PfR Internet Presence - Implementing Cost based Policies

Navigation

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

Contents

Introduction

This solution guide describes how to configure and apply Cisco IOS Performance Routing (PfR) cost policies. A PfR policy can be configured to optimize traffic based on the monetary cost of the exit links. The PfR Cost Based Optimization feature provides financial benefits by directing traffic to lower cost links, while at the same time honoring other configured policies such as delay, loss, and utilization.

There are two main methods of billing:

- **Fixed-rate billing** is used when the ISP bills one flat rate for a link regardless of bandwidth usage. If fixed-rate billing only is configured on the exit links, all exits are considered equal with regard to cost-optimization and other policy parameters (such as delay, loss, and utilization) are used to determine if the prefix or exit link is in-policy.

  Configured under external interfaces using: *cost-minimization {fixed fee cost}*

- **Tier-based billing** is used when the ISP bills at a tiered rate based on the percentage of exit link utilization. Each cost tier is configured separately with an associated monetary cost and a percentage of bandwidth utilization that activates the tier is defined. The lowest cost tier for an exit using tier-based billing is charged each month regardless of the bandwidth actually utilized. An allowance is made for bursting in the algorithm used to determine the tier-based billing. In this situation, bursting is defined as short periods of high bandwidth usage that would be expensive under fixed-rate billing.

  Configured under external interfaces using: *cost-minimization {tier percentage fee fee}*

Cost Based Optimization can be applied to links that are billed using a fixed or tiered billing method. Load balancing based on cost can also be achieved.

Link Policies

Overview

PfR link policies are a set of rules that are applied against PfR-managed external links (an external link is an interface on a border router on the network edge). Link policies define the desired performance characteristics of the links. Instead of defining the performance of an individual traffic class entry that uses the link (as in traffic class performance policies), link policies are concerned with the performance of the link as a whole. Link policies are applied both to exit (egress) links and entrance (ingress) links. The following link policy types describe the different performance characteristics that can be managed using link policies.

Link Utilization Policy

A traffic load (also referred to as utilization) policy consists of an upper threshold on the amount of traffic that a specific link can carry. Cisco IOS PfR supports per traffic class load distribution. Every 20 seconds, by default, the border router reports the link utilization to the master controller, after an external interface is configured for a border router. Both exit link traffic and entrance link traffic load thresholds can be configured as a PfR policy. If the exit or entrance link utilization is above the configured threshold, or the default threshold of 75-percent, the exit or entrance link is in an out-of-policy (OOP) state and PfR starts the monitoring process to find an alternative link for the traffic class. The link utilization threshold can be manually configured either as an absolute value in kilobytes per second (kbps) or as a percentage. A load utilization policy for an individual interface is configured on the master controller under the border router configuration.

When configuring load distribution, we recommend that you set the interface load calculation on external interfaces to 30-second intervals with the load-interval interface configuration command. The default calculation interval is 300 seconds. The load calculation is configured under interface configuration mode on the border routers.

Range Policy

A range policy is defined to maintain all links within a certain utilization range, relative to each other in order to ensure that the traffic load is distributed. For example, if a network has multiple exit links, and there is no financial reason to choose one link over another, the optimal choice is to provide an even load distribution across all links. The load-sharing provided by traditional routing protocols is not always evenly distributed, because the load-sharing is flow-based rather than performance- or policy-based. Cisco PfR range functionality allows you to configure PfR to maintain the traffic utilization on a set of links within a certain percentage range of each other. If the difference between the links becomes too great, PfR will attempt to bring the link back to an in-policy state by distributing traffic classes among the available links. The master controller sets the maximum range utilization to 20 percent for all PfR-managed links by default, but the utilization range can be configured using a maximum percentage value.

Both exit link and entrance link utilization ranges can be configured as a PfR policy.

Cost Policy

PfR support for cost-based optimization was introduced in Cisco IOS Release 12.3(14)T, 12.2(33)SRB, and later releases. Cost-based optimization allows you to configure policies based on the monetary cost of each exit link in your network. To implement PfR cost-based optimization the PfR master controller is configured to send traffic over exit links that provide the most cost-effective bandwidth utilization, while still maintaining the desired performance characteristics. The load balancing algorithm is modified to allow for more efficient bandwidth utilization while minimizing the link cost.

Cost Policy Billing Models

Link Utilization Rollup Calculations

The first step in determining the billing fee for each exit link per month is to calculate the link utilization rollup values.

Link utilization rollup values are the averages of the link utilization readings taken at regular intervals (sampling period) from the ingress and egress interfaces at the border routers for a given rollup period. For example, if a sampling period was set to 60 minutes, and the rollup was set at 1440 minutes (24 hours), we would have 24 ingress and 24 egress link utilization samples used for calculating the link utilization rollup. An average is taken for each set of ingress and egress samples from that rollup period to get a link utilization rollup value for the ingress and egress links.

