PfR Basic Load Balancing using BGP

Navigation

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

Contents

PfR Solution Used

This solution describes a basic load sharing for an enterprise dual attached to two IP-VPN Service Providers. The load-balancing used here is taking place between external interfaces and will be based on the top talker prefixes.

- Learning phase: automatic. Top talkers based on Netflow reports from the Border Routers
- Measuring phase: mode both, which means passive monitoring based on Netflow and active probes when choosing new exits.
- Policies:

  utilization: Link utilization for each exit interface on the Border Routers should not exceed 90% of the defined bandwidth
  range: traffic should be load-balanced so that the average bandwidth range between external interfaces should be maintain within 10%.

- Enforcement: load-balancing based on prefixes and BGP used on the uplinks toward the Service Providers. PfR will enforce the path by modifying the BGP local-pref attribute for controlled prefixes.

PfR Features that Enable Load Balancing

Link Utilization

Usage of this policy sets an upper threshold on the amount of traffic a specific link can carry. For example, if the upper threshold for a link is 90 % of total bandwidth, and it is currently running at 95 % of bandwidth, the link is Out-of-Policy (OOP). Cisco PfR will attempt to bring the link back into policy by repeatedly moving prefixes from the over-used link onto other exit links.

Range

Usage of this policy keeps some or all WAN links within a certain utilization range, relative to each other in order to ensure fair load-sharing across all concerned links. The range functionality allows the network administrator to instruct Cisco PfR to keep the usage on a set of exit links with in a certain percentage range of each other. If the difference between the links becomes too great, Cisco PfR will attempt to bring the link back in to policy by distributing data traffic among the available exit links.

