   Prefix-list: 10.10.4.0/24, 10.10.3.0/24, 10.10.2.0/24, 10.10.1.0/24
```

```
 MC-peer: 10.2.11.11         Desc: R11
   Target-list: 10.1.11.254
   Prefix-list: 10.1.11.0/24

 MC-peer: 10.2.12.12         Desc: R12
   Target-list: 10.1.12.254
   Prefix-list: 10.1.12.0/24

 MC-peer: 10.2.13.13         Desc: R13
   Target-list: 10.1.13.254
   Prefix-list: 10.1.13.0/24

R10#
```

Note:

- R10 has only one peer which is the hub R3
- But R10 learns all prefixes from all sites, including the other spokes. This information is reflected by the hub MC (R3).

You can display the state of active probes on the MC.
Remember that IP SLA probes are generated by the BRs, which report back the results to the MC.

```
R10#sh pfr master active-probes target-discovery
PfR Master Controller active-probes (TD)
Border  = Border Roter running this probe
MC-Peer = Remote MC associated with this target
Type    = Probe Type
Target  = Target Address
TPort   = Target Port
N - Not applicable

Destination Site Peer Addresses:

MC-Peer           Targets
10.10.10.3        10.10.10.2
10.2.11.11        10.1.11.254
10.2.12.12        10.1.12.254
10.2.13.13        10.1.13.254

The following Probes are running:

Border          Idx  State     MC-Peer            Type     Target           TPort
10.2.10.10       9   TD-Actv   10.10.10.3         jitter   10.10.10.2       5000
10.2.10.10       8   TD-Actv   10.10.10.3         jitter   10.10.10.2       5000
10.2.10.10       9   TD-Actv   10.10.10.3         jitter   10.10.10.2       5000
10.2.10.10       8   TD-Actv   10.10.10.3         jitter   10.10.10.2       5000


R10#
```

Note:

- Multiple probe types generated: echo for critical applications (no need for jitter information) and jitter probes (voice/video traffic).
- Target Discovery will only configure and generate jitter probes for fast mode.
- The probe packets will automatically take on the DSCP properties of the traffic class. In the case that the traffic class type has multiple DSCP settings (which is very likely when the traffic class is learned rather than explicitly specified), a unique probe for each DSCP value will be generated.
    - Voice:
        - DSCP EF (46, 0x2E)
        - TOS = 0xB8, 184
    - Business:
        - DSCP AF31 (26, 0x1A)
        - TOS = 0x68, 104
- Note that probes can be generated between spokes if there voice or critical traffic between the spokes. When there is no traffic, or if voice/critical traffic stops, then probe generation between spokes will also stop.

Notes specific to Target Discovery probes

- Traffic classes going to the same remote site share probe statistics
- Only one set of probes will be run on each BR even thought multiple applications to the same site are being monitored by PfR
- Probes to TD targets are always of type jitter
- If there are multiple targets advertised by a remote site, one probes to all targets will be run simultaneously

Verify Traffic Class and Statistics

As soon as you see:

```
R3#
*Sep 20 20:24:43.515: %OER_MC-5-NOTICE: Prefix Learning WRITING DATA
*Sep 20 20:24:43.589: %OER_MC-5-NOTICE: Prefix Learning STARTED
R3#
```

You should be able to see the traffic classes on the Master Controller. You will need a few cycles before having all applications in INPOLICY state.

Check if applications are automatically learnt under each learn-list

```
R3#sh pfr master learn list

 Learn-List seq 10 refname LEARN_VOICE_VIDEO
   Configuration:
    Traffic-Class Access-list: VOICE_VIDEO
    Filter: BRANCH_PREFIX
    Aggregation-type: prefix-length 24
```

