BUILDING HIGHLY AVAILABLE IP AND MPLS NETWORKS

BRKIPM-3011

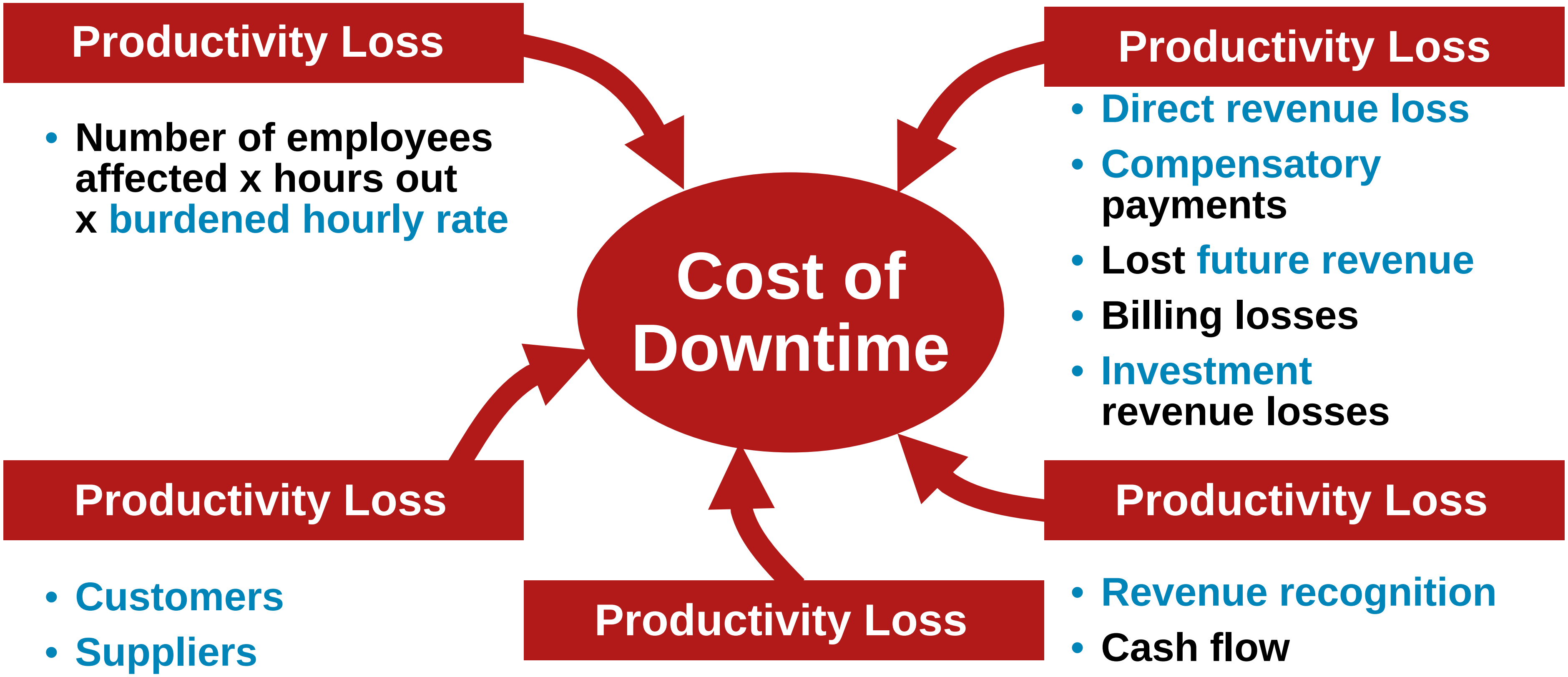
Markus Hies

Cisco Networkers 2007

HOUSEKEEPING

- We value your feedback, don’t forget to complete your online session evaluations after each session and complete the Overall Conference Evaluation which will be available online from Friday.
- Visit the World of Solutions on Level -01!
- Please remember this is a ‘No Smoking’ venue!
- Please switch off your mobile phones!
- Please remember to wear your badge at all times including the Party!
- Do you have a question? Feel free to ask them during the Q&A section or write your question on the Question form given to you and hand it to the Room Monitor when you see them holding up the Q&A sign.

What Are the Costs of Downtime?



Productivity Loss

- Number of employees affected x hours out x burdened hourly rate

Productivity Loss

- Direct revenue loss
- Compensatory payments
- Lost future revenue
- Billing losses
- Investment revenue losses

Cost of Downtime

Productivity Loss

- Customers
- Suppliers
- Financial markets
- Business partners
- Careers lost

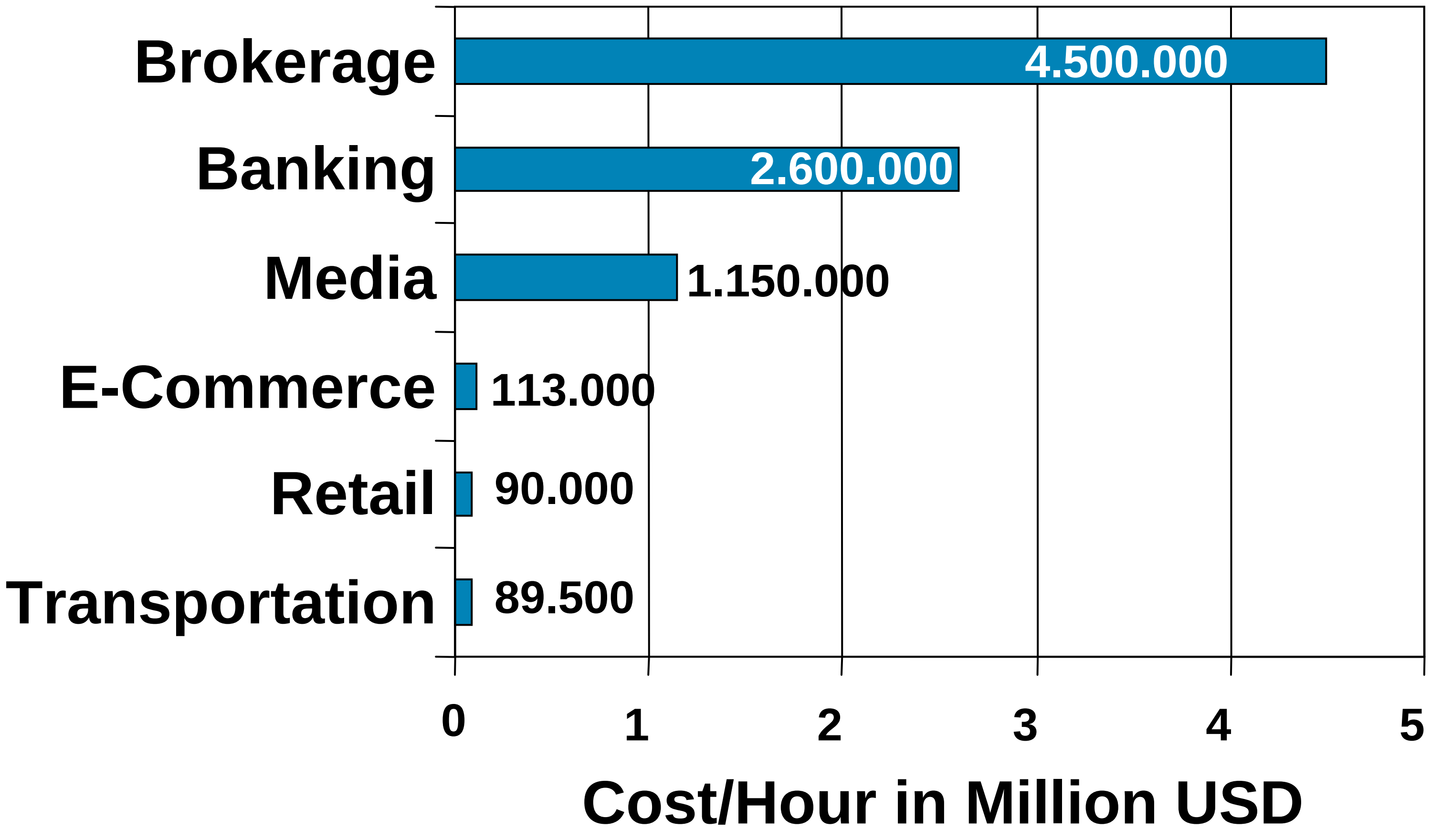
Productivity Loss

- Temporary employees
- Equipment rental
- Overtime costs
- Travel expenses

Productivity Loss

- Revenue recognition
- Cash flow
- Lost discounts
- Payment guarantees
- Credit rating
- Stock price

The Cost of Network Downtime per Hour



Source: Yankee Report: The Road to a Five-Nines Network, 2004

“24x7 availability is designed in — not bought, is expensive and requires a strategy and plan.”

“Surviving in a 24 hours world”,
Gartner, 2001

Agenda

- **High Availability Fundamentals:**

  Definition, MTBF, MTTR, Calculate vs Measured Availability

- **System Level Resiliency**

  SSO, NSF, NSR, Warm Reload/Upgrade, ISSU

- **Network Level Resiliency**

  IP Event Dampening, BFD

  Fast Convergence, Fast Rerouting

- **Embedded Management**

  MPLS Diagnostic Expert

  Generic Online Diagnostic

  Embedded Event Manager

- **High Availability Best Practises**

  The Culture of Availability

  Trouble Ticket Availability Measures (Cisco NAIS Service)

- **References**

- **Summary**

© 2007 Cisco Systems, Inc. All rights reserved. Cisco Public

HIGH AVAILABILITY FUNDAMENTALS

Availability Definitions

$$\text{Availability} = \frac{\text{MTBF}}{\text{MTBF + MTTR}}$$

You can simply read, “The uptime divided by the total time” to create the percentage time your network is operational

- Mean Time Between Failure (MTBF)

  When does it fail?

- Mean Time To Repair (MTTR)

  How long does it take to fix?

© 2007 Cisco Systems, Inc. All rights reserved.

Calculation of Availability of Complex Systems

- **MTBF -> Calculate**

  Cisco uses Industry standards to compute Hardware MTBF can be calculated

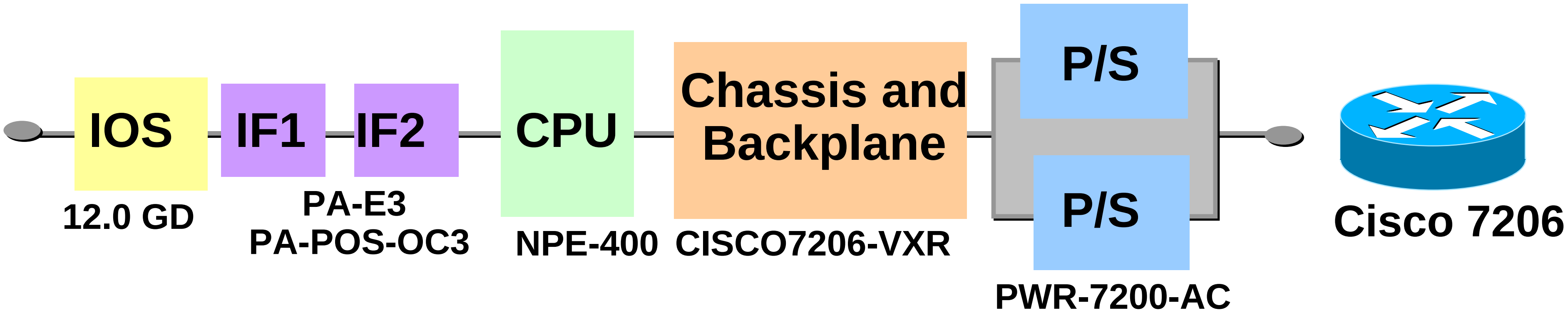
- **MTTR -> Estimate**

  Can be reasonable estimated (e.g. Reboot, exchange chassis/LC)

$$A_{Series} = \prod_{k=1}^{N} A_K = A_1 \text{ x } A_2 \text{ x } \cdots \text{x } A_N$$

$$A_{Parallel} = 1 - \prod_{K=1}^{N} (1 - A_k) = 1 - (1\text{-}A_1) \text{ x } \cdots \text{ x } (1\text{-}A_N)$$

Device Availability Calculation: Cisco 7206



$$\text{IOS} = \frac{30.000}{30.000 + 0.1} = 0.999997$$

$$\text{CPU} = \frac{490.000}{490.000 + 4} = 0.999992$$

$$\text{IF1} = \frac{1.120.000}{1.120.000 + 4} = 0.999996$$

$$\text{BB} = \frac{460.000}{460.000 + 8} = 0.999983$$

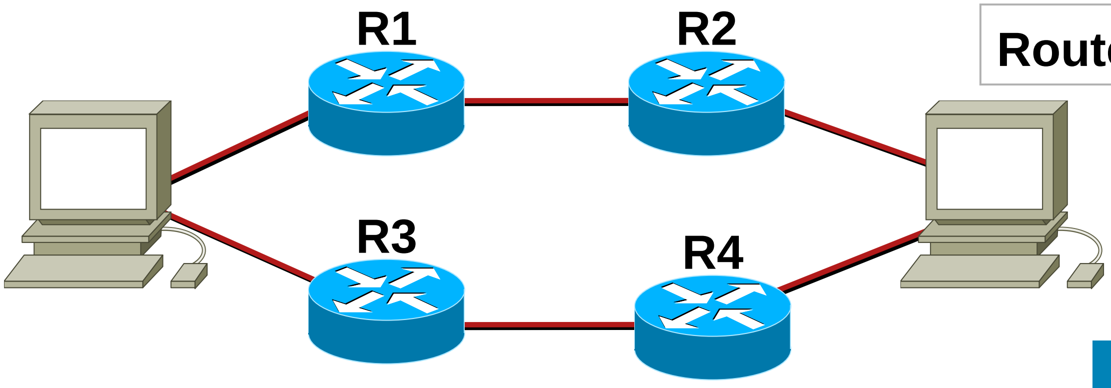
$$\text{IF2} = \frac{600.000}{600.000 + 4} = 0.999993$$

$$\text{P/S} = \frac{750.000}{750.000 + 4} = 0.999995$$

System Availability = 0.999997 * ....*0.999983*(1-(1-0.999995)^2)
= 0.999961 = 99.9961%

Calculated MTBF Values from Cisco Database

Network Availability Calculation



Router R1,R2,R3, R4: 0.999961

but not considered:
-Links (WAN, LAN)
-Computer NICs
-Computer OS
-Computer Applications

1

Availability of R1 and R2 in Series
= (0.999961*0.999961) = 0.99992175

2

Availability of Parallel Network Path (R1-R4)
= 1 - ((1-0.999921)(1 - 0.999921)) = 0.999999994

3

Network Availability = 99.9999%
Only Based on Device Availability Values

Availability Calculation vs Measurement

- **Calculation based on:**
  - –component MTBF and MTTR

    different underlying models, simulations
  - –network design (redundancy)
- **Estimation based on:**
  - –HW/SW exchange processes (MTTR)
  - –Resiliency features (e.g. Fast Convergence for MTTR)
- **Measurement based on:**
  - –ICMP Reachability (E2E, Device)
  - –Cisco Cisco IOS IP Service Level Agreement (IP SLA)

    network performance measurement and diagnostics tool
  - –Trouble Ticket Analysis
  - –Outage Logs Analysis
  - –History Method: observe shipping/RMA and project for MTBF

What You Measure Will Improve

What Is High Availability?

Availability	DPM	Downtime per Year (24x365)		
99.000%	10000	3 Days	15 Hours	36 Minutes
99.500%	5000	1 Day	19 Hours	48 Minutes
99.900%	1000		8 Hours	46 Minutes
99.950%	500		4 Hours	23 Minutes
99.990%	100			53 Minutes
99.999%	10			5 Minutes
99.9999%	1			30 seconds

Reactive

Proactive

Predictive

"High Availability"

- HA is hard work, NO Silver Bullets
- Adding a "9" can cost significantly more
- Two ways to state availability of a network:
    - Percentage Method
    - DPM Method = Defects per Million (Hours of Running Time)

© 2007 Cisco Systems, Inc. All rights reserved.

Cisco High Availability Focus: End-to-End

System Level Resiliency

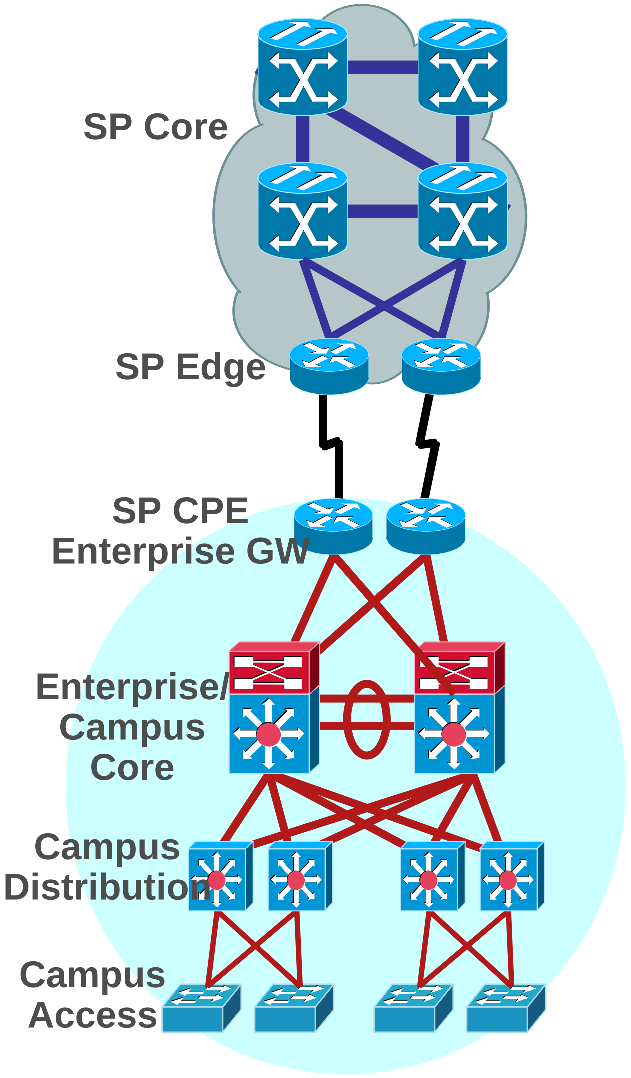
- at critical network edges
- increase MTBF using reliable and rebust HW/SW designed for HA
- minimize MTTR for system failures (resilient HW)
- Mitigate planned outages by providing hitless software upgrades

Network Level Resiliency

- Fast rerouting in the core and where redundant paths exist
- Deliver features for fast network convergence, protection & restoration

Embedded Management & Automation

- embedded management with active devices
- intelligent event management for proactive maintenance
- Automation and configuration management to reduce human errors



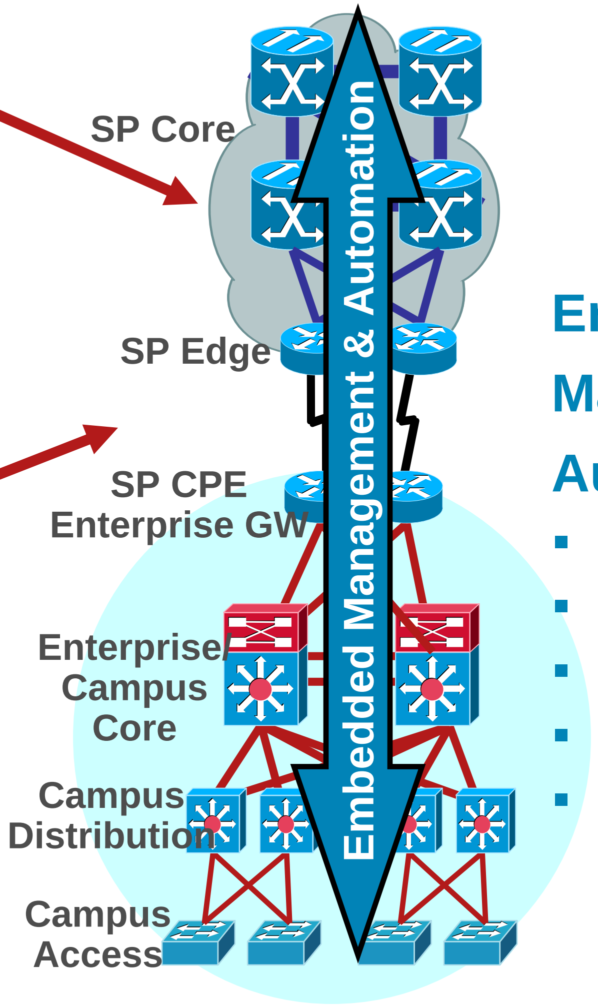
Cisco HA Feature Toolbox

Network Level Resiliency

- NSF Awareness
- IP Event Dampening
- Bi-Directional Forwarding Detection (BFD)
- Fast Convergence
  - BGP Convergence Optimalization
  - iSPF Optimalization (OSPF, IS-IS)
  - Multicast Subsecond Convergence
- Fast Rerouting (IP and MPLS)

System Level Resiliency

- Control/Data Plane Resiliency:
  - HSA, RPR, RPR+, Stateful NAT/IPSec/FW,
  - NSF /w SSO including MPLS
  - BGP Nonstop Routing
  - Control Plane Policing, GLBP, HSRP,
  - Warm Reload
- Planned Outages: ISSU, Warm Upgrade
- Link Resiliency:
  - Line Card Redundancy with Y-Cable
  - Link Bundeling (Etherchannel/POS-Channel)



Embedded Management & Automation:

- CiscoWorks
- MPLS OAM (ISC)
- EEM
- GOLD
- .....

covered in detail
covered in brief

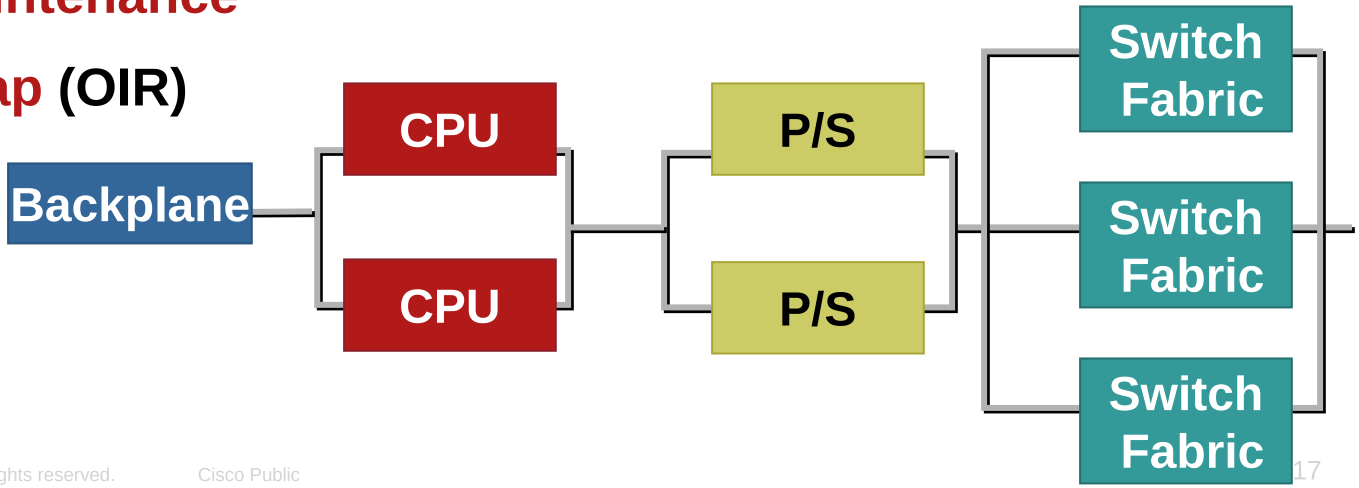
© 2007 Cisco Systems, Inc. All rights reserved.