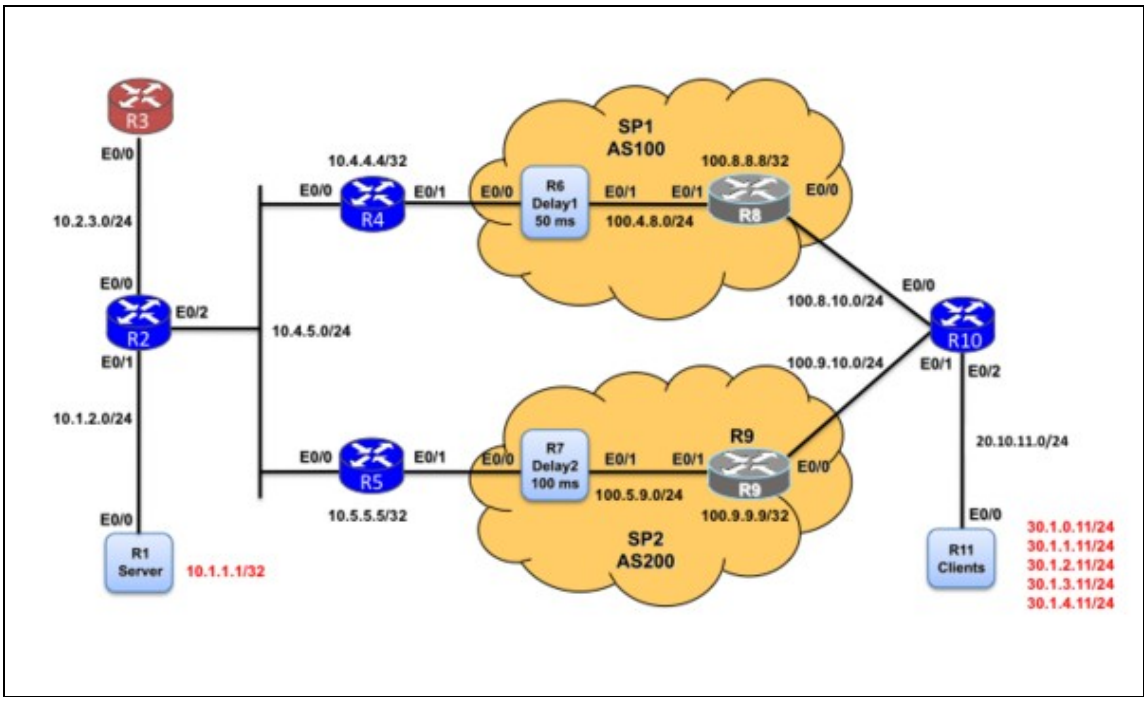
Traffic Class Performance

Usage of this policy enables the customer to define multiple paths that a set of traffic (ie voice traffic) could use as long as all the paths maintain the performance SLA ?s that are needed forth at set of traffic. Hence, a policy that determines voice traffic to have a delay threshold of less than 250 msecs can utilize multiple paths in the network if available, as long as all the paths deliver the traffic within its performance bounds.

PfR Network Topology Used

The central site has two Border Routers, connected to two separate Service Providers using eBGP. R2, R4 and R5 are iBGP peers.

- R3 is the Master Controller
- R4 and R5 the Border Routers
- Traffic Simulator tool is used between R1 and R11 to emulates traffic
- R6 and R7 are delay generators that add delay/loss to the path through SP1 and SP2. By default, 100 ms through SP1 and 50 ms through SP2.
- R1 and R11 are packet generators that send/receive traffic (http, ssh, etc).



PfR Components Configuration

Master Controller Configuration	Comments
! key chain pfr key 0 key-string cisco ! pfr master max-range-utilization percent 10 logging ! border 10.4.5.5 key-chain pfr interface Ethernet0/1 external max-xmit-utilization percentage 90 interface Ethernet0/0 internal	• max-range-utilization percent 10 : Start Load balancing when exits links utilization differ more than 10%. • logging : Enable PfR Syslogs (can be checked using show logging). • max-xmit-utilization percentage 90 : If utilization on an external link exceeds 90% of the configured bandwidth, PfR will detect a Load OOP condition. • mode route control: PfR control the routes

`` ! border 10.4.5.4 key-chain pfr interface Ethernet0/1 external max-xmit-utilization percentage 90 interface Ethernet0/0 internal ! learn throughput delay periodic-interval 0 monitor-period 1 mode route control periodic 180 resolve range priority 1 resolve utilization priority 5 variance 20 no resolve delay ! ```	• resolve: Check range then utilization. If utilization on an external link is more than 20% greater than the range across all of the links, PfR will detect a Range OOP condition. • periodic 180: policies are reevaluated every 180s (optional)

Border Router Configuration	Comments
`` ! key chain pfr key 0 key-string cisco ! pfr border local Ethernet0/0 master 10.2.3.3 key-chain pfr ! ! interface Ethernet0/0 description --INTERNAL-- bandwidth 10000 ip address 10.4.5.4 255.255.255.0 ip ospf 100 area 0 load-interval 30 ! interface Ethernet0/1 description --ISP1-- bandwidth 500 ip address 100.4.8.4 255.255.255.0 load-interval 30 ! ```	The Master Controller needs to be configured to talk to the border routers. The IP address used is the one referred by the ?local <interface>? configuration on the BR. Netflow configuration on internal and external interfaces is not needed for PfR to operate. Just configure here as a mean to verify flows that are crossing this BR. External interface Don?t forget to configure external interfaces with the contracted bandwidth so that PfR knows how to share load. Also set the load interval to 30 seconds for better accuracy.

Flexible Netflow

While configuring Netflow is not a mandatory task for PfR to work, it allows to have a good understanding of the traffic flows across the border routers. The following configuration is just an example of a flow monitor definition.