Monthly Sustained Utilization Calculation

After the link utilization rollup calculation is performed, the monthly sustained utilization is calculated. The specific details of tier-based billing models vary by ISP. However, most ISPs use some variation of the following algorithm to calculate what an enterprise should pay in a tiered billing plan:

- Gather periodic measurements of egress and ingress traffic carried on the enterprise connection to the ISP network and aggregate the measurements to generate a rollup value for a rollup period.
- Calculate one or more rollup values per billing period.
- Rank the rollup values for the billing period into a stack from the largest value to the smallest.
- Discard the top X percent (5% percent is the default) to accommodate bursting (Any bandwidth above the sustained monthly utilization.
- Apply the highest remaining rollup value in the stack, referred to as the sustained Monthly Target Link Utilization (MTLU), to a tiered structure to determine a tier associated with the rollup value.
- Charge the customer based on a set cost associated with the identified tier.

The monthly sustained utilization rollup calculations can be configured to use one of the following three techniques:

- **Combined**: the monthly sustained utilization calculation is based on a combination of the egress and ingress rollup samples on a single sorted stack, the highest X rollup values are discarded, and the next highest rollup value is the MTLU.

  Configured under the external interface: *cost-minimization calc combined*

- **Separate**: the egress and ingress rollup samples for a link are sorted into separate stacks and the highest X rollup values for each stack are discarded. The highest remaining rollup value of the two stacks is selected as the MTLU.

  Configured under the external interfaces: *cost-minimization calc separate*

- **Summed**: the egress and ingress rollup samples are added together. The summed values of each rollup sample are placed into one stack, the top X rollup values are discarded, leaving the next highest rollup value as the MTLU.

  Configured under the external interfaces: *cost-minimization calc sum*

Example:

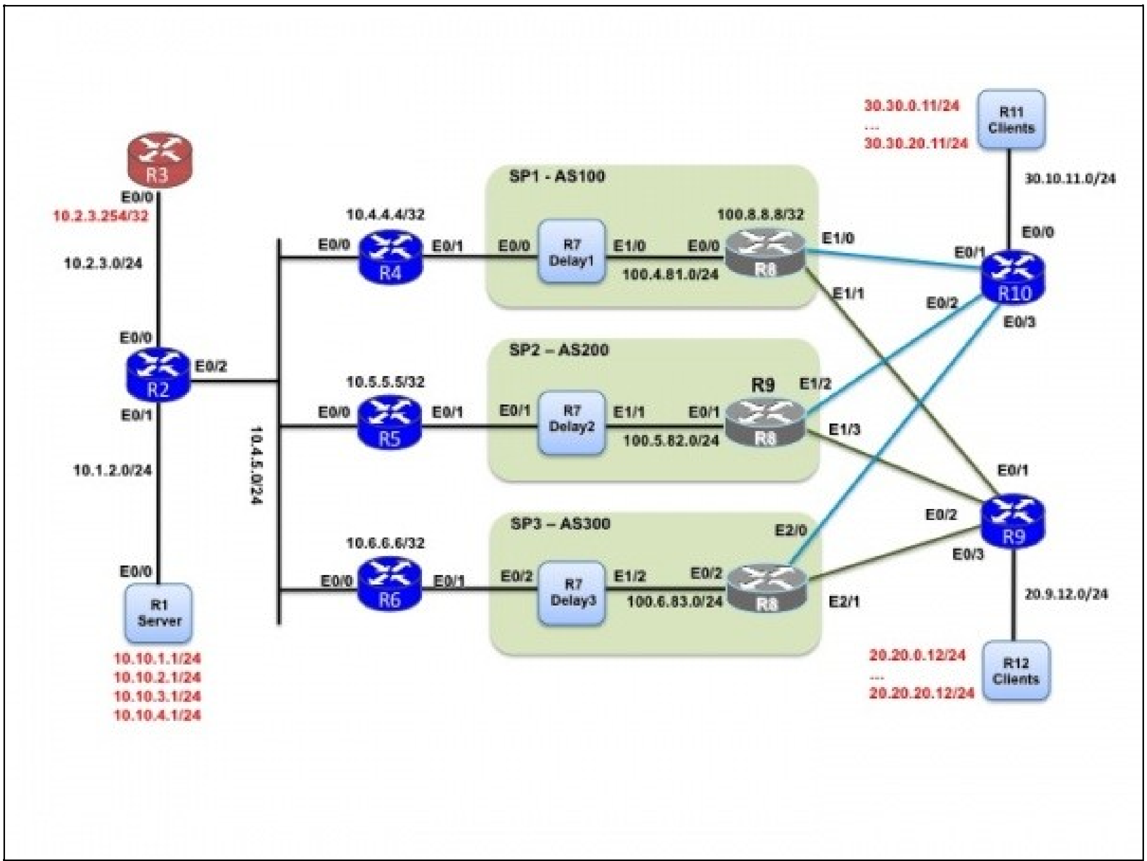
In the following example, we use the separate technique, which means we have all samples sorted in two different columns. We also define an absolute value of 5 highest value to discard. Therefore, PfR will remove the highest 5 rollup values in each column. After the discarded values, the next highest value is 52 and this becomes the *Sustained Monthly Utilization.*