SYSTEM LEVEL RESILIENCY:

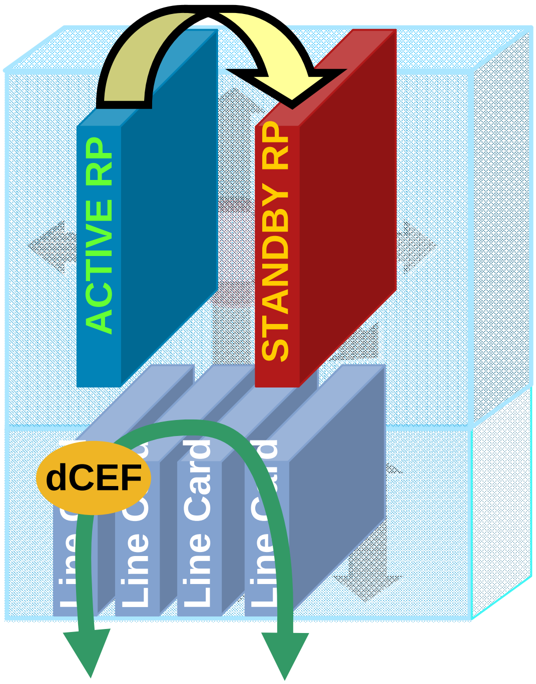
Stateful Switchover (SSO)

Improving Hardware Availability

- Reliable Hardware
- Load sharing redundancy
- Active/standby redundancy (processor, power, fans, line-cards)
- Active/standby fault detection
- Card MTBF (100,000 hrs)
- ECC Memory
- Separate control and forwarding plane
- Spares and Maintenance
- Robust hot swap (OIR)



Dual Route Processor Resiliency

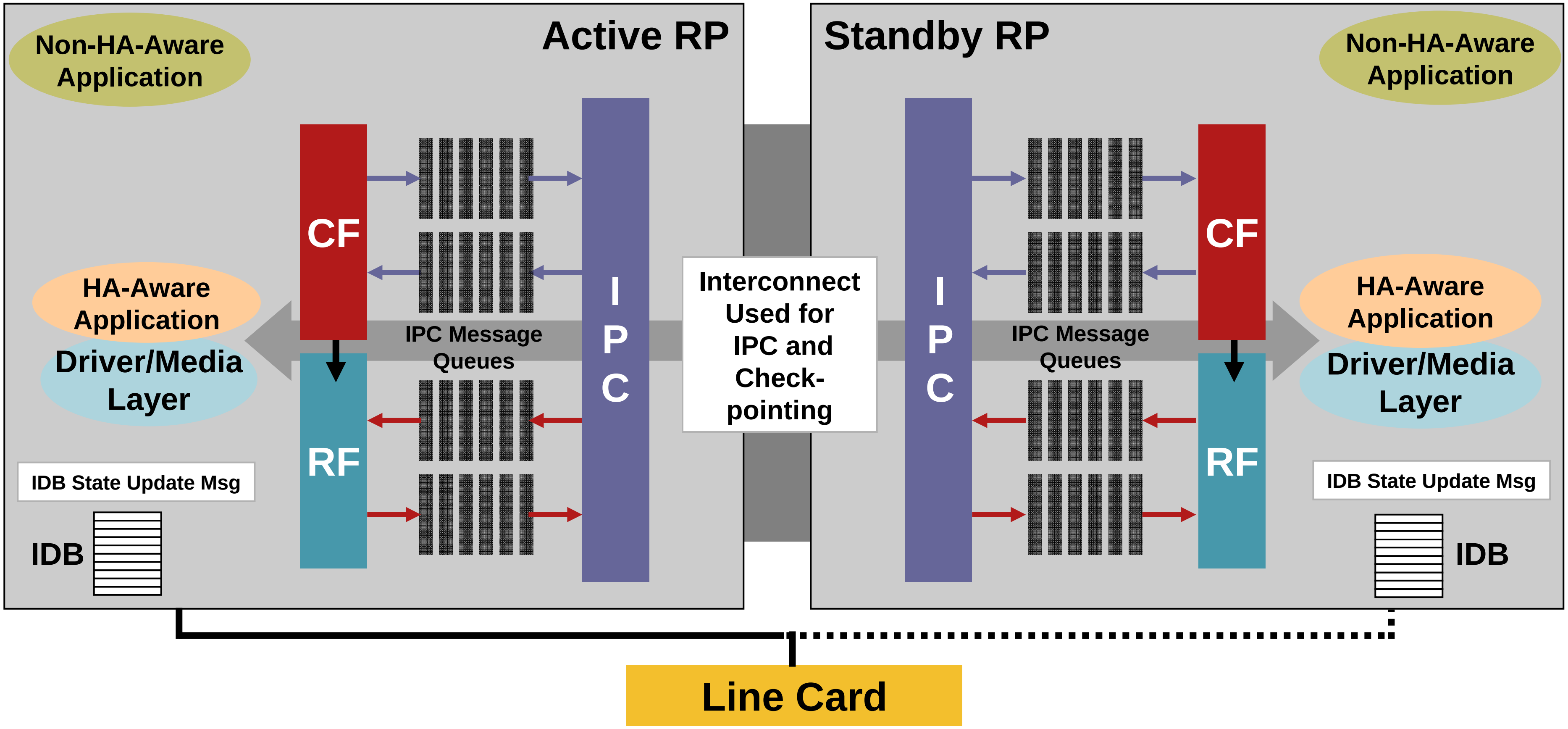


- **Cold Redundancy (2001)**
  - **HSA High System Availability (identify failure)**
  - **RPR RP Redundancy (preload/boot standby RP)**
- **Warm Redundancy (2002)**
  - **RPR+ (no reset/reload of LCs)**
- **Hot Redundancy (2004)**
  - **GR / NSF / NSR w/ SSO Graceful Restart / Non-Stop Forwarding/Routing with Stateful Switchover**
- **In-Service-Software-Upgrade: ISSU (2005)**

- **Standby RP takes control of router after hardware/software fault on active RP**
- **SSO allows standby RP to take immediate control and maintain connectivity protocols**
- **NSF continues to forward packets until route convergence is complete, need neigbor awareness**
- **NSR works with SSO to synchronize routing information between active and standby**
- **GR (graceful restart) reestablish routing information without churning the network**
- **Ultimate Goal: achieve 0% packet loss**

© 2007 Cisco Systems, Inc. All rights reserved.
Cisco Public

SSO Infrastructure

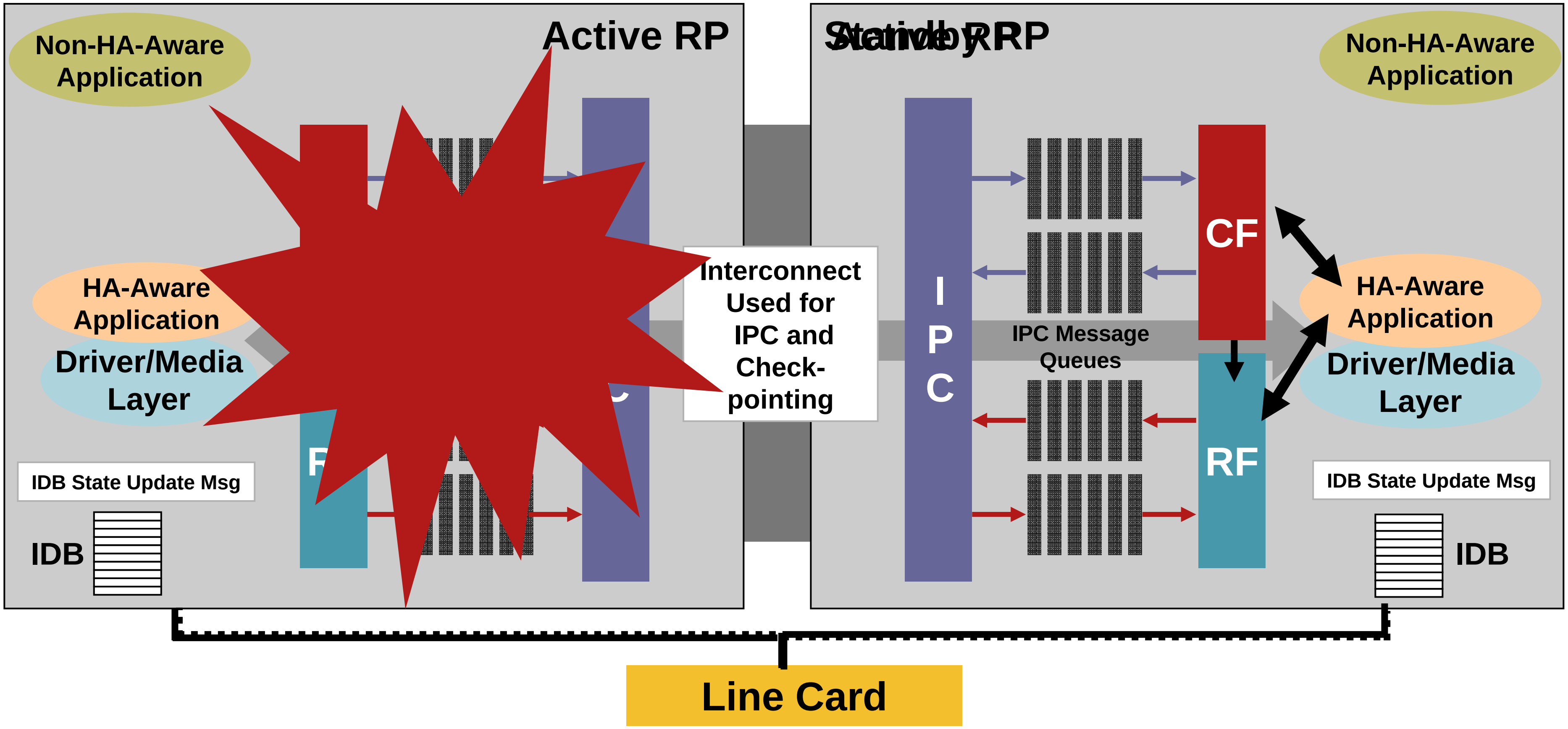


- **RF (Redundancy Facility): monitoring and reporting of RP transitions**
- **CF (Checkpointing Facility): allows clients to send state updates from Active to Standby**
- **IPC (Inter-Process Communication): transport for CF, RF and Config Sync**
- **Driver/Media layer: platform independent/dependent code to maintain IDB state**
- **Config Sync: maintains the same configuration on the Standby as on the Active**

© 2007 Cisco Systems, Inc. All rights reserved.

SSO Architecture with Stateful L2 Protocols (PPP, FR, HDLC …)

animated



- Failure on active RP initializes through RF messages a switchover
- L2 Information is maintained across switchover using the CF messages
- Line Cards are connected to new Active RP
- adjacent devices do not see a link failure/flap during switchover

SSO Supported Protocols and Applications

Line Protocols and Features

- ATM
- APS
- Frame Relay
- HDLC
- PPP
- SRPLink negotiation
- VLANs, VTP, trunks, DTP
- Spanning tree
- UDLD
- SPAN/RSPAN
- 802.1x
- Port security

•Traffic storm control

•L2 protocol tunneling

•Flow control

•LACP/PAGP

•MAC move notification

•ARP

•Diagnostics

•DAI, IPSG, Port Security

•...................

Other Applications

- Access control lists
- QoS policers
- IP Multicast entries
- FIB/CEF Table
- Adjacency Table
- MAC-address Table
- Routing Protocols

Line Card Drivers

•Platform dependend

•Loaded with IOS image

•Linecard status information

For a complete list check release notes of the platform/ IOS Release and the Feature Navigator.

© 2007 Cisco Systems, Inc. All rights reserved.

Enabling and monitoring SSO

```
RouterA(config)# redundancy
RouterA(red-config)# mode ?
rpr       Route Processor Redundancy
rpr-plus  Route Processor Redundancy Plus
sso       Stateful Switchover (Hot Standby)
```

```
show redundancy [all | arbitration | clients | counters | history | negotiation |
   switchover | standby-cpu | states | trace | trace all]
```

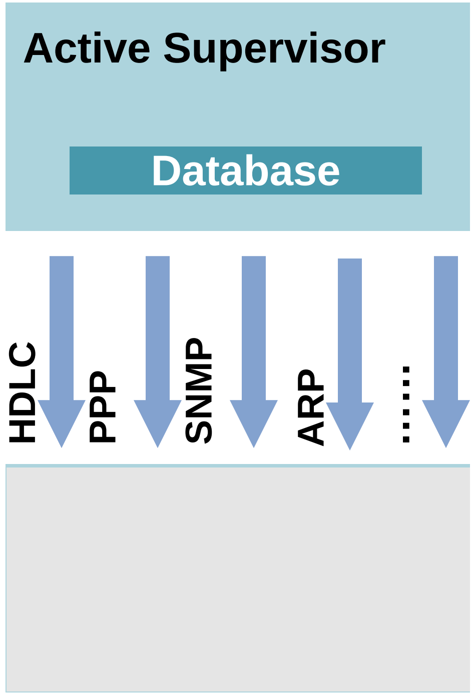
Cisco 12000 syntax options with 12.0(S)

```
router# show redundancy states
my state = 13 -ACTIVE
peer state = 8 -STANDBY HOT
<snip>
client count = 13
Redundancy Mode = SSO
```

© 2007 Cisco Systems, Inc. All rights reserved.

SSO Operation Example

```
Router# show redundancy client
clientID = 0 clientSeq = 0      RF_INTERNAL_MSG
clientID = 25 clientSeq = 130 CHKPT RF
clientID = 27 clientSeq = 132 C12K RF COMMON Client
clientID = 30 clientSeq = 135 Redundancy Mode RF
clientID = 22 clientSeq = 140 Network RF Client
clientID = 24 clientSeq = 150 CEF RRP RF Client
clientID = 49 clientSeq = 225 HDLC
clientID = 21 clientSeq = 320 PPP RF
clientID = 34 clientSeq = 330 SNMP RF Client
<snip>
```



```
router# show redundancy switchover history
Index Prev Active Curr Active     Swact Reason    Swact Time
1          1             0               unsupported     8:03:52 UTC Thu Nov 29 2003
2          0             1               unsupported     08:07:00 UTC Thu Nov 29 2003
```

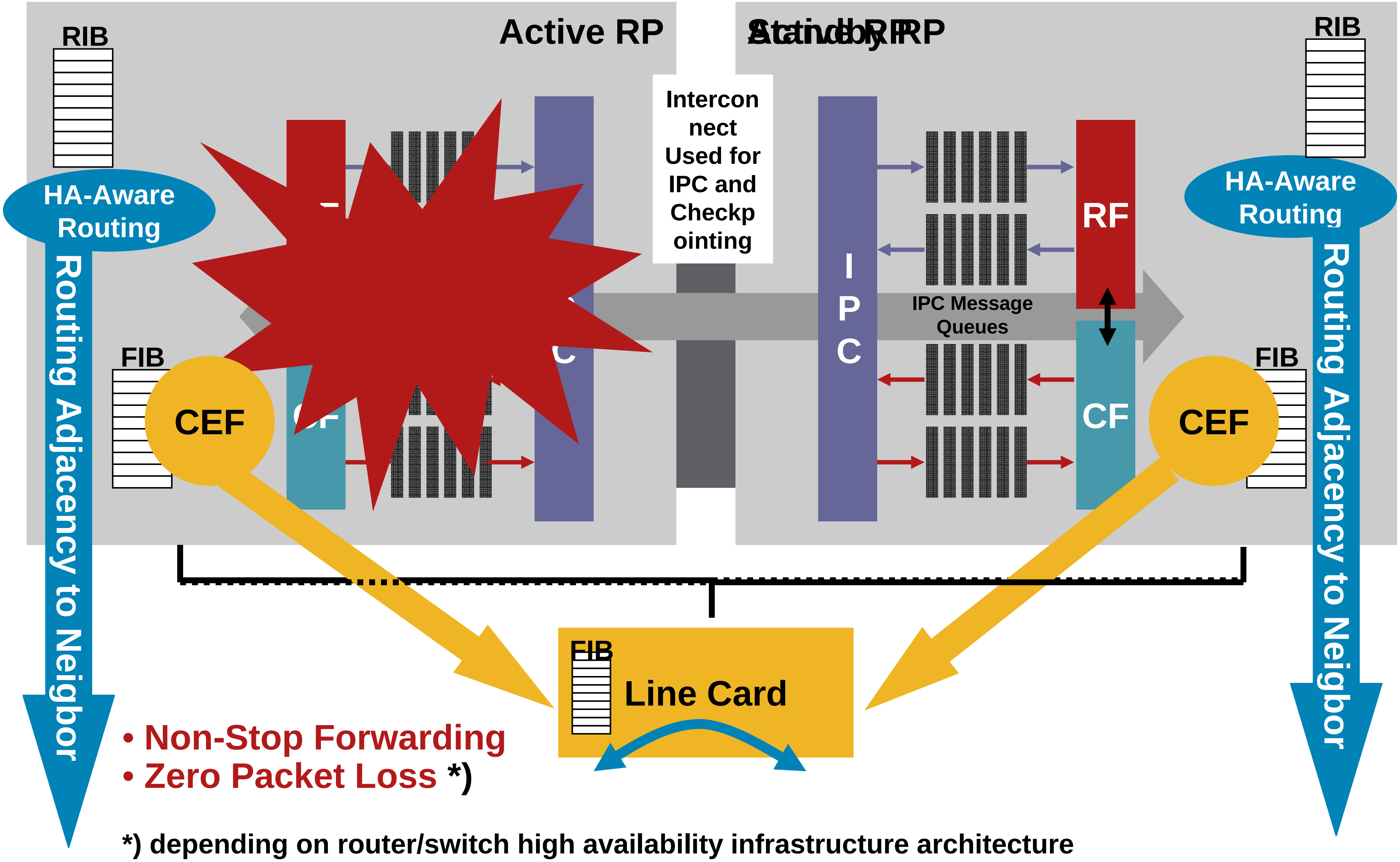
© 2007 Cisco Systems, Inc. All rights reserved.

SYSTEM LEVEL RESILIENCY:

Non-Stop Forwarding (NSF)

animated

NSF Architecture with HA-aware Routing



- Non-Stop Forwarding
- Zero Packet Loss *)

*) depending on router/switch high availability infrastructure architecture

Requirements and Enhancements for NSF-aware Routing Protocol

Requirements:

- Switchover MUST be completed before dead/hold timer expires

  Else peers will reset the adjacency and reroute the traffic

- FIB MUST remain unchanged during switchover

  Current routes marked as "stale" during restart

  "Cleaned" once convergence is complete

  Transient routing loops or black holes MAY be introduced if the network topology changes before the FIB is updated

- Adjacencies MUST NOT be reset when switchover is complete

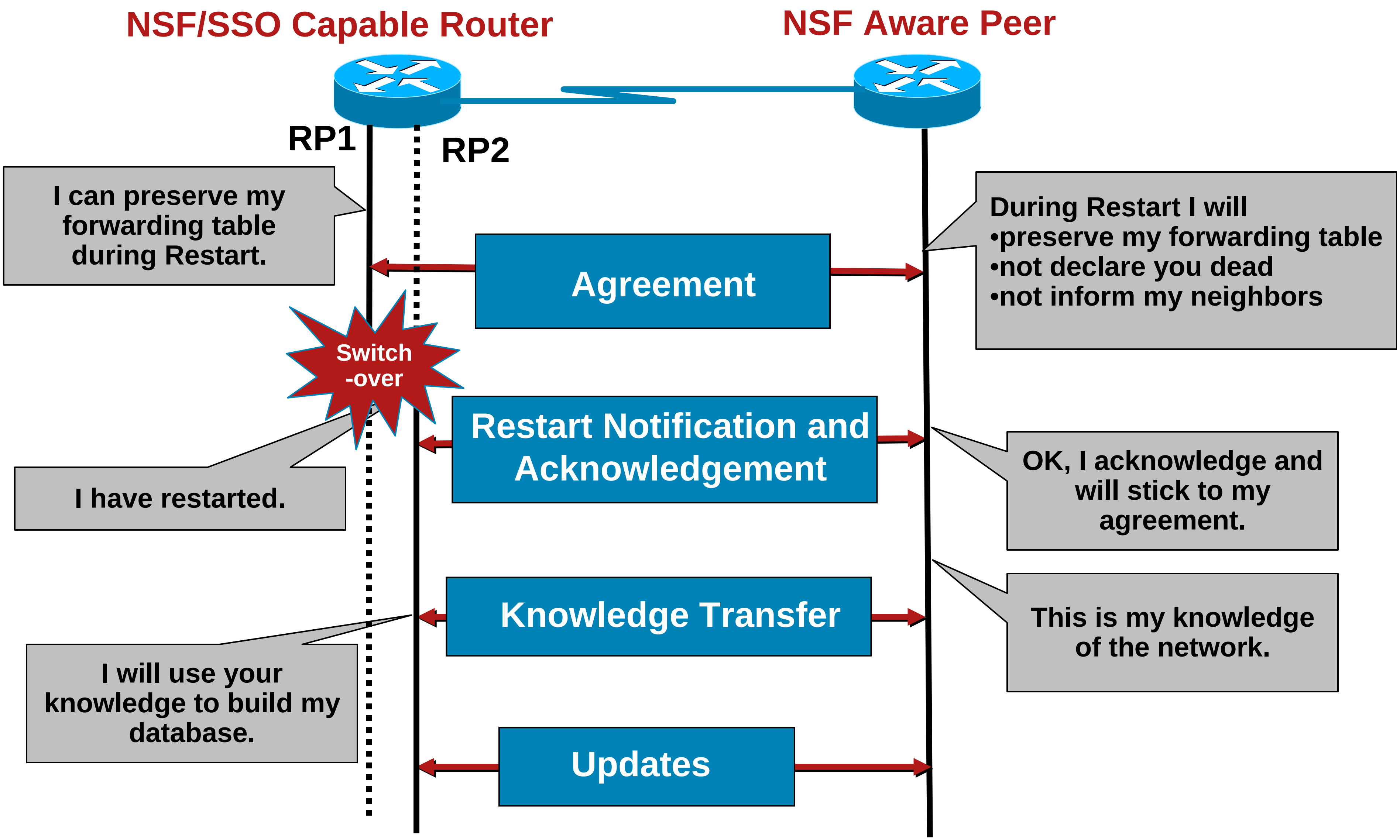
  Protocol state is not maintained

Enhancements:

- Neighbors must know that NSF router can still forward packets: NSF aware, as opposed to NSF capable
- Enhancements to ISIS, OSPF, EIGRP, BGP and LDP designed to prevent route flapping

© 2007 Cisco Systems, Inc. All rights reserved.