Flow Record Definition

```
!
```

```
flow record MYRECORD
 match ipv4 protocol
 match ipv4 source address
 match ipv4 destination address
 match transport source-port
 match transport destination-port
 match interface input
 collect interface output
 collect counter bytes
 collect counter packets
!
```

Flow Monitor Definition

```
flow monitor MYMONITOR
 record MYRECORD
!
```

And then apply the FNF Monitor on the interface

```
interface Ethernet0/0
 ip flow monitor MYMONITOR input
!
```

Master Controller Verification ? show pfr master

Goal:

- Verify that the MC is active
- Verify that the learning process is enabled on the Master Controller
- Display the policy settings as well as the learning parameters and global timers.

```
MC#sh pfr master
OER state: ENABLED and ACTIVE
  Conn Status: SUCCESS, PORT: 3949
  Version: 3.0
  Number of Border routers: 2
  Number of Exits: 2
  Number of monitored prefixes: 6 (max 5000)
  Max prefixes: total 5000 learn 2500
  Prefix count: total 6, learn 6, cfg 0
  PBR Requirements met
  Nbar Status: Inactive

Border           Status   UP/DOWN             AuthFail  Version
10.4.5.5         ACTIVE   UP       00:15:21          0  3.0
10.4.5.4         ACTIVE   UP       00:15:16          0  3.0

Global Settings:
  max-range-utilization percent 10 recv 0
  mode route metric bgp local-pref 5000
  mode route metric static tag 5000
  trace probe delay 1000
```

```
  logging
  exit holddown time 60 secs, time remaining 0

Default Policy Settings:
  backoff 300 3000 300
  delay relative 50
  holddown 300
  periodic 180
  probe frequency 56
  number of jitter probe packets 100
  mode route control
  mode monitor both
  mode select-exit good
  loss relative 10
  jitter threshold 20
  mos threshold 3.60 percent 30
  unreachable relative 50
  resolve range priority 1 variance 0
  resolve utilization priority 5 variance 20

Learn Settings:
  current state : STARTED
  time remaining in current state : 77 seconds
  throughput
  delay
  no inside bgp
  monitor-period 1
  periodic-interval 0
  aggregation-type prefix-length 24
  prefixes 100 appls 100
  expire after time 720
MC#
```

What to check:

- Both Border Routers are up and running
- Number of Monitored Prefixes: 6 are learned
- Global Settings: max-range-utilization is 10 (percentage of bandwidth difference between all exit interfaces)
- All default policy settings are displayed
- Learn is started (current state : STARTED)

The PfR state should be "enabled and active" and the learning state "started". You can also notice that there are a few default values already defined for several thresholds.

Netflow Statistics

As explained before, explicitly enabling Netflow is not required for PfR to run but is a good practice to check active flows crossing the Border Routers, verify the ingress/egress interfaces used (must be internal to external or vice-versa).

In this example, we check if traffic is flowing through the border router R4 (by default R4 is the exit point according to BGP):

```
R4#
```

```
R4# sh flow monitor MYMONITOR cache format table
  Cache type:                               Normal
  Cache size:                                 4096
  Current entries:                              68
  High Watermark:                               68

  Flows added:                                  73
  Flows aged:                                    5
    - Active timeout      (  1800 secs)          0
    - Inactive timeout    (    15 secs)          5
    - Event aged                                 0
    - Watermark aged                             0
    - Emergency aged                             0

IPV4 SRC ADDR    IPV4 DST ADDR    TRNS SRC PORT  TRNS DST PORT  INTF INPUT            IP PROT  int
===============  ===============  =============  =============  ====================  =======  ===
10.4.5.2         224.0.0.5                    0              0  Et0/0                      89  Nul
10.4.5.5         224.0.0.5                    0              0  Et0/0                      89  Nul
10.2.3.3         10.4.5.4                  3949          52511  Et0/0                       6  Nul
10.1.1.1         30.1.1.11                   22           1011  Et0/0                       6  Et0
10.2.3.3         10.4.5.4                  3949          15621  Et0/0                       6  Nul
10.1.1.1         30.1.2.11                   25           2008  Et0/0                       6  Et0
10.1.1.1         30.1.3.11                   21           3002  Et0/0                       6  Et0
10.1.1.1         30.1.0.11                  262             10  Et0/0                       6  Et0
10.1.1.1         30.1.4.11                   22           1012  Et0/0                       6  Et0
10.1.1.1         30.1.5.11                   25           2009  Et0/0                       6  Et0
10.1.1.1         30.1.1.11                  263             10  Et0/0                       6  Et0
10.1.1.1         30.1.2.11                   22           1013  Et0/0                       6  Et0
10.1.1.1         30.1.0.11                  264             10  Et0/0                       6  Et0
10.1.1.1         30.1.4.11                   22           1014  Et0/0                       6  Et0
10.1.1.1         30.1.5.11                   22           1015  Et0/0                       6  Et0
10.1.1.1         30.1.1.11                  265             10  Et0/0                       6  Et0
10.1.1.1         30.1.2.11                   22           1016  Et0/0                       6  Et0
10.1.1.1         30.1.0.11                  266             10  Et0/0                       6  Et0
10.1.1.1         30.1.4.11                   22           1017  Et0/0                       6  Et0
10.1.1.1         30.1.5.11                   22           1018  Et0/0                       6  Et0
10.1.1.1         30.1.3.11                   80           3003  Et0/0                       6  Et0

[snip}

10.1.1.1         30.1.5.11                  205             10  Et0/0                       6  Et0
10.1.1.1         30.1.1.11                  206             10  Et0/0                       6  Et0
10.1.1.1         30.1.5.11                   22           1060  Et0/0                       6  Et0
10.1.1.1         30.1.2.11                  207             10  Et0/0                       6  Et0

R4#
```

