Understanding Performance Routing (PfR)

Ron Trunk, CCIE

Chesapeake Netcraftsmen

Agenda

- PfR Overview
- Deployment
- Performance
- Conclusion

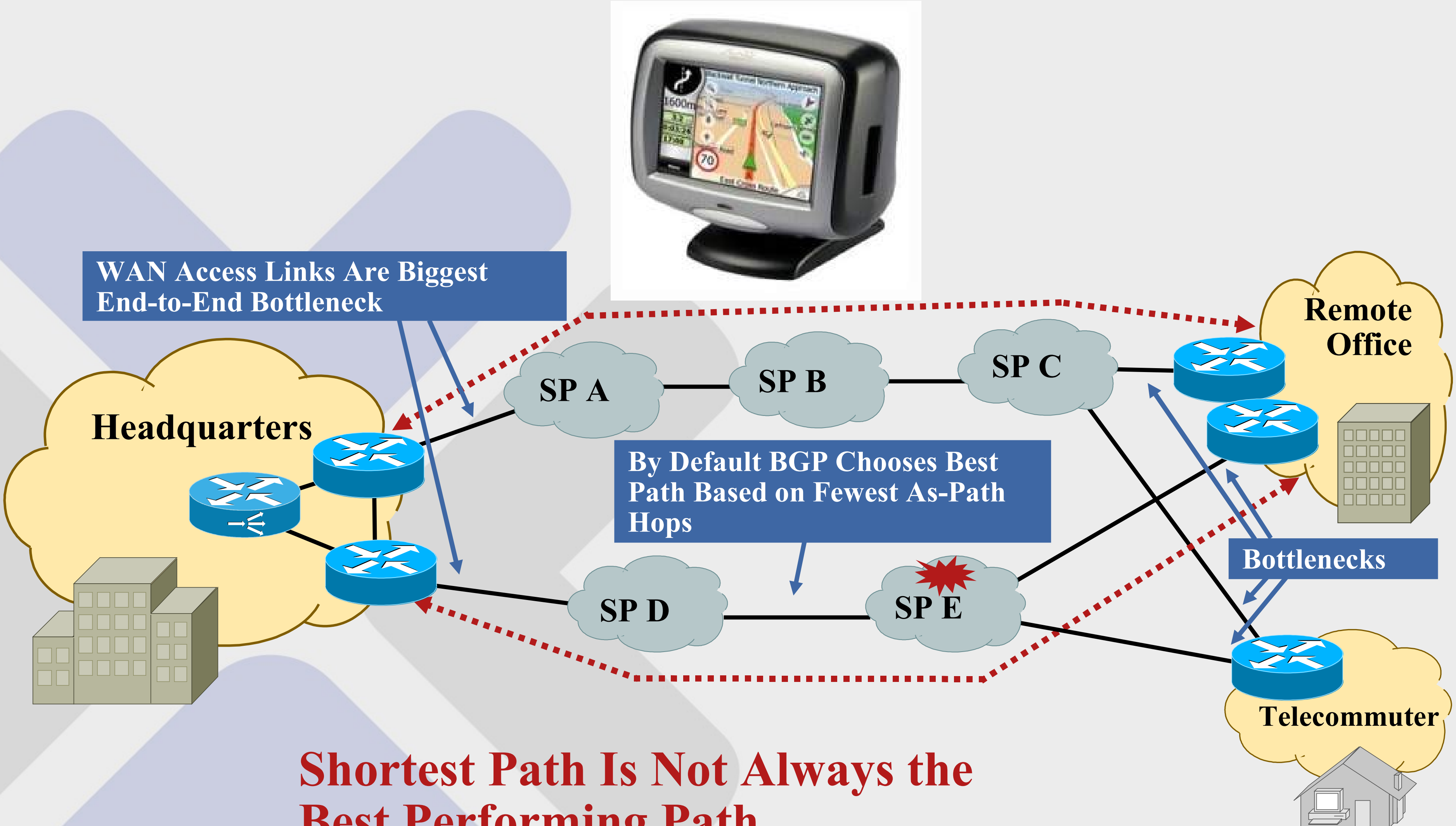
What Is Performance Routing (PfR)?

- Traditional routing protocols select shortest path
  - Shortest = Highest BW
-
- Sometimes "shortest path" is not best path
  - Congestion
  - Delay
  - Co$t
  -
- PfR selects the best performing path
  -

Where Is PfR Used?

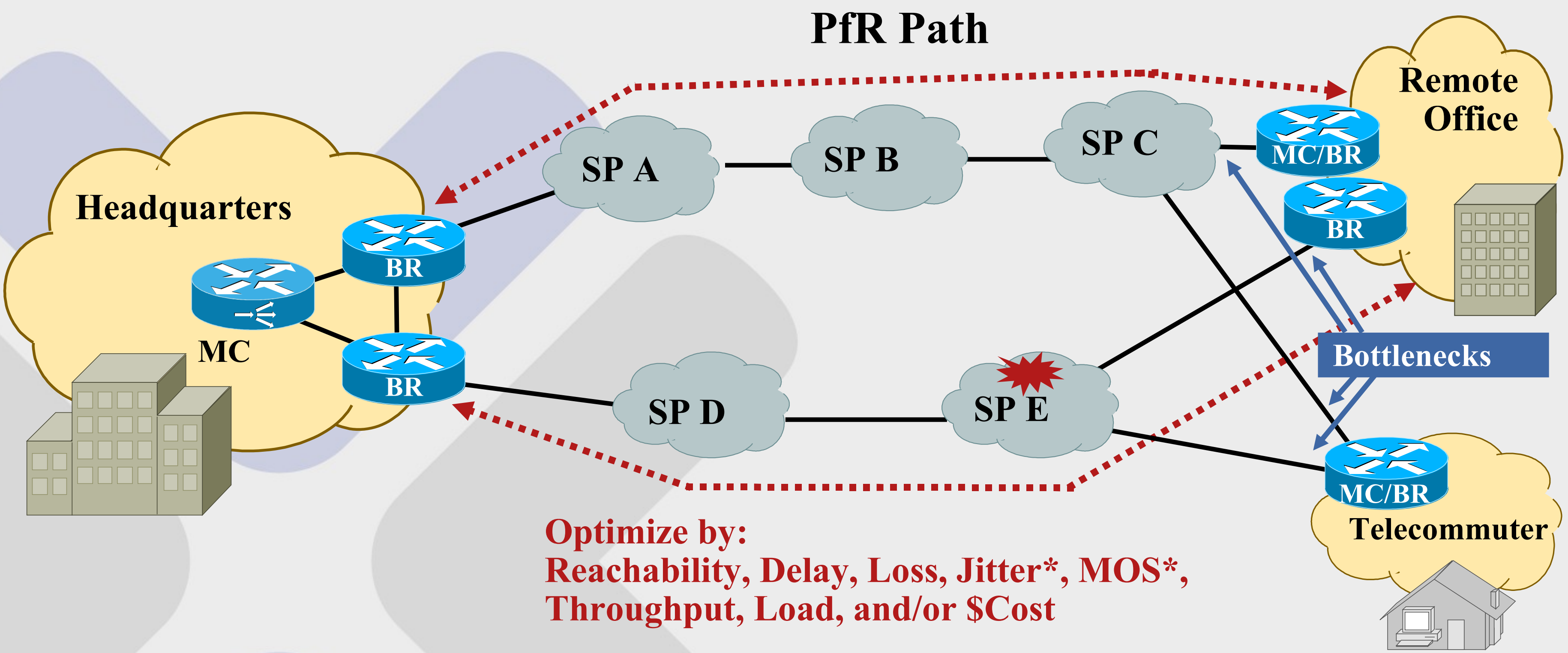
- WAN Edge
- Internet Edge
- More than one available path
-

Best Path Selection per Prefix, Two or More Paths



Shortest Path Is Not Always the Best Performing Path

PfR Best Path



Optimize by:
Reachability, Delay, Loss, Jitter*, MOS*,
Throughput, Load, and/or $Cost

PfR Components

- BR—Border Router
- MC—Master Controller (decision maker)

PfR and OER
What's the Difference?

- **PfR has a broader technology scope**
- **PfR will greatly expand application intelligence**
- **PfR will leverage OER and other Cisco IOS® technologies to enable adaptive routing throughout the enterprise**

OER

- Prefix
- WAN edge
- Network selection
- Exit routing

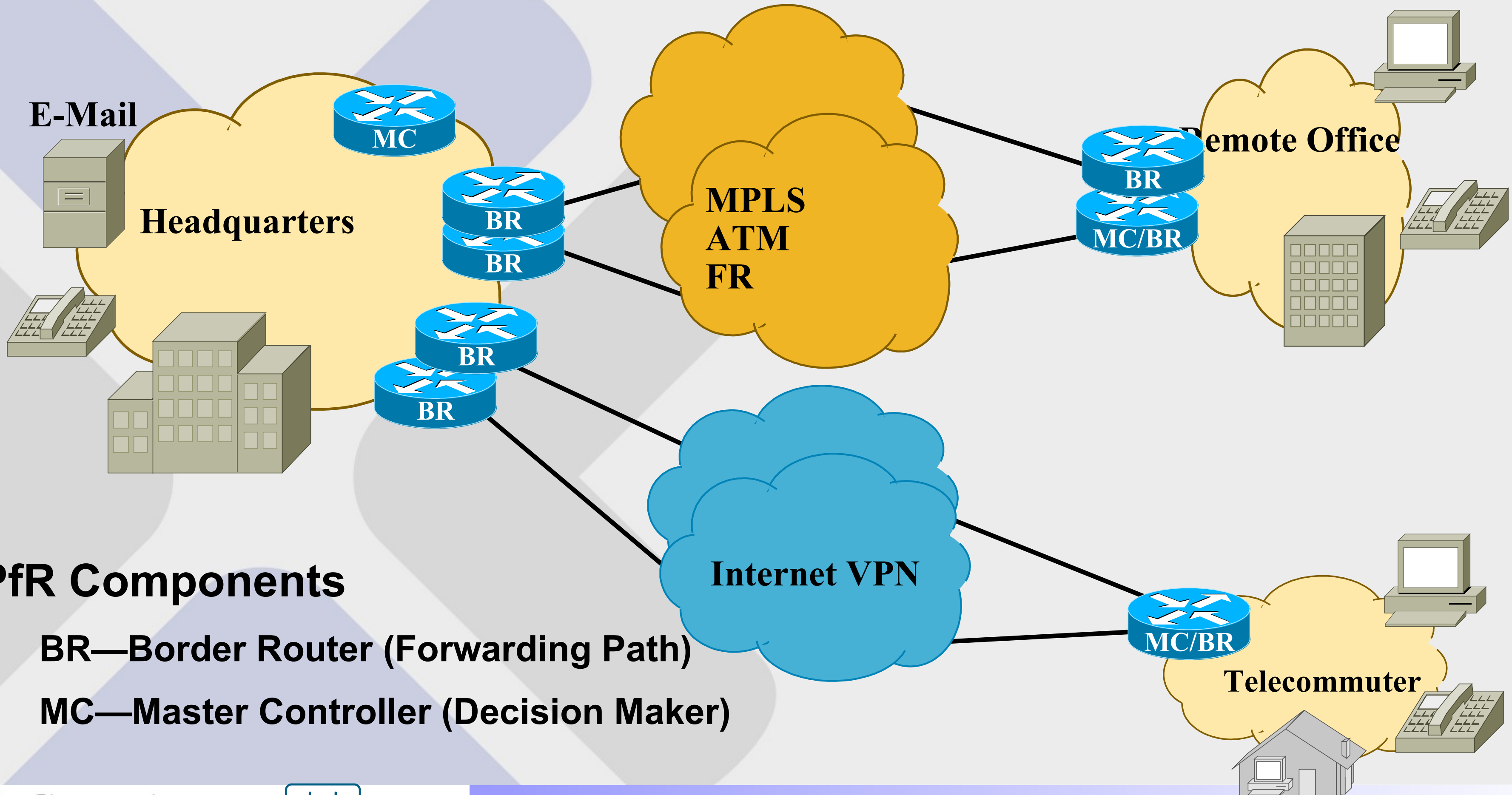
PfR

- Applications
- Private IP (MPLS)
- Path selection
- Networkwide

Performance Routing (PfR)