Egress Rollup	**Ingress Rollup**	Rollups are Sorted from Highest Bandwidth to Lowest Bandwidth in Billing Period
96	40	
85	35	
70	34	
65	32	
60	30	
52	26	This value becomes the highest value after discarded the 5 highest rollup.
50	25	
48	23	
40	22	
35	20	
34	19	

PfR Network Topology Used

The central site has three Border Routers, connected to three separate Service Providers using eBGP. R2, R4, R5 and R6 are iBGP peers. For an Internet Presence solution, it may be recommended to have a dedicated Master Controller given the possible high number of prefixes that have to be optimized and managed.

- R2, R4, R5 and R6 are iBGP peers in AS 100
- R3 is the Master Controller
- R4, R5 and R6 are the Border Routers
- Traffic Simulator tool is used between R1 and R11, R12 to emulates traffic
- R1, R11 and R12 are traffic generators (to send/receive http, ssh, etc.).



Checking Statistics and Flows

Explicitly enabling Netflow **is not required for PfR to run** but here we enable Flexible NetFlow to check active flows crossing the Border Routers, verify the ingress/egress interfaces used (must be internal to external or vice-versa).

Configuring Flexible Netflow

The following configuration is just an example of a flow monitor definition in order to monitor active flows based on the 5-tuples + source interface and collect IP Source and Destination Address, ports and DSCP values.

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 ipv4 dscp
```

```
 collect interface output
 collect counter bytes
 collect counter packets
!
```

Flow Monitor Definition

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

And then apply the Flexible NetFlow Monitor on the Border Routers as well as R2:

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

Checking flows on R2

Here is the output on R2 which sees all flows:

```
R2#sh flow monitor MYMONITOR cache format table
 R2#shflow
  Cache type:                               Normal
  Cache size:                                 4096
  Current entries:                             208
  High Watermark:                              208

  Flows added:                                 208
  Flows aged:                                    0
    - Active timeout      (  1800 secs)          0
    - Inactive timeout    (    15 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            IP PROT  int
===============  ===============  =============  =============  ====================  =======  ===
10.10.3.1        30.30.16.11                 80           2037  Et0/1                       6  Et0
10.10.1.1        30.30.20.11               7000           7001  Et0/1                       6  Et0
10.10.1.1        30.30.8.11                7000           7005  Et0/1                       6  Et0
10.10.2.1        30.30.9.11                  25           1028  Et0/1                       6  Et0
10.10.1.1        30.30.17.11               7000           7004  Et0/1                       6  Et0
10.10.2.1        20.20.8.12                  25           1016  Et0/1                       6  Et0
10.10.3.1        30.30.18.11                 80           2038  Et0/1                       6  Et0
10.10.4.1        20.20.20.12                 80           2045  Et0/1                       6  Et0
10.10.1.1        20.20.12.12               7000           7005  Et0/1                       6  Et0
10.10.4.1        20.20.6.12                  80           2001  Et0/1                       6  Et0
10.10.1.1        30.30.5.11                7000           7008  Et0/1                       6  Et0
10.10.2.1        20.20.5.12                  25           1029  Et0/1                       6  Et0
10.10.1.1        30.30.11.11               7000           7006  Et0/1                       6  Et0
10.10.2.1        30.30.4.11                  25           1027  Et0/1                       6  Et0
10.10.1.1        20.20.4.12                7000           7003  Et0/1                       6  Et0
10.10.1.1        20.20.15.12               7000           7009  Et0/1                       6  Et0
10.10.4.1        20.20.19.12                 80           2048  Et0/1                       6  Et0

[SNIP]
```

Display Routing Table (Central Site)

Let's have a look at the routing table before applying a PfR configuration. Central site has prefixes 10.10.0.0/16 in Autonomous System 100, remote sites have prefixes 20.20.0.0/16 in Autonomous System 200 and 30.30.0.0/16 in Autonomous System 300. For clarity, only interesting part matching the destination prefixes of the routing tables are displayed. The servers subnets (inside prefixes) are 10.10.1.0/24, 10.10.2.0/24, 10.10.3.0/24 and 10.10.4.0/24. (Remember that these prefixes must be in the BGP table in case of inbound optimization).

Note: there is a BGP policy in place to enforce the path through R6. This is to demonstrate that even if there is a BGP local preference policy in place, PfR is able to override this when needed.

- A Local Preference of 50 is assigned on R4 for 20.20.0.0/16 and 30.30.0.0/16 routes
- A Local Preference of 100 is assigned on R5 for 20.20.0.0/16 and 30.30.0.0/16 routes
- A Local Preference of 200 is assigned on R6 for 20.20.0.0/16 and 30.30.0.0/16 routes

On the Border Router R4:

```
R4#sh bgp
BGP table version is 1476, 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, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 * i 10.10.1.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.2.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.3.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.4.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 *>i 20.20.0.0/16     100.6.83.1               0    200      0 300 20 i <-- REMOTE AS200
 *                    100.4.81.1                     50      0 100 20 i
 *>i 30.30.0.0/16     100.6.83.1               0    200      0 300 30 i <-- REMOTE AS300
 *                    100.4.81.1                     50      0 100 30 i

[SNIP]

R4#
```

On the Border Router R5:

```
R5#sh bgp
BGP table version is 1442, 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, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 * i 10.10.1.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.2.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.3.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.4.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 *>i 20.20.0.0/16     100.6.83.1               0    200      0 300 20 i <-- REMOTE AS200
 *                    100.5.82.1                    100      0 200 20 i
 *>i 30.30.0.0/16     100.6.83.1               0    200      0 300 30 i <-- REMOTE AS300
 *                    100.5.82.1                    100      0 200 30 i

[SNIP]

R5#
```

On the Border Router R6:

```
R6#sh bgp
BGP table version is 1436, local router ID is 10.6.6.6
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 * i 10.10.1.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.2.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.3.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 * i 10.10.4.0/24     10.4.5.2                21    100      0 i <-- INSIDE
 * i                  10.4.5.2                21    100      0 i
 *>                   10.4.5.2                21         32768 i
 *>  20.20.0.0/16     100.6.83.1                    200      0 300 20 i <------ HERE
 *>  30.30.0.0/16     100.6.83.1                    200      0 300 30 i <------ HERE

[SNIP]

R6#
```

PfR Configuration

Main configuration commands are the following:

Border Router External Interfaces:

- *cost-minimization nickname*: configures a nickname for a border router interface within a cost-based optimization policy on a master controller. This is useful to quickly have the show command outputs (see later).
- *cost-minimization calc {combined | separate | sum}*: choose the mode you want to use between combined, separate and summed (as explained in the previous chapter Monthly Sustained Utilization Calculation). Separate is the default here.
- *cost-minimization {fixed fee <cost>| tier <percentage> fee <fee>}*: choose between a fixed cost billing cycle or a tier-based billing cycle. If you choose a tier-based model, then define the various tiers and the associated fees.
- *cost-minimization sampling period <minutes> [rollup <minutes>]*: define the rollup period and the sampling interval. For this test, we have defined a very short rollup period and sampling interval which are not representatives of a real-life scenario.

Global policies:

- resolve cost priority X: configure the cost policy
- no resolve xx: disable other policies to avoid optimization conflicts.

```
!
pfr master
 max-range-utilization percent 100
 logging
 !
 border 10.4.5.6 key-chain pfr
  interface Ethernet0/1 external
   ! ----------------------------------------
   ! Nickname for R6-E0/1 is COST-ISP3
   ! Rollup period is 10 min, sampling per minute
   ! Tier-based model:
   !   up to 80% -> fee=50
   !   then fee=300
   ! ----------------------------------------
   cost-minimization nickname COST-ISP3
   cost-minimization tier 100 fee        300
   cost-minimization tier  80 fee         50
   cost-minimization sampling period 1 rollup 10
  interface Ethernet0/0 internal
 !
 border 10.4.5.5 key-chain pfr
  interface Ethernet0/1 external
   ! ----------------------------------------
   ! Nickname for R6-E0/1 is COST-ISP2
   ! Rollup period is 10 min, sampling per minute
```

```
    ! Tier-based model:
    !   up to 20% -> fee=200
    !   from 20% to 60 -> fee=250
    !   then fee=300
    ! ----------------------------------------
    cost-minimization nickname COST-ISP2
    cost-minimization tier 100 fee        300
    cost-minimization tier  60 fee        250
    cost-minimization tier  20 fee        200
    cost-minimization sampling period 1 rollup 10
   interface Ethernet0/0 internal
  !
  border 10.4.5.4 key-chain pfr
   interface Ethernet0/1 external
    ! ----------------------------------------
    ! Nickname for R6-E0/1 is COST-ISP1
    ! Rollup period is 10 min, sampling per minute
    ! Tier-based model:
    !   up to 20% -> fee=200
    !   from 20% to 60 -> fee=250
    !   then fee=300
    ! ----------------------------------------
    cost-minimization nickname COST-ISP1
    cost-minimization tier 100 fee        300
    cost-minimization tier  60 fee        250
    cost-minimization tier  20 fee        200
    cost-minimization sampling period 1 rollup 10
   interface Ethernet0/0 internal
  !
  learn
   expire after time 300
  max prefix total 10000 learn 10000
  mode monitor passive
 !---------------------
 ! Enable cost based policy only
 !---------------------
  resolve cost priority 1
  no resolve delay
  no resolve range
  no resolve utilization
  !
```

Verifying PfR Cost Minimization Policies

After cost-minimization policies are configured and applied to traffic the show command below allow you to verify that the policy configuration is working as expected.

- sh pfr master exits: will give an overall view and a summary of all metrics for all border routers and external interfaces. This command has been defined explicitly to check the link usage in case of load balancing or cost optimization policies (new command available from 15.2(2)T).

- show pfr master cost-minimization border <ip-address> <interface>
- show pfr master cost-minimization nickname <name>

  both commands display the same cost-minimization information. Example here:

sh pfr master cost border 10.4.5.4 Ethernet 0/1 (or sh pfr master cost-minimization nickname COST-ISP1)
sh pfr master cost border 10.4.5.5 Ethernet 0/1 (or sh pfr master cost-minimization nickname COST-ISP2)
sh pfr master cost border 10.4.5.6 Ethernet 0/1 (or sh pfr master cost-minimization nickname COST-ISP3)

Overall: the first step is to check the bandwidth usage on the Border Router external interfaces. There is a new command that summarize all informations relative to external interface bandwidth utilization. You can also use the *sh pfr master border detail* command.