Relationship Building of NSF-aware Routing Protocols



I can preserve my forwarding table during Restart.

During Restart I will
- preserve my forwarding table
- not declare you dead
- not inform my neighbors

I have restarted.

OK, I acknowledge and will stick to my agreement.

This is my knowledge of the network.

I will use your knowledge to build my database.

© 2007 Cisco Systems, Inc. All rights reserved.

NSF IGP Routing Protocols Extensions:

Enabling NSF in Routing Protocols:

```
router eigrp / ospf / isis
   nsf
   <protocol specific timer/interval configuration>
```

Relevant Standards and Drafts

- The mechanisms used to provide continuous forwarding in the event of a route processor switchover are not completely standardized
- 2 different OSPF implementations: Cisco’s OSFP NSF vs IETF Graceful OSPF Restart
- Cisco’s NSF implementation for ISIS (ietf option) follows the specification described in RFC 3847 Restart Signaling for Intermediate System to Intermediate System (IS-IS) -> stateful solution providing NSR exists also

For Details see NSF Deployment Guide:
http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper0900aecd801dc5e2.shtml

© 2007 Cisco Systems, Inc. All rights reserved.

BGP Graceful Restart

- IETF: draft-ietf-idr-restart-13.txt

  submitted to IESG as proposed standard (expires 1/2007)

- Provides a graceful recovery mechanism for a restarting BGP process

- Implementation on Cisco IOS/XR:

  12.0S, 12.2T, 12.2S, XR 2.0

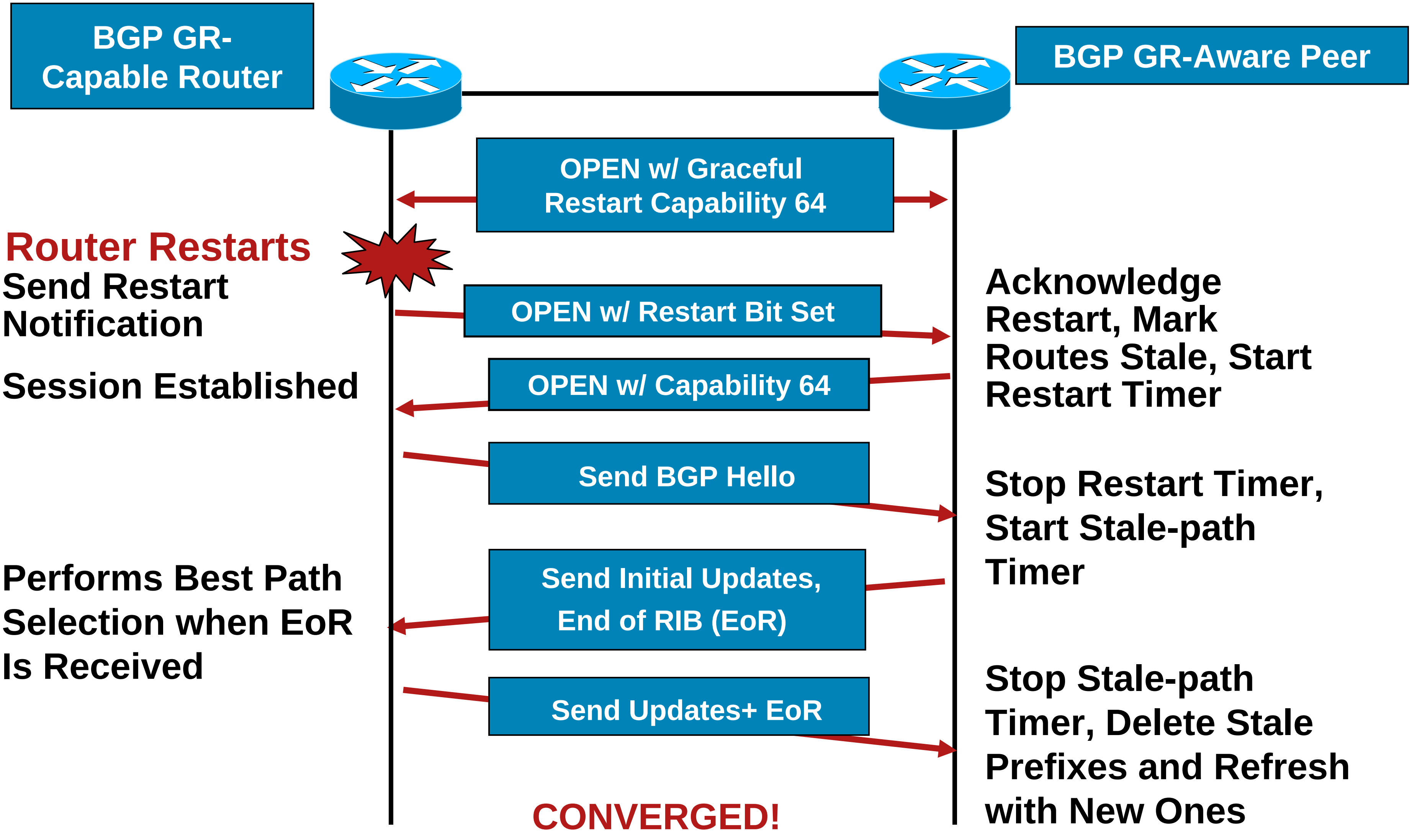
  release and device dependent

- Requires a graceful restart aware neighbor

- Graceful restart capable routers are 7300, 7500, 7600, 12000, 10000, CRS-1

© 2007 Cisco Systems, Inc. All rights reserved. Cisco Public

Graceful Restart BGP Operation



BGP GR-Capable Router

BGP GR-Aware Peer

OPEN w/ Graceful Restart Capability 64

Router Restarts

Send Restart Notification

OPEN w/ Restart Bit Set

Acknowledge Restart, Mark Routes Stale, Start Restart Timer

OPEN w/ Capability 64

Session Established

Send BGP Hello

Stop Restart Timer, Start Stale-path Timer

Send Initial Updates, End of RIB (EoR)

Performs Best Path Selection when EoR Is Received

Send Updates+ EoR

Stop Stale-path Timer, Delete Stale Prefixes and Refresh with New Ones

CONVERGED!

© 2007 Cisco Systems, Inc. All rights reserved.

BGP Graceful Restart Commands

```
router bgp 100
   bgp graceful-restart
   bgp graceful-restart restart-time 120
   bgp graceful-restart stalepath-time 360
```

- **Restart timers:** max time peer waits for reconection BGP session (def. 120s, adv. to peer)
- **Stalepath timers:** upper limit, how long peer will continue to use stale routes after re-establishing BGP (def. 360s, used internally)
- **BGP hold time**: 180 seconds (3 x 60 sec keep-alive)

```
show ip bgp neigbor
  BGP neighbor is 10.10.104.1,  remote AS 100, internal link
  BGP version 4, remote router ID 10.10.104.1
  BGP state = Established, up for 00:00:10
  Last read 00:00:09, hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
  Route refresh: advertised and received(new)
  Address family IPv4 Unicast: advertised and received
  Graceful Restart Capabilty: advertised and received
  Remote Restart timer is 140 seconds
  Address families preserved by peer:
  IPv4 Unicast
```

Indicates Neighbor
Is NSF Aware

Towards Zero Downtime: Route Processor High Availability Feature Evolution

Technology/Router	Link Flap	Route Flap	Cisco 7500 Series Router	Cisco 12000 Series Router	Cisco 10000 Series Router	Cisco 6500/7600 Series Routers
Single Route Processor Reboot	YES	YES	8:00 Minutes	2:32 Minutes	2:45 Minutes	
Cold Redundancy (Cisco RPR)	YES	YES	2:06 Minutes	1:20 Minutes	0:26 Seconds	3:00 Minutes
Warm Redundancy (Cisco RPR+)	NO	YES	0:30 Seconds	0:08 Seconds	0:14 Seconds	0:15 – 0:30 Seconds
Hot Redundancy Cisco NSF with SSO (Available Since Cisco IOS Software Release 12.0(22)S)	NO	NO	~0:06 Seconds	0 Seconds	~0:01.63 Seconds	< 0:05 Seconds

This is a SSO/NSF switchover on a system with 200 ATM PVCs, 100 defined channels, 100K + BGP routes, 30K OSPF routes, traffic

© 2007 Cisco Systems, Inc. All rights reserved.

MPLS and IP NSF/SSO-Coexistence

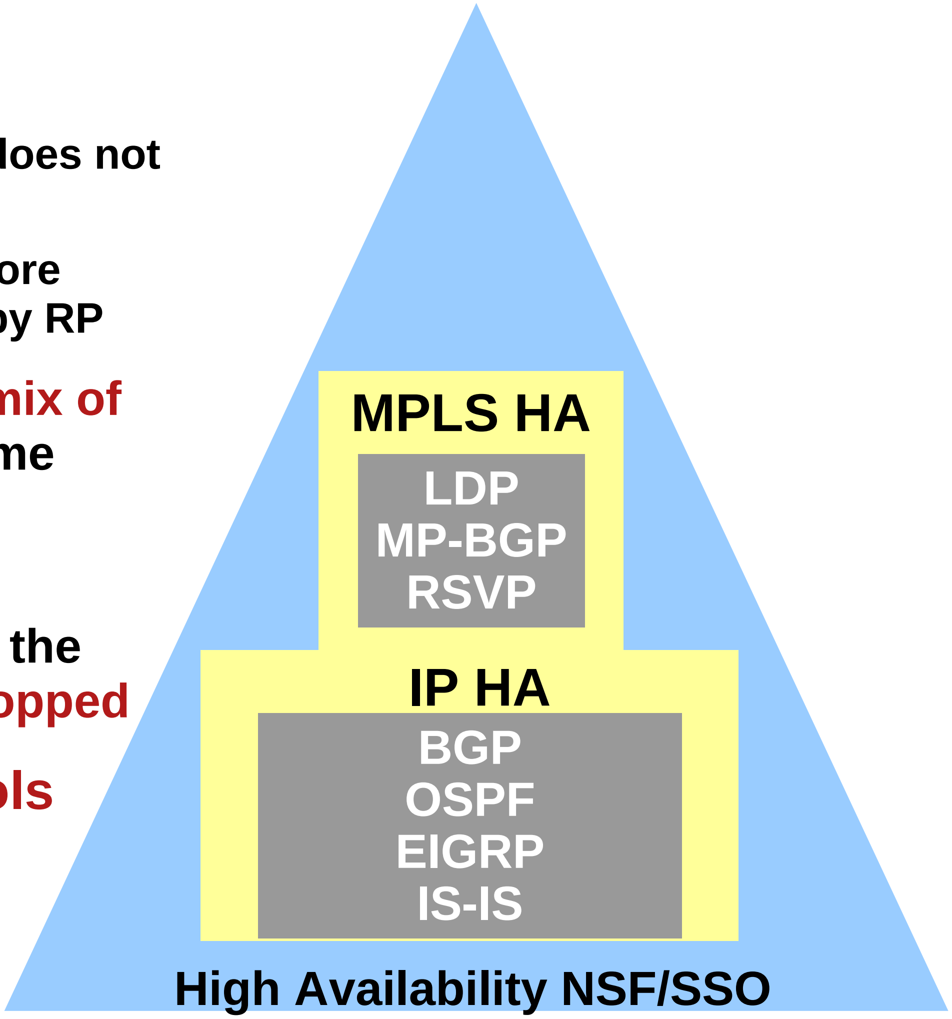
- IP HA speeds up MPLS recovery

  No waiting for the route processor

  No loss of Layer 2 connectivity, so it does not need to be re-established

  MPLS with IP SSO begin rebuilding more quickly after switchover to the standby RP

- SSO coexistence feature allows the mix of SSO and non-SSO features at the same time
- During the IP NSF switchover, MPLS forwarding entries are removed from the linecards and MPLS forwarding is stopped
- Need to enhance the Key Protocols used in MPLS Control Plane to minimize the disruption in MPLS forwarding plane



© 2007 Cisco Systems, Inc. All rights reserved.

MPLS HA Components and Key Elements

- MPLS HA—LDP NSF/SSO
  1. Checkpointing local label bindings to backup RP

     On devices with route processor redundancy
  2. LDP graceful restart capability

     On participating PEs, RRs, and P routers
  3. Checkpoint refreshed/new local label bindings

- MPLS HA—BGP VPNv4 NSF/SSO
  1. MPLS VPN checkpointing capability
  2. BGP graceful restart capability

For Details see:
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122s/122snwft/release/122s25/fshaov.htm

MPLS VPN HA: Putting it Together

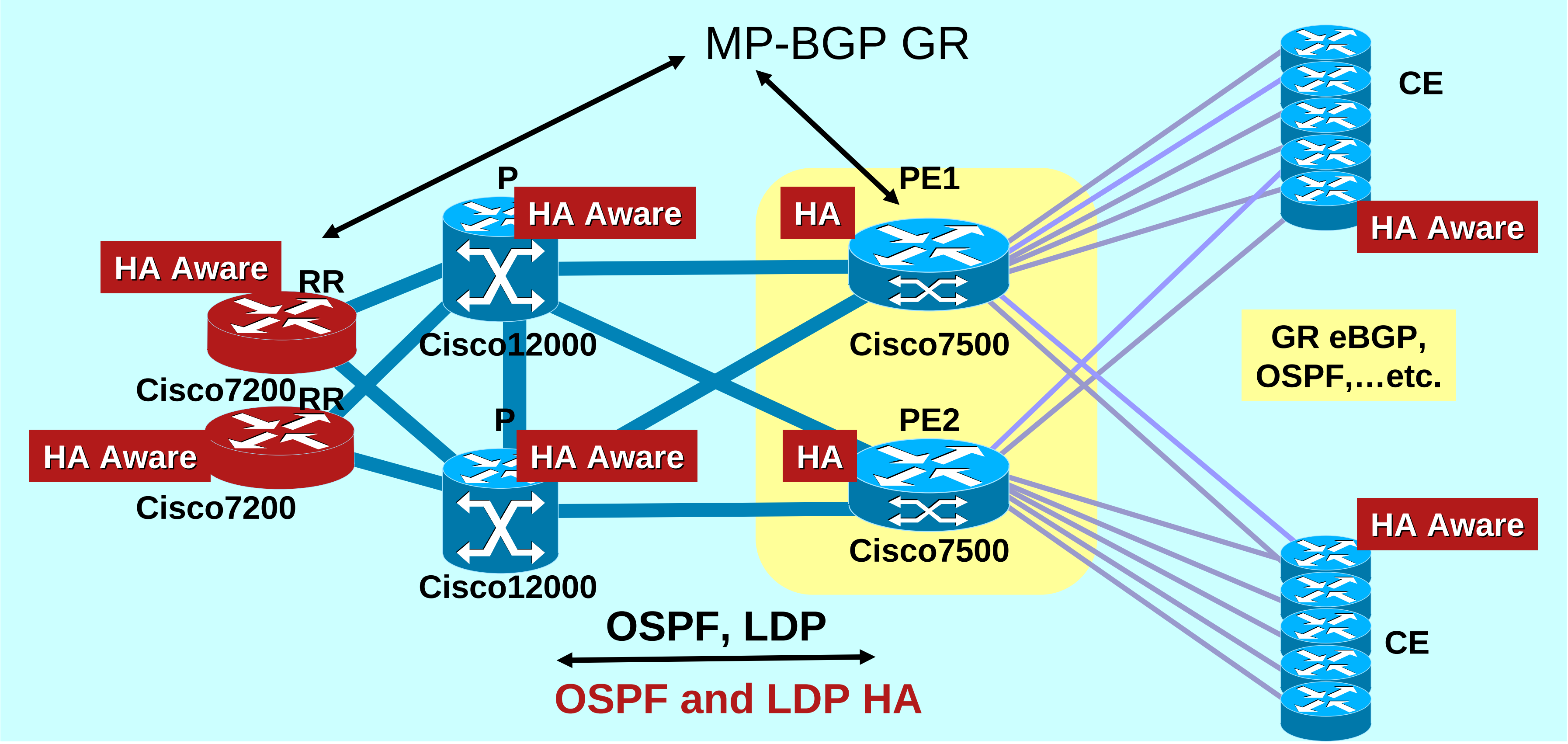
```
redundancy
  mode sso
!
mpls ldp graceful-restart
mpls ip
mpls label protocol ldp
mpls ldp router-id Loopback0 force
mpls ldp advertise-tags
!
router ospf 10
  nsf
network 8.8.8.8 0.0.0.0 area 0
!
router bgp 1
  bgp graceful-restart restart-time 120
  bgp graceful-restart stalepath-time 360
  bgp graceful-restart
```

```
show ip bgp labels
show ip bgp vpnv4 all labels
debug ip bgp vpnv4 checkpoint
debug ip bgp vpnv4 nsf
show mpls ldp checkpoint
```

Deploying MPLS HA Example

HA Aware Devices:
Graceful Restart Functionality for All the Related Protocols CEs, Ps, and RRs

HA Capable Devices
Full NSF/SSO Functionality (PE1 and PE2)

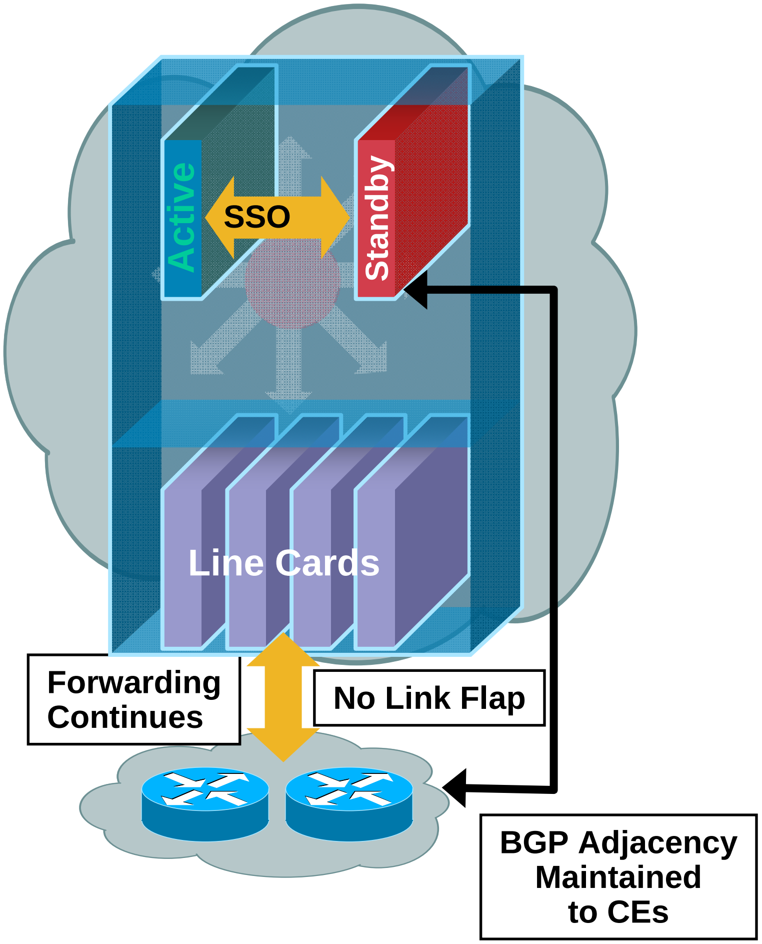


SYSTEM LEVEL RESILIENCY:

Non-Stop Routing (NSR)

Non-Stop Routing (NSR)

- NSR and NSF are not the same
- NSR in a nutshell
  - –Provides forwarding and preserves routing during Active RP failover to Standby RP like NSF
  - –does not require any protocol extension like NSF
  - –does not require software upgrades on peer routers (NSF-aware)
  - –TCP and applications (BGP/LDP) are maintained and stateful switchover is achieved
- IOS Support for NSR:
  - –ISIS NSR (stateful NSF!, Cisco Version)
  - –BGP NSR (introduced with 12.2(28)SB for Cisco 10000 PRE2)
  - –LDP NSR (IOS-XR in 2007)



BGP NonStop Routing with SSO

- Cisco BGP NSR SSO

  provides a transparent BGP failover mechanism for PE routers engage in eBGP peering with CE routers that do not support the graceful restart mechanism

- Simplified deployment for service providers

  Only PEs need to be upgraded to support NSR (incremental deployment)

  CEs are not touched (i.e., no software upgrade required)