We clearly see that traffic is crossing the BR between internal and external interfaces (E0/0 and E0/1). Next step is to check the Traffic Classes on the Master Controller.

Verify Load balancing

As soon as you see:

```
MC#
*Sep  3 16:01:22.924: %OER_MC-5-NOTICE: Prefix Learning WRITING DATA
*Sep  3 16:01:22.966: %OER_MC-5-NOTICE: Prefix Learning STARTED
MC#
```

You should be able to see the traffic classes on the Master Controller. You will need a few cycles before having all prefixes in INPOLICY state. You will start getting OOP (out of policy messages) from PfR regarding range, because R4 (10.4.5.4) is currently the only exit point while R5 (10.4.5.5) has no traffic:

```
MC#
*Sep  3 16:01:39.232: %OER_MC-5-NOTICE: Range OOP BR 10.4.5.4, i/f Et0/1, percent 76. Other BR 10.
*Sep  3 16:01:39.232: %OER_MC-5-NOTICE: Exit 10.4.5.4 intf Et0/1 OOP, Tx BW 383, Rx BW 54, Tx Loac
MC#
```

Traffic Classes

On the Master Controller, you have all the Traffic Classes (in this case prefixes) learnt as well as statistics: Using show oer master traffic-class

```
MC#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 (packets-per-million), 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
------------------------------------------------------------------------------
30.1.0.0/24              N    N    N          N           N N
                       INPOLICY*      @83          10.4.5.4 Et0/1           U
              52       52        0        0        0        0       67        7
              51       51        0        0        N        N        N        N

30.1.1.0/24              N    N    N          N           N N
                       INPOLICY       @82          10.4.5.5 Et0/1           BGP
             104      102        0        0        0        0       55        6
             102      102        0        0        N        N        N        N

30.1.2.0/24              N    N    N          N           N N
                       INPOLICY       @84          10.4.5.5 Et0/1           BGP
             104      100        0        0        0        0       55        6
             100      100        0        0        N        N        N        N

30.1.3.0/24              N    N    N          N           N N
                       INPOLICY*      @82          10.4.5.4 Et0/1           U
              52       52        0        0        0        0       66        7
              51       51        0        0        N        N        N        N

30.1.4.0/24              N    N    N          N           N N
                       INPOLICY       @77          10.4.5.5 Et0/1           BGP
             104      102        0        0        0        0       53        6
             105      105        0        0        N        N        N        N

30.1.5.0/24              N    N    N          N           N N
                       INPOLICY*      @76          10.4.5.4 Et0/1           U
              52       52        0        0        0        0       67        7
```

```
              51        51         0         0         N         N         N         N

MC#
```

A closer look at the results

If we look at 2 specific prefixes 30.1.0.0/24 and 30.1.1.0/24, we can see that PfR controls one, while the other is under the parent route control.