```
MC# 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
  --- ------------ --------------- ----------- ----- --------------- ---- ----------- ---- ----
    6              10.4.5.6        Et0/1           2 100.6.83.6        24 Cost & Util   E  UP
    5              10.4.5.5        Et0/1           2 100.5.82.5        24 Cost & Util   E  UP
    4              10.4.5.4        Et0/1           2 100.4.81.4        24 Cost & Util   E  UP

Global Exit Policy:
===================
    Range Egress:      In Policy - Max difference 43% between Exits 4 & 6 - Policy 100%
   Range Ingress:  Out of Policy - Max difference 12% between Exits 6 & 4 - Policy 0%
     Util Egress:      In Policy
    Util Ingress:      In Policy
            Cost:      In Policy

Exits Performance:
==================
                   Egress                                                  Ingress
    ---------------------------------------------------- ------------------------------------
  ID Capacity  MaxUtil    Usage   %      RSVP POOL    OOP Capacity  MaxUtil    Usage   %  OOP
  --- -------- -------- -------- --- -------------- ----- -------- -------- -------- --- -----
   6     2000     1800     1164  58            N/A   Cost     2000     2000        0   0  N/A
   5     2000     1800      330  16            N/A   Cost     2000     2000        0   0  N/A
   4     2000     1800      305  15            N/A   Cost     2000     2000      241  12  N/A

TC and BW Distribution:
=======================
                       # of TCs                  BW (kbps)             Probe   Active
     Name/ID   Current Controlled InPolicy    Controlled       Total  Failed  Unreach
                                                                      (count)  (fpm)
        ----   ----------------------------   ----------------------  ------  --------
           6       31          31       31        1130          1164       0         0
           5        6           6        6         317           330       0         0
           4        5           5        5         330           305       0         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:        1       41

       Total number of TCs:       42
MC#
```

Notes:

- Global Exit Policy: we do not care about the range policy because we use the cost policy. Outgoing range we defined as 100% which explains why it is marked as In Policy.
- Exits Performance: this part is more interesting here because it gives the bandwidth repartition between all external interfaces. No surprise here, R6 has more bandwidth because it's fee is lower (50) than the other up to 80%. Note that we only have egress bandwidth optimization defined, and R4 is the only ingress BR used (follows BGP rules).
- TC and BW Distribution: gives the Traffic Classes repartition on all exits.

Details per ISP: the next step would be to check the details for each ISP

```
MC# sh pfr master cost-minimization nickname COST-ISP1
 pM - per Month, pD - per Day
-------------------------------------------------------------------------------------------
 Nickname  : COST-ISP1        Border: 10.4.5.4         Interface: Et0/1
 Calc type : Separate
 End Date  : 0
 Summer time: Disabled
 Fee       : Tier Based
             Tier 1: 100, fee:        300
             Tier 2:  60, fee:        250
             Tier 3:  20, fee:        200
 Period    : Sampling 1, Rollup 10
 Discard   : Type Absolute, Value 5

 Rollup Information:
 Total(pM)       Discard(pM)      Remaining(pM)   Collected(pM)
 4320            5                1222            11

 Current Rollup Information:
   MomentaryTgtUtil:          400 Kbps    CumRxBytes:         13243971
  StartingRollupTgt:          400 Kbps    CumTxBytes:         17672897
   CurrentRollupTgt:          400 Kbps    TimeRemain:         00:02:59

 Rollup Utilization (Kbps):
 Egress Utilization Rollups (Descending order)

 1   : 0           2   : 1108         3   : 758          4   : 686
 5   : 319         6   : 319          7   : 318          8   : 318
 9   : 316         10  : 0            11  : 0            12  : 0
 Ingress Utilization Rollups (Descending order)