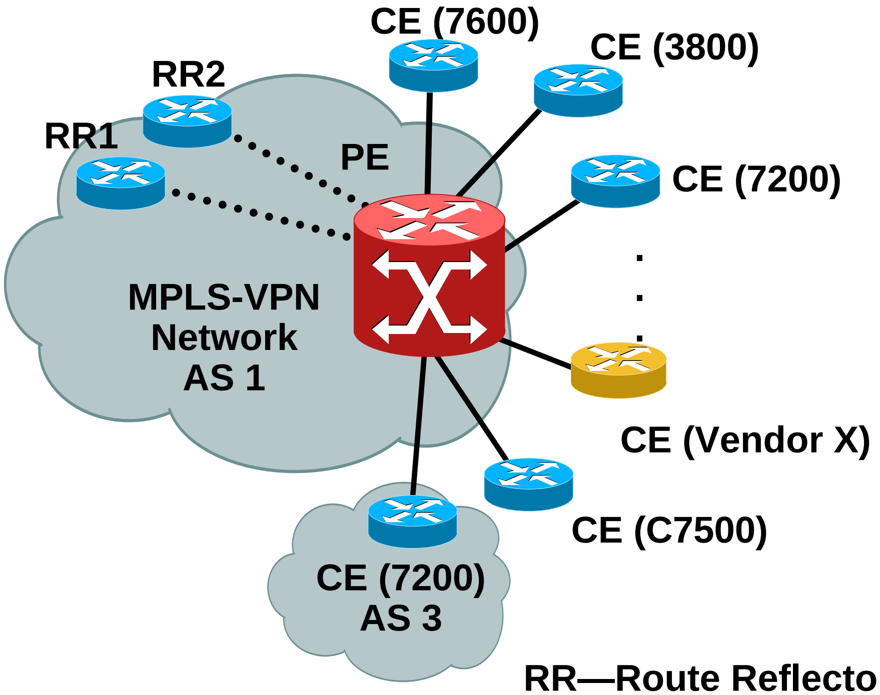
- Scaling optimizations

  PE uses NSR with CEs that are not NSF-aware

  PE uses NSF (Graceful Re-start) with NSF-aware CEs

  iBGP sessions to RRs use NSF (Graceful Restart)

PE Focused Deployment Scenario



RR—Route Reflector
PE—Provider Edge
CE—Customer Edge

NSR – PE Configuration

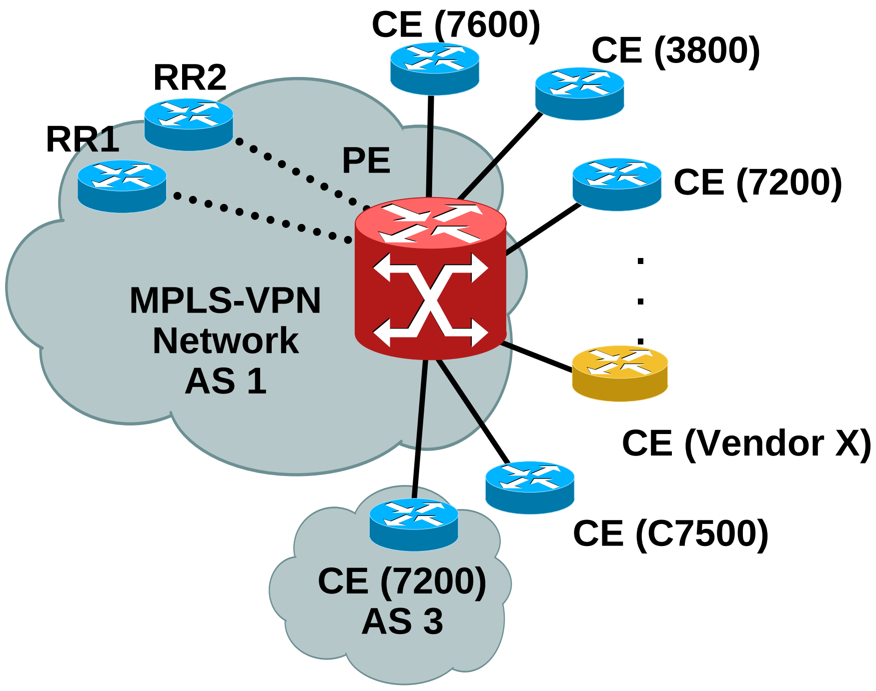
Configuration:

```
neighbor x.x.x.x ha-mode sso
```

- x.x.x.x IP address of neighbor router
- used to configure a BGP neighbor to support SSO
- supported for BGP peer, BGP peer group, and BGP session template configurations

Example:

```
router bgp 1
 bgp graceful-restart restart-time 120
 bgp graceful-restart stalepath-time 360
 bgp graceful-restart
 neighbor 10.1.1.1 remote-as 1
!
<snip>
!
 address-family ipv4 vrf Customer1
  neighbor 10.3.3.3 remote-as 3
  neighbor 10.3.3.3 ha-mode sso
  neighbor 10.3.3.3 activate
  neighbor 10.3.3.3 as-override
 exit-address-family
 !
```



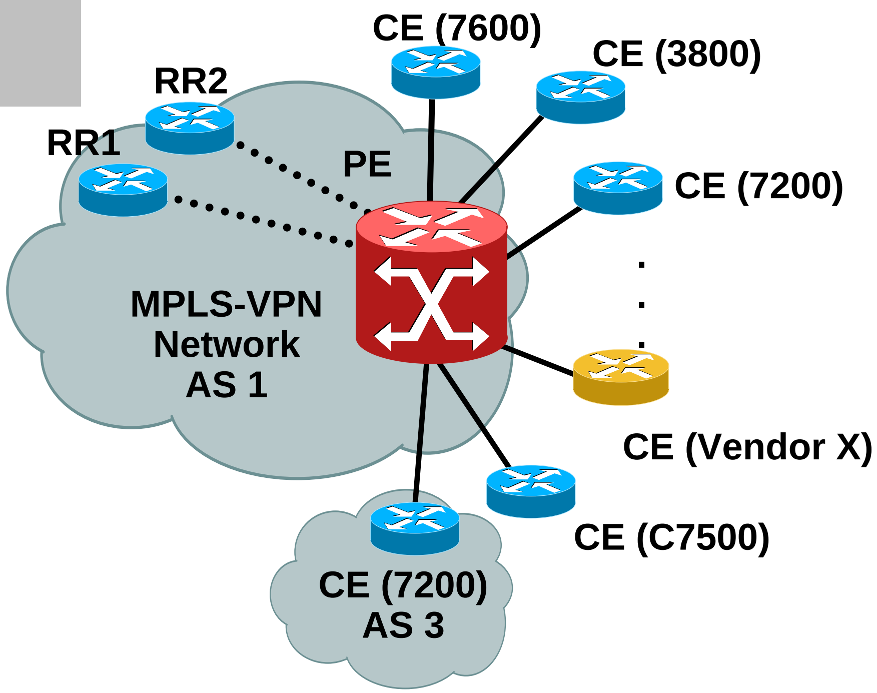
10.1.1.1	RR—Route Reflector
10.2.2.2	PE—Provider Edge
10.3.3.3	CE—Customer Edge

NSR – CE Configuration

```
router bgp 3
 neighbor 10.2.2.2 remote-as 1
!
```

NOTE: No special BGP code or configuration (i.e., NSF-awareness) needed on the CE side to take advantage of the Non Stop Routing capabilities of the PE

PE Focused Deployment Scenario



10.1.1.1	RR—Route Reflector
10.2.2.2	PE—Provider Edge
10.3.3.3	CE—Customer Edge

Verifying BGP Support for NSR with SSO

```
Router# show ip bgp vpnv4 all sso summary

        stateful switchover support enabled for 40 neighbors
```

displays the number of BGP neighbors that are in SSO mode

```
Router# show ip bgp vpnv4 all neighbors 10.3.3.3

BGP neighbor is 10.3.3.3, vrf vrf1, remote AS 3,
external link

<snip>

Stateful switchover support enabled

<snip>

SSO Last Disable Reason: Application Disable (Active)
```

displays VPN information from the BGP indicating whether SSO is enabled or disabled and displays information about the last BGP session that lost SSO capability

For details see the BGP NSR Feature Guide:
http://www.cisco.com/en/US/products/ps6566/products_feature_guide09186a008067a607.html

Troubleshooting BGP NSR with SSO

```
debug ip bgp sso {events | transactions} [detail]
```

displays BGP-related SSO events

displays debugging information for BGP-related interactions between the active RP and the standby RP

useful for monitoring or troubleshooting BGP sessions on a PE router during an RP switchover or during a planned ISSU

```
debug ip tcp ha {events | transactions} [detail]
```

displays TCP HA events or debugging information for TCP stack interactions between the active RP and the standby RP

useful for troubleshooting SSO-aware TCP connections

```
show tcp [line-number] [tcb address]
```

displays the status of TCP connections.

Output includes SSO capability flag to indicate the reason that the SSO property failed on a TCP connection.

```
show tcp ha connections
```

Displays details of TCP connections that support BGP NSR with SSO (number of connections and connection-ID-to-TCP mapping data)

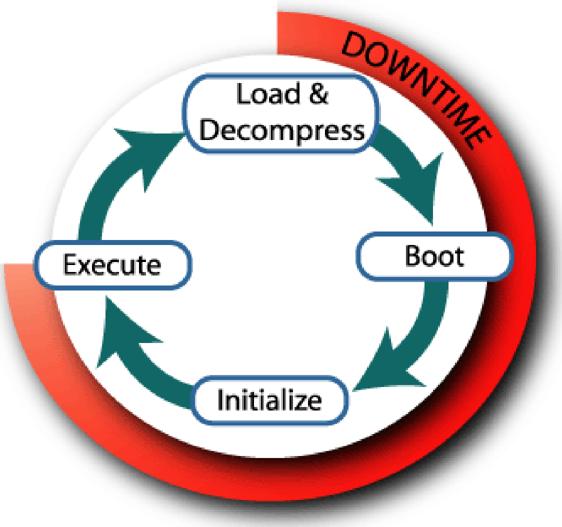
SYSTEM LEVEL RESILIENCY:

Warm Reload / Warm Upgrade

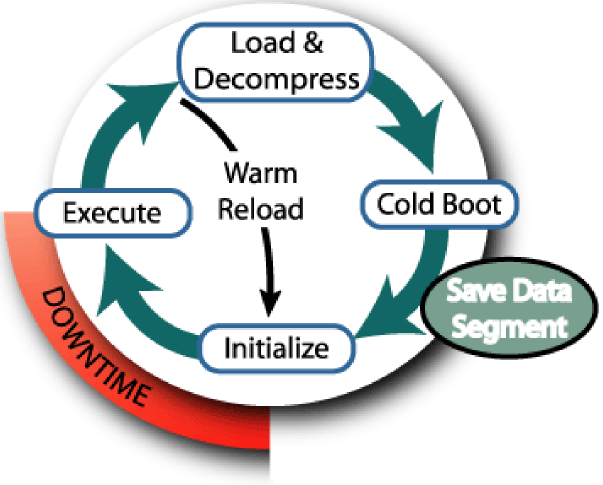
© 2007 Cisco Systems, Inc. All rights reserved.

Cisco IOS Warm Reload

Boot Process Prior to Warm Reload



Warm Reload Process



Enables significant reduction in device reboot time by lowering the mean time to repair (MTTR) for software failures

- Executing begins during re-run from the start address with previously saved, pre-initialized variables
- Particularly applicable to single processor systems

Warm Reload Details

- Savings from reading and decompressing of image
- Additional memory consumption to store a compressed copy of initialized variables in read-only section – typically 1-2 MB
- Useful in case of software design error:

  Software-induced crash

  Requires restart to repair
- Hardware failure will force the ‘cold’ reboot
- If the router reboots for the same reason within 5 minutes it will ‘cold’ reboot

```
Router(config)# warm-reboot <count> <uptime>
```

count - maximum number of warm reboots allowed (default 5, value 1-50)

uptime - minimum time (minutes) between initial system configuration and an exception before a warm reboot is attempted (default value is 5 minutes)

© 2007 Cisco Systems, Inc. All rights reserved.

Improved Availability of Single Processor Routers and Switches

- NPE-G1 Setup

  Normal reload: 3:43 minutes

  Warm reload: 0:32 minutes

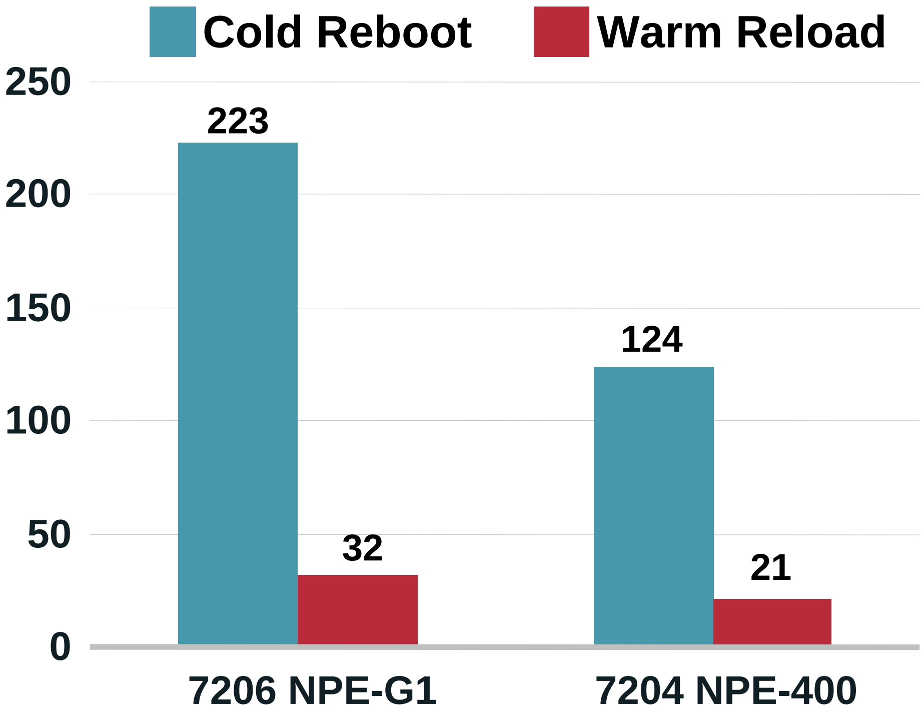
  Reduced downtime by 86%

- NPE-400 Setup

  Normal reload: 2:04 minutes

  Warm reload: 0:21 minutes

  Reduced downtime by 83%



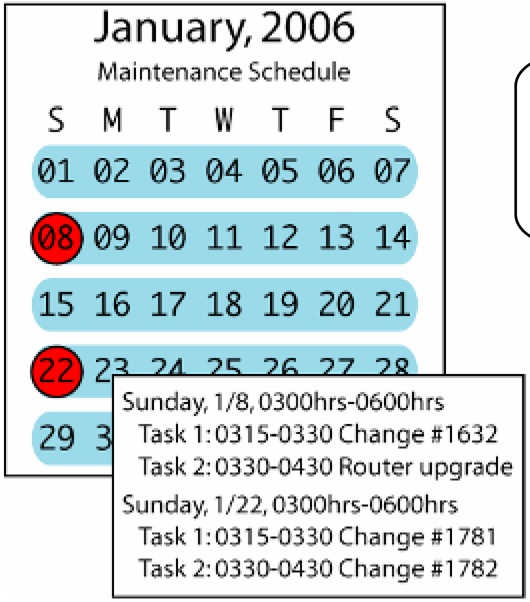
introduced with 12.2(18)S and 12.3(2)T

http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper0900aecd801778e8.shtm

Check Feature Navigator for Support on other platforms: 1xxx, 2xxx, 3xxx, 7xxx

© 2007 Cisco Systems, Inc. All rights reserved.

Various Methods for Minimizing Downtime Due to Planned Software Upgrades



January, 2006

Maintenance Schedule

S	M	T	W	T	F	S
01	02	03	04	05	06	07
08	09	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	3					

Sunday, 1/8, 0300hrs-0600hrs

Task 1: 0315-0330 Change #1632

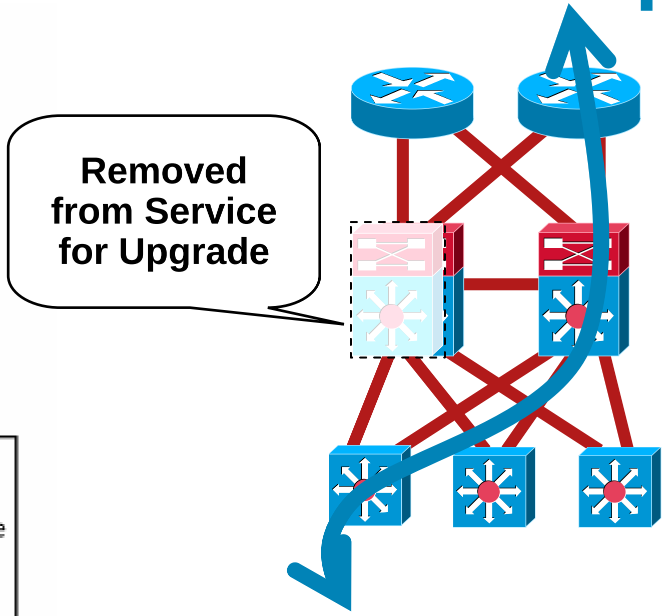
Task 2: 0330-0430 Router upgrade

Sunday, 1/22, 0300hrs-0600hrs

Task 1: 0315-0330 Change #1781

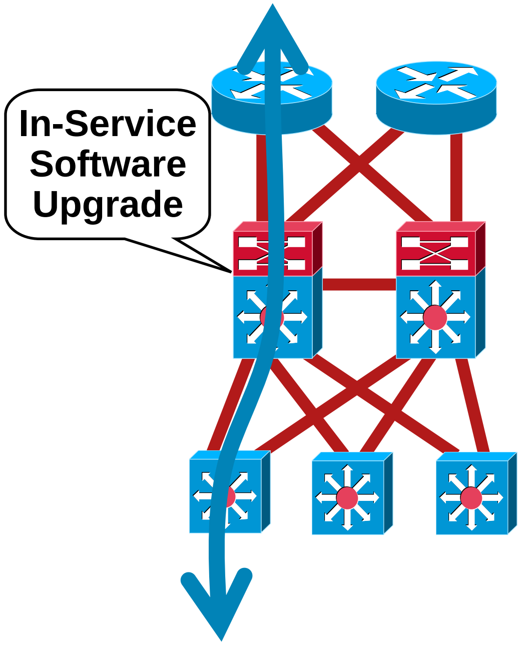
Task 2: 0330-0430 Change #1782

Service Impact Mitigated by Scheduling



Removed from Service for Upgrade

Service Impact Mitigated by Network Design



In-Service Software Upgrade

Service Impact Mitigated by ISSU or Warm Upgrade

Typical procedure performed countless times by Cisco customers

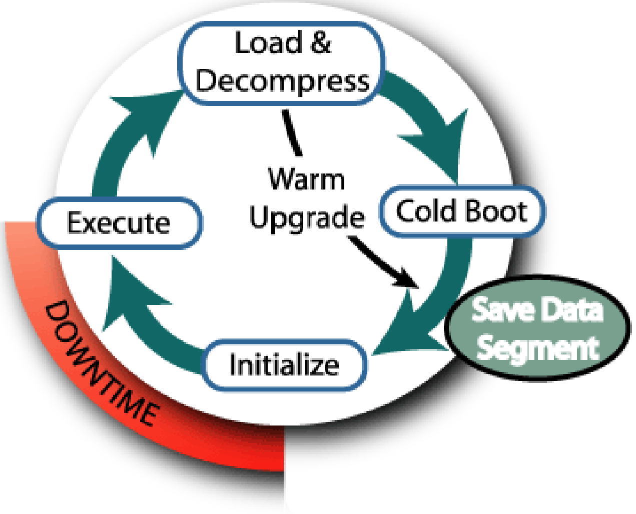
- –Download Cisco IOS Software from Cisco.com
- –Transfer to device’s file system
- –Set to reload using new software
- –Users see service impact during the reload
- –Or: If network resiliency available, impact equal to reconvergence time

© 2007 Cisco Systems, Inc. All rights reserved.

Basic Upgrade Process Improved with Warm Upgrade

- Builds on Warm Reload, can be used in conjunction with Warm Reload
- reduce downtime for planned upgrades and downgrades
- Enables router to read and decompress the new Cisco IOS Software image and then to transfer control to it, while packet forwarding is continued
- If upgrade fails, the current instance of Cisco IOS Software will continue to run, unless the image is partially or fully erased
- Requires router to have sufficient free memory to read and decompress the new image, while the current instance of Cisco IOS Software is running

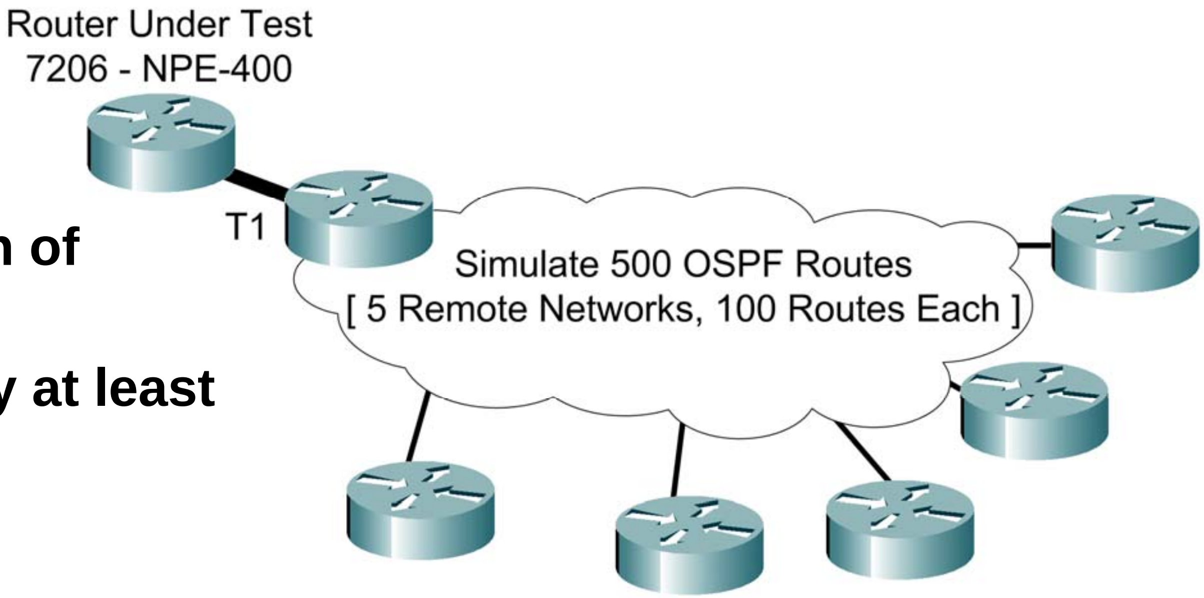
Warm Upgrade process



```
Router# reload warm file disk2:c7200-js-mz.122-18.S3
```

Warm Upgrade Reduces Service Impact

- **Reduction in duration of packet loss**
- **Downtime reduced by at least 27 seconds**
- **68 % Less Service Impact Time**



	Without Warm Upgrade	With Warm Upgrade
Reload Start	0:00	0:00
Packet Loss Seen	0:00	0:27
Reload Complete	2:50	1:00
OSPF Adjacency Restored	3:20	1:30
Traffic Flow Restored to All 500 Destinations	3:35	1:35

© 2007 Cisco Systems, Inc. All rights reserved.