Exit Selection Criteria

Reachability, Delay, Loss, Jitter, MOS, Load, $Cost



PfR Components

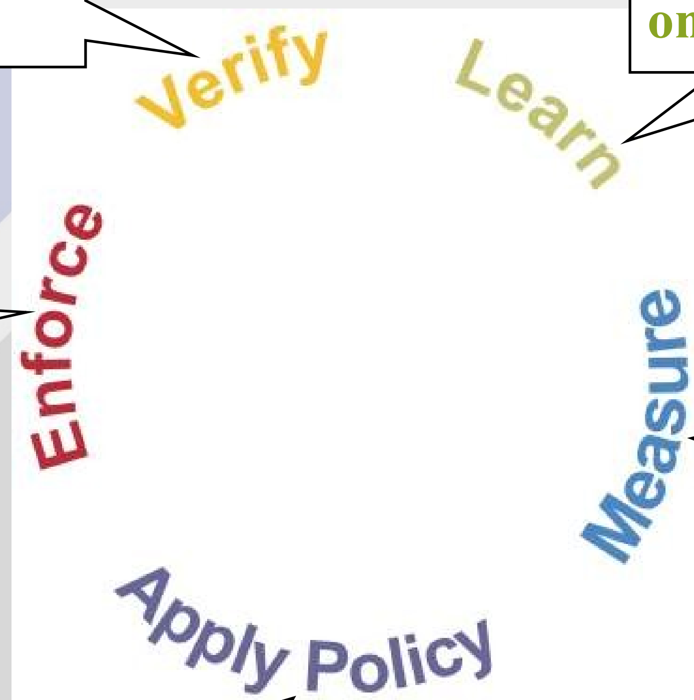
- BR—Border Router (Forwarding Path)
- MC—Master Controller (Decision Maker)

Performance Routing Policy Engine

Verify Enforcement and Performance

Learn Applications on the Network

Reroute Traffic



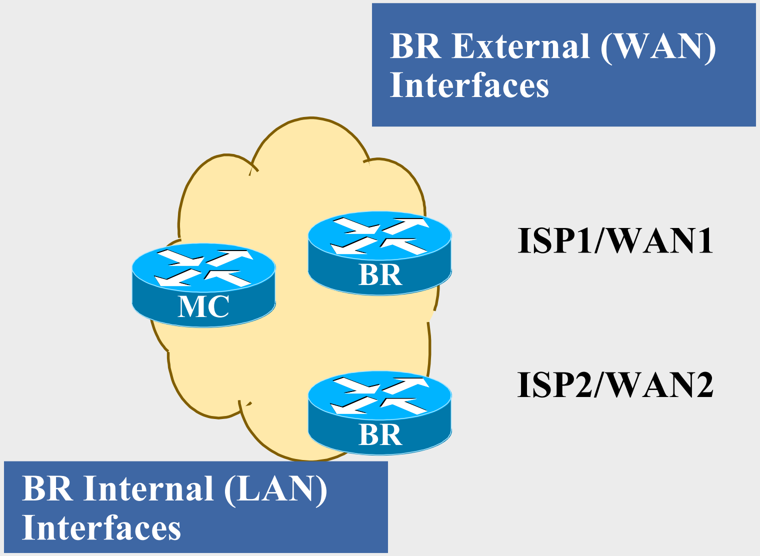
Measure Application Performance

Measure Alternate Paths
Apply Performance Policies to Measurements

Component Description

Master Controller (MC)

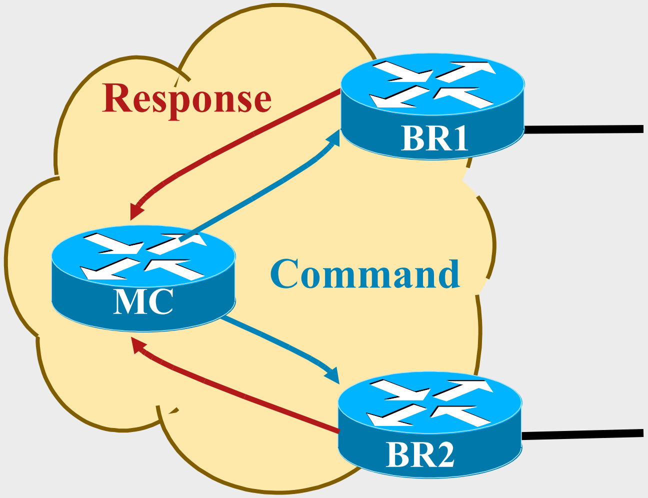
- Cisco IOS software feature
- Apply policy, verification, reporting
- Standalone or collocated with BR
- No routing protocol required
- Not required to be in forwarding path



- **Border Router (BR)**
  - Cisco IOS software feature in forwarding router
  - Learn, measure, enforcement
  - NetFlow collector
  - Probe source (IP SLA client)

Information Flow

- MC controls all operation
  - Issues commands to BRs
  - Contains traffic class/link data
  - Reports events
  - Reports measurements
  - Makes policy decisions
- BR responds to MC commands
  - Sends responses to MC
  - Uses NetFlow, IP SLA, BGP, static, PBR
  - Measures traffic class performance
  - Measures link performance
  - Enforces performance-based routing



PfR Operates on Traffic Classes

<table>
<tr><th colspan="2">Type</th><th>Example</th></tr>
<tr><td colspan="2">Destination Prefix
(Mandatory)</td><td>10.0.0.0/8
20.1.1.0/24</td></tr>
<tr><td rowspan="3">Application
(Optional)</td><td>ACL</td><td>10.1.1.0/24 dscp ef
10.1.1.0/24 dst-port 50</td></tr>
<tr><td>Well-Known</td><td>10.1.1.0/24 telnet
20.1.0.0/16 ssh</td></tr>
<tr><td>NBAR</td><td>10.1.1.0/24 nbar RTP
20.1.1.0/24 nbar citrix</td></tr>
</table>

Multiple ways to identify a traffic class.

Measuring Traffic Class Performance

- Passive
  - PfR **NetFlow** monitoring of traffic classes
  - Flows do not need to be on symmetrical paths provided that all exit/entry points are PfR-managed
- Active
  - PfR enables IP SLA feature
  - Probes sourced from BRs
  - icmp probes learned or configured
  - tcp, udp, jitter need `ip sla responder`
- Both mode
  - Attempts to measure performance passively with NetFlow and only launches IP SLA probes as needed
- Fast mode
  - Probes all paths all the time

Delay
Loss
Reachability
Egress BW
Ingress BW

Delay
Reachability
Jitter 12.4T
MOS 12.4T
Loss 12.4(15)T

PfR Policy

Traffic Class		Link	
Performance	Security	Performance	Administrative
▪ Delay ▪ Loss ▪ Reachability ▪ MOS ▪ Jitter	▪ Sinkhole ▪ Blackhole	▪ Load balancing ▪ Max utilization	▪ Link grouping ▪ $Cost

Scope	Global or per Policy

Selecting "Best" Traffic Class Path

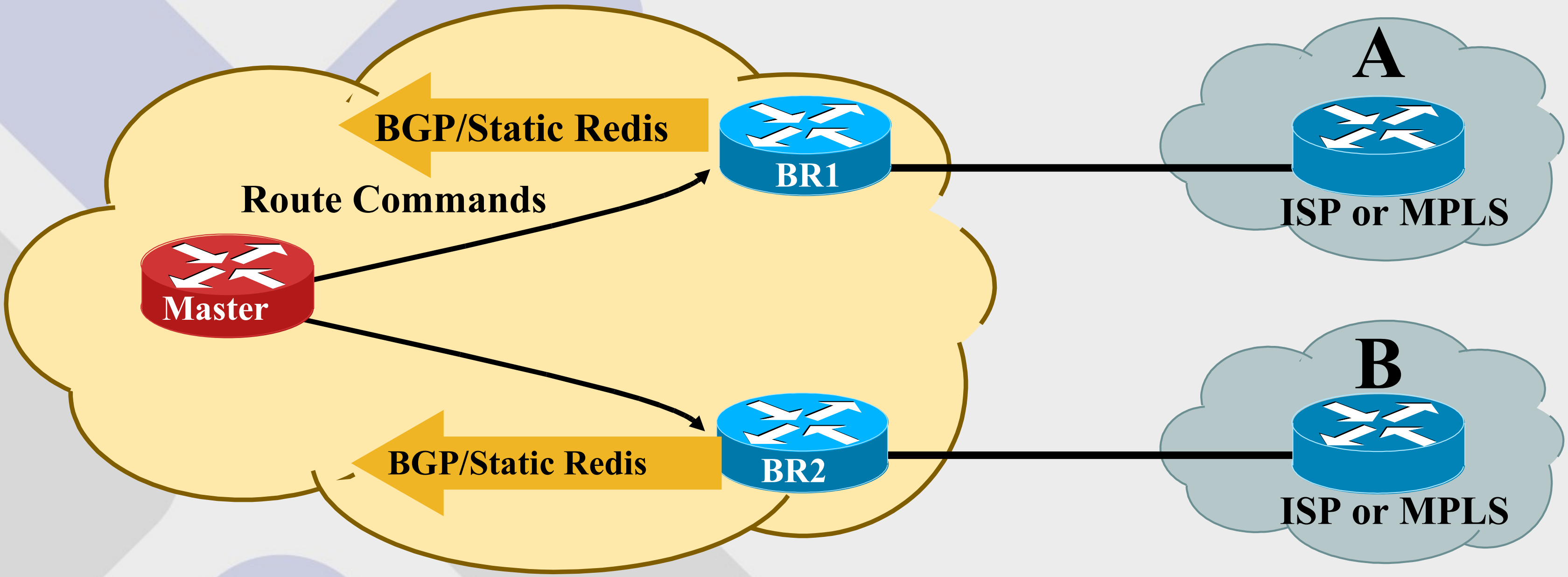
- 1. Ignore paths without sufficient capacity
- 2. Select best performing path based on priority and variance
- 3. If tie, keep current or select random

Selecting “Best” Traffic Class Path

Link	Utilization	Delay (ms) Priority 1	Jitter (ms) Priority 2
~~Serial1~~	~~89%~~	~~100~~	~~30~~
Serial2	50%	113	30
Serial3	60%	119	32
~~Serial4~~	~~40%~~	~~150~~	~~20~~

How Best Exit Path Is Enforced

- MC tells BR to insert prefix in BGP or static table
- MC tells BR to insert application/DSCP in policy route



- Modifying BGP local preference
  - Local preference must be highest
- Installing a static route at the exit
  - Redistribute static should be configured