 1   : 0           2   : 283          3   : 269          4   : 240
 5   : 239         6   : 239          7   : 238          8   : 237
 9   : 209         10  : 0            11  : 0            12  : 0
MC#
```

Notes:

- Calc type : Separate (default)
- Period: Sampling 1, Rollup 10 - 10 minutes rollup with a sampling interval of 1 minute (this is just for the test and is not a recommended value for real life deployment).
- Discard: Type Absolute, Value 5 - the top 5 values are discarded in each column (egress, ingress).
- Rollup Utilization: gives you the rollup values

```
MC# sh pfr master cost-minimization nickname COST-ISP2
 pM - per Month, pD - per Day
--------------------------------------------------------------------------------------------
 Nickname  : COST-ISP2        Border: 10.4.5.5         Interface: Et0/1
 Calc type : Separate
 End Date  : 0
 Summer time: Disabled
 Fee       : Tier Based
             Tier 1: 100, fee:        300
             Tier 2:  60, fee:        250
             Tier 3:  20, fee:        200
 Period    : Sampling 1, Rollup 10
 Discard   : Type Absolute, Value 5

 Rollup Information:
 Total(pM)       Discard(pM)     Remaining(pM)   Collected(pM)
 4320            5               1222            11

 Current Rollup Information:
   MomentaryTgtUtil:          400 Kbps    CumRxBytes:             5579
  StartingRollupTgt:          400 Kbps    CumTxBytes:         16557729
   CurrentRollupTgt:          400 Kbps    TimeRemain:         00:02:59

 Rollup Utilization (Kbps):
 Egress Utilization Rollups (Descending order)

 1   : 0            2   : 476          3   : 358          4   : 317
 5   : 306          6   : 302          7   : 300          8   : 300
 9   : 105          10  : 0            11  : 0            12  : 0
 Ingress Utilization Rollups (Descending order)

 1   : 0            2   : 0            3   : 0            4   : 0
 5   : 0            6   : 0            7   : 0            8   : 0
 9   : 0            10  : 0            11  : 0            12  : 0
MC#
```

```
MC# sh pfr master cost-minimization nickname COST-ISP3
 pM - per Month, pD - per Day
--------------------------------------------------------------------------------------------
 Nickname  : COST-ISP3        Border: 10.4.5.6         Interface: Et0/1
 Calc type : Separate
 End Date  : 0
 Summer time: Disabled
 Fee       : Tier Based
```

```
            Tier 1: 100, fee:        300
            Tier 2:  80, fee:         50
 Period    : Sampling 1, Rollup 10
 Discard   : Type Absolute, Value 5

 Rollup Information:
 Total(pM)       Discard(pM)     Remaining(pM)   Collected(pM)
 4320            5               1222            11

 Current Rollup Information:
   MomentaryTgtUtil:         1600 Kbps    CumRxBytes:             5820
  StartingRollupTgt:         1600 Kbps    CumTxBytes:         59928524
   CurrentRollupTgt:         1600 Kbps    TimeRemain:         00:02:59

 Rollup Utilization (Kbps):
 Egress Utilization Rollups (Descending order)

 1   : 0            2   : 1147         3   : 1141         4   : 1105
 5   : 1103         6   : 1087         7   : 744          8   : 738
 9   : 592          10  : 0            11  : 0            12  : 0
 Ingress Utilization Rollups (Descending order)

 1   : 0            2   : 0            3   : 0            4   : 0
 5   : 0            6   : 0            7   : 0            8   : 0
 9   : 0            10  : 0            11  : 0            12  : 0
MC#
```

Verifying Enforcement

The first step is to look at the traffic classes and check that everything is in policy and what is the BR and external interface used. Only part of the output is displayed below with a focus on a few prefixes only:

```
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
--------------------------------------------------------------------------------
20.20.0.0/24              N    N    N           N           N N
                          INPOLICY        0          10.4.5.5 Et0/1           BGP
             111      106        0        0      337      140       52        5
               N        N        N        N        N        N

20.20.8.0/24              N    N    N           N           N N
                          INPOLICY        0          10.4.5.6 Et0/1           BGP
```

```
            153      155        0        0      715      414       37        4
              N        N        N        N        N        N

20.20.16.0/24             N    N    N           N           N N
                          INPOLICY        0          10.4.5.4 Et0/1           BGP
             52       52        0        0      121       77       67        7
              N        N        N        N        N        N

30.30.0.0/24              N    N    N           N           N N
                          INPOLICY        0          10.4.5.6 Et0/1           BGP
            153      154        0        0      477      401       35        4
              N        N        N        N        N        N

30.30.8.0/24              N    N    N           N           N N
                          INPOLICY        0          10.4.5.5 Et0/1           BGP
            104      104        0        0        0      125       52        6
              N        N        N        N        N        N

[SNIP]
```

Let's focus on the following prefixes:

- 20.20.0.0/24: exit is R5
- 20.20.8.0/24: exit is R6
- 20.20.16.0/24: exit is R4
- 30.30.0.0/24: exit is R6
- 30.30.8.0/24: exit is R5

The next step is to look at the BGP routing table. By default the path for a specific prefix is enforced by setting the Local Preference to 5000 on the BR/Interface that was chosen by the PfR Master Controller.