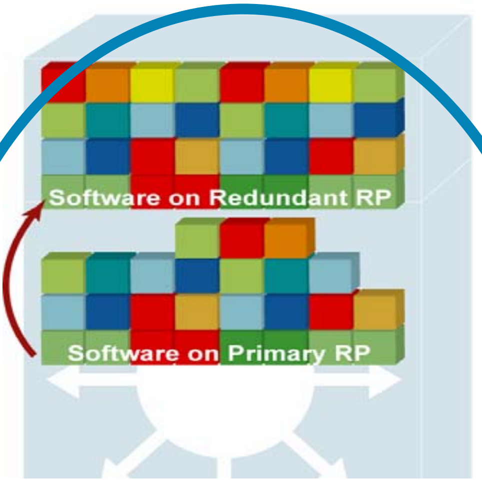
SYSTEM LEVEL RESILIENCY:

In-Service Software Upgrade (ISSU)

Cisco's In-Service Software Upgrade (ISSU)

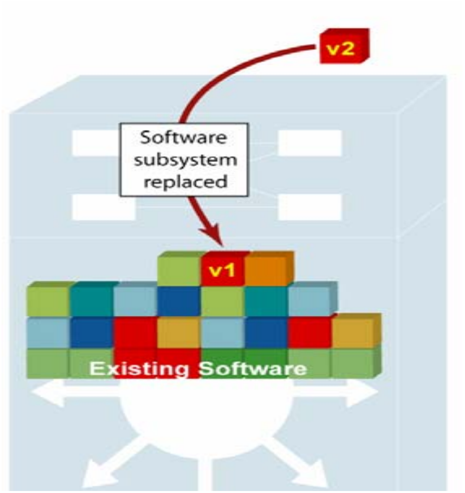
- **targeting planned downtime (software upgrades, maintenance)**
- **strategy spans all Cisco IOS product lines**
- **ranging from full image upgrades to granular, selective software maintenance (upgrade vs patch vs component upgrade)**

Focus in this session



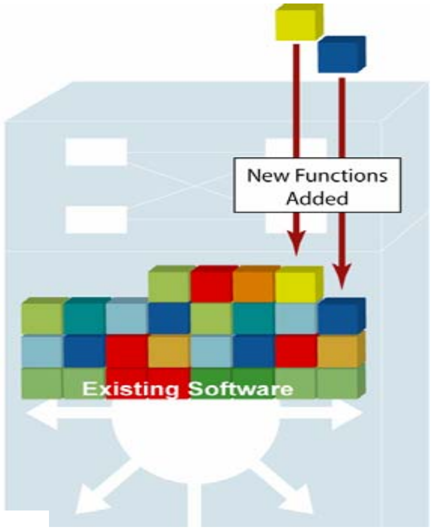
UPGRADE

- **True upgrade including new features and function**
- **Full image upgrade**



PATCH

- **Selective maintenance**
- **Patch a component**



Component upgrade

- **Add new features to existing base**

© 2007 Cisco Systems, Inc. All rights reserved.

Cisco IOS Full Image ISSU Steps

animated

🟥 = RP Is Active

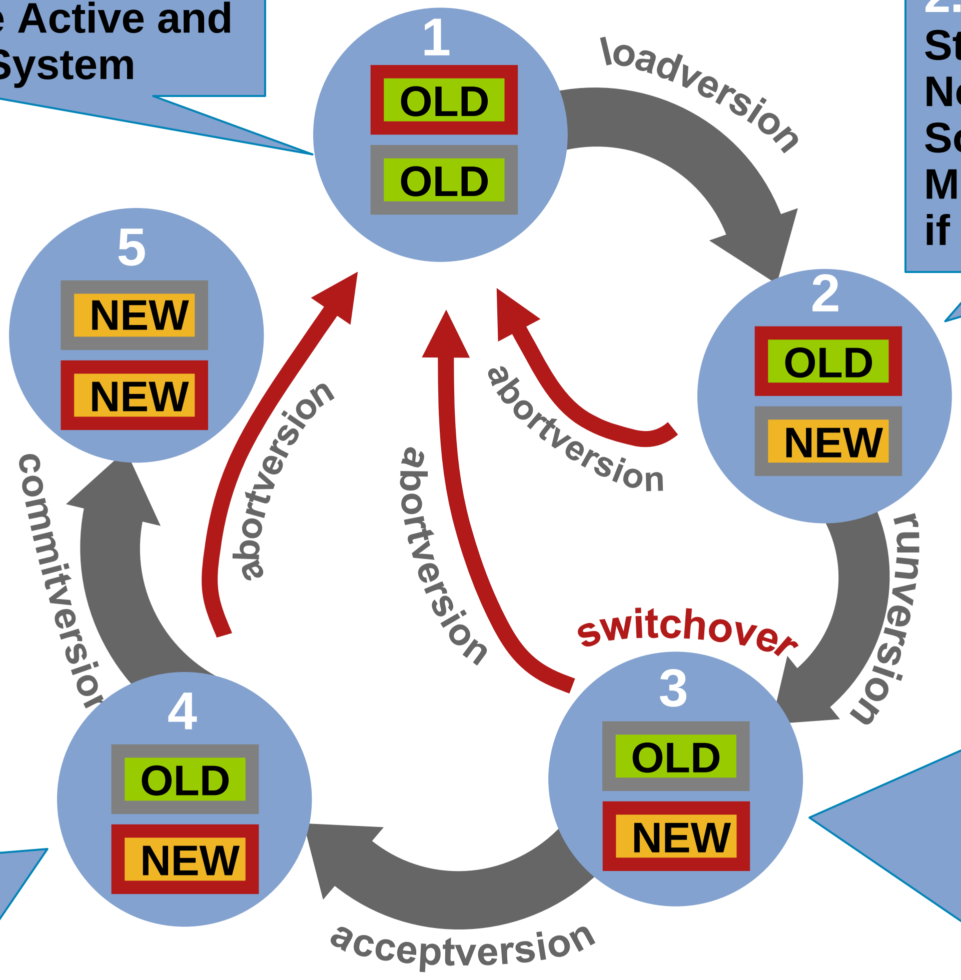
□ = RP Is Standby

OLD = Old Cisco IOS

NEW = New Cisco IOS

1. Prepare for ISSU:
Copying New Cisco IOS Version to Both the Active and Standby RP's File System

2. Load Standby
Standby RP Resets Now Running New Software, still in SSO Mode, automatic abort if image incompatible



3. Switchover and run new version

Switchover Occurs

Standby Becomes Active

Old Active RP Is Reset and Becomes Standby Running Old Software

Still in SSO Mode

4. Stop Autorollback and Check Network

Must issue acceptversion before rollback timer expires

Remain in the state while checking, not for long term

5. Complete Process and Commit

commit new version, standby will reset and load with new SW.

ISSU Commands for Full Software Upgrade

Optional Parameter

- **issu loadversion**

```
r1# issu loadversion b stby-disk0:c10k2-p11-mz.2.20040830 force
```

“force" used to override the automatic rollback when new version is detected to be incompatible (e.g. fast software upgrade in RPR mode, service impacting if running ISSU between incompatible releases)

- **issu runversion**

```
r1# issu runversion b stby-disk0:c10k2-p11-mz.2.20040830
```

Switches to the redundant RP with the new image and loads lines cards, parses the config, etc.

- **issu acceptversion**

```
r1# issu acceptversion b disk0:c10k2-p11-mz.2.20040830
```

- **issu commitversion**

```
r1# issu commitversion a stby-disk0:c10k2-p11-mz.2.20040830
```

Will cause the Standby RP to be reset and reloaded with the new software version and come up in the highest HA mode attainable, which should be SSO, since the images are the same

- **issu abortversion**

```
r1# issu abortversion a stby-disk0:c10k2-p11-mz.2.20040830
```

When issued prior to runversion—resets and reload the Standby; When issued after runversion—switches to old version, loads lines cards, parses config, etc.; result is two service outages

Show ISSU State Detail

After “issu runversion”

```
router#sh issu state det
                          Slot = B
                      RP State = Active
                    ISSU State = Run Version
                 Boot Variable = disk0:c10k2-p11-mz.2.20040830,12;
                                 disk0:c10k2-p11-mz.1.20040830,1;
                Operating Mode = SSO
               Primary Version = disk0:c10k2-p11-mz.2.20040830
             Secondary Version = disk0:c10k2-p11-mz.1.20040830
               Current Version = disk0:c10k2-p11-mz.2.20040830
                          Slot = A
                      RP State = Standby
                    ISSU State = Run Version
                 Boot Variable = disk0:c10k2-p11-mz.1.20040830,1;
                Operating Mode = SSO
               Primary Version = disk0:c10k2-p11-mz.2.20040830
             Secondary Version = disk0:c10k2-p11-mz.1.20040830
               Current Version = disk0:c10k2-p11-mz.1.20040830
```

Slot B Is Active

Bootvar Adjusted

New Version “2”

Old Version “1”

```
router# show issu rollback-timer
        Rollback Process State = In progress
        Configured Rollback Time = 45:00
        Automatic Rollback Time = 29:03
```

Which IOS features are ISSU capable?

- ISSU builds on NSF/SSO support for IOS features
- NSF/SSO capable feature preserved following an ISSU upgrade

  HA system infrastructure components

  Forwarding (CEF)

  Connectivity features (ATM, FR, HDLC, PPP, MLPPP)

  Routing and IP services features (BGP, OPSF, ISIS, EIGRP, ARP, HSRP)

  MPLS features (LDP, MPLS/VPN, InterAS, CsC)

  Management Protocol (SNMP)

- Majority of IOS features do NOT require stateful information synch

  just need configuration synchronization between RPs

- Other features requiring stateful information synchronization support HA co-existence

  These features will restart following ISSU (as in a system reboot)

  ISSU architecture allows ISSU support for additional features to be added in a incremental fashion over future software releases

© 2007 Cisco Systems, Inc. All rights reserved.

ISSU Compatibility Matrix

ISSU compatibility for all capable Cisco IOS software assigned

Compatible

C

Base-level system infrastructure and all optional HA-aware sub-systems are compatible, ISSU between these versions will succeed with minimal service impact

Base-level compatible

B

One or more of the optional HA-aware sub-systems are not compatible

ISSU between these versions will succeed, however, some sub-systems will not be able to maintain state during the transition

Careful consideration of the impact this may have on operation and service is required before an in-service upgrade should be attempted

Incompatible

I

There exists core set of system infrastructure that must be able to interoperate in a stateful manner for SSO to function correctly

If any “required” features or protocols is not interoperable, then the two versions of the Cisco IOS images are declared “incompatible”, ISSU not possible between these versions.

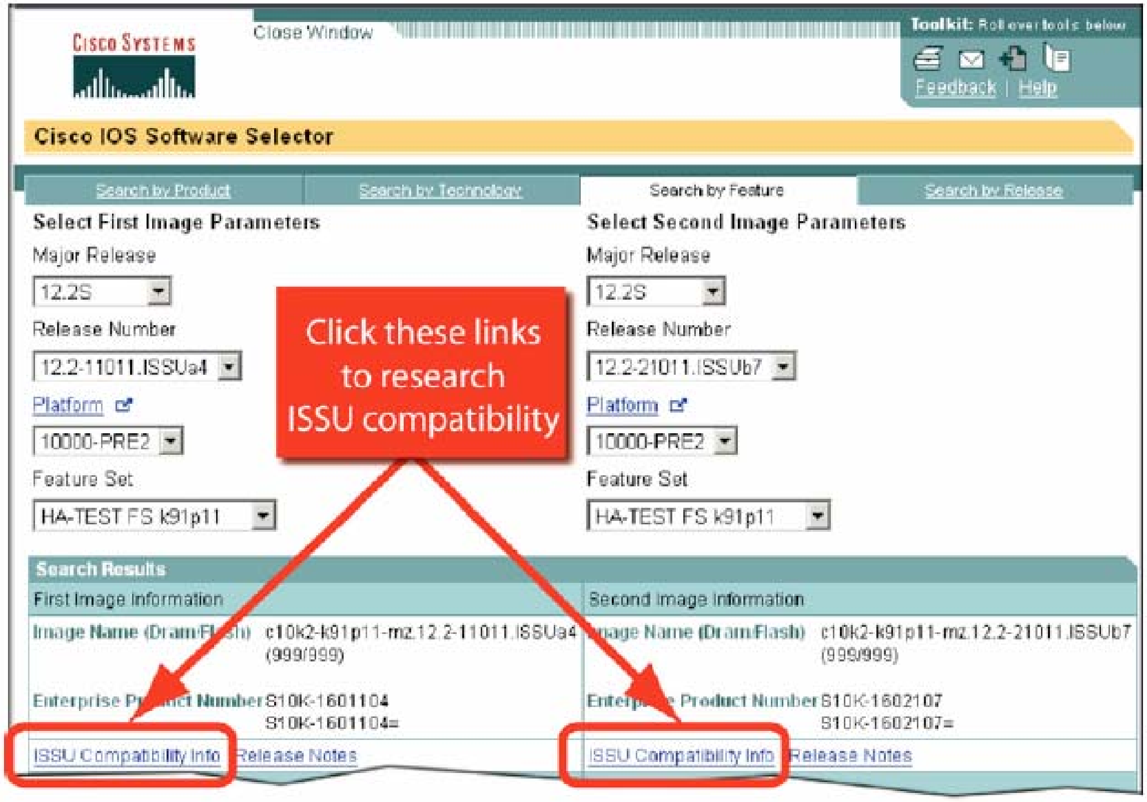
```
Router# show issu comp-matrix
```

display the compatibility matrix data between 2 software versions on a system

Compatibility Verification Using Cisco Feature Navigator

ISSU application on Cisco Feature Navigator www.cisco.com/go/fn

- **Select an ISSU-capable image**
- **Identify which images are compatible with that image**
- **Compare two images and understand the compatibility level (C/B/I)**
- **Compare two images and see the client compatibility for each ISSU client**
- **Provide links to release notes for the image**



Select First Image Parameters

Major Release: 12.2S

Release Number: 12.2-11011.ISSUa4

Platform: 10000-PRE2

Feature Set: HA-TEST FS k91p11

Select Second Image Parameters

Major Release: 12.2S

Release Number: 12.2-21011.ISSUb7

Platform: 10000-PRE2

Feature Set: HA-TEST FS k91p11

Search Results

First Image Information		Second Image Information	
Image Name (DramFlash)	c10k2-k91p11-mz.12.2-11011.ISSUa4 (999/999)	Image Name (DramFlash)	c10k2-k91p11-mz.12.2-21011.ISSUb7 (999/999)
Enterprise Product Number	S10K-1601104 S10K-1601104=	Enterprise Product Number	S10K-1602107 S10K-1602107=
ISSU Compatibility Info	Release Notes	ISSU Compatibility Info	Release Notes

Click these links to research ISSU compatibility

ISSU Best Practices

- Avoid manual switchovers.
- Avoid card OIR (online insertion and removal).
- Copy Cisco IOS Software prior to Cisco IOS ISSU.
- Do not change redundancy mode during the Cisco IOS ISSU process.
- MDR and line-card versioning is required.
- Ensure adequate local file system capacity.
- Minimize duration of the Cisco IOS ISSU process.
- Use maintenance windows.
- Do not implement new features while Cisco IOS ISSU is in progress.
- Disable unsupported features and functions when performing a “downgrade.”

Slide not in Printouts, only in pdf

© 2007 Cisco Systems, Inc. All rights reserved.

Cisco IOS and IOS-XR ISSU Availability

- ISSU is a **process or procedure**
- Based on an **architecture for high availability**

	Cisco ISR	Cat 3750	C7200	Cat 4500	Cat 6500	C7300	C7500	C7600	C10k	C12k	CRS-1
Full Image ISSU	■	□	□	▣	■	□	□	■	▣	□	□
Subsystem ISSU	□	□	□	□	▣ Mod IOS	□	□	□	□	▣ XR	▣ XR
Enhanced FSU (SSO)	□	□	□	□	□	□	□	■	□	□	□
FSU (RPR+)	□	▣	□	▣	▣	▣ 7304 Only	▣	▣	▣	▣	□
Warm Upgrade	▣	□	▣	□	□	▣	□	□	□	□	□

▣ Available ■ Planned □ Not Planned

NETWORK LEVEL RESILIENCY

Network Level Resiliency

NSF awareness

IP Event Dampening

Bi-Directional Forwarding Detection

Fast Convergence

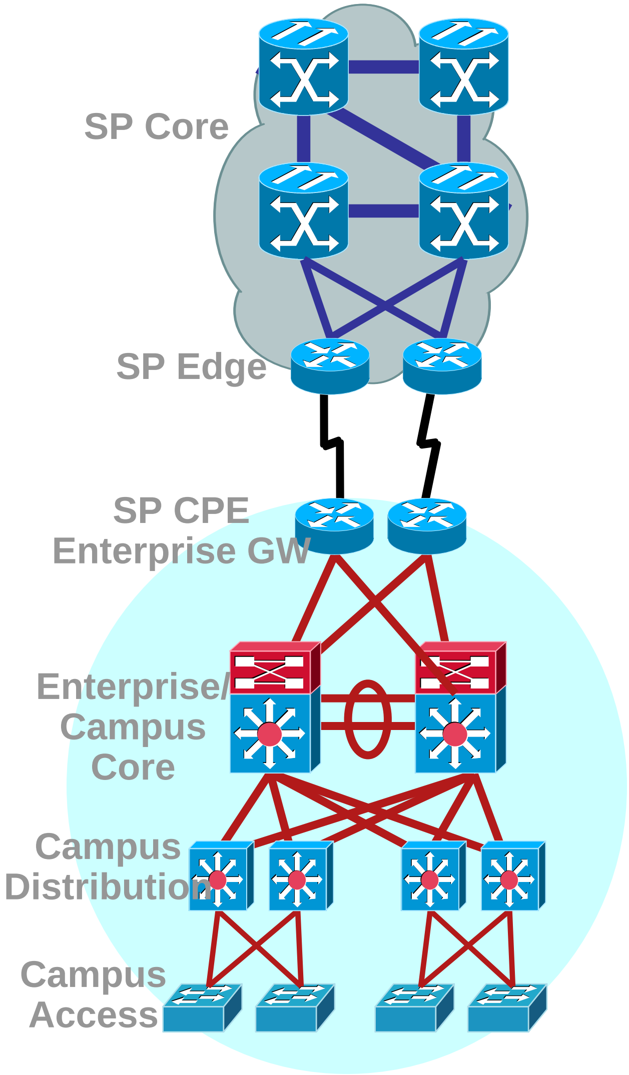
- iSPF Optimization (OSPF, IS-IS)
- BGP Optimization
- FC and NSF/SSO Coexistence

GR Shutdown

Fast ReRoute (FRR)

- MPLS FRR
- IP FRR

....



© 2007 Cisco Systems, Inc. All rights reserved.

© 2004 Cisco Systems, Inc. All rights reserved.