Installing a Dynamic PBR route-map at the Exit

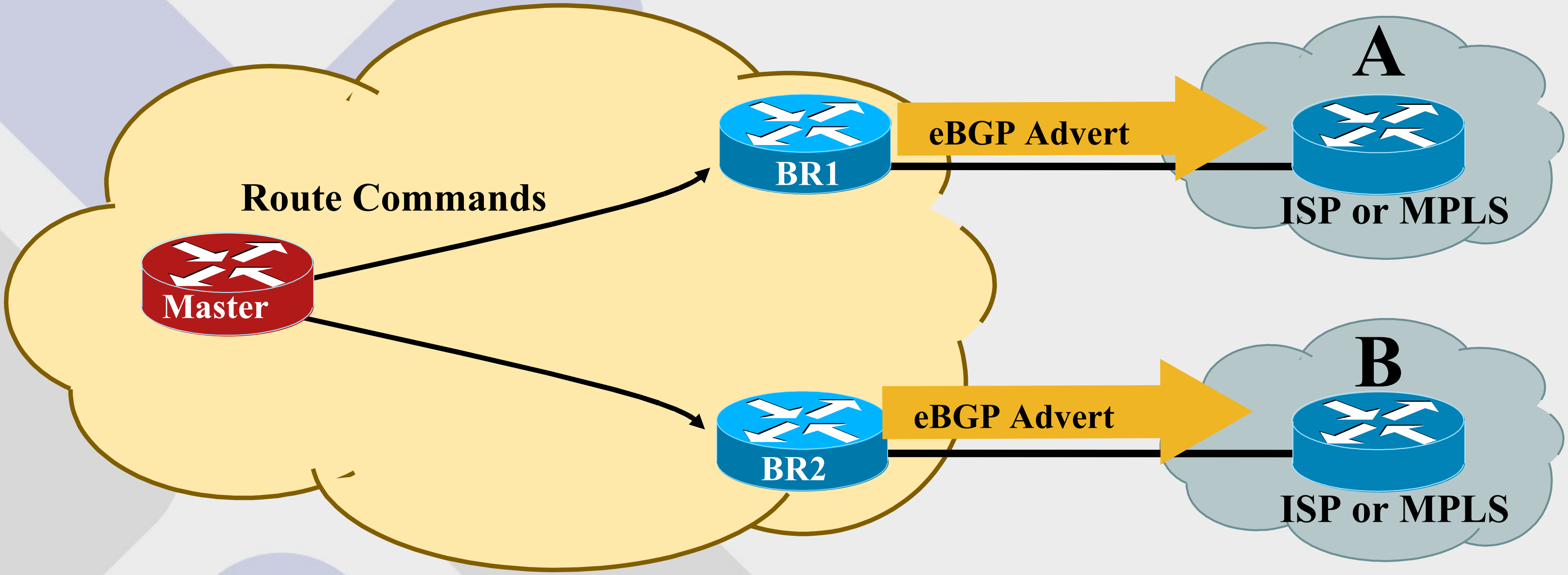
Direct Link or GRE Tunnel Between BRs Necessary

How Best Entrance Path Is Chosen

- Measurements gathered for all entrances
- Measurements applied in priority order
- Identify entrances to downgrade
- Downgrade entrance using BGP advertisement
  - AS path prepend
  - Append downgrade BGP community

How Best Entrance Path Is Enforced

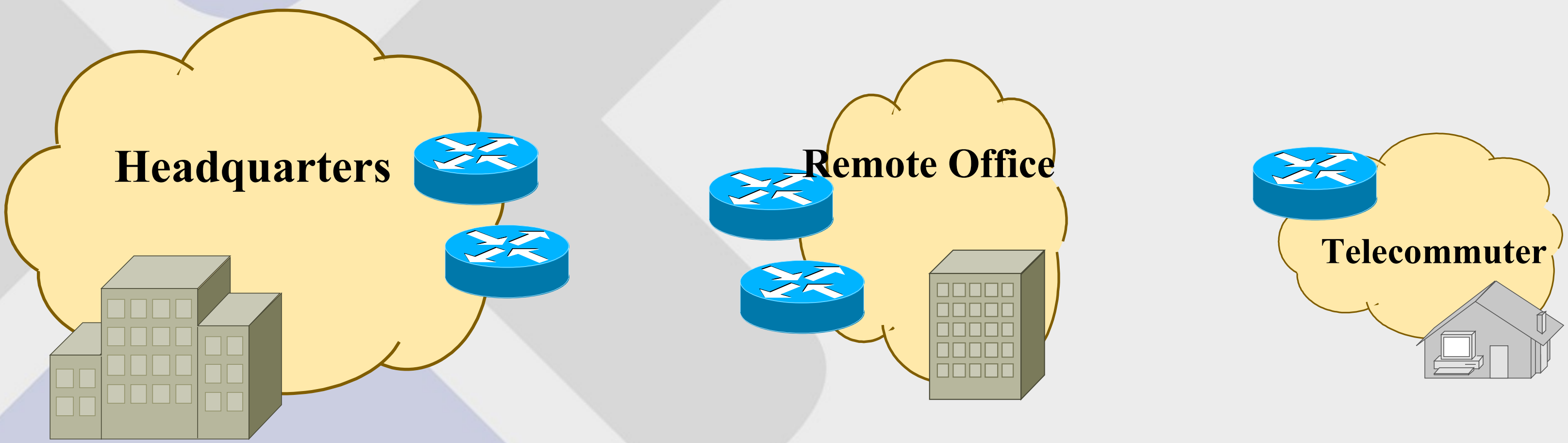
- **Needed for inbound load balancing**
- **MC tells BR to modify eBGP advertisement**



- Modifying eBGP
  - Prepend AS hop(s)
  - Append BGP downgrade community

PfR Typical Customers

- Large, medium, and small enterprises with mission-critical Internet presence
- Enterprises with redundant WAN networks
- Enterprises with remote offices
- Home office with dual internet connections



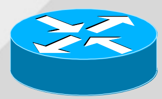
PfR Platform Support

Cisco
1800 ISR

1700*
12.4, 12.4T

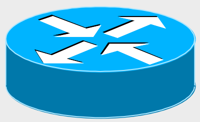
Cisco 2800
ISR

2600*
12.4, 12.4T



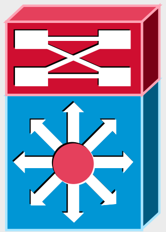
Cisco
3800 ISR

3640*/3660*/3700*
12.4, 12.4T



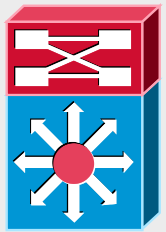
Cisco
7200-NPE-G2
Current Highest Performing PfR Device**

12.4, 12.4T



Cisco
6500***

12.2(33)SXH



Cisco 7600

12.2(33)SRB

*Announced/reached end-of-sale (EoS)
**Cisco 7301 with fixed NPE-G1 also supports PfR
***Only BR function supported, no support for MC

Key Features of Cisco PfR Manager by Fluke Networks

- **Executive-level** reports
- Network health reports
- Fully **Web-based**
- **Reports** and **alerts** on network events
- Historical and **trending** graphical reports
- **Troubleshooting** analysis
- **Easy** traffic class and policy configuration
- Technical support 24 hours a day
- Same design as NetFlow and IP SLA monitoring products

FLUKE networks

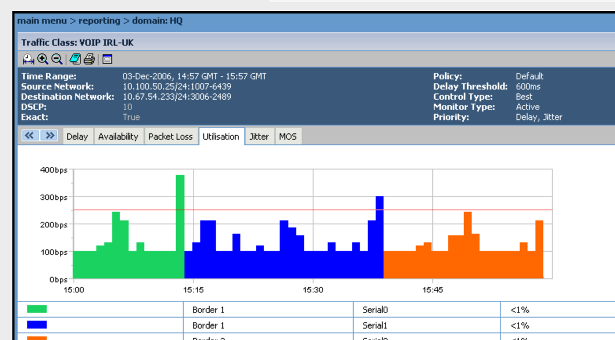
main menu > reporting > domain: HQ

HQ: Exit Links

Status	Exit Link	Device	Last Change	Threshold Utilisation	Total Utilisation
	Serial0	Border 1	02 Dec 2006, 09:12 (2 days)	80%	50%
	Serial1	Border1	02 Dec 2006, 09:12 (2 days)	30%	17%
	Serial0	Border2	04 Dec 2006, 12:05 (32 minutes)	0%	0%

HQ: Traffic Classes

Status	Description	Policy	Current Exit	Last Change	Delay	Util	Avail	Loss	Jitter	MOS
	VOIP IRL-UK	Critical VOIP	Border1/Serial0	02 Dec 2006, 09:12 (2 Days)	10ms	10%	100%	0%	5ms	4.85
	SAP-1	Priority 2 SAP	Border1/Serial1	04 Dec 2006, 11:15 (1:22 Hrs)	180ms	5%	100%	-	-	-
	HTTP	LOW Priority HTML	Border2/Serial0	02 Dec 2006, 09:12 (2 Days)	100 ms	30%	99.99%	-	-	-



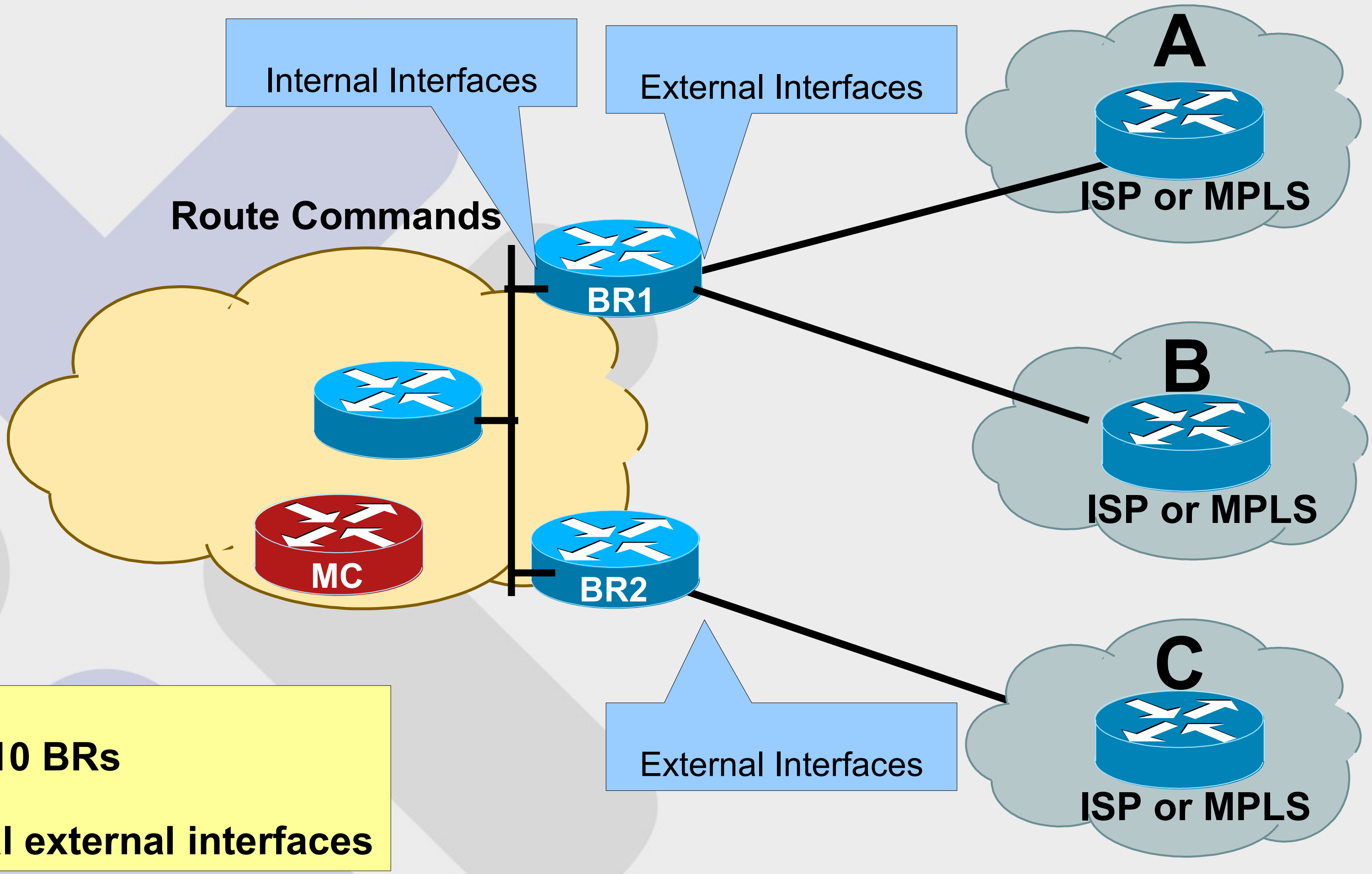
Agenda

- PfR Overview
- Deployment
- Performance
- Conclusion
- Q and A
- Backup Slides—Troubleshooting