```
R2#sh bgp
BGP table version is 184, 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, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path

 *>i 20.20.0.0/24     100.5.82.1               0   5000      0 200 20 i <-- HERE: exit is through
 *>i 20.20.0.0/16     100.6.83.1               0    200      0 300 20 i <-- PARENT ROUTE
 *>i 20.20.1.0/24     100.6.83.1               0   5000      0 300 20 i
 *>i 20.20.2.0/24     100.6.83.1               0   5000      0 300 20 i
 *>i 20.20.3.0/24     100.6.83.1               0   5000      0 300 20 i
 *>i 20.20.4.0/24     100.4.81.1               0   5000      0 100 20 i
 *>i 20.20.5.0/24     100.6.83.1               0   5000      0 300 20 i
 *>i 20.20.6.0/24     100.6.83.1               0   5000      0 300 20 i
 *>i 20.20.7.0/24     100.4.81.1               0   5000      0 100 20 i
 *>i 20.20.8.0/24     100.6.83.1               0   5000      0 300 20 i <-- HERE: exit is through
[SNIP]
 *>i 20.20.16.0/24    100.4.81.1               0   5000      0 100 20 i <-- HERE: exit is through
 *>i 30.30.0.0/24     100.6.83.1               0   5000      0 300 30 i <-- HERE: exit is through
 *>i 30.30.0.0/16     100.6.83.1               0    200      0 300 30 i <-- PARENT ROUTE
```

```
 *>i 30.30.1.0/24     100.6.83.1               0    5000      0 300 30 i
 *>i 30.30.2.0/24     100.6.83.1               0    5000      0 300 30 i
 *>i 30.30.3.0/24     100.6.83.1               0    5000      0 300 30 i
 *>i 30.30.4.0/24     100.6.83.1               0    5000      0 300 30 i
 *>i 30.30.5.0/24     100.6.83.1               0    5000      0 300 30 i
 *>i 30.30.6.0/24     100.6.83.1               0    5000      0 300 30 i
 *>i 30.30.7.0/24     100.5.82.1               0    5000      0 200 30 i
 *>i 30.30.8.0/24     100.5.82.1               0    5000      0 200 30 i <-- HERE: exit is through
[SNIP]
 *>i 30.30.16.0/24    100.6.83.1               0    5000      0 300 30 i

[SNIP]

R2#
```

Then, we can check on each BR the BGP routes that are enforced by the Master Controller:

```
R4#sh pfr border routes bgp
BGP table version is 178, 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, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
OER Flags: C - Controlled, X - Excluded, E - Exact, N - Non-exact, I - Injected

   Network          Next Hop        OER    LocPrf Weight Path
*>i20.20.0.0/24     100.5.82.1      XN      5000      0 200 20 i <-- CONTROLLED BY R5
*>i20.20.1.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.2.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.3.0/24     100.6.83.1      XN      5000      0 300 20 i
*> 20.20.4.0/24     100.4.81.1      CEI       50      0 100 20 i
*>i20.20.5.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.6.0/24     100.6.83.1      XN      5000      0 300 20 i
*> 20.20.7.0/24     100.4.81.1      CEI       50      0 100 20 i
*>i20.20.8.0/24     100.6.83.1      XN      5000      0 300 20 i
[SNIP]
*> 20.20.16.0/24    100.4.81.1      CEI       50      0 100 20 i <-- LOCALLY CONTROLLED
[SNIP]
*>i30.30.0.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.1.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.2.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.3.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.4.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.5.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.6.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.7.0/24     100.5.82.1      XN      5000      0 200 30 i
*>i30.30.8.0/24     100.5.82.1      XN      5000      0 200 30 i

[SNIP]

R4#
```

Notes:

- The ?X? under the OER column for the 20.20.0.0/24 route on R4 means that the route is not locally controlled. Meaning that the local preference 5000 is being injected from another router (in that case

R5). When the ?X? attribute is set, the exact vs. non-exact is meaningless.
- The prefixe 20.20.16.0/24 is locally controlled by R4. The ?exact? means that the 20.20.16.0/24 route is in the BGP table and there are no more specific subnets underneath. This route is also marked as 'injected', because only the parent route 20.20.0.0/16 was in the BGP routing table before PfR kicked in.
- Also note that in the LocPrf column, the local preference value displayed for locally controlled prefixes is the one that is manually configured in the BGP configuration (in R4 case, local preference of 50 was assigned to 20.20.0.0/16 and 30.30.0.0/16 routes).