```
    Learn type: throughput
    Session count: 2000 Max count: 10000
    Policies assigned: 10
    Status: ACTIVE
   Stats:
    Traffic-Class Count: 4
    Traffic-Class Learned:
     Appl Prefix 10.1.11.0/24 ef   256
     Appl Prefix 10.1.13.0/24 ef   256
     Appl Prefix 10.1.10.0/24 ef   256
     Appl Prefix 10.1.12.0/24 ef   256
 Learn-List seq 20 refname LEARN_CRITICAL
   Configuration:
    Traffic-Class Access-list: CRITICAL
    Filter: BRANCH_PREFIX
    Aggregation-type: prefix-length 24
    Learn type: throughput
    Session count: 2000 Max count: 10000
    Policies assigned: 20
    Status: ACTIVE
   Stats:
    Traffic-Class Count: 4
    Traffic-Class Learned:
     Appl Prefix 10.1.12.0/24 af31 256
     Appl Prefix 10.1.13.0/24 af31 256
     Appl Prefix 10.1.11.0/24 af31 256
     Appl Prefix 10.1.10.0/24 af31 256
 Learn-List seq 30 refname LEARN_BEST_EFFORT
   Configuration:
    Traffic-Class Access-list: BEST_EFFORT
    Filter: BRANCH_PREFIX
    Aggregation-type: prefix-length 24
    Learn type: throughput
    Session count: 2000 Max count: 10000
    Policies assigned: 30
    Status: ACTIVE
   Stats:
    Traffic-Class Count: 4
    Traffic-Class Learned:
     Appl Prefix 10.1.12.0/24 defa 256
     Appl Prefix 10.1.13.0/24 defa 256
     Appl Prefix 10.1.11.0/24 defa 256
     Appl Prefix 10.1.10.0/24 defa 256
R3#
```

Verify learnt Traffic Classes

On the Master Controller, you have all the Traffic Classes (learnt as well as statistics): Use show pfr master traffic-class to see the state of the application:

```
R3#sh pfr master traffic-class
OER Prefix Statistics:
 Pas - Passive, Act - Active, S - Short term, L - Long term, Dly - Delay (ms),
 P - Percentage below threshold, Jit - Jitter (ms),
 MOS - Mean Opinion Score
 Los - Packet Loss (percent/10000), Un - Unreachable (flows-per-million),
 E - Egress, I - Ingress, Bw - Bandwidth (kbps), N - Not applicable
 U - unknown, * - uncontrolled, + - control more specific, @ - active probe all
```

```
 # - Prefix monitor mode is Special, & - Blackholed Prefix
 % - Force Next-Hop, ^ - Prefix is denied

DstPrefix           Appl_ID Dscp Prot     SrcPort     DstPort SrcPrefix
           Flags              State     Time            CurrBR  CurrI/F Protocol
         PasSDly  PasLDly   PasSUn   PasLUn  PasSLos  PasLLos      EBw      IBw
         ActSDly  ActLDly   ActSUn   ActLUn  ActSJit  ActPMOS  ActSLos  ActLLos
--------------------------------------------------------------------------------
10.1.10.0/24              N defa  256           N           N 0.0.0.0/0
                          INPOLICY       47        10.10.10.5 Tu200           PBR
             110      110        0        0        0        0      171       13
             108      108        0        0        3        0        0        0

10.1.10.0/24              N af31  256           N           N 0.0.0.0/0
                          INPOLICY      @47        10.10.10.4 Tu100           PBR
              57       57        0        0        0        0       84        6
              57       56        0        0        3        0        0        0

10.1.10.0/24              N   ef  256           N           N 0.0.0.0/0
                          INPOLICY      @16        10.10.10.4 Tu100           PBR
               U        U        0        0        0        0        1        1
              56       56        0        0        3        0        0        0

10.1.11.0/24              N defa  256           N           N 0.0.0.0/0
                          INPOLICY       62        10.10.10.5 Tu200           PBR
             110      111        0        0        0        0      173       12
             108      107        0        0        3        0        0        0

10.1.11.0/24              N af31  256           N           N 0.0.0.0/0
                          INPOLICY      @39        10.10.10.4 Tu100           PBR
              57       57        0        0        0        0       77        6
              57       56        0        0        3        0        0        0

10.1.11.0/24              N   ef  256           N           N 0.0.0.0/0
                          INPOLICY      @12        10.10.10.4 Tu100           PBR
               U        U        0        0        0        0        1        1
              55       55        0        0        3        0        0        0

10.1.12.0/24              N defa  256           N           N 0.0.0.0/0
                          INPOLICY        5        10.10.10.4 Tu100           PBR
              59       59        0        0        0        0      148       12
               U       55        0        0        0        U        0        0

10.1.12.0/24              N af31  256           N           N 0.0.0.0/0
                          INPOLICY      @23        10.10.10.4 Tu100           PBR
              59       59        0        0        0        0       83        6
              55       55        0        0        3        0        0        0