Design Questions

1. Do I have redundant WAN connections?
   - Internet, IPSEC/GRE, MPLS, ATM, Frame Relay
   - Configure as PfR external interfaces
1. Which routers terminate the WAN?
   - These are PfR border routers
1. What routing protocols over WAN?
   - BGP, static covered by PfR
   - All others, cfg static with redistribution and filtering
1. Which router is PfR master controller?
   - Up to 5000 prefixes, dedicated 7200 or 3800 MC
   - Up to 20K prefixes with NPE-G2
   - For a few to few hundred prefixes, configure MC on BR

Design Basics



Up to 10 BRs

20 total external interfaces

Designing Your Policy

1. **What policy is important?**
   - Exit performance
     - Delay, loss, reachability, throughput
     - Jitter, MOS
   - Entrance performance—12.4T
     - Delay, loss, reachability, throughput
   - Load distribution
   - Cost minimization ($cost)
   - Primary/backup link groups
   - Path discovery (for troubleshooting)
   - Security 12.4T
   - Default priority is performance then load

Design Questions

1. **Determine interesting traffic class by:**

   Configure prefix

   Configure application

   Configure full ACL

   Learn interesting prefixes

   Learn interesting traffic classes

   Learn eBGP advertised prefixes (inbound optimization)

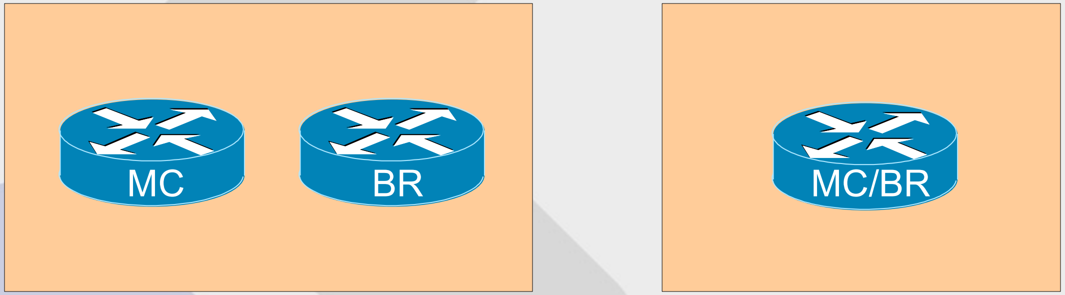
   Learn application

Routing Table Interaction

- For static routing, you must configure "Parent" Routes
  - Static equal cost routes
  - Points to external interface or next hop
- PfR injects additional routes with longer prefixes to steer traffic
- For BGP, parent routes must exist in BGP table
- If more than one BR, they must be IGP peers
- For application control, BRs must have direct link (or GRE tunnel)

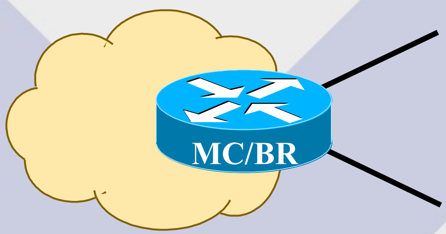
Which Router is the Master Controller?

- Up to a few hundred prefixes
  - Configure MC and BR on same router
- Up to 5000 prefixes
  - Use 3800 or 7200
- Up to 20,000 prefixes
  - 7200 with NPE-G2

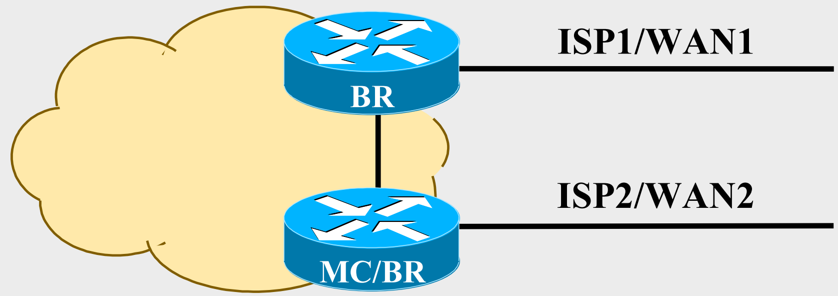


Typical Deployments

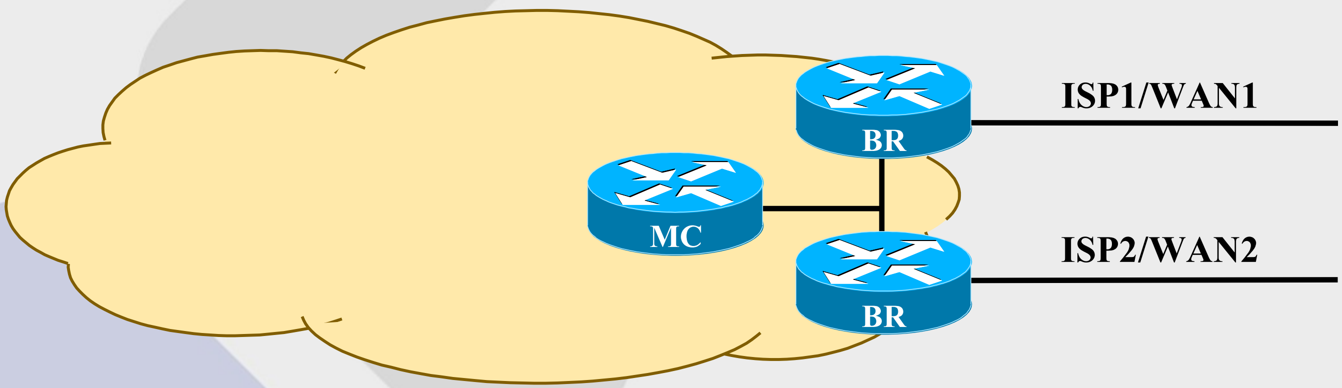
1. SOHO/Broadband



2. Remote Office



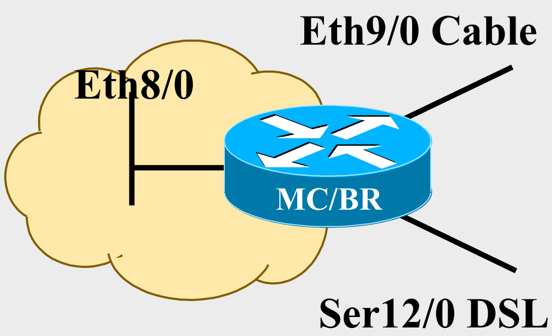
3. Headquarters/Content/Hosting/Data Centers



BR—Border Router, MC—Master Controller

SOHO/Broadband Deployment

- Cable and DSL WAN interfaces
  - Eth8/0—OER Internal
  - Eth9/0—OER External
  - Ser12/0—OER External
- ISR router terminates WAN
  - ISR is OER BR
- Static default routing
- 10 to 100 traffic classes
  - ISR is also MC
  - 12.4
- Performance is most important
  - Use OER default policy (performance overload)
- Learn throughput and delay to get prefixes



BR—Border Router, MC—Master Controller

SOHO/Broadband Configuration

```
key chain key1
 key 1
  key-string oer
oer master
 logging
 mode route control
 max prefix total 100
 backoff 90 3000 300
 border 10.10.10.1 key-chain key1
  interface Ethernet8/0 internal
  interface Ethernet9/0 external
    max-xmit-utilization absolute 1000
  interface Ser12/0 external
    max-xmit-utilization absolute 300
 learn
  throughput
  delay
  monitor-period 1
  periodic-interval 0
oer border
 logging
 local Ethernet8/0
 master 10.10.10.1 key-chain key1
interface Ethernet8/0
 ip address 10.10.10.1 255.255.255.0
interface Ethernet 9/0
 load-interval 30
interface Serial 12/0
 load-interval 30
```

Enable Logging

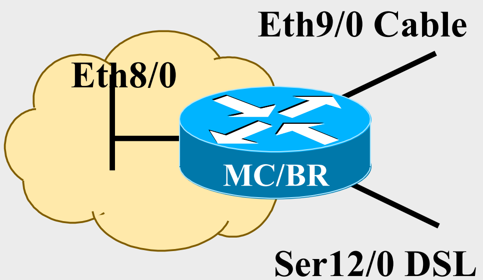
Enforce Routing Changes

Authentication Required

Limit Cable and DSL Throughput

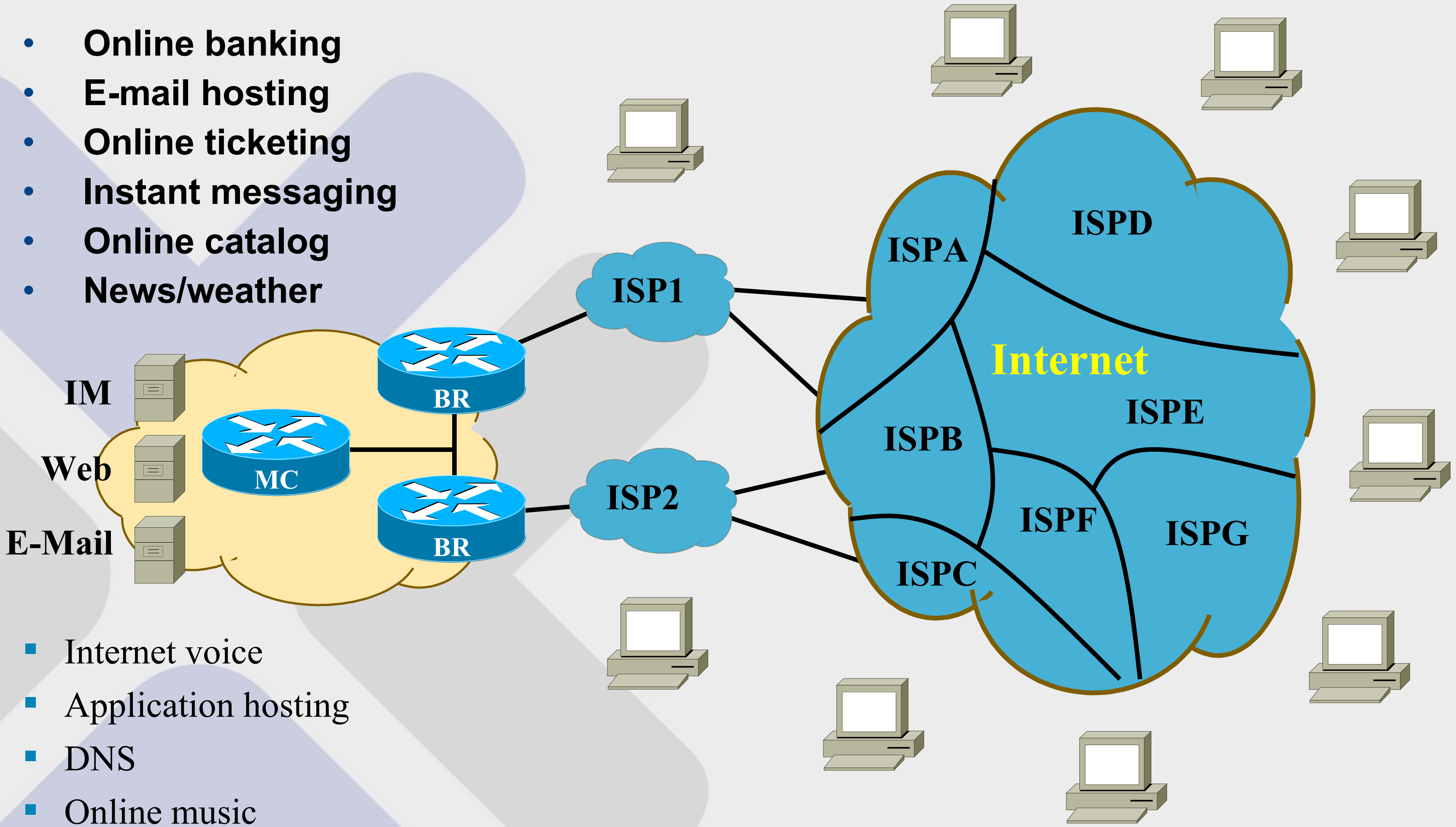
Learn Delay and Throughput Prefixes Every Minute