NETWORK LEVEL RESILIENCY:

IP EVENT DAMPENING

© 2007 Cisco Systems, Inc. All rights reserved. Cisco Public

IP Event Dampening: Concept

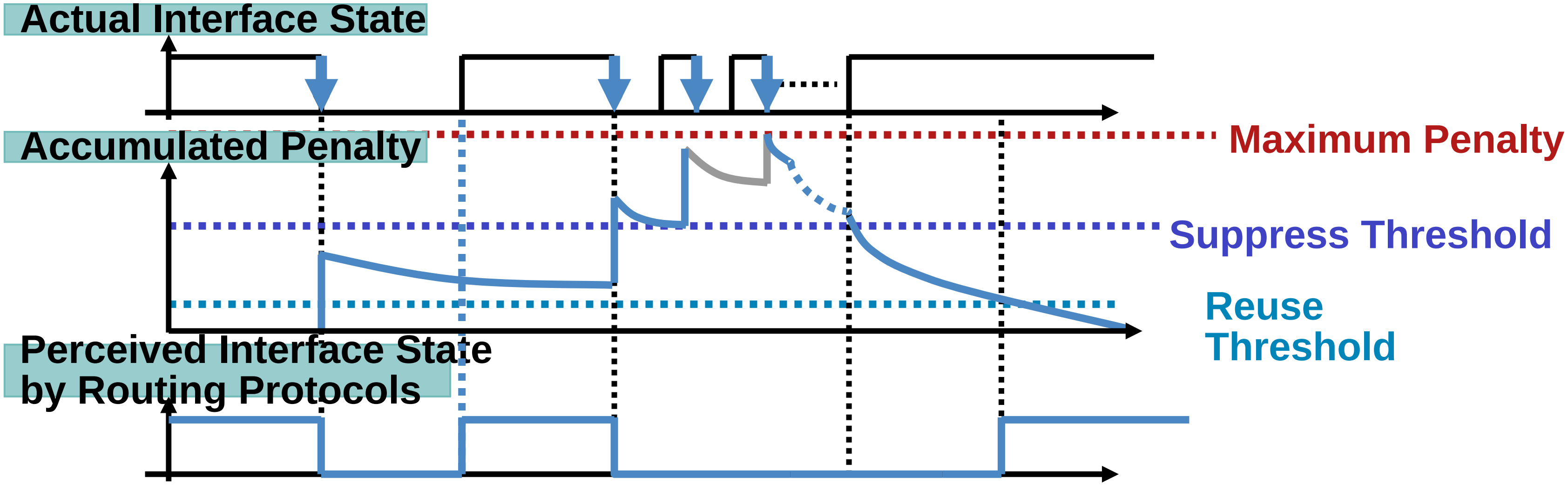
- IP Event Dampening **logically isolates unstable links**:

  reducing packet loss & routing CPU overhead

  reducing network oscillations

  isolating unstable network elements

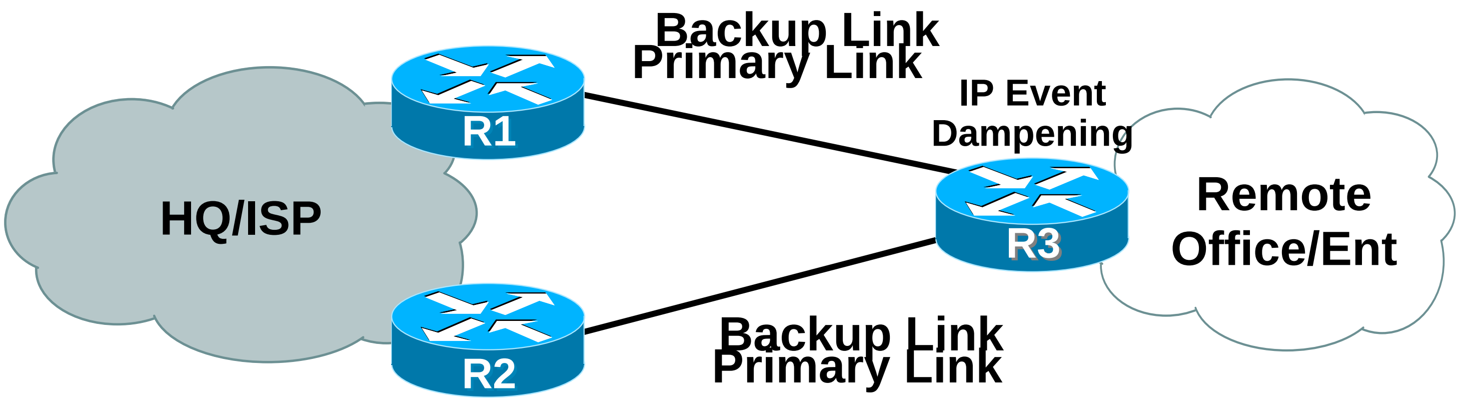
- Takes **concept of BGP route-flap dampening** to interface level
- **Tracks interface flapping**, applying a “penalty” to a flapping interface
- Puts the interface **in “down” state from routing protocol perspective** if the penalty is over a threshold tolerance
- Uses **exponential decay algorithm** to decrease the penalty over time and brings the interface back to “up” state



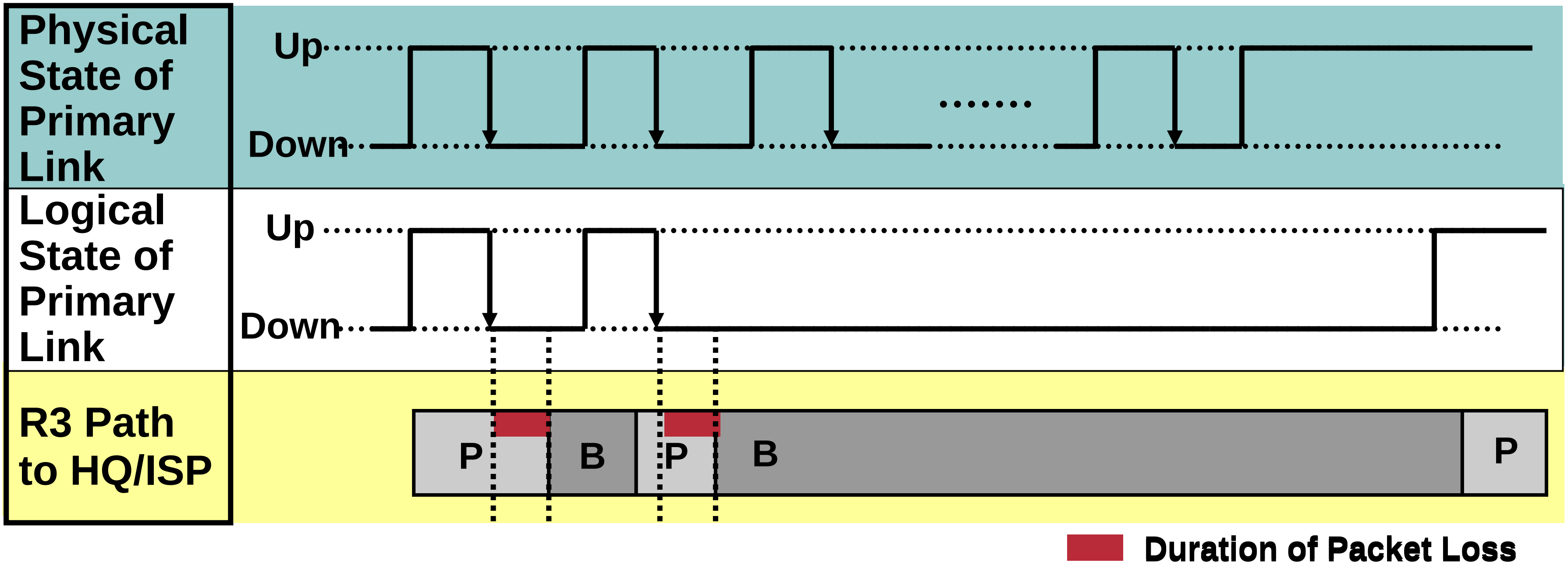
© 2007 Cisco Systems, Inc. All rights reserved.

animated

IP Event Dampening: Deployment


Backup Link
Primary Link
IP Event Dampening
HQ/ISP
R1
R2
R3
Remote Office/Ent
Backup Link
Primary Link

IP Event Dampening Absorbs Link-Flapping Effects on Routing Protocols


Physical State of Primary Link
Up
Down
....
Logical State of Primary Link
Up
Down
R3 Path to HQ/ISP
P
B
P
B
P
🟥 Duration of Packet Loss

IP Event Dampening: Configuration

```
interface Serial 0
    dampening [half-life reuse] [suppress max-suppress [[restart-penalty]]
```

- **Penalty: numeric value applied to the interface each time it flaps**
- **Half-life: time that must elapse without a flap to reduce penalty by half**
- **Reuse: <penalty limit interface is reintroduced to routing**
- **Suppress: >penalty limit interface is suppressed from routing**
- **Max-Suppress: Maximum time an interface can be suppressed**
- **Restart-Penalty: initial penalty applied to interface when system boots**
- **Defaults: dampening 15 1000 2000 60 0**
- **Supports all IP routing protocols**
  - **Static routing, RIP, EIGRP, OSPF, IS-IS, BGP**
  - **Subinterface Restriction: Applies to all subinterfaces on physical interfaces**
  - **Virtual Templates not supported**
- **Available in 12.0(22)S, 12.2(13)T, 12.2(14)S, 12.2(18)SXD**
- **Platforms: 1700, 1800, 2600, 2800, 3600, 3700, 3800, 7200, 7300, 7500, 7600, 10000, 12000 and Catalyst Platforms**
- **Check Feature Navigator for more details.**

© 2007 Cisco Systems, Inc. All rights reserved.

NETWORK LEVEL RESILIENCY:

Bi-Directional Forwarding Detection (BFD)

The Problem with Convergence

Process of Network Convergence

- Failure Detection, Information Dissemination, Repair

Failure Detection most problematic and inconsitent

- varying methods to detect loss of routing adjacency in different routing protocols (subsecond hello of routing protocols very CPU intensiv)
- slow neighbor failure detection by IGP built-in hellos is main reason for delayed IGP Convergence
- link-layer failure detection depending on physical media and L2 encapsulation
- intervening devices hide link-layer failure from routing protocols
- POS (SDH/Sonet) has become benchmark to detect/react to media or protocol failures (~50 msec)

Need for single standardized method of link/device/protocol failure detection at any protocol layer over any media

BFD – Bidirectional Forwarding Detection

IETF Working Group for BFD since 2004

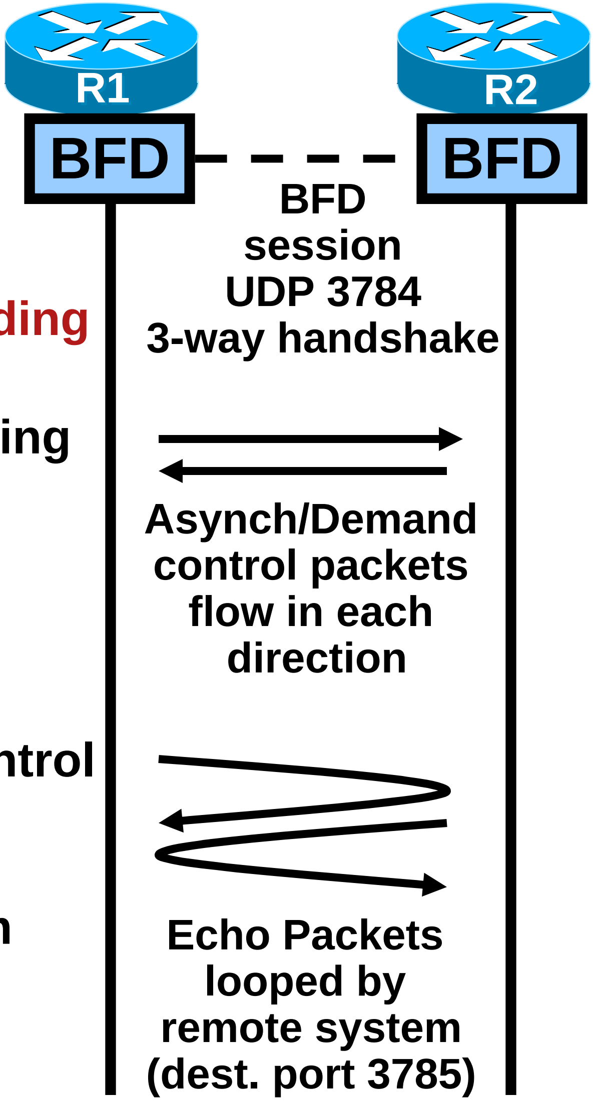
- http://www.ietf.org/html.charters/bfd-charter.html
- 6 drafts: Generic, Base, Multihop, MPLS, MIB, v4v6-1hop

Goals:

- detect faults in the bidirectional path between forwarding engines, interfaces and data links with low latency
- operates independently of media, data protocols, routing protocols
- single mechanism for liveness detection (lightweight protocol, easy-to-parse)

Different Modes

- Asynchronus mode: periodically transmitting BFD control packets
- Demand mode: after establishment of BFD session, control packets on demand ,target at low-end platform
- Echo Function: loop back of echo packets through forwarding path (HW implementation)

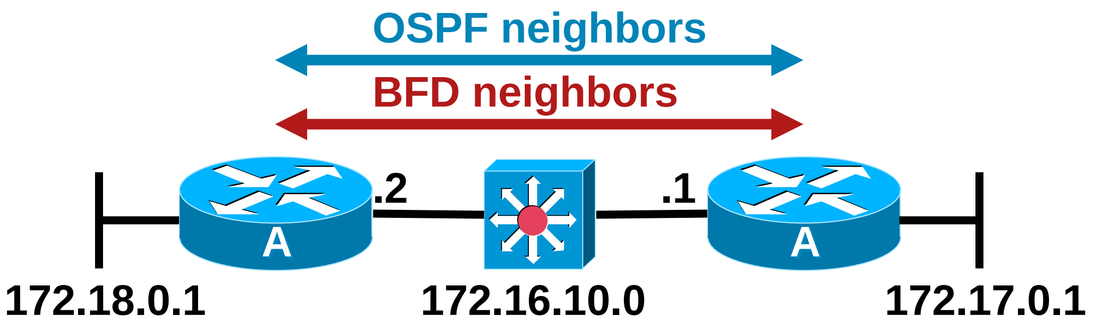


BFD session UDP 3784 3-way handshake

Asynch/Demand control packets flow in each direction

Echo Packets looped by remote system (dest. port 3785)

BFD Operation with OSPF



OSPF neighbors

BFD neighbors

.2

.1

A

A

172.18.0.1

172.16.10.0

172.17.0.1

```
Configuration Router A:
!
interface fastethernet 0/1
   ip address 172.16.10.1 255.255.255.0
   bfd interval 50 min_rx 50 multiplier 3
!
router ospf 123
  bfd interface fastethernet 0/1
```

```
RouterA# show bfd neighbors details
OurAddr  NeighAddr  LD/RD RH Holdown(mult) State Int
172.16.10.1 172.16.10.2    1/2    1   532 (3 )          Up    Fa0/1
.............
Registered protocols: OSPF
Uptime: 02:18:49
```

- OSPF Hellos still needed for control plane verification, discovery, …
- OSPF process registers neighbors on BFD enabled interfaces with BFD process
- BFD monitors liveliness of forwarding plane
- Swiftly notifies OSPF of BFD session failures
- Upon notification, OSPF brings down neighbor and recalculates routes

BFD: Support, Scaling and Performance

IOS/XR Support:

IOS: 12.0(31)S, 12.2(18)SXE, 12.4(4)T

IOS-XR: 3.2

Centralized platforms:

+2% CPU @ 100 BFD session

Distributed platforms: Cisco 12000

no CPU impact on RP

+2% LC-CPU @ 100 session / LC
with 150msec detection time

Protocol Support:

ISIS, OSPF, EIGRP, BGP-SingleHop

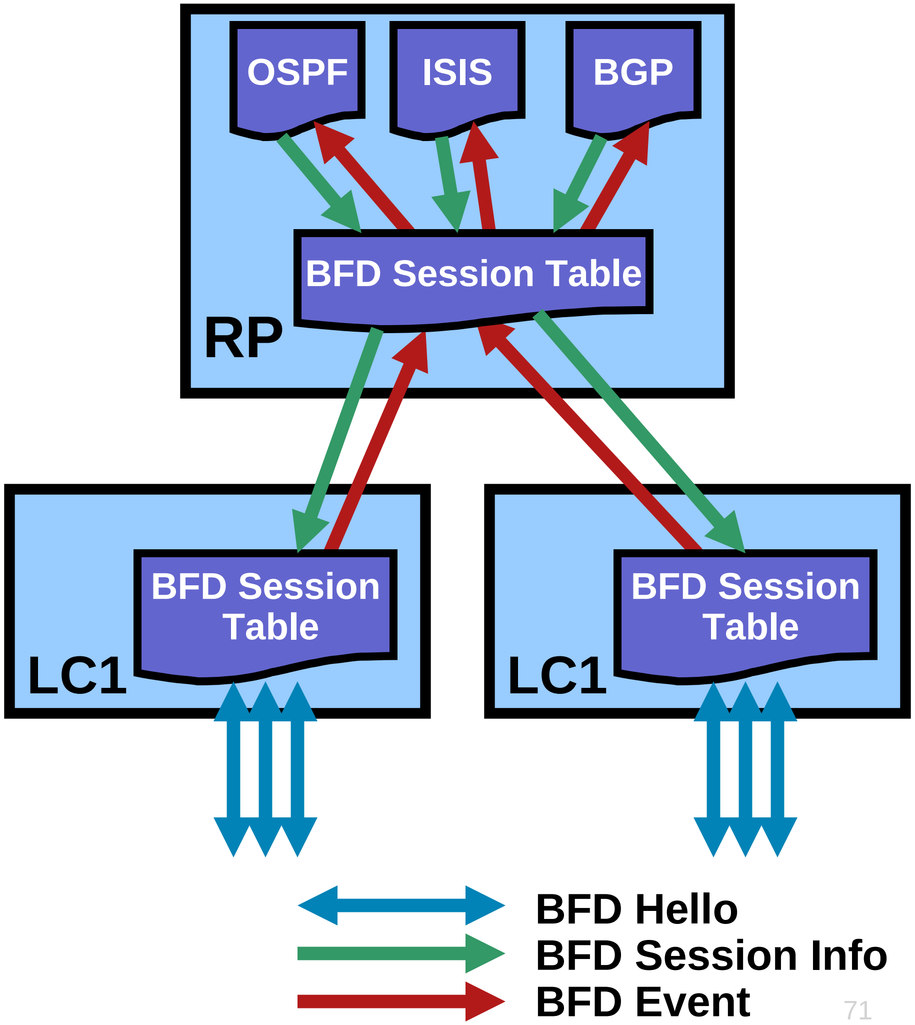
(BGP-Multihop, LDP, IPv6)

Internetworking

IP Event Dampening

(NSF/SSO/GR of BFD)

Cisco 12000 BFD Architecture

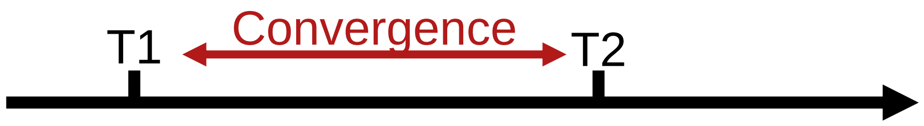


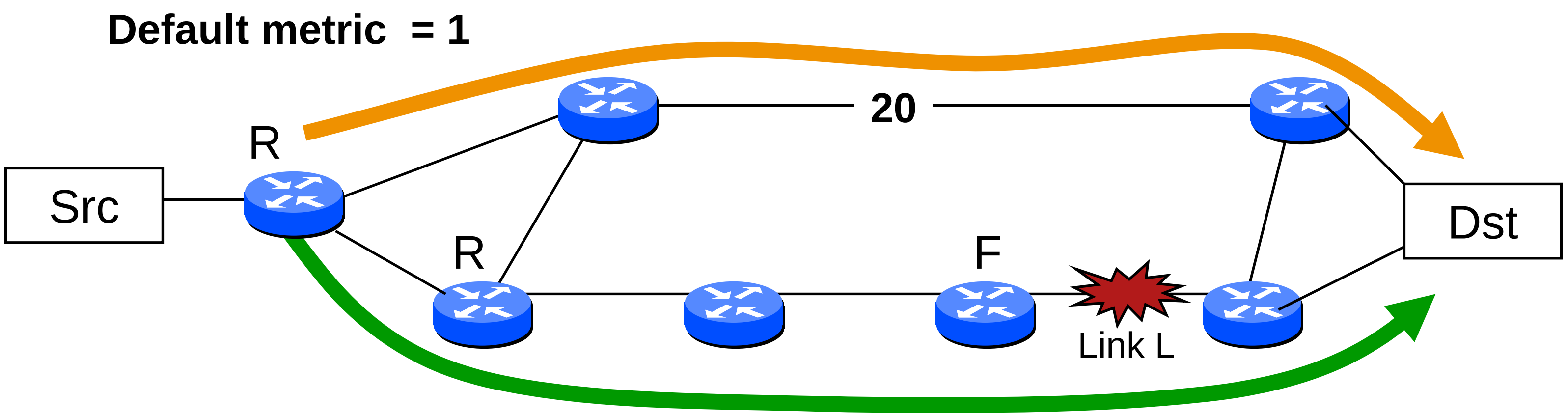
© 2007 Cisco Systems, Inc. All rights reserved.

NETWORK LEVEL RESILIENCY:

FAST CONVERGENCE and FAST REROUTING

Fast Convergence Objectives





- **Loss of Connectivity: T2 – T1, called “convergence”**
- **How fast should Fast Convergence be?:**
  - **Sub-Second: requirements for most IP networks**
  - **Sub-200ms: a few applications are sensitive to LoC <= 200ms**
  - **Sub-50ms: business requirement for some IP networks**

For the first 500 IGP (OSPF/ISIS) Prefixes and all BGP prefixes whose next-hop is within the first 500 IGP prefixes assuming the BGP routes are stable

Fast Convergence Summary

- NSF /w SSO: preserves Traffic Forwarding
  - routing information is recovered dynamically in the background
- Fast Convergence: quickly redirect flow of traffic on alternate path
  - Quicker detection of failures: signaling POS to IS-IS < 10 msec
  - Faster announcement of failure throughout the network: opt. flooding
  - Prioritized update of the routing table: important prefixes
  - Caching of redistributed routes:
  - Accelerated computation of the new network topology: iSPF
  - No compromise in stability: exponential backoff timers
- Failure scenarios:
  - GR/NSF covers redundant RP failure only
  - FC covers all failure scenarios: link failure, node failure, ...
- For more information: IGP and BGP Fast Convergence (BRKIPM-3004)

NSF and IGP Fast Hello Coexistence?

- NSF/SSO and FC have conflicting goals:

  NSF: maintain flow of traffic through failure router

  FC: fast redirect of flow of traffic away from failure router

- Deployment scenarios are often different:

  NSF: SP edge

  FC: IGP FC focus on Core (edge)

- NSF/SSO Testing with various IGP Timer settings

  Testbed with 3 SUT: Cisco 12000, 10000, 7500 / 12.0(22)S

  ISIS with 5000 routes: hello: 1 sec (multiplier: 3)

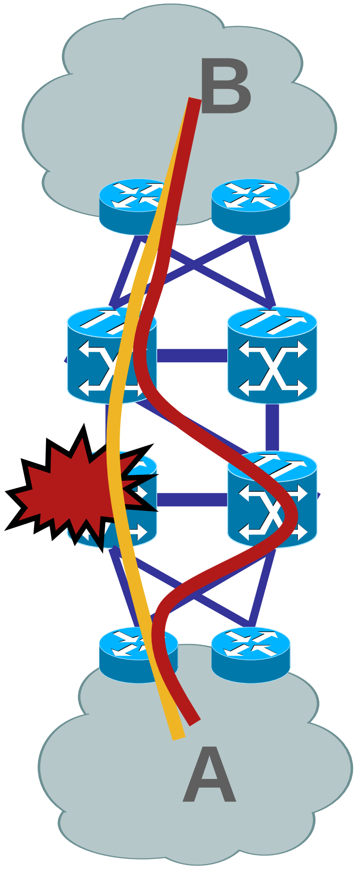
  OSPF with 5000 routes: hello: 2 sec (dead: 8 sec)

  NSF/SSO still operates properly with these timers

  first hello send ~2 sec after switchover (neighbor view ~ 3 sec)

  Conservative setting of timers > 4sec required

  For details see:
  http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper09186a00801dce40.shtml



Traffic flow with

🟨 NSF/SSO

🟥 Fast Converg.