From the command ?show oer master traffic-class", let?s focus on the prefix 30.1.0.0/24 to understand the interesting values reported here:

```
MC#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 (packets-per-million), 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
-------------------------------------------------------------------------------
30.1.0.0/24               N    N    N           N           N N
                        INPOLICY*       29          10.4.5.4 Et0/1           U
              52        52         0         0         0         0       66         7
              51        51         0         0         N         N        N         N

[snip]

MC#
```

In looking at the detail display of the prefix, several items bear notice:

- Line1: prefix
- Line2: State of INPOLICY*?The asteric (*) indicates this prefix is uncontrolled by PfR (the parent route controls routing) , but is currently inpolicy.

- **Line3: Passive results**
- Line3 (PasSDly, PasLDly): short-term and long-term passive delay, measured from the TCP Syn/Ack. As explained previously, TCP Syn/Ack are used to check the reachability of the prefix and to collect the delay and loss information. The delay is around 50 ms, which is the delay through ISP1 (BR R4).
- Line3 (PasSUn, PasLUn): short-term and long-term statistics for Unreachable.
- Line3 (PasSLos, PasLLos): short-term and long-term statistics for Loss.
- Line3 (EBw/IBw): the bandwidth for this prefix in egress and ingress direction.

- **Line4: Active results**
- Line4 (ActSDly, ActLDly): short-term and long-term active delay, measured by active probes that are automatically learned.
- Line4 (ActSUn, ActLUn): short-term and long-term Unreachable results.
- Line4 (ActSLos, ActLLoss): short-term and long-term Loss results.

Now, let?s focus on the prefix 30.1.1.0/24:

```
MC#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 (packets-per-million), 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]

30.1.1.0/24               N    N    N           N           N N
                          INPOLICY      @37          10.4.5.5 Et0/1           BGP
             104      103        0        0        0        0       53        6
             103      103        0        0        N        N        N        N

[snip]

MC#
```

In looking at the detail display of the prefix, several items bear notice:

- Line1: prefix
- Line2: State of INPOLICY?this prefix is controlled by PfR and is currently inpolicy. BGP is used to enforce the path.
- Line2: The ?at sign? (@) on the Time Remaining value means the prefix is being actively probed. The numerical value is a countdown timer indicating when this state will expire.

- **Line3: Passive results**
- Line3 (PasSDly, PasLDly): short-term and long-term passive delay, measured from the TCP Syn/Ack. As explained previously, TCP Syn/Ack are used to check the reachability of the prefix and to collect the delay and loss information. The delay is around 100 ms, which is the delay through ISP2 (BR R5).
- Line3 (PasSUn, PasLUn): short-term and long-term statistics for Unreachable.

- Line3 (PasSLos, PasLLos): short-term and long-term statistics for Loss.
- Line3 (EBw/IBw): the bandwidth for this prefix in egress and ingress direction.

- **Line4: Active results**
- Line4 (ActSDly, ActLDly): short-term and long-term active delay, measured by active probes that are automatically learned.
- Line4 (ActSUn, ActLUn): short-term and long-term Unreachable results.
- Line4 (ActSLos, ActLLoss): short-term and long-term Loss results.