MC and BR on Same Router



Mission-Critical Internet Presence

- Online banking
- E-mail hosting
- Online ticketing
- Instant messaging
- Online catalog
- News/weather



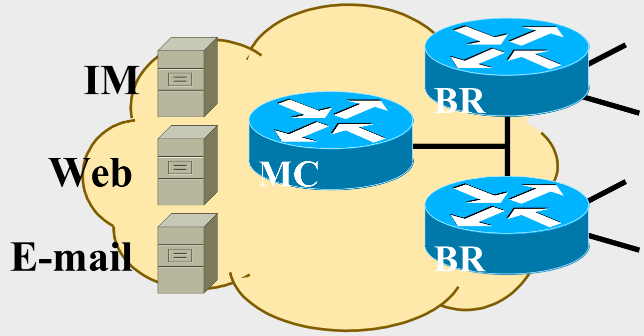
- Internet voice
- Application hosting
- DNS
- Online music
- Online video

Internet Presence Deployment

1. DS3 interfaces
   - Ser12/0, Ser13/0, etc.
1. Cisco 7200 and Cisco 3800 are typical BR/MC with BR terminating WAN connections
2. BGP routing
   - BRs must be iBGP peers
   - Default routing -or-
   - Partial routes -or-
   - Full routes

Same PfR Configuration for All

1. Support of up to 20,000 prefixes (with Cisco 7200-NPE-G2)
   - 12.4T/14.4M
   - Entrance Optimization
1. Customers differ on policy priority
2. Learn prefixes by throughput and delay



Internet Presence Configuration

Default Policy: Performance Then Load

MC 10.1.1.1

```
key chain key1
 key 1
  key-string oer
oer master
 logging
 mode route control
 mode select-exit best
 backoff 90 3000 300
 periodic 600
 border 10.1.1.2 key-chain key1
  interface Ethernet8/0 internal
  interface Serial12/0 external
  interface Serial13/0 external
 border 10.1.1.3 key-chain key1
  interface Ethernet 8/0 internal
  interface Serial12/0 external
  interface Serial13/0 external
 learn
  throughput
  delay
  monitor-period 1
  periodic-interval 0
  prefixes 500
  expire after time 240
```

Choose Best Exit Regardless of In or Out of Policy

Revaluate Exit 10 Minutes

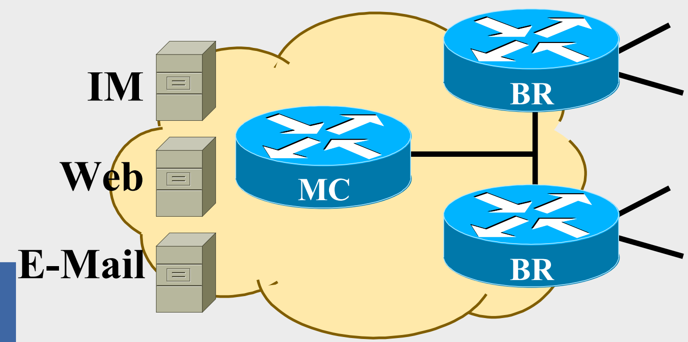
Learn 500 Prefixes

Delete Prefix if Not Relearned in4 Minutes

BR 10.10.10.2

BR 10.10.10.3

```
key chain key1
 key 1
  key-string oer
oer border
 logging
 local loopback 1
 master 10.10.10.1 key-chain key1
interface ser12/0
 load-interval 30
interface ser13/0
 load-interval 30
```



Internet Presence Configuration

Outbound Load Balancing Only

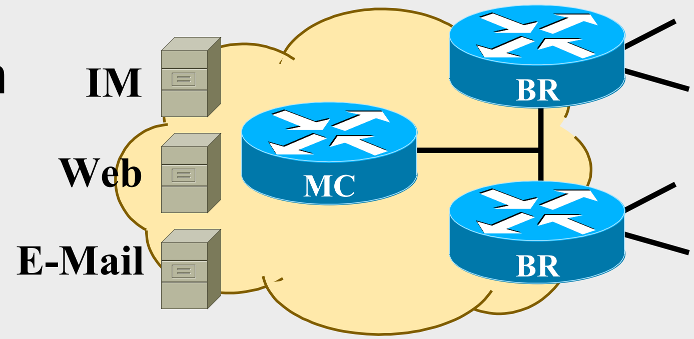
- **Add to default policy configuration**

Disable Periodic Prefix Evaluation

MC 10.1.1.1

```
oer master
 no periodic
 resolve utilization priority 1 variance 5
 resolve range priority 2
 no resolve delay
 no resolve loss
 max-range-utilization percent 10
 border 10.1.1.2
  interface Serial12/0 external
   max-xmit-utilization percent 70
  interface Serial13/0 external
   max-xmit-utilization percent 70
 border 10.1.1.3
  interface Serial12/0 external
   max-xmit-utilization percent 70
  interface Serial13/0 external
   max-xmit-utilization percent 70
```