```
R5#sh pfr border routes bgp
BGP table version is 156, 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, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
OER Flags: C - Controlled, X - Excluded, E - Exact, N - Non-exact, I - Injected

   Network          Next Hop        OER    LocPrf Weight Path
*> 20.20.0.0/24     100.5.82.1      CEI      100      0 200 20 i <-- LOCALLY CONTROLLED
*>i20.20.1.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.2.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.3.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.4.0/24     100.4.81.1      XN      5000      0 100 20 i
*>i20.20.5.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.6.0/24     100.6.83.1      XN      5000      0 300 20 i
*>i20.20.8.0/24     100.6.83.1      XN      5000      0 300 20 i <-- CONTROLLED BY R6
[SNIP]
*>i30.30.0.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.1.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.2.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.3.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.4.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.5.0/24     100.6.83.1      XN      5000      0 300 30 i
*>i30.30.6.0/24     100.6.83.1      XN      5000      0 300 30 i
*> 30.30.7.0/24     100.5.82.1      CEI      100      0 200 30 i
*> 30.30.8.0/24     100.5.82.1      CEI      100      0 200 30 i

[SNIP]

R5#
```

Notes:

- The ?X? under the OER column for the 20.20.8.0/24 route on R5 means that the route is not locally controlled. Meaning that the local preference 5000 is being injected from another router (in that case R6). When the ?X? attribute is set, the exact vs. non-exact is meaningless.
- The prefixe 20.20.0.0/24 is locally controlled by R5. The ?exact? means that the 20.20.0.0/24 route is in the BGP table and there are no more specific subnets underneath. This route is also marked as 'injected', because only the parent route 20.20.0.0/16 was in the BGP routing table before PfR kicked in.
- Also note that in the LocPrf column, the local preference value displayed for locally controlled prefixes is the one that is manually configured in the BGP configuration (in R5 case, local preference of 100 was assigned to 20.20.0.0/16 and 30.30.0.0/16 routes).

```
R6#sh pfr border routes bgp
BGP table version is 185, local router ID is 10.6.6.6
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
OER Flags: C - Controlled, X - Excluded, E - Exact, N - Non-exact, I - Injected

   Network          Next Hop        OER    LocPrf Weight Path
*>i20.20.0.0/24     100.5.82.1      XN      5000      0 200 20 i <-- CONTROLLED BY R5
*> 20.20.1.0/24     100.6.83.1      CEI      200      0 300 20 i
*> 20.20.2.0/24     100.6.83.1      CEI      200      0 300 20 i
*> 20.20.3.0/24     100.6.83.1      CEI      200      0 300 20 i
*>i20.20.4.0/24     100.4.81.1      XN      5000      0 100 20 i
*> 20.20.5.0/24     100.6.83.1      CEI      200      0 300 20 i
*> 20.20.6.0/24     100.6.83.1      CEI      200      0 300 20 i
*>i20.20.7.0/24     100.4.81.1      XN      5000      0 100 20 i
*> 20.20.8.0/24     100.6.83.1      CEI      200      0 300 20 i <-- LOCALLY CONTROLLED
[SNIP]
*>i20.20.16.0/24    100.4.81.1      XN      5000      0 100 20 i <-- CONTROLLED BY R4
[SNIP]
*> 30.30.0.0/24     100.6.83.1      CEI      200      0 300 30 i
*> 30.30.1.0/24     100.6.83.1      CEI      200      0 300 30 i
*> 30.30.2.0/24     100.6.83.1      CEI      200      0 300 30 i
*> 30.30.3.0/24     100.6.83.1      CEI      200      0 300 30 i
*> 30.30.4.0/24     100.6.83.1      CEI      200      0 300 30 i
*> 30.30.5.0/24     100.6.83.1      CEI      200      0 300 30 i
*> 30.30.6.0/24     100.6.83.1      CEI      200      0 300 30 i
*>i30.30.7.0/24     100.5.82.1      XN      5000      0 200 30 i
*>i30.30.8.0/24     100.5.82.1      XN      5000      0 200 30 i

[SNIP]

R6#
```

Notes:

- The ?X? under the OER column for the 20.20.0.0/24 route on R6 means that the route is not locally controlled. Meaning that the local preference 5000 is being injected from another router (in that case R5). When the ?X? attribute is set, the exact vs. non-exact is meaningless.
- The prefixe 20.20.8.0/24 is locally controlled by R6. The ?exact? means that the 20.20.8.0/24 route is in the BGP table and there are no more specific subnets underneath. This route is also marked as 'injected', because only the parent route 20.20.0.0/16 was in the BGP routing table before PfR kicked in.
- Also note that in the LocPrf column, the local preference value displayed for locally controlled prefixes is the one that is manually configured in the BGP configuration (in R4 case, local preference of 200 was assigned to 20.20.0.0/16 and 30.30.0.0/16 routes).

Conclusion

Configuring and using PfR cost policies is straight forward. If you want to review more information about PfR Cost Policies, see the Configuration Guide