As the mode in ?both?, active probes are automatically calculated and generated:

```
MC#sh pfr master act
MC#sh pfr master active-probes
        OER Master Controller active-probes
Border   = Border Router running this Probe
State    = Un/Assigned to a Prefix
Prefix   = Probe is assigned to this Prefix
Type     = Probe Type
Target   = Target Address
TPort    = Target Port
How      = Was the probe Learned or Configured
N - Not applicable

The following Probes exist:

State      Prefix             Type     Target          TPort   How     Codec
Assigned   30.1.4.0/24        echo     30.1.4.11           N  Lrnd         N
Assigned   30.1.5.0/24        echo     30.1.5.11           N  Lrnd         N
Assigned   30.1.2.0/24        echo     30.1.2.11           N  Lrnd         N
Assigned   30.1.0.0/24        echo     30.1.0.11           N  Lrnd         N
Assigned   30.1.1.0/24        echo     30.1.1.11           N  Lrnd         N
Assigned   30.1.3.0/24        echo     30.1.3.11           N  Lrnd         N

The following Probes are running:

Border          State    Prefix             Type     Target          TPort
10.4.5.4        ACTIVE   30.1.3.0/24        echo     30.1.3.11           N
10.4.5.5        ACTIVE   30.1.3.0/24        echo     30.1.3.11           N
10.4.5.4        ACTIVE   30.1.1.0/24        echo     30.1.1.11           N
10.4.5.5        ACTIVE   30.1.1.0/24        echo     30.1.1.11           N
10.4.5.4        ACTIVE   30.1.2.0/24        echo     30.1.2.11           N
10.4.5.5        ACTIVE   30.1.2.0/24        echo     30.1.2.11           N


MC#
```

The policy applied to a specific prefix can be seen by the ?show oer master prefix XX policy?:

```
MC#sh pfr master prefix 30.1.1.0/24 policy
Default Policy Settings:
  backoff 300 3000 300
```

```
  delay relative 50
  holddown 300
  periodic 180
  probe frequency 56
  number of jitter probe packets 100
  mode route control
  mode monitor both
  mode select-exit good
  loss relative 10
  jitter threshold 20
  mos threshold 3.60 percent 30
  unreachable relative 50
  resolve utilization priority 1 variance 20
  resolve range priority 2 variance 0
MC#
```

Bandwidth used on exit links

The next step step is to verify the accuracy of the load-balancing scheme on both Border Routers R4 and R5.

```
MC#sh pfr master border detail
Border           Status   UP/DOWN             AuthFail  Version
10.4.5.4         ACTIVE   UP       00:52:21          0  3.0
 Et0/0           INTERNAL UP
 Et0/1           EXTERNAL UP

 External            Capacity      Max BW   BW Used    Load Status          Exit Id
 Interface             (kbps)       (kbps)    (kbps)    (%)
 ---------           --------      ------   ------- ------- ------           ------
 Et0/1           Tx       500         450       192      39 UP                    2
                 Rx                   500        49       9
------------------------------------------------------------------------------------
Border           Status   UP/DOWN             AuthFail  Version
10.4.5.5         ACTIVE   UP       00:52:21          0  3.0
 Et0/0           INTERNAL UP
 Et0/1           EXTERNAL UP

 External            Capacity      Max BW   BW Used    Load Status          Exit Id
 Interface             (kbps)       (kbps)    (kbps)    (%)
 ---------           --------      ------   ------- ------- ------           ------
 Et0/1           Tx       500         450       175      33 UP                    1
                 Rx                   500         0       0
MC#
```

Verify Enforcement

BGP Route Table on R2

R2 is an iBGP peer for both border routers and as such as the BGP route table. PfR being not active, the parent routes are BGP based on R2, R4 and R5 and R4 is the preferred exit point as seen below:

```
R2#sh ip bgp
BGP table version is 27, local router ID is 10.2.2.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
r i10.4.5.0/24      10.5.5.5                 0    100      0 i
r>i                 10.4.4.4                 0    100      0 i
*>i20.10.11.0/24    100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i20.11.11.11/32   100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i30.1.0.0/24      100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i30.1.1.0/24      100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i30.1.2.0/24      100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i30.1.3.0/24      100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i30.1.4.0/24      100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i30.1.5.0/24      100.4.8.8                0    100      0 100 20 i
   Network          Next Hop            Metric LocPrf Weight Path
* i                 100.5.9.9                0    100      0 200 20 i
r>i100.4.8.0/24     10.4.4.4                 0    100      0 i
r>i100.5.9.0/24     10.5.5.5                 0    100      0 i
*>i100.8.10.0/24    100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i100.9.10.0/24    100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
R2#
```