Link OOP if :% util > Lowest + 10 or % util > 70

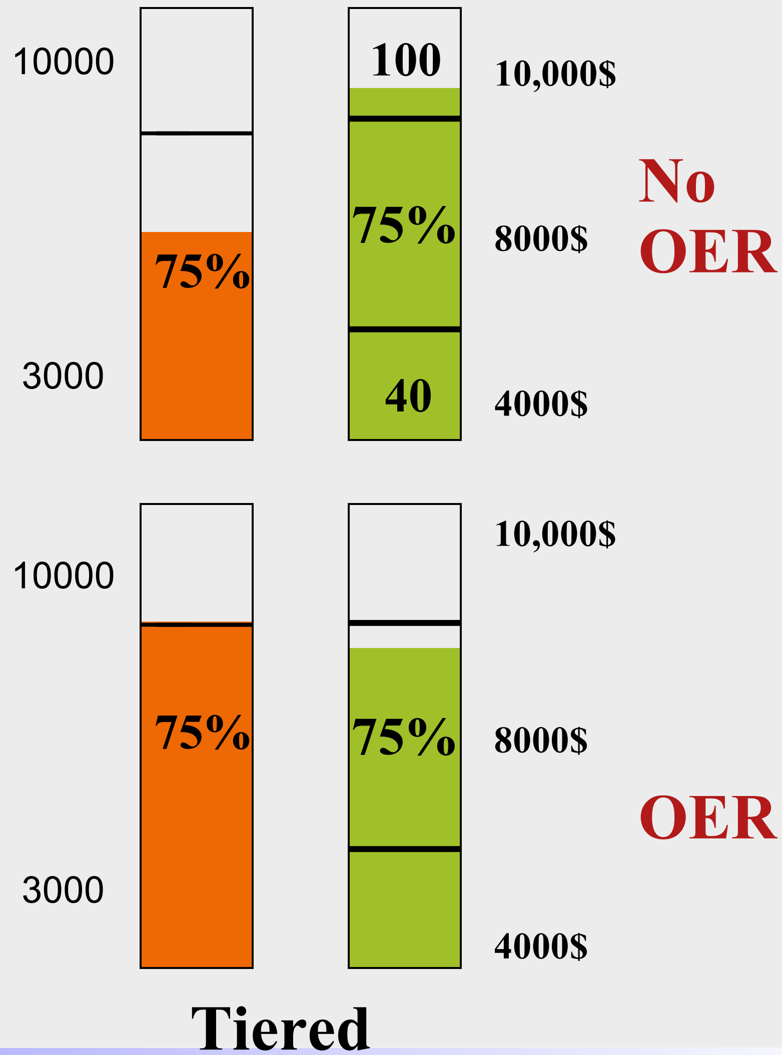


Internet Presence Configuration

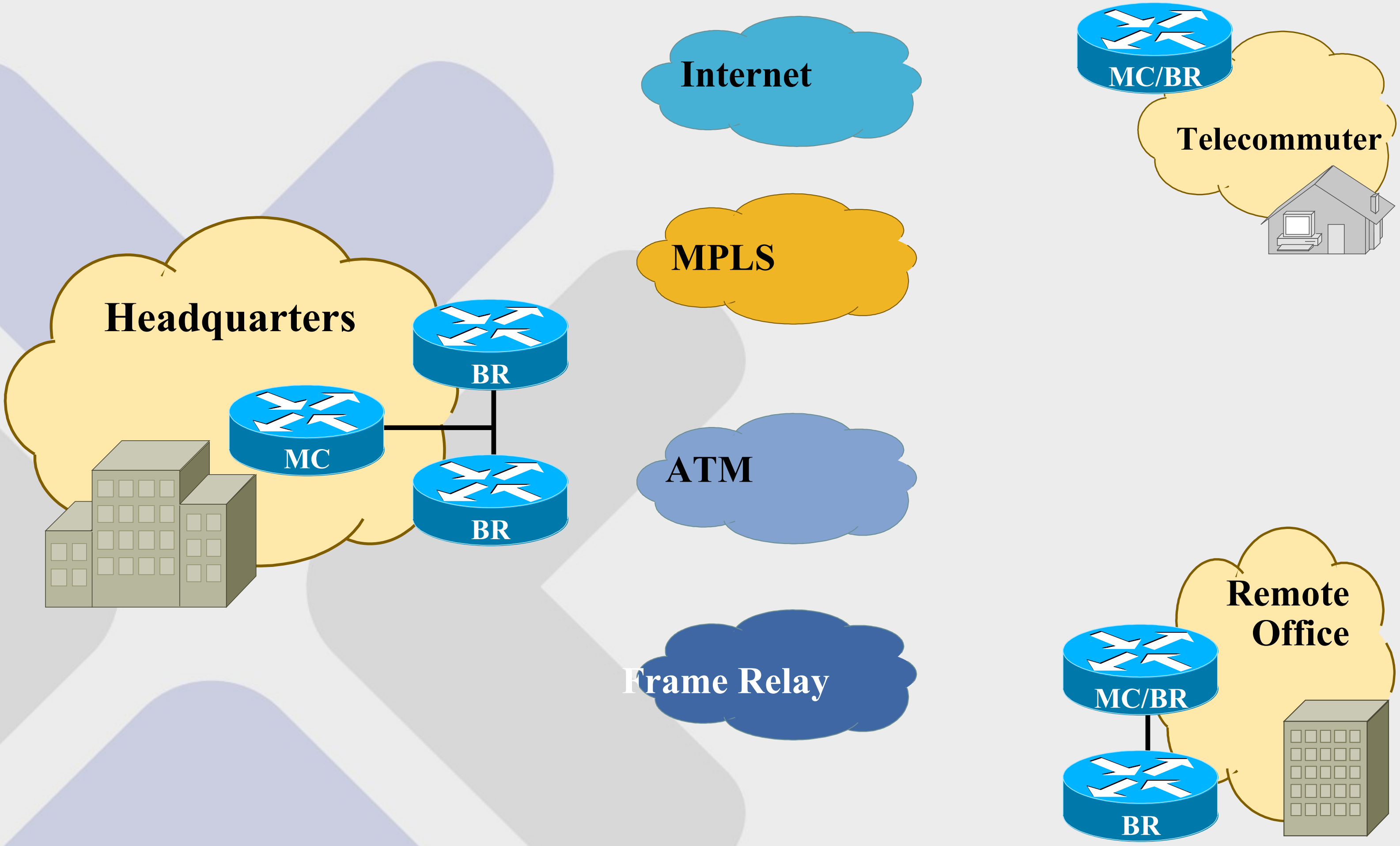
$Cost Minimization Only

MC 10.1.1.1

```
oer master
 no periodic
 resolve cost priority 1
 no resolve delay
 no resolve utilization
 border 10.1.1.2
  interface Serial12/0 external
   cost-minimization tier 100 fee        10000
   cost-minimization tier  75 fee         8000
   cost-minimization tier  40 fee         4000
   cost-minimization end day-of-month 31
  interface Serial13/0 external
   cost-minimization tier 75 fee        3000
   Cost-minimization tier 100          10000
```



Enterprise VPN Deployment

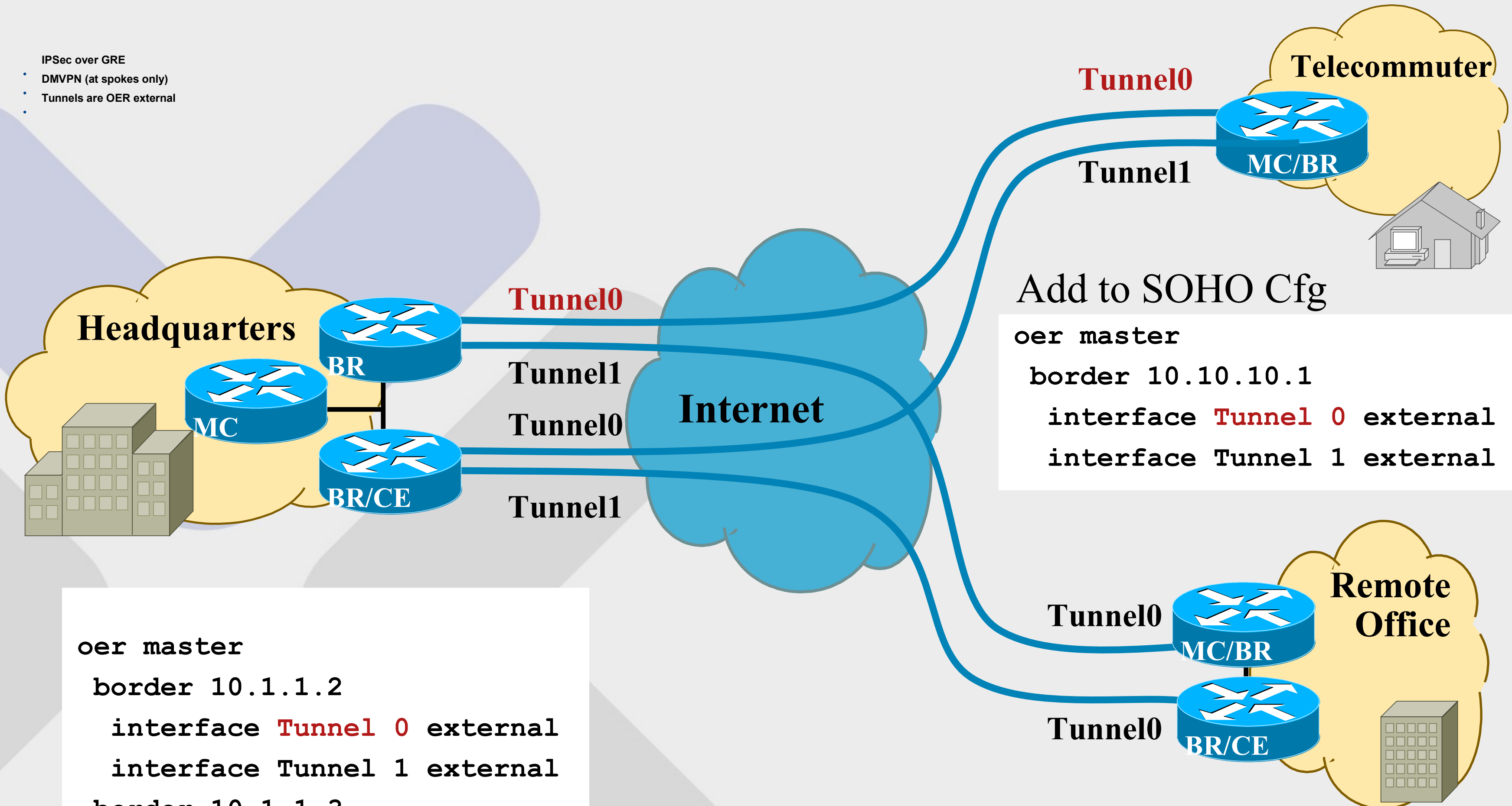


BR—Border Router, MC—Master Controller

Enterprise VPN Deployment

Dual IPSec/GRE Tunnels

- IPSec over GRE
- DMVPN (at spokes only)
- Tunnels are OER external



Add to SOHO Cfg

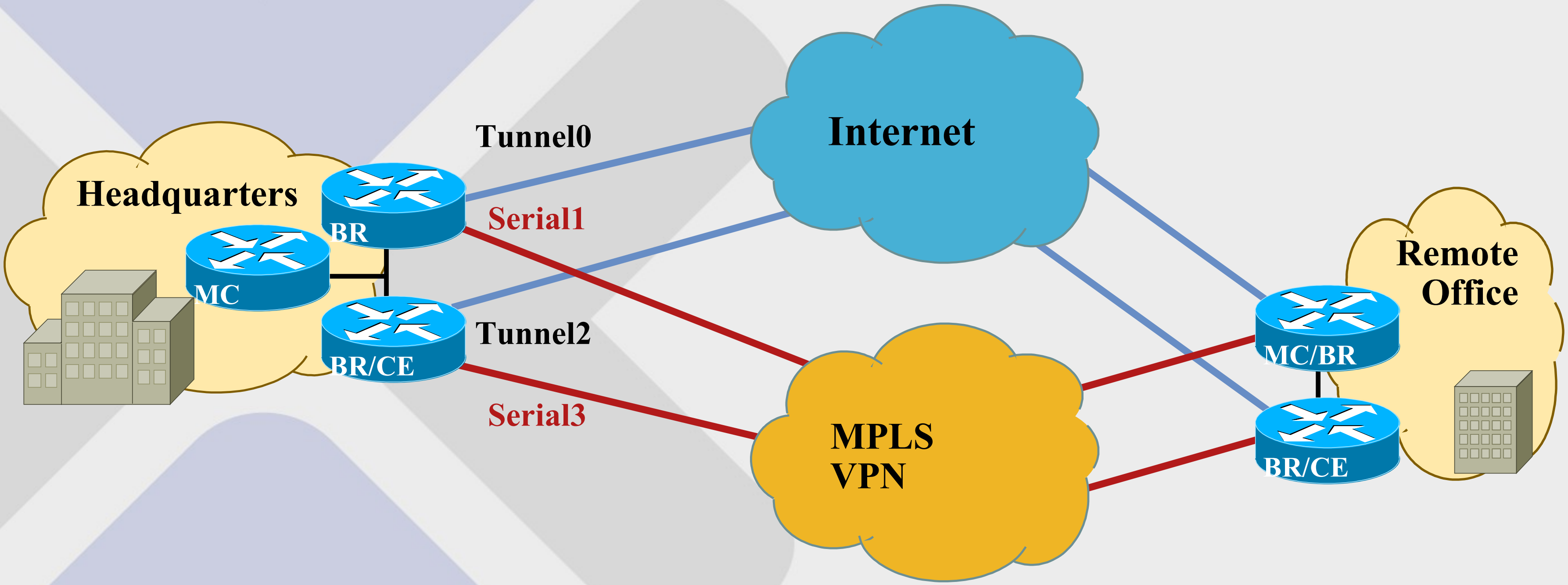
```
oer master
 border 10.10.10.1
  interface Tunnel 0 external
  interface Tunnel 1 external
```

```
oer master
 border 10.1.1.2
  interface Tunnel 0 external
  interface Tunnel 1 external
 border 10.1.1.3
  interface Tunnel 0 external
  interface Tunnel 1 external
```

Enterprise VPN Deployment

MPLS Primary with IPSec/GRE Backup

- **Application 1: Primary MPLS, backup IPSEC**
- **Application 2: I**
- **Backup then performance policy**



BR—Border Router, MC—Master Controller

Enterprise VPN Deployment

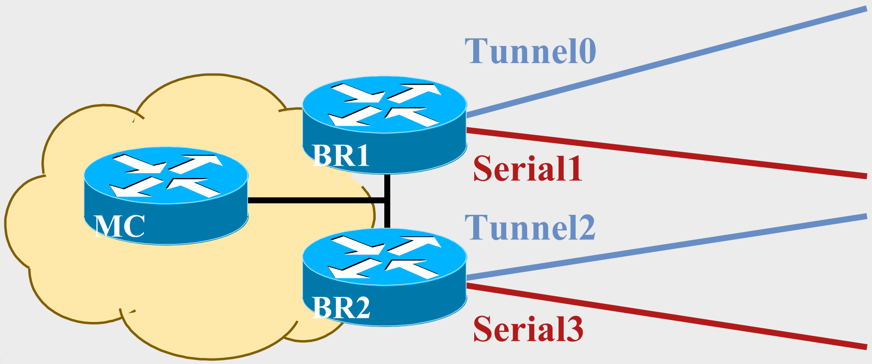
MPLS Primary and IPSec/GRE Backup Configurations*

Group Links

```
oer master
  border 1.1.1.1 key-chain key1
    interface Serial1 external
       link-group RED
    interface Tu0 external
       link-group BLUE
    interface eth1/1 internal

 border 1.1.1.2 key-chain key2
    interface Serial3 external
      link-group RED
    interface Tu2 external
      link-group BLUE
    interface et3/1 internal
```