10.1.12.0/24              N   ef  256           N           N 0.0.0.0/0
                          INPOLICY       @4        10.10.10.4 Tu100           PBR
               U        U        0        0        0        0        1        1
              56       55        0        0        3        0        0        0

10.1.13.0/24              N defa  256           N           N 0.0.0.0/0
                          INPOLICY       42        10.10.10.5 Tu200           PBR
             110      111        0        0        0        0      174       12
             107      107        0        0        3        0        0        0

10.1.13.0/24              N af31  256           N           N 0.0.0.0/0
                          INPOLICY      @29        10.10.10.4 Tu100           PBR
              58       57        0        0        0        0       88        6
              56       56        0        0        3        0        0        0

10.1.13.0/24              N   ef  256           N           N 0.0.0.0/0
```

```
                        INPOLICY      @32        10.10.10.4 Tu100           PBR
              U        U        0        0        0        0        1        1
             56       55        0        0        3        0        0        0

R3#
```

Notes:

- Note the DSCP values under the DSCP field column.
- The TCs belonging to CRITICAL and VOICE_VIDEO learn-list are being forced out R4 as it has a delay within the defined threshold to the remote branch and because the preferred path is SP1 (note that R5 would be a valid choice too without the link-group policy).
- The TCs belonging to BEST_EFFORT are supposed to be load-balanced across both Border Routers and exit links.
- Note the path enforcement used.

You can also display more details to check the probe used and latest statistics per Traffic Class:

```
R3#sh pfr master traffic-class detail
OER Prefix Statistics:
 Pas - Passive, Act - Active, S - Short term, L - Long term, Dly - Delay (ms),
 P - Percentage below threshold, Jit - Jitter (ms),
 MOS - Mean Opinion Score
 Los - Packet Loss (percent/10000), Un - Unreachable (flows-per-million),
 E - Egress, I - Ingress, Bw - Bandwidth (kbps), N - Not applicable
 U - unknown, * - uncontrolled, + - control more specific, @ - active probe all
 # - Prefix monitor mode is Special, & - Blackholed Prefix
 % - Force Next-Hop, ^ - Prefix is denied

DstPrefix           Appl_ID Dscp Prot     SrcPort     DstPort SrcPrefix
           Flags             State     Time            CurrBR  CurrI/F Protocol
         PasSDly  PasLDly   PasSUn   PasLUn  PasSLos  PasLLos      EBw      IBw
         ActSDly  ActLDly   ActSUn   ActLUn  ActSJit  ActPMOS  ActSLos  ActLLos
--------------------------------------------------------------------------------

[SNIP]

--------------------------------------------------------------------------------

Prefix: 10.1.10.0/24  Protocol: 256  Port: [0, 0] [0, 0] DSCP: af31
   State: INPOLICY    Time Remaining: @66
   Policy: MYMAP 20

   Most recent data per exit
   Border          Interface           PasSDly    PasLDly    ActSDly    ActLDly
  *10.10.10.4      Tu100                    57         57         55         56
   10.10.10.5      Tu200                     0          0        107        107

   Most recent voice data per exit
   Border          Interface         ActSJit  ActPMOS  ActSLos  ActLLos
  *10.10.10.4      Tu100                   3        0        0        0
   10.10.10.5      Tu200                   3        0        0        0

   Latest Active Stats on Current Exit:
   Type     Target          TPort Attem Comps    DSum     Min     Max     Dly
   jitter   10.1.10.254      5000     1   100    5689      44     131      56
   jitter   10.1.10.254      5000     1   100    5517      44     131      55
   jitter   10.1.10.254      5000     1   100    5530      44     131      55
```

```
   jitter   10.1.10.254      5000     1   100    5573      44     131      55
   jitter   10.1.10.254      5000     1   100    5621      44     131      56


   Latest Active Voice Stats on Current Exit:
   Type     Target          TPort     Codec Attem Comps  JitSum      MOS
   jitter   10.1.10.254     5000      g729a     1   100     354     4.06
   jitter   10.1.10.254     5000      g729a     1   100     317     4.06
   jitter   10.1.10.254     5000      g729a     1   100     329     4.06
   jitter   10.1.10.254     5000      g729a     1   100     376     4.06
   jitter   10.1.10.254     5000      g729a     1   100     373     4.06