After a few cycles, PfR controls the prefixes and modifies BGP local-pref for prefixes that have to go to R5 instead of R4.

```
R2#sh ip bgp
BGP table version is 52, local router ID is 10.2.2.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, x best-external
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
r>i10.4.5.0/24      10.4.4.4                 0    100      0 i
r i                 10.5.5.5                 0    100      0 i
*>i20.10.11.0/24    100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i20.11.11.11/32   100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
* i30.1.0.0/24      100.5.9.9                0    100      0 200 20 i
*>i                 100.4.8.8                0    100      0 100 20 i
*>i30.1.1.0/24      100.5.9.9                0   5000      0 200 20 i
* i30.1.2.0/24      100.5.9.9                0    100      0 200 20 i
*>i                 100.4.8.8                0    100      0 100 20 i
*>i30.1.3.0/24      100.5.9.9                0   5000      0 200 20 i
*>i30.1.4.0/24      100.4.8.8                0    100      0 100 20 i
* i                 100.5.9.9                0    100      0 200 20 i
*>i30.1.5.0/24      100.5.9.9                0   5000      0 200 20 i
   Network          Next Hop            Metric LocPrf Weight Path
```

```
r>i100.4.8.0/24     10.4.4.4                 0    100      0 i
r>i100.5.9.0/24     10.5.5.5                 0    100      0 i
*>i100.8.10.0/24    100.4.8.8                0    100      0 100 20 i
* i                  100.5.9.9                0    100      0 200 20 i
*>i100.9.10.0/24    100.4.8.8                0    100      0 100 20 i
* i                  100.5.9.9                0    100      0 200 20 i
R2#
```

- The prefix 30.1.0.0/24 uncontrolled by PfR has a default local-preference of 100 (which is the default local-pref) toward exit BR R4.
- The prefix 30.1.1.0/24 controlled by PfR has a local-preference of 5000 (this is the default value assigned by PfR and can be changed in the PfR global configuration) toward exit BR R5.

Border Routers

Let?s have a look at the border routers.

Let?s begin with R4:

```
R4#sh pfr border routes bgp
BGP table version is 22, local router ID is 10.4.4.4
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, x best-external
Origin codes: i - IGP, e - EGP, ? - incomplete
OER Flags: C - Controlled, X - Excluded, E - Exact, N - Non-exact, I - Injected

   Network          Next Hop        OER    LocPrf Weight Path
*>i30.1.1.0/24      100.5.9.9       XN      5000      0 200 20 i
*>i30.1.2.0/24      100.5.9.9       XN      5000      0 200 20 i
*>i30.1.4.0/24      100.5.9.9       XN      5000      0 200 20 i
R4#
```

The ?X? under the OER column for the 30.1.1.0/24 route on R4 means that the route is not locally controlled. Meaning that the local preference 5000 is being injected from another router. When the ?X? attribute is set, the exact vs. non-exact is meaningless.

If we look at the route on R5, we can see that it is locally controlled, and the exact route is controlled. The ?exact? means that the 30.1.1.0/24 route was found in the BGP table and there are no more specific subnets underneath:

```
R5#sh pfr border routes bgp
BGP table version is 22, local router ID is 10.5.5.5
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, x best-external
Origin codes: i - IGP, e - EGP, ? - incomplete
OER Flags: C - Controlled, X - Excluded, E - Exact, N - Non-exact, I - Injected

   Network          Next Hop        OER    LocPrf Weight Path
*> 30.1.1.0/24      100.5.9.9       CE              0 200 20 i
*> 30.1.2.0/24      100.5.9.9       CE              0 200 20 i
```

```
*> 30.1.4.0/24      100.5.9.9       CE                0 200 20 i
R5#
```

Conclusion

One of the basic solution provided by PfR is to be able to load-balance traffic among multiple exit interface. The configuration is straightforward and very efficient. Based on that, a customer can evolve to a more complex Performance Routing solution based on Traffic Class performance and not only link utilization.