Fast IGP Convergence Current Status

- Link/node down event is detected as fast as possible
  - Failure Detection (POS today, BFD emerging) < ~ 20ms
  - Origination < ~ 10ms
  - Queueing, Serialization, Propagation < 30ms
- Propagating the change in the network as soon as possible
  - Flooding < 5 * 2ms = 10ms
- Recalculate the paths (run SPF) as soon as possible
  - SPF < n * 40us
- Install the new routes in the routing/forwarding table
  - FIB update: p * 100us
  - FIB Distribution Delay: 50ms
    - ~ 100ms + p * 0.1 ms
- 500 important prefixes: ~ 150ms
- Worst-case over 100 iterations of most important prefixes:

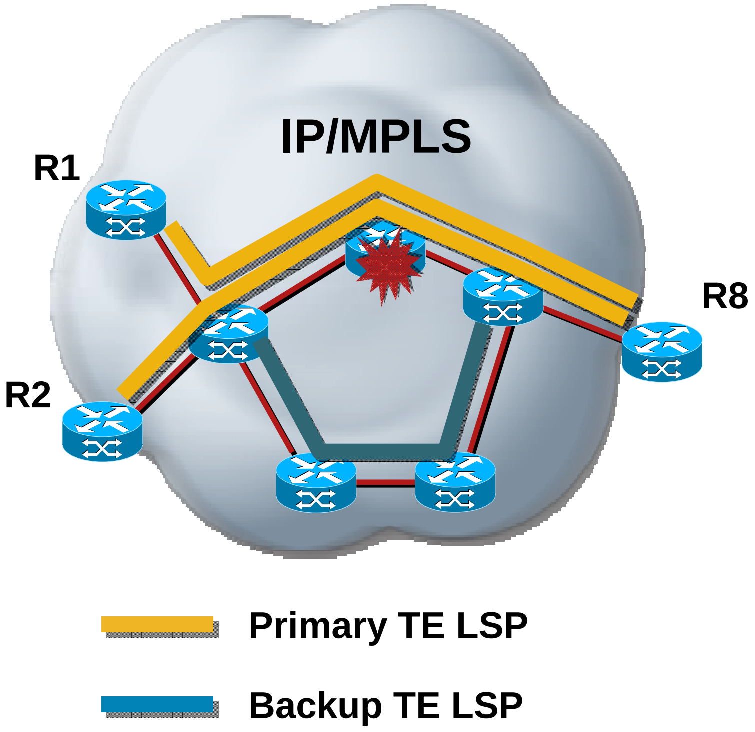
  ~280ms for 1500 nodes and 2500 prefixes
- Sub-50ms impossible today -> need Fast Reroute

© 2007 Cisco Systems, Inc. All rights reserved.

MPLS Fast Re-Route (FRR)

Key Element of Fast Reroute:

- Pre-computation of path
- Local action (to avoid propagation/distribution)
- Tunneling (to avoid propagation/distribution)



MPLS FRR:

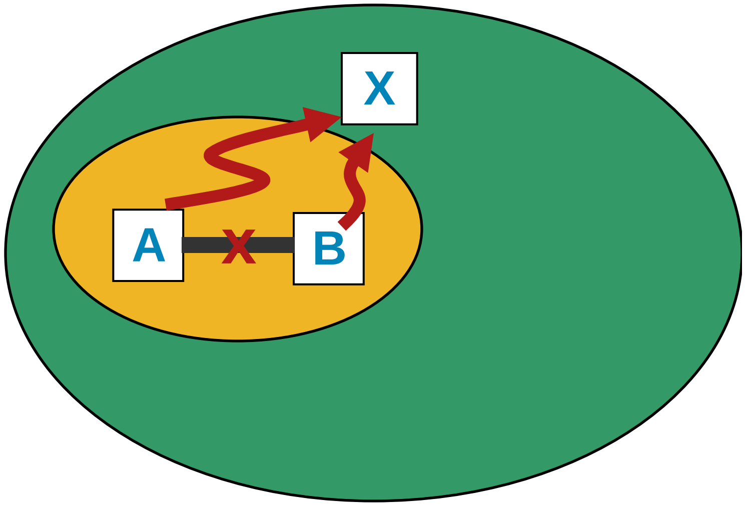
- fast recovery against node/link failures
- Scalable 1:N protection
- Greater protection granularity
- Cost-effective alternative to optical protection
- Bandwidth protection

IP Fast Reroute (IPFRR) Concepts

- **Limited Area of failure**

  –**Failure of Link A <--> B and topology change impacts only subset of network (orange layer, confirmed by FC project)**

  –**Outside this area subset routing is consistent (green layers)**



- **Find a consisten point in the network (X)**

  – **X is not impacted by the failure**

  – **X can be reached independent of failure**

  – **X forwards traffic to any destination /wo AB**

  – **From X all packets flow to their destination while avoiding the failure (and without knowledge of the failure)**

🟢 **Consistent routing**

🟠 **Impacted area of topology change**

- **Several proposal to IETF**

  Release Point, Downstream Routes, Loop-Free Alternates, U-Turns, Not-Via Addresses

- **Cisco proposal consists of**

  Loop Free Alternates (aka: Downstream Routes)

  Not-Via Addresses

  Ordered-SPF Algorithm

see session BRKIPM-3017 for details

© 2007 Cisco Systems, Inc. All rights reserved.

EMBEDDED MANAGEMENT

© 2007 Cisco Systems, Inc. All rights reserved.

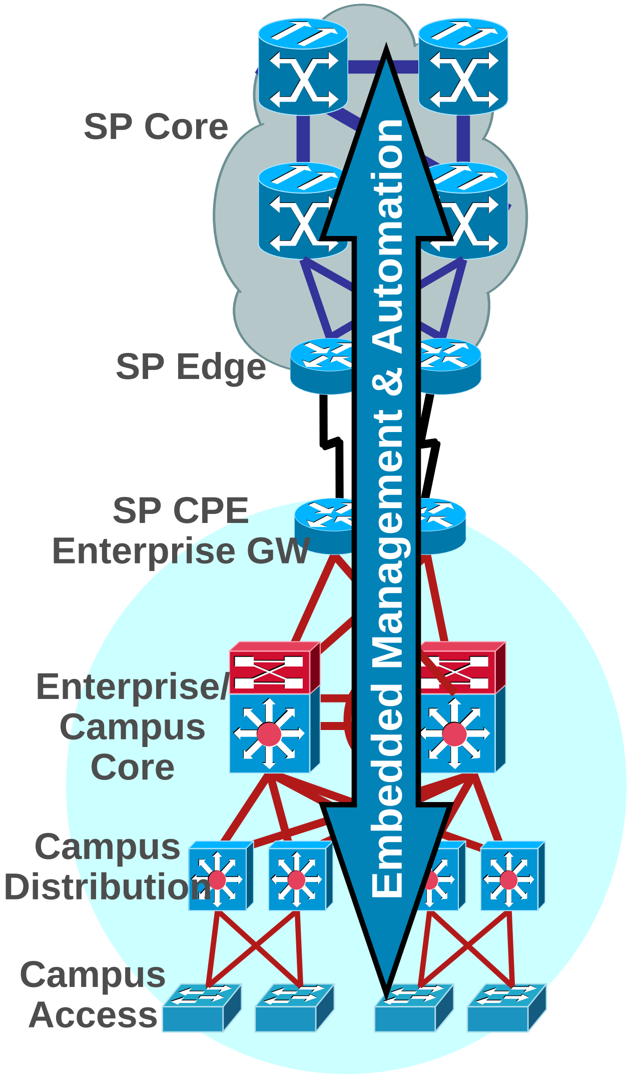
Embedded Management & Automation

- LDP Autoconfig
- MPLS OAM Toolbox:

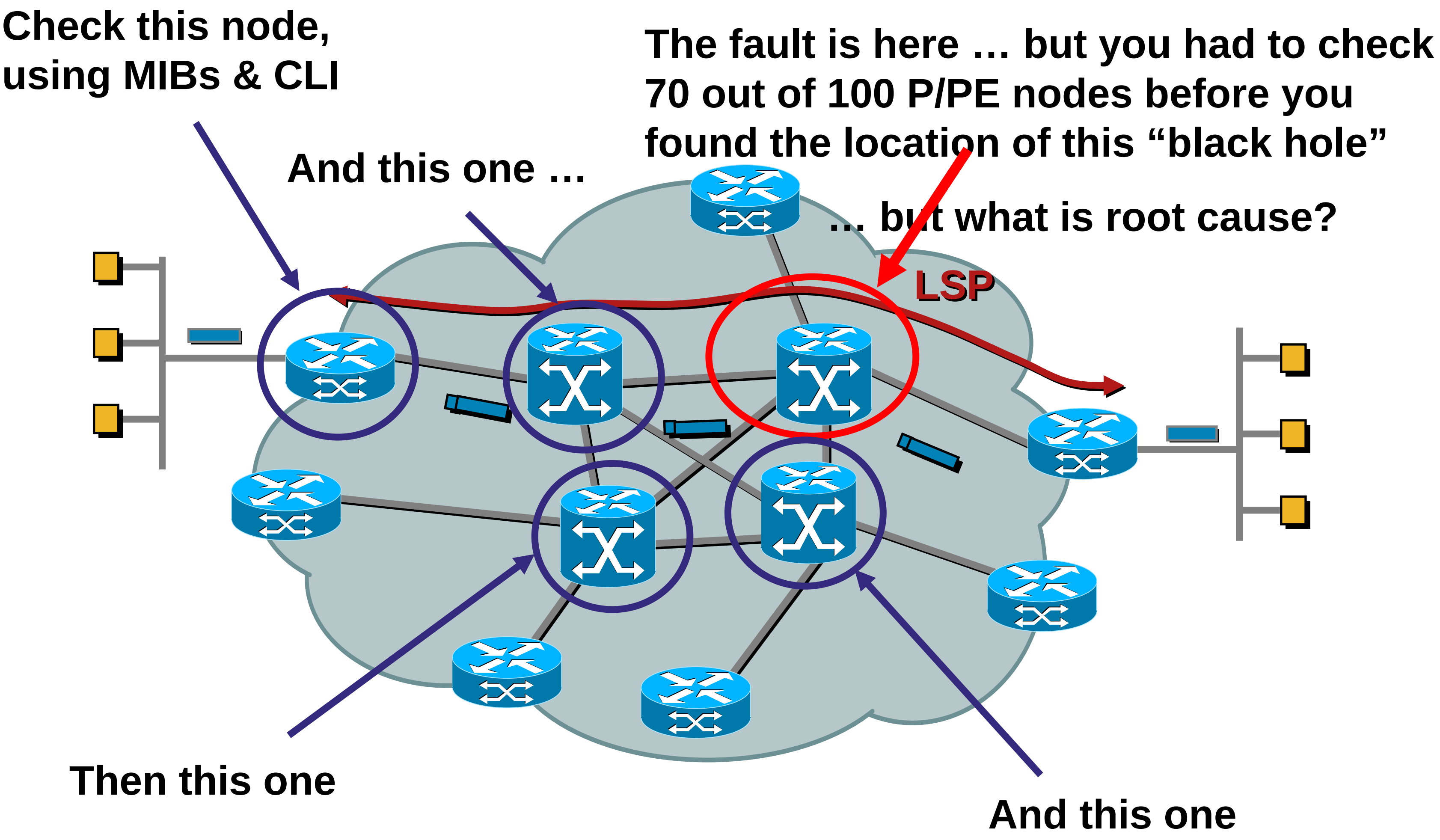
  MPLS Ping

  MPLS Traceroute
- Device Reachability:

  ICMP, IP SLA,
- SNMP, RMON, Syslog
- Component Outage Online Measurement (COOL)
- Embedded Event Manager (EEM)
- Generic Online Diagnostics (GOLD)
- Internet Solution Center (ISC)
- Cisco MPLS Diagnostics Expert (MDE)



Troubleshooting MPLS/VPN: Fault in MPLS Core LSP “blackhole” (Real Life) Example



Check this node, using MIBs & CLI

And this one …

The fault is here … but you had to check 70 out of 100 P/PE nodes before you found the location of this “black hole”

… but what is root cause?

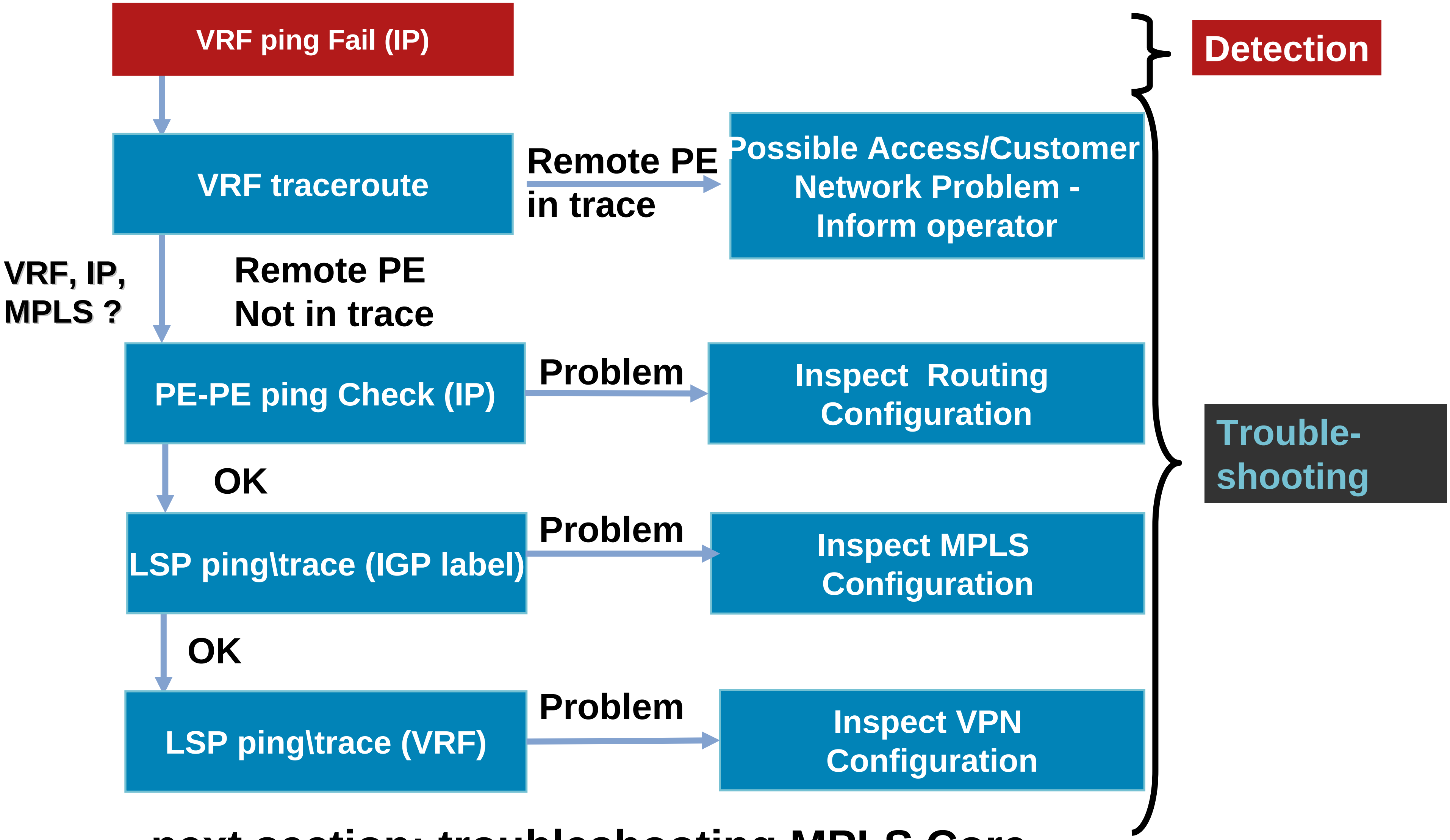
Then this one

And this one

- Real Life Example … took 17 hours of outage to find and diagnose
- Root Cause: partial HW failure on LC affecting LSP forwarding

© 2007 Cisco Systems, Inc. All rights reserved.

Troubleshooting Workflow - VRF data plane



VRF ping Fail (IP)

VRF traceroute

Remote PE in trace

Possible Access/Customer Network Problem - Inform operator

VRF, IP, MPLS ?

Remote PE Not in trace

PE-PE ping Check (IP)

Problem

Inspect Routing Configuration

OK

LSP ping\trace (IGP label)

Problem

Inspect MPLS Configuration

OK

LSP ping\trace (VRF)

Problem

Inspect VPN Configuration

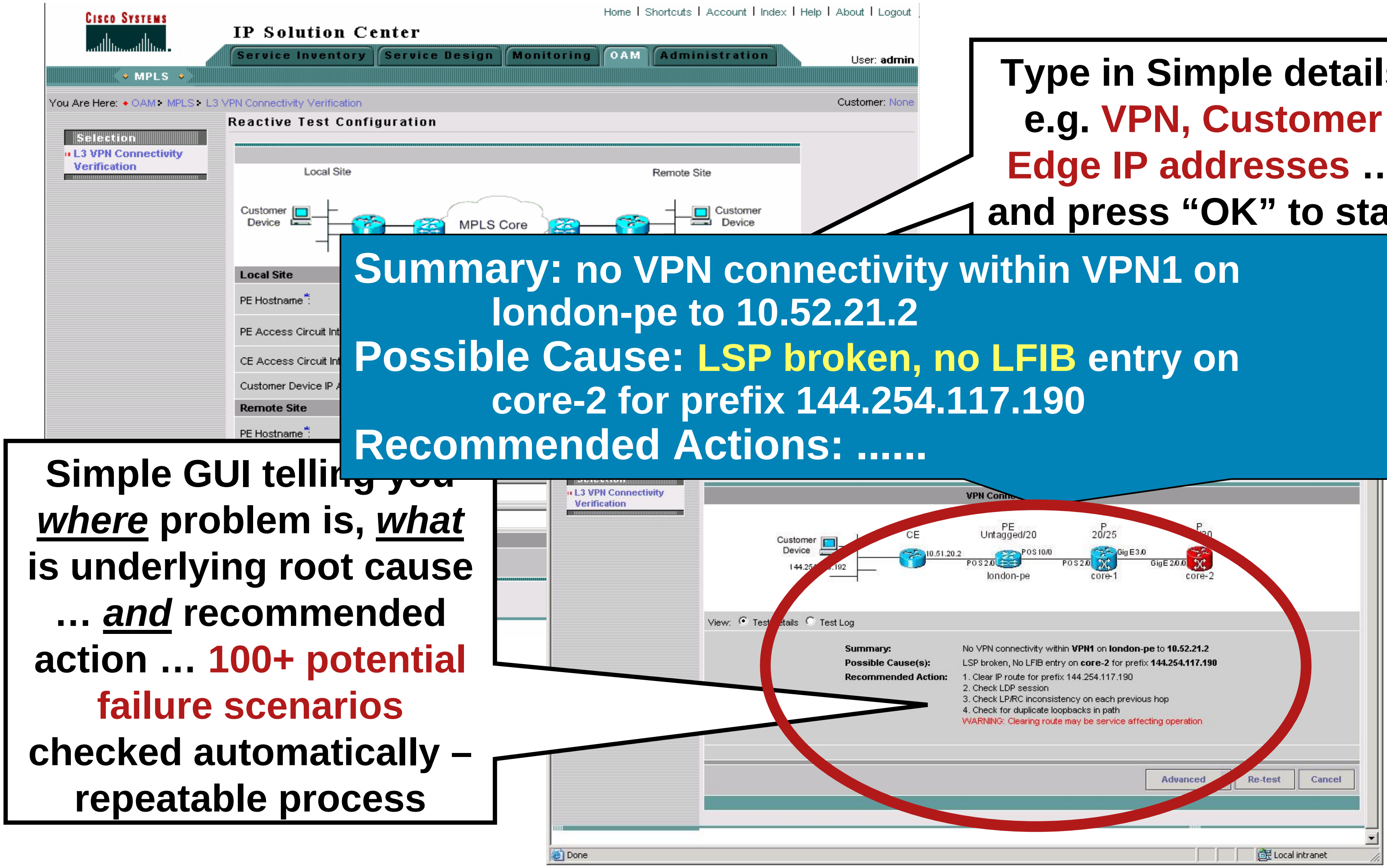
Detection

Trouble-shooting

next section: troubleshooting MPLS Core

© 2007 Cisco Systems, Inc. All rights reserved.

MPLS Diagnostics Expert (MDE): Intelligent MPLS Instrumentation



View: ● Test Details ○ Test Log

Summary: No VPN connectivity within **VPN1** on **london-pe** to **10.52.21.2**

Possible Cause(s): LSP broken, No LFIB entry on **core-2** for prefix **144.254.117.190**

Recommended Action:
1. Clear IP route for prefix 144.254.117.190
2. Check LDP session
3. Check LP/RC inconsistency on each previous hop
4. Check for duplicate loopbacks in path

WARNING: Clearing route may be service affecting operation

Type in Simple details e.g. VPN, Customer Edge IP addresses … and press "OK" to start

Summary: no VPN connectivity within VPN1 on london-pe to 10.52.21.2
Possible Cause: LSP broken, no LFIB entry on core-2 for prefix 144.254.117.190
Recommended Actions: ......

Simple GUI telling you *where* problem is, *what* is underlying root cause … *and* recommended action … 100+ potential failure scenarios checked automatically – repeatable process

www.cisco.com/go/mde

© 2007 Cisco Systems, Inc. All rights reserved.

Automated Failure Scenarios with Cisco MDE

- **Edge: > 30 Unique Scenarios**
  - –Config issues e.g. Route Target Mismatches between Ingress/Egress PE
  - –Interface not associated with VRF; VRF route limit exceeded
  - –Inconsistencies – e.g. Route installed into BGP table but not VRF
  - –Mismatches between FIB/LFIB; Routes not distributed into MP-BGP
- **Core: > 30 Unique Scenarios**
  - –Config issues – *"finger trouble"* e.g. CEF/VPNv4 Address family disabled
  - –Label allocation/installation issues; RP/LC inconsistencies
  - –LSP Blackholes; Packets too big for Interface MTU
- **Access Circuits: > 40 Unique Scenarios**
  - –Config issues - Interface admin down, Line protocol down
  - –CE/PE connectivity – including automated execution of ATM & FR OAM
  - –Packets being dropped in switched (ATM/FR) circuit
- **Core failure diagnosis depends upon LSP ping & LSP traceroute – Edge & Access Don't**
- **The 80/20 Rule**

Generic OnLine Diagnostics (GOLD)
What is it?

- GOLD defines a common framework for diagnostics operations across Cisco platforms running Cisco IOS Software.
- It checks the health of hardware components and verifies proper operation of the system data and control planes.
- Provides a common CLI and scheduling for field diagnostics including :
  - –Bootup Tests (includes online insertion)
  - –Health Monitoring Tests (background non-disruptive)
  - –User Scheduled and On-Demand Tests (disruptive and Non-disruptive)
  - –SNMP/CLI access to data via Management Interface
  - –Deployment tool

© 2007 Cisco Systems, Inc. All rights reserved.