Prefix performance history records
 Current index 39, S_avg interval(min) 5, L_avg interval(min) 60

Age       Border          Interface       OOP/RteChg Reasons
Pas: DSum  Samples  DAvg  PktLoss  Unreach   Ebytes   Ibytes     Pkts    Flows
Act: Dsum Attempts  DAvg    Comps  Unreach   Jitter LoMOSCnt   MOSCnt
00:00:11  10.10.10.4      Tu100
Pas: 4410       76    58        0        0   363642    59782     1397      145
Act: 5689        1    56      100        0        3        0        1        0
00:00:22  10.10.10.4      Tu100
Pas:    0        0     0        0        0        0        0        0        0
Act:11047        2    55      200        0        3        0        2        0
00:00:30  10.10.10.4      Tu100
Pas:    0        0     0        0        0        0        0        0        0
Act: 5573        1    55      100        0        3        0        1        0


[SNIP]

R3#
```

A closer look at the results

You can also have all information related to a specific traffic class with the following command:

```
R3#sh pfr master traffic-class performance ip any 10.1.10.0/24 dscp ef
==========================================================================================

Traffic-class:
 Destination Prefix : 10.1.10.0/24            Source Prefix    : 0.0.0.0/0
 Destination Port   : N                       Source Port      : N
 DSCP               : ef                      Protocol         : 256
 Application Name:  : N/A

 General:
   Control State                   : Controlled using PBR
   Traffic-class status            : INPOLICY
   Current Exit                    : BR 10.10.10.4 interface Tu100, Tie breaker was None
   Time on current exit            : 0d 0:29:43
   Time remaining in current state : @71 seconds
   Last uncontrol reason           : Couldn't choose exit in prefix timeout
   Time since last uncontrol       : 4d 2:18:11
   Traffic-class type              : Learned
   Target-Discovery origin         : 10.2.10.10
   Improper config                 : None
```

```
 Last Out-of-Policy event:
   No Out-of-Policy Event

 Average Passive Performance Current Exit: (Average for last 5 minutes)
   Unreachable            : 0% -- Threshold: 50%
   Delay                  : 0 msecs -- Threshold: 150 msecs
   Loss                   : 0 -- Threshold: 50000
   Egress BW              : 1 kbps
   Ingress BW             : 1 kbps
   Time since last update : 0d 0:0:7

 Average Active Performance Current Exit: (Average for last 5 minutes)
   Unreachable            : 0% -- Threshold: 50%
   Delay                  : 57 msec -- Threshold: 150 msec
   Loss                   : 0 -- Threshold: 50000
   Jitter                 : 397 msec -- Threshold: 3000 msec

 Last Resolver Decision:
   BR              Interface    Status       Reason       Performance Threshold
   --------------- ------------ ------------ ------------ ----------- ---------
   10.10.10.5      Tu200        Eliminated   Link Group   N/A          N/A
   10.10.10.4      Tu100        Best Exit    Unreachable  N/A          N/A

==========================================================================================
R3#
```

Note:

- You can check that the last resolver reason was link-group. This TC was moved to DMVPN1 because that is the preferred path for voice and critical traffic.

Monitoring Load Balancing.

Traffic other than voice/video and critical application is load balanced across all external interfaces. To visualize the accuracy of the load balancing, the following command can be used:

```
R3#sh pfr master exits
==========================================================================================
PfR Master Controller Exits:

General Info:
=============
  E - External
  I - Internal
  N/A - Not Applicable
                                                                                    Up/
   ID Name         Border          Interface   ifIdx IP Address      Mask Policy      Type Down
  --- ------------ --------------- ----------- ----- --------------- ---- ----------- ---- ----
    2              10.10.10.5      Tu200           8 10.0.200.5        24 Util          E  UP
    1              10.10.10.4      Tu100           8 10.0.100.4        24 Util          E  UP

Global Exit Policy:
===================
            Cost:      In Policy

Exits Performance:
==================
```

```
                  Egress                                                  Ingress
    ------------------------------------------------------ ------------------------------------
 ID Capacity  MaxUtil    Usage   %      RSVP POOL    OOP Capacity  MaxUtil    Usage   %  OOP
 --- -------- -------- -------- --- -------------- ----- -------- -------- -------- --- -----
  2     1000      900      782  78            N/A   Util     1000     1000      202  20  N/A
  1     1000      900      785  78            N/A   Util     1000     1000      188  18  N/A