Specify Link Preference

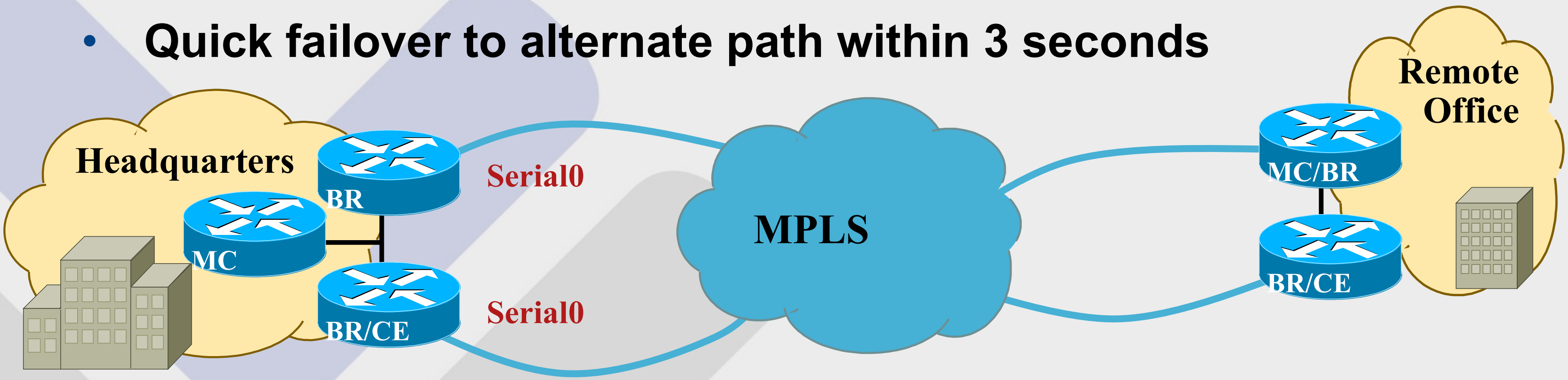
```
oer-map MAP 10
match Appl1
set delay threshold 100
set link-group RED fallback BLUE


oer-map MAP 20
match Appl2
set link-group BLUE
```



*PfR also supported with ISDN and 3G wireless interfaces

Enterprise VPN Deployment
Fast Failover and Load Balancing

- Simultaneous probing on all exits
- Quick failover to alternate path within 3 seconds



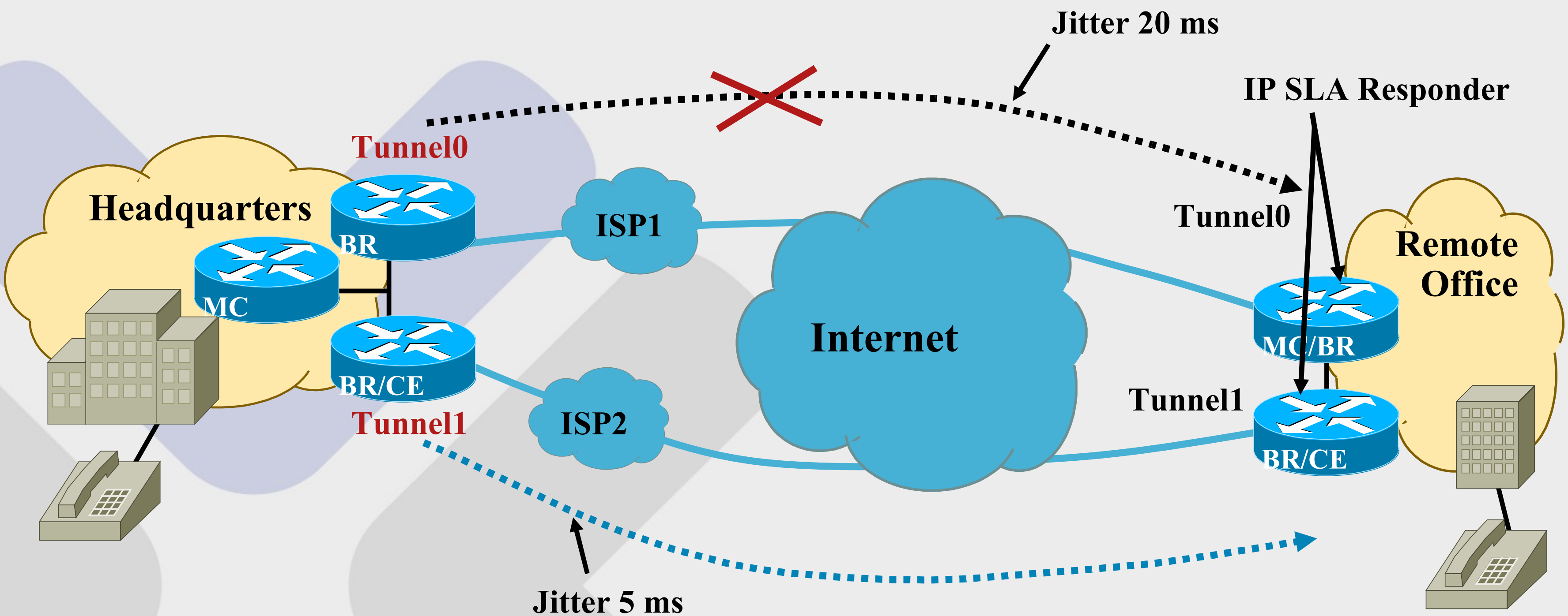

```
Oer master
  max-range-utilization percent 10
  learn
    list sequence 10 refname REM_OFC
      traffic-class prefix-list REM_OFC_LIST
      throughput

Ip prefix-lst REM_OFC_LIST permit 10.1.0.0/16
Ip prefix-lst REM_OFC_LIST deny 0.0.0.0/0
```

```
oer-map MAP 10
  match oer learn list REM_OFC
  set mode monitor fast
  set unreachable threshold 5
  set active-probe echo 10.1.1.1
  set active-probe echo 10.1.1.2
  set probe frequency 2
  set resolve range priority 1
```

BR—Border Router, MC—Master Controller

Enterprise VPN Deployment

Optimize Voice Traffic Between Two Sites



- Select exit with highest percentage of estimated MOS above threshold
  - Tunnel1–5 out of 100 sample had MOS < 4.00 ← better
  - Tunnel0–20 out of 100 sample had MOS < 4.00

BR—Border Router, MC—Master Controller

Enterprise VPN Deployment

Optimize Voice Traffic Between Two Sites

Identify Voice Traffic

Packets marked with DSCP bits

```
ip access-list extended VOICE-LIST
 permit ip any 10.1.1.0 0.0.0.255 dscp
    ef
```

Or

UDP port range

```
ip access-list extended VOICE-LIST
 permit udp any 10.1.1.0 0.0.0.255
  range x y
```

Configure Voice Policy

```
oer-map MAP 10
 match traffic-class access-list
  VOICE-LIST
 set active-probe jitter 10.1.1.1
  target-port 2000 codec g729a
 set probe frequency 2
 set mos percent 20 threshold 4.00
 set resolve mos priority 1
 set mode monitor fast
```

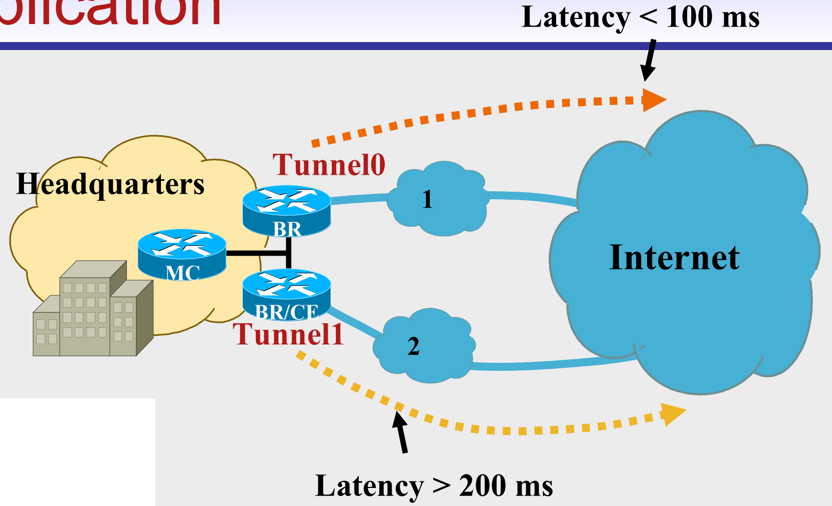
Far End configuration

```
Ip sla responder
```

Enterprise VPN Deployment

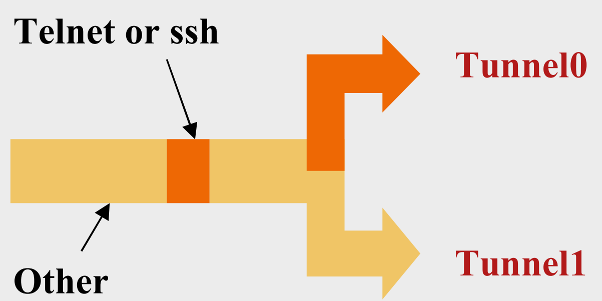
Optimize Application

- **Traffic to branch office**
  - Latency sensitive application—telnet, ssh
  - Latency tolerant—other



Learn Application Traffic

```
Ip prefix-list BRANCH_PFX permit 10.1.0.0/16
!
oer master
 learn
  list sequence 10 refname BRANCH_APPL
   traffic-class application telnet ssh filter BRANCH_PFX
   throughput
  list sequence 20 refname BRANCH_PFX
   traffic-class prefix-list BRANCH_PFX
   throughput
```



Configure Policy

```
oer-map MAP 10
  match oer learn list BRANCH_APPL
  set delay threshold 100
  set resolve delay priority 1 variance 5
```

```
oer-map MAP 20
 match oer learn list BRANCH_PFX
 set delay threshold 400
 set resolve utilization priority 1
   variance 5
```

Enterprise VPN Deployment

Optimize Application—Define Your Own Application

Define Application Using access-list

```
Ip access-list extended APPL1_DEF
  permit tcp any eq 200 any
  permit tcp any any eq 200

Ip access-list extended APPL2_DEF
 permit ip any any dscp af12
```

Add Application Definition to OER Database

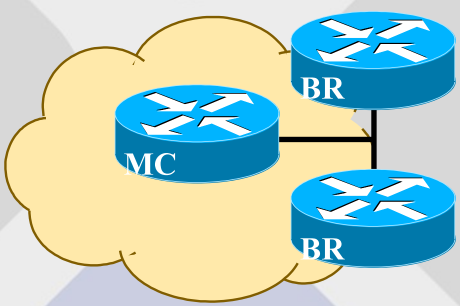
```
Oer master
 application define APPL1 access-list APPL1_DEF
 application define APPL2 access-list APPL2_DEF
```

Learning User Defined Applications

```
Oer master
 learn
  list seq 30 refname LISTA
   traffic-class application APPL1
  list seq 40 refname LISTB
   traffic-class application APPL2
```

Apply Policy to Learned Application

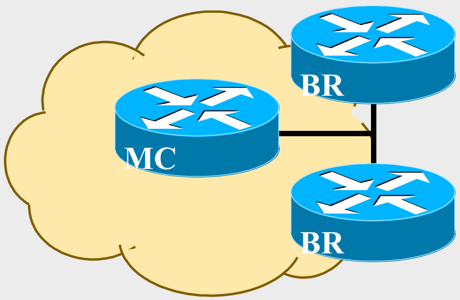
```
oer-map MAP 10
 match traffic-class learn list LISTA
 set resolve delay priority 1 variance 5
oer-map MAP 20
 match traffic-class learn list LISTB
 set resolve range priority 1
```



Enterprise VPN Deployment

Optimize Application Identified by NBAR*

- **Use NBAR to identify application traffic**
- **NBAR is activated automatically on BR**



Learning NBAR Identified Applications

```
Oer master
 learn
  list seq 30 refname LISTA
   traffic-class application nbar rtp-audio
  list seq 40 refname LISTB
   traffic-class application nbar citrix
```