TC and BW Distribution:
=======================
                       # of TCs                  BW (kbps)            Probe   Active
     Name/ID   Current Controlled InPolicy    Controlled       Total  Failed  Unreach
                                                                      (count) (fpm)
        ----   ----------------------------   ----------------------  ------  --------
           2        3          3        3          514          782 4294966900          0
           1        9          9        9          505          785 4294963138          0

Exit Related TC Stats:
======================
                                 Priority
                             highest       nth
                            ------------------
  Number of TCs with range:        0         0
   Number of TCs with util:        0         0
   Number of TCs with cost:        0         0

       Total number of TCs:       16
R3#
```

Notes:

- You can check the bandwidth usage on all external interfaces
- You can also check the number of TCs on each link.

Check Border Routers (BRs)

Active Probing

Check Active probing status on Border Router R4:

```
R4#sh pfr border active-probes
        OER Border active-probes
Type      = Probe Type
Target    = Target IP Address
TPort     = Target Port
Source    = Send From Source IP Address
Interface = Exit interface
Att       = Number of Attempts
Comps   = Number of completions
N - Not applicable

Type     Target          TPort Source          Interface           Att   Comps
DSCP
jitter   10.1.12.254      5000 10.0.100.4      Tu100                 1     100
0
```

```
jitter   10.1.12.254      5000 10.0.100.4      Tu100                276   27477
104
jitter   10.1.13.254      5000 10.0.100.4      Tu100                285   28419
104
jitter   10.1.11.254      5000 10.0.100.4      Tu100                285   28391
104
jitter   10.1.10.254      5000 10.0.100.4      Tu100                286   28551
104
jitter   10.1.12.254      5000 10.0.100.4      Tu100                160   15982
184
jitter   10.1.10.254      5000 10.0.100.4      Tu100                160   15920
184
jitter   10.1.11.254      5000 10.0.100.4      Tu100                162   16200
184
jitter   10.1.13.254      5000 10.0.100.4      Tu100                163   16300
184


R4#
```

Check Active probing status on Border Router R5:

```
R5#sh pfr border active-probes
        OER Border active-probes
Type      = Probe Type
Target    = Target IP Address
TPort     = Target Port
Source    = Send From Source IP Address
Interface = Exit interface
Att       = Number of Attempts
Comps   = Number of completions
N - Not applicable

Type     Target          TPort Source          Interface           Att   Comps
DSCP
jitter   10.1.10.254      5000 10.0.200.5      Tu200                 0       0
0
jitter   10.1.13.254      5000 10.0.200.5      Tu200                 0       0
0
jitter   10.1.12.254      5000 10.0.200.5      Tu200                279   27673
104
jitter   10.1.13.254      5000 10.0.200.5      Tu200                289   28810
104
jitter   10.1.11.254      5000 10.0.200.5      Tu200                289   28756
104
jitter   10.1.10.254      5000 10.0.200.5      Tu200                291   29100
104
jitter   10.1.12.254      5000 10.0.200.5      Tu200                161   15937
184
jitter   10.1.10.254      5000 10.0.200.5      Tu200                165   16407
184
jitter   10.1.11.254      5000 10.0.200.5      Tu200                165   16406
184
jitter   10.1.13.254      5000 10.0.200.5      Tu200                167   16700
184


R5#
```

You can check the IP SLA configuration:

```
R5#sh ip sla configuration
IP SLAs Infrastructure Engine-III
Entry number: 27885
Owner: Optimized Edge Routing (OER)
Tag:
Operation timeout (milliseconds): 4000
Type of operation to perform: udp-jitter
Target address/Source address: 10.1.13.254/10.0.200.5
Target port/Source port: 5000/0
Type Of Service parameter: 0xB8
Codec Type: g729a
Codec Number Of Packets: 100
Codec Packet Size: 32
Codec Interval (milliseconds): 20
Advantage Factor: 0
Verify data: No
Vrf Name:
Control Packets: enabled
Schedule:
   Operation frequency (seconds): 8  (not considered if randomly scheduled)
   Next Scheduled Start Time: Start Time already passed
   Group Scheduled : FALSE
   Randomly Scheduled : FALSE
   Life (seconds): Forever
   Entry Ageout (seconds): never
   Recurring (Starting Everyday): FALSE
   Status of entry (SNMP RowStatus): Active
Threshold (milliseconds): 4000
Distribution Statistics:
   Number of statistic hours kept: 2
   Number of statistic distribution buckets kept: 1
   Statistic distribution interval (milliseconds): 20
Enhanced History:
Percentile:
Entry number: 27887

[SNIP]

R4#
```

Notes:

- Each probe has its own entry number.
- These probes are managed by PfR (means there is no configuration done directly on the router)
- You can check the probe type (echo or jitter)
- For jitter you can also check the codec used, as well as packet size, interval, etc

Detailed statistics can be displayed using the following command:

```
R5#sh ip sla statistics 27885
IPSLAs Latest Operation Statistics

IPSLA operation id: 27885
Type of operation: udp-jitter
        Latest RTT: 108 milliseconds
Latest operation start time: 18:26:46 CET Mon Feb 3 2014
Latest operation return code: OK
RTT Values:
        Number Of RTT: 100              RTT Min/Avg/Max: 101/108/140 milliseconds
Latency one-way time:
        Number of Latency one-way Samples: 100
        Source to Destination Latency one way Min/Avg/Max: 100/105/118 milliseconds
        Destination to Source Latency one way Min/Avg/Max: 0/2/31 milliseconds
Jitter Time:
        Number of SD Jitter Samples: 99
        Number of DS Jitter Samples: 99
        Source to Destination Jitter Min/Avg/Max: 0/4/14 milliseconds
        Destination to Source Jitter Min/Avg/Max: 0/3/26 milliseconds
Over Threshold:
        Number Of RTT Over Threshold: 0 (0%)
Packet Loss Values:
        Loss Source to Destination: 0
        Source to Destination Loss Periods Number: 0
        Source to Destination Loss Period Length Min/Max: 0/0
        Source to Destination Inter Loss Period Length Min/Max: 0/0
        Loss Destination to Source: 0
        Destination to Source Loss Periods Number: 0
        Destination to Source Loss Period Length Min/Max: 0/0
        Destination to Source Inter Loss Period Length Min/Max: 0/0
        Out Of Sequence: 0      Tail Drop: 0
        Packet Late Arrival: 0  Packet Skipped: 0
Voice Score Values:
        Calculated Planning Impairment Factor (ICPIF): 12
        MOS score: 4.03
Number of successes: 179
Number of failures: 0
Operation time to live: Forever


R5#
```

Path Enforcement

How does PfR modify Paths on the Border Routers:

PfR will create dynamic route-maps (dynamic PBR) on both Borders Routers (R4 and R5) to forward any flows matching the application:

- over the PFR configured internal interfaces to the forwarding BR if it's the non-forwarding BR.
- or over the external interfaces to the WAN if it is the forwarding BR.

In-order for the PBR based route control to be successful, the BRs have to be adjacent either through a direct connection or one hop away. This can be verified with the following command:

```
R3#sh pfr master
OER state: ENABLED and ACTIVE
  Conn Status: SUCCESS, PORT: 3949
  Version: 3.3
  Number of Border routers: 2
  Number of Exits: 2
  Number of monitored prefixes: 16 (max 5000)
  Max prefixes: total 5000 learn 2500
  Prefix count: total 16, learn 12, cfg 0
  PBR Requirements met                                        <---- HERE
  Nbar Status: Inactive

Border           Status                UP/DOWN             AuthFail  Version  DOWN Reason
10.10.10.5       ACTIVE                UP       1w2d          0  3.3
10.10.10.4       ACTIVE                UP       1w2d          0  3.3

[SNIP]
```

What to check:

- Look at PBR Requirement - should be met.

You can also check the Border Routers topology:

```
R3#sh pfr master border topology
Local Border   Local Interface   Remote Border  Remote Interface  Neighbor type
-------------------------------------------------------------------------------------------
10.10.10.5     Ethernet0/2       10.10.10.4     Ethernet0/2       Directly Connected
10.10.10.4     Ethernet0/2       10.10.10.5     Ethernet0/2       Directly Connected
PBR Requirements met
R3#
```

Dynamic route maps can be verified directly on the Border Routers:

```
R4#sh route-map dynamic detail
route-map OER_INTERNAL_RMAP, permit, sequence 0, identifier 1006633014
  Match clauses:
    ip address (access-lists): oer#54
      Extended IP access list oer#54
          268435455 permit ip any 10.1.13.0 0.0.0.255 dscp ef (2284 matches)
          536870911 permit ip any 10.1.13.0 0.0.0.255 dscp af31 (30461 matches)
          2147483647 deny ip any any (55143486 matches)
  Set clauses:
    ip next-hop 10.0.100.13
    interface Tunnel100
  Policy routing matches: 384616 packets, 48225450 bytes
route-map OER_INTERNAL_RMAP, permit, sequence 1, identifier 2600468535
  Match clauses:
    ip address (access-lists): oer#55
      Extended IP access list oer#55
          67108863 permit ip any 10.1.11.0 0.0.0.255 dscp ef (2292 matches)
          134217727 permit ip any 10.1.11.0 0.0.0.255 dscp af31 (30470 matches)
          2147483647 deny ip any any (54756190 matches)
```

```
  Set clauses:
    ip next-hop 10.0.100.11
    interface Tunnel100
  Policy routing matches: 387109 packets, 50487734 bytes
route-map OER_INTERNAL_RMAP, permit, sequence 2, identifier 1845493816
  Match clauses:
    ip address (access-lists): oer#56
      Extended IP access list oer#56
          67108863 permit ip any 10.1.10.0 0.0.0.255 dscp ef (2274 matches)
          134217727 permit ip any 10.1.10.0 0.0.0.255 dscp af31 (30473 matches)
          2147483647 deny ip any any (54371078 matches)
  Set clauses:
    ip next-hop 10.0.100.10
    interface Tunnel100
  Policy routing matches: 384141 packets, 48205222 bytes
route-map OER_INTERNAL_RMAP, permit, sequence 3, identifier 2835349561
  Match clauses:
    ip address (access-lists): oer#57
      Extended IP access list oer#57
          16777215 permit ip any 10.1.12.0 0.0.0.255 dscp ef (2300 matches)
          33554431 permit ip any 10.1.12.0 0.0.0.255 dscp af31 (29750 matches)
          67108863 permit ip any 10.1.12.0 0.0.0.255 dscp default (59833 matches)
          2147483647 deny ip any any (53927113 matches)
  Set clauses:
    ip next-hop 10.0.100.12
    interface Tunnel100
  Policy routing matches: 443570 packets, 101940153 bytes
route-map OER_INTERNAL_RMAP, permit, sequence 4, identifier 1342177339
  Match clauses:
    ip address (access-lists): oer#59
      Extended IP access list oer#59
          268435455 permit ip any 10.1.11.0 0.0.0.255 dscp default (30277 matches)
          536870911 permit ip any 10.1.13.0 0.0.0.255 dscp default (29916 matches)
          1073741823 permit ip any 10.1.10.0 0.0.0.255 dscp default (33010 matches)
          2147483647 deny ip any any (385202 matches)
  Set clauses:
    ip next-hop 10.10.45.5
    interface Ethernet0/2
  Policy routing matches: 93518 packets, 84717376 bytes
Current active dynamic routemaps = 1
R4#
```

What to check:

- 2 possible next-hops: either the direct WAN tunnel interface or the other BR
- oer#54, 55, 56 and 57 points directly to the tunnel end-points of each branch offices.
- oer#59 points to the other bR (R5). Note that the traffic matched is only DSCP 0 because voice and critical traffic have DMVPN1 has their preferred path

More Information

IWAN Cisco Validated designs can be found here.

Remote Sites Using Local Internet Access Technology Design Guide - December 2013

- This CVD complements the VPN Remote Site Design Guide and does not use Frontdoor-VRF for the DMVPN configs

Conclusion

Performance Routing is an advanced technology that allows multiple deployments, one of them being the one described in this document. The use of an overlay topology (DMVPN) over any given transport allows flexibility while keeping the configuration exactly the same for PfR.

The configuration is straightforward and very efficient. Based on that, a customer can evolve to a more complex Performance Routing solution or can fine-tune the Traffic Class definition.