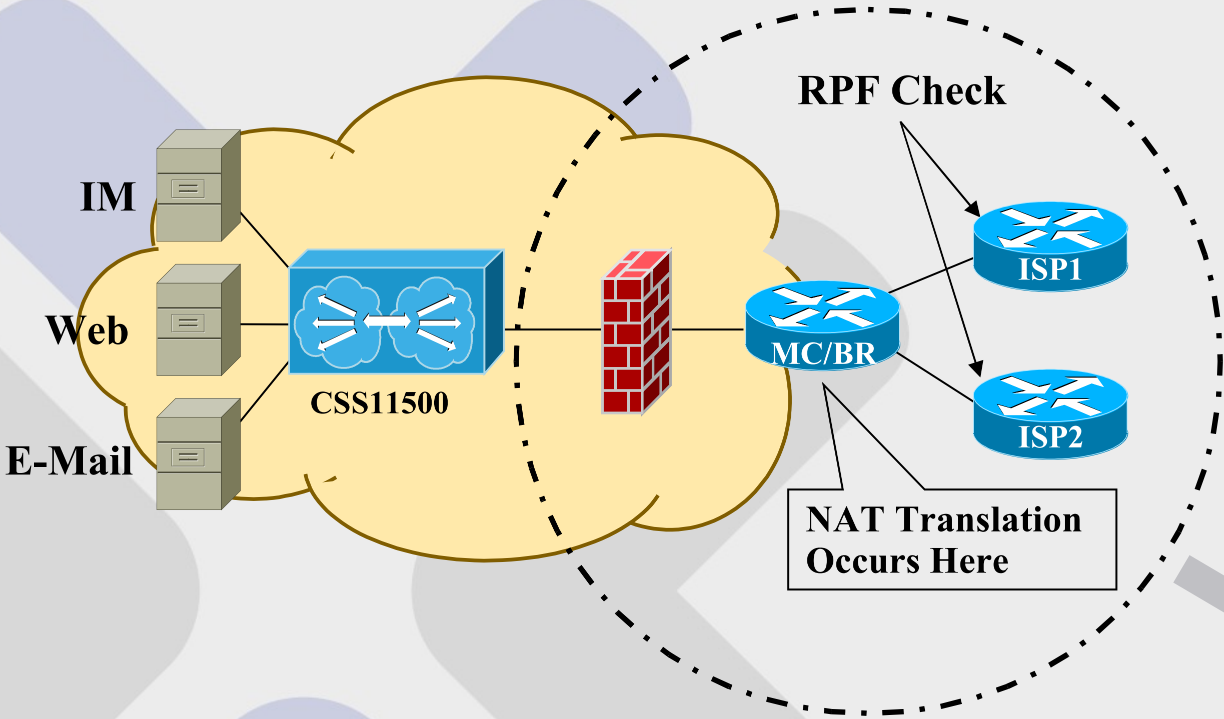
Configure NBAR Identified Applications

```
Ip prefix-list LIST1 permit 10.1.1.0/24
Ip prefix-list LIST1 permit 10.1.2.0/24

Oer-map MAP 10
 match traffic-class application nbar citrix prefix-list LIST1
```

* To be released in 12.5 (1st) T

PfR with NAT

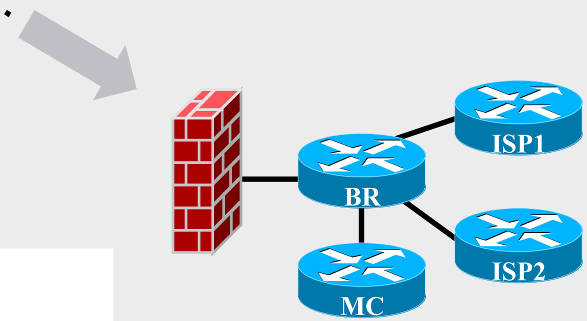
MC/BR Router Combined



- **PfR and NAT**
  - Existing flow continues on same exit; no sessions are dropped
  - New flow goes out via new exit
  - Avoids problems if ISP is performing RPF checking

Minimal Configuration Change

```
interface virtual-template 1
ip nat inside source <x> interface Virtual-Template 1 overload oer
```



With Separate MC and BR

PfR with NAT—Configuration Example

Identify Traffic to be NAT Translated

```
access-list 1 permit 10.1.0.0
     0.0.255.255
route-map isp-1 permit 10
 match ip address 1
 match interface Se1/0
route-map isp-2 permit 10
 match ip address 1
 match interface Se2/0
```

```
interface Eth3/0
  ip nat inside
interface Se1/0
  ip nat outside
interface Se2/0
  ip nat outside
```

OER Internal Interface

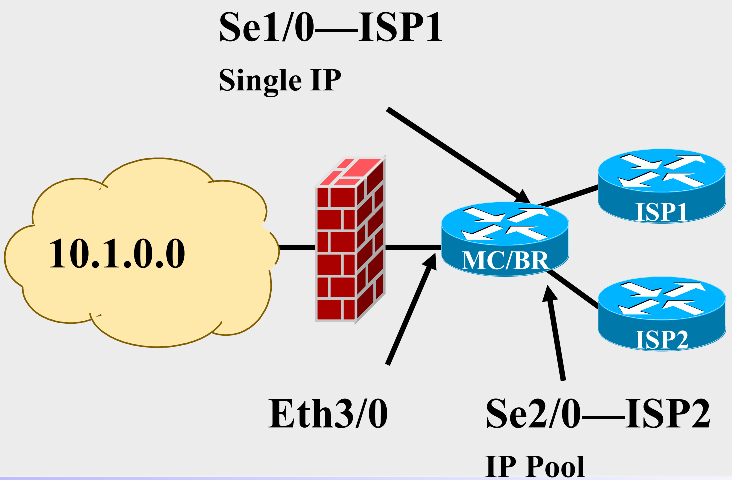
OER External Interface

Single IP

```
interface virtual-template 1
ip nat inside source route-map isp-1 interface
   Virtual-Template1 overload oer
```

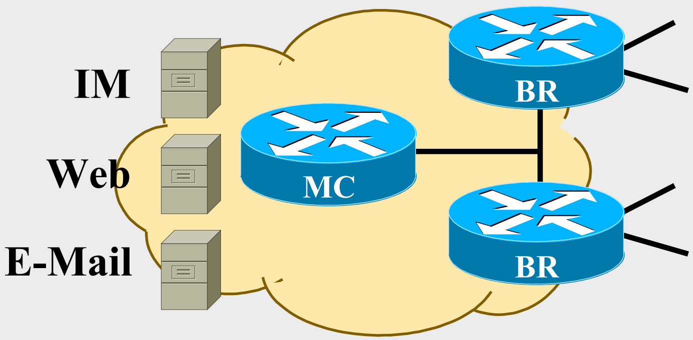
IP Pool

```
ip nat pool ISP-2 <min-ip-addr> <max-ip-addr>
   prefix-length <len>
ip nat inside source route-map isp-2 pool ISP-2
   oer
```



Security Considerations

- Deploy MC behind firewall
- Separate private VLAN for MC and BR
- Private addressing for MC and BR communication
- No routing on MC




```
no ip routing
no router …
```

Routing Not Required on MC

PfR MC Redundancy

- **What if MC goes down?**

  Routing defaults to normal as if PfR was not configured

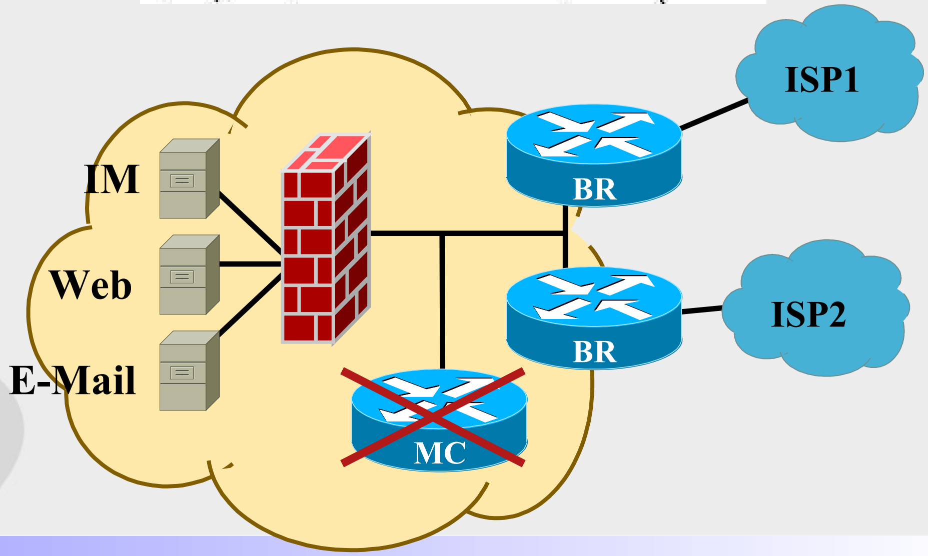
- **Still need MC redundancy?**

  Available

  Stateless redundancy without configuration synchronization available using HSRP

  On roadmap

  Stateless redundancy with synchronized configuration and stateful redundancy



Chesapeake NETCRAFTSMEN

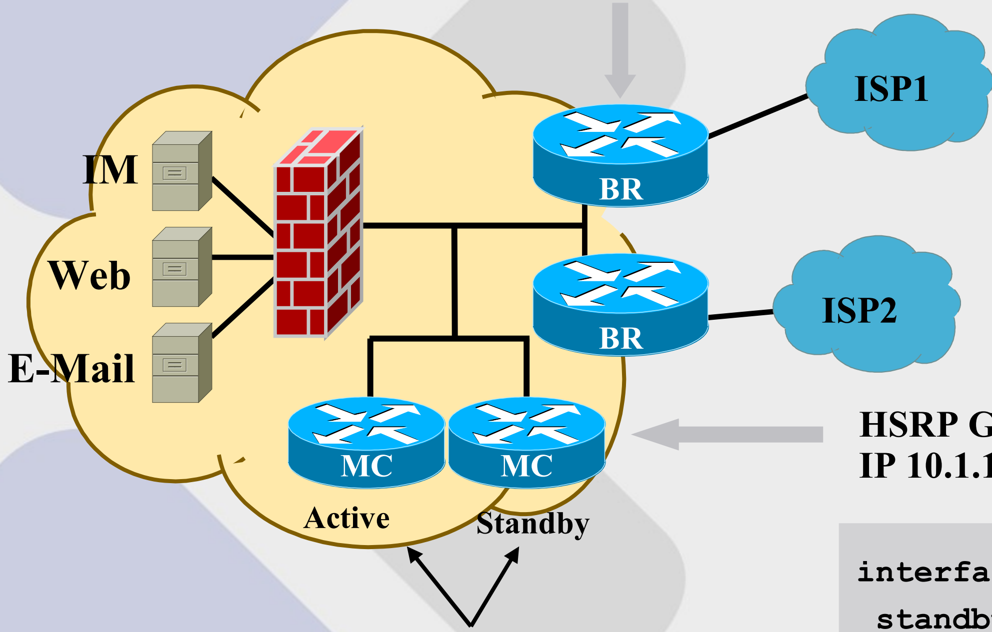
CISCO PARTNER Premier Certified

PfR MC Redundancy

- **Stateless redundancy using HSRP**

On Border Configure HSRP Group IP as MC IP

```
oer border
 master 10.1.1.100 key-chain oer
```



HSRP Group IP 10.1.1.100

```
interface Ethernet0/0
 standby 100 ip 10.1.1.100
```

Duplicate Configuration on both MC