Generic OnLine Diagnostics

Diagnostics Operations

Boot-Up diagnostics

```
6k(config)# diagnostic bootup level
complete
```

Run During System Bootup, Line Card OIR or Supervisor Switchover
Makes sure faulty hardware is taken out of service

Runtime diagnostics

Health-Monitoring

```
6K(config)# diagnostic monitor module 5 test 2
6K(config)# diagnostic monitor interval module
5 test 2 00:00:15
```

Non-disruptive tests run in the background
Serves as HA trigger

On-Demand

```
6k# diagnostic start module 4 test 8
Module 4: Running test(s) 8 may disrupt
normal system operation
Do you want to continue? [no]: y
6k#diagnostic stop module 4
```

All diagnostics tests can be run on demand, for *troubleshooting purposes. It can also be used as a pre-deployment tool.*

Scheduled

```
6k(config)# diagnostic schedule module 4
test 1 port 3 on Jan 3 2005 23:32
6k(config)#diagnostic schedule module 4
test 2 daily 14:45
```

Schedule diagnostics tests, for *verification and troubleshooting purposes*

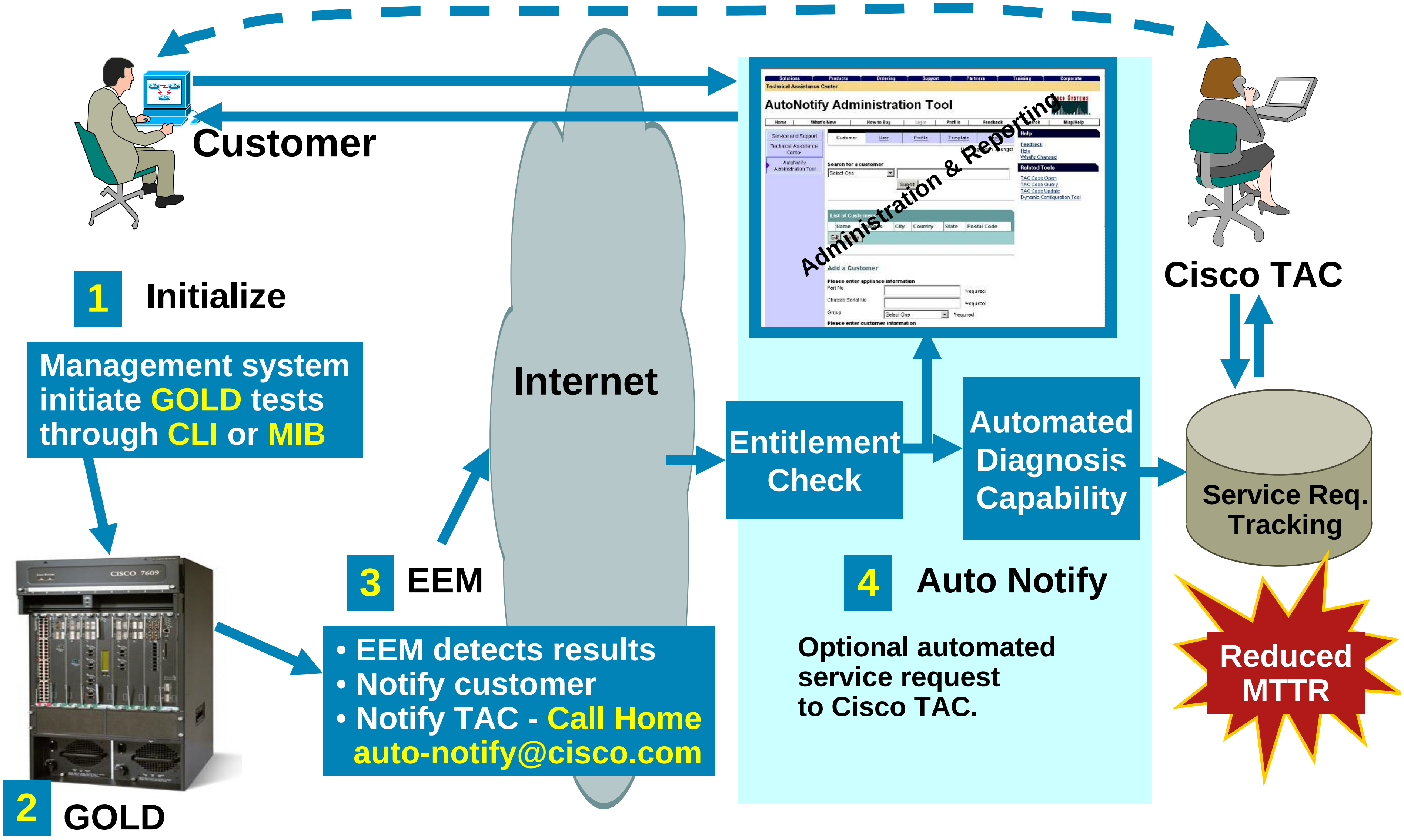
Generic OnLine Diagnostics:
GOLD Test Suite

- **Bootup Diagnostics**
  - forwarding Engine Learning Tests (Sup/DFC)
  - L2 Tests (Channel, BPDU, Capture)
  - L3 Tests (IPv4, IPv6, MPLS)
  - Span and Multicast Tests
  - CAM Lookup Tests (FIB, NetFlow, QoS CAM)
  - Port Loopback Test (all cards)
  - Fabric Snake Tests
- **Health Monitoring Diagnostics**
  - SP-RP Inband Ping Test (Sup's SP/RP, EARL(L2&L3), RW engine)
  - Fabric Channel Health Test (Fabric enabled line cards)
  - MacNotification Test (DFC line cards)
  - Non Disruptive Loopback Test
  - Scratch Registers Test (PLD & ASICs)
- **On-Demand Diagnostics**
  - Exhaustive Memory Test
  - Exhaustive TCAM Search Test
  - Stress Testing
  - All bootup and health monitoring tests can be run on-demand
- **Scheduled Diagnostics**
  - All bootup and health monitoring tests can be scheduled
  - Scheduled Switch-over

Functional Testing combined with components monitoring to detect fault in passive components (connector, solder joint etc.) and active components (ASICs, PLDs etc.)

Catalyst 6500

GOLD / EEM / Call Home / Auto Notify



1 Initialize

Management system initiate **GOLD** tests through **CLI** or **MIB**

2 GOLD

3 EEM

- EEM detects results
- Notify customer
- Notify TAC - **Call Home** **auto-notify@cisco.com**

4 Auto Notify

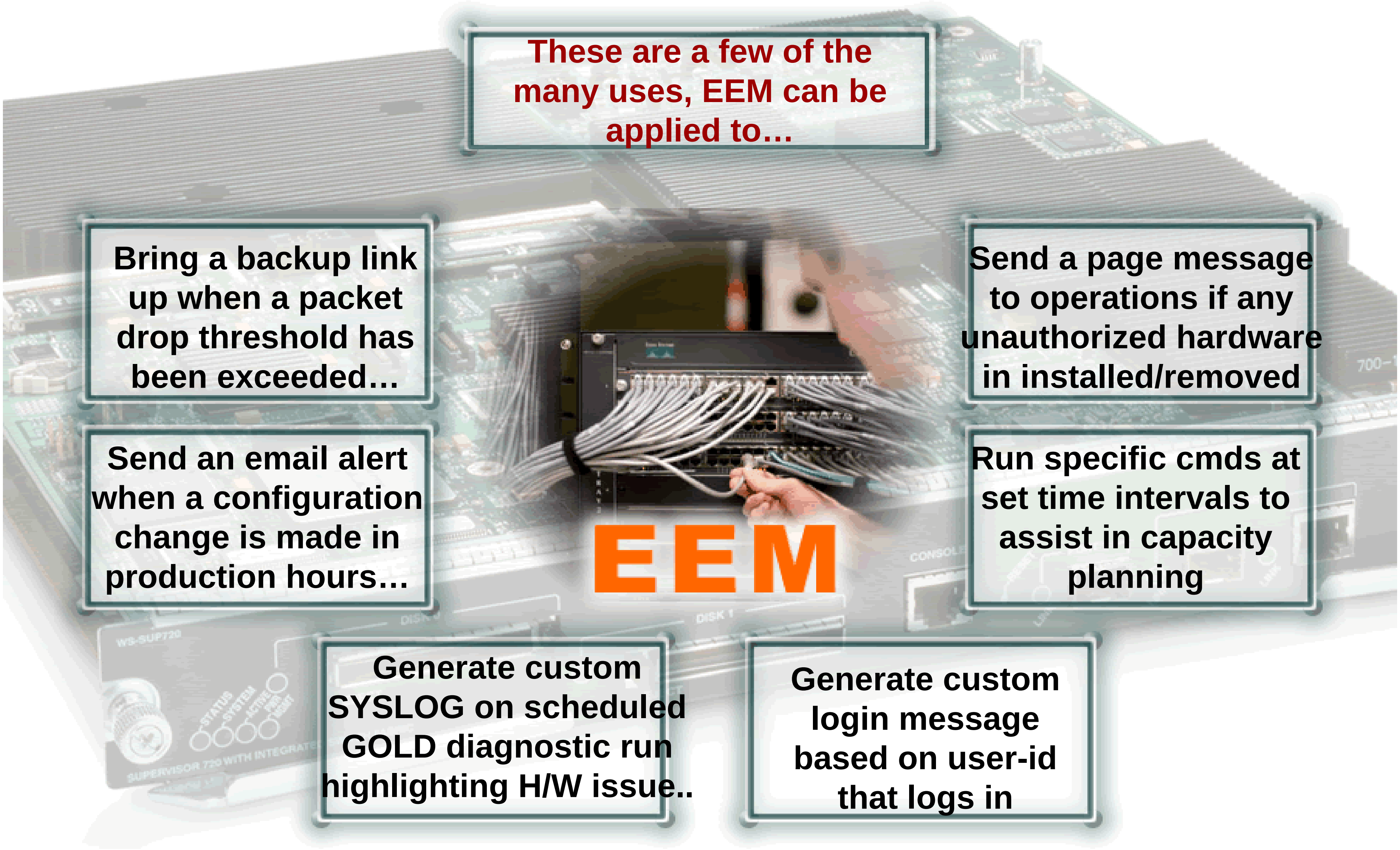
Optional automated service request to Cisco TAC.

© 2007 Cisco Systems, Inc. All rights reserved.

Embedded Event Manager

- EEM is an IOS enhancement running on CPU
- Combination of processes designed to monitor key system parameters such as CPU utilization, interface counters, SNMP and SYSLOG events.
- It acts on specific events or thresholds/counters that are exceeded…
- Available on 12.0S, 12.2S, 12.2SX, 12.3T, 12.4 and 12.4T and various platforms, check Feature Navigator

Embedded Event Manager
How can it be used?



These are a few of the many uses, EEM can be applied to…

Bring a backup link up when a packet drop threshold has been exceeded…

Send an email alert when a configuration change is made in production hours…

Send a page message to operations if any unauthorized hardware in installed/removed

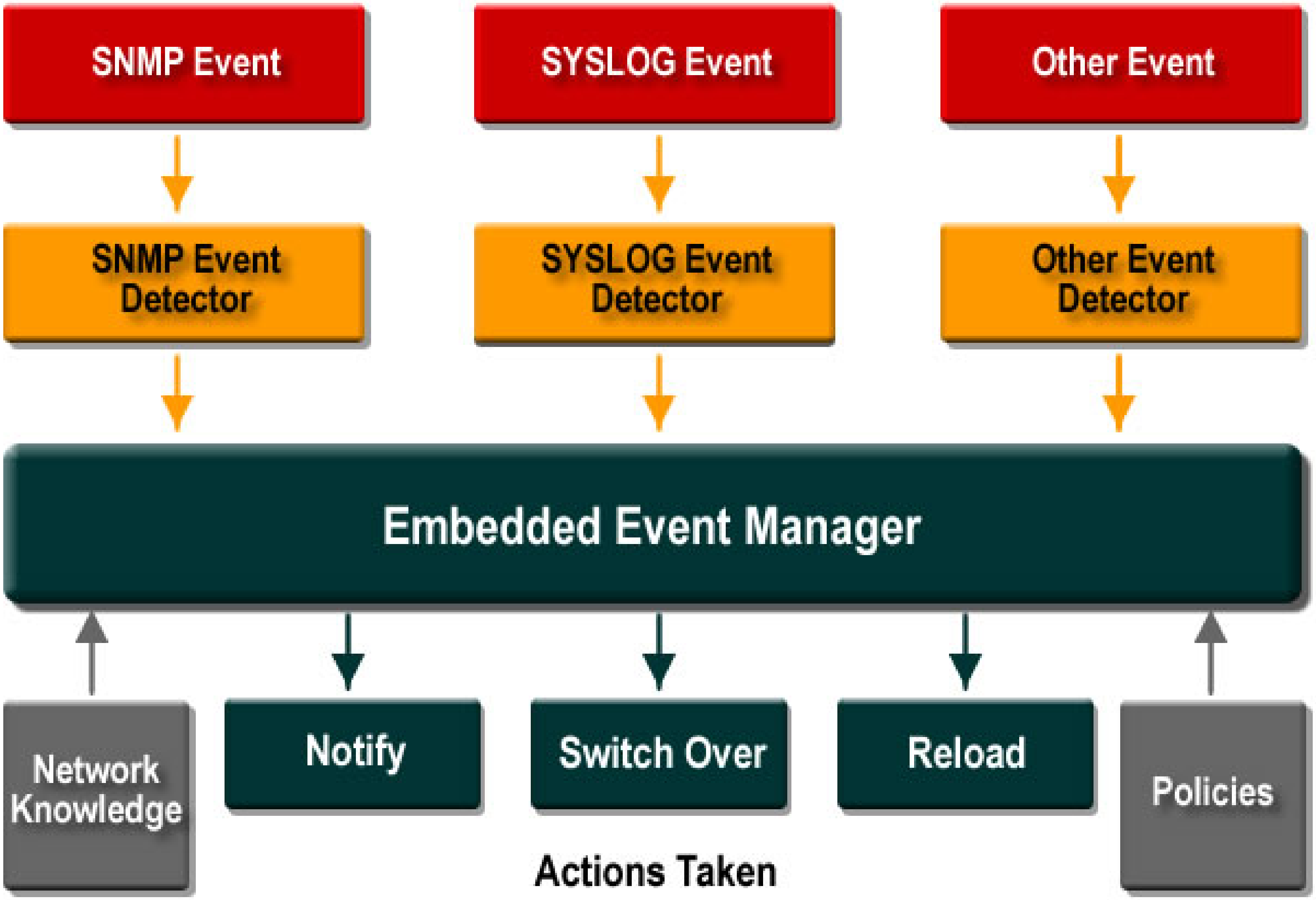
Run specific cmds at set time intervals to assist in capacity planning

Generate custom SYSLOG on scheduled GOLD diagnostic run highlighting H/W issue..

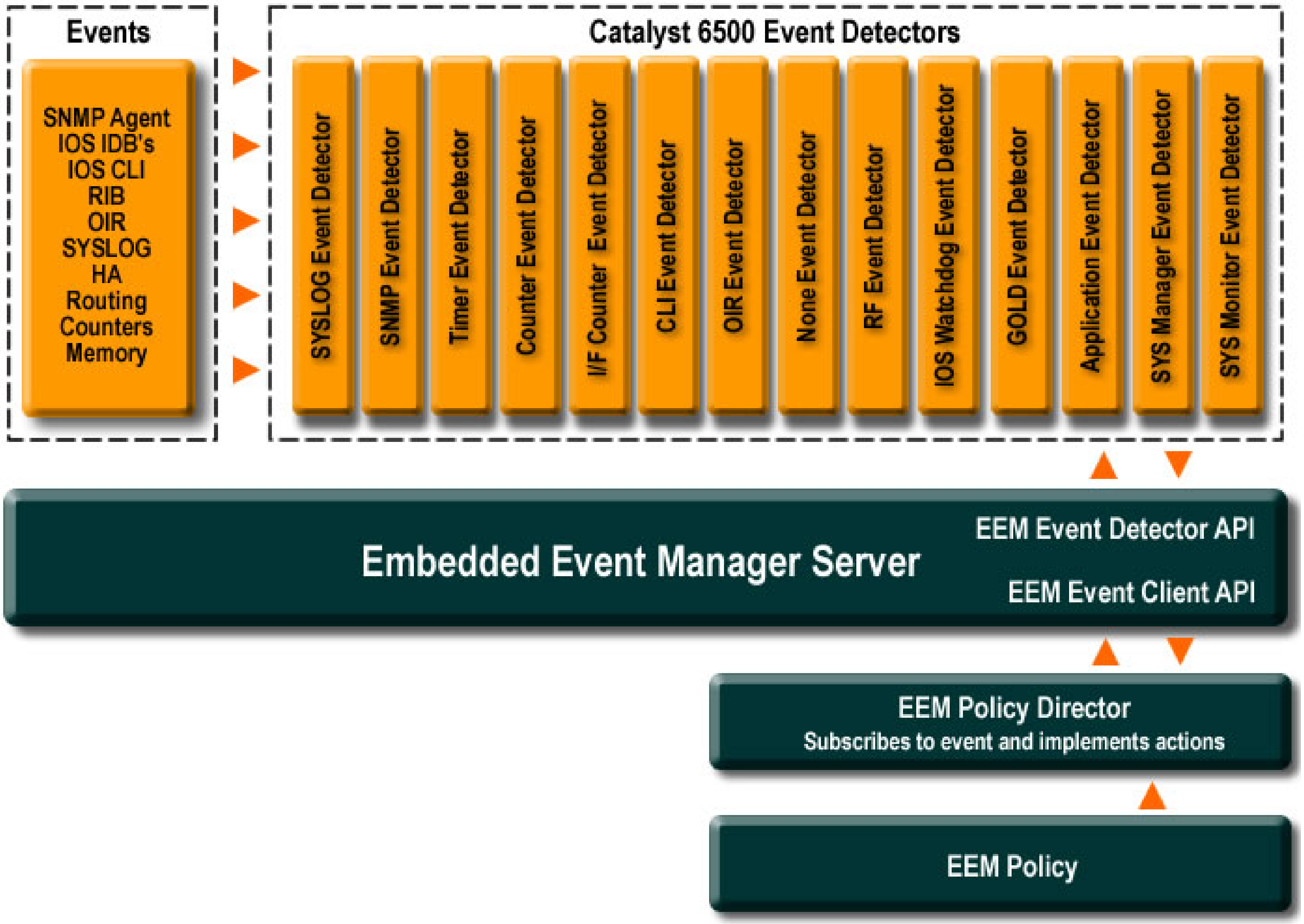
Generate custom login message based on user-id that logs in

Embedded Event Manager

Basic EEM Architecture

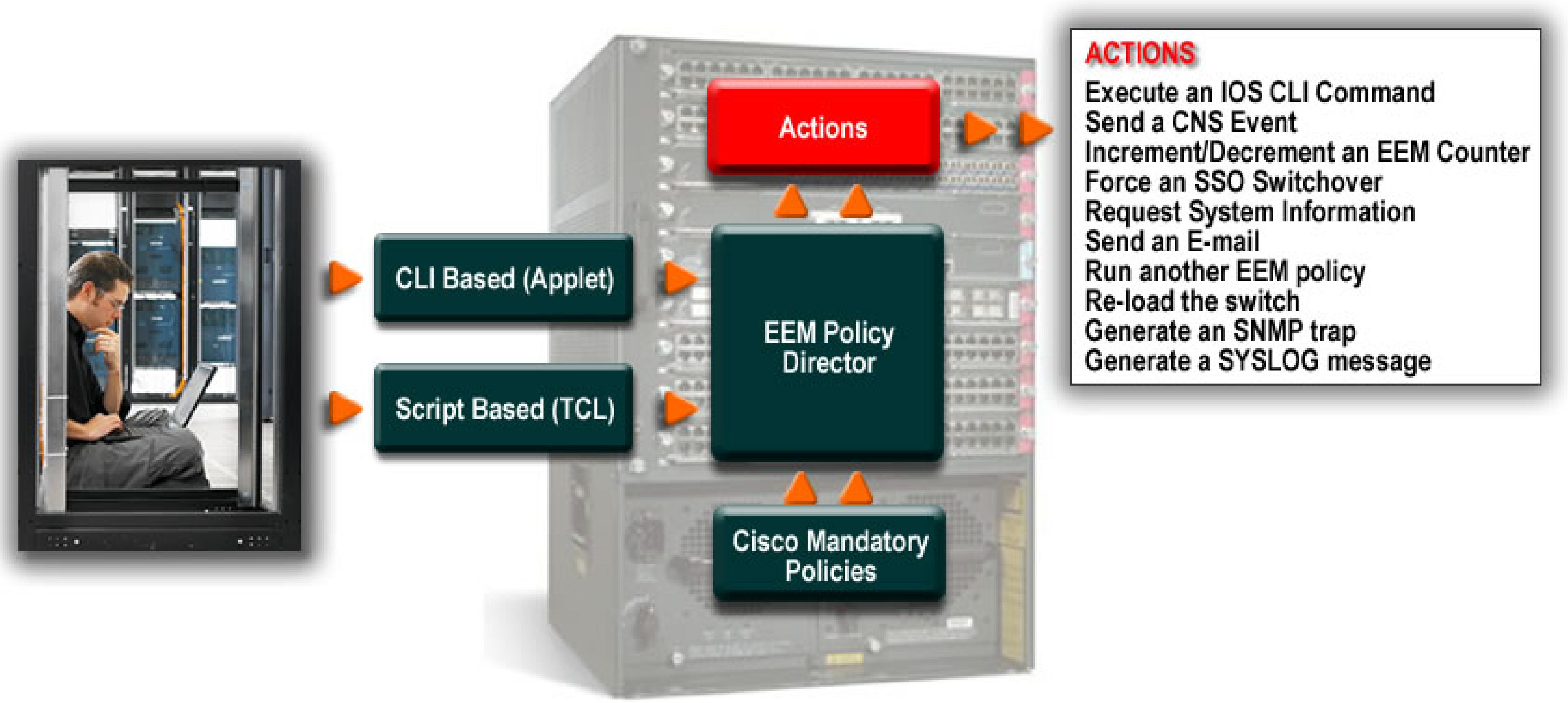


Embedded Event Manager
Detailed Architecture and Event Detectors



Embedded Event Manager Policies

- Policies defined via:
  - - CLI (known as an applet) or
  - - TCL script
- Policies loaded onto a local file system
- Policies can generate a variety of actions



CLI Based (Applet)

Script Based (TCL)

EEM Policy Director

Cisco Mandatory Policies

Actions

ACTIONS

Execute an IOS CLI Command
Send a CNS Event
Increment/Decrement an EEM Counter
Force an SSO Switchover
Request System Information
Send an E-mail
Run another EEM policy
Re-load the switch
Generate an SNMP trap
Generate a SYSLOG message

Embedded Event Manager
CLI Commands – applet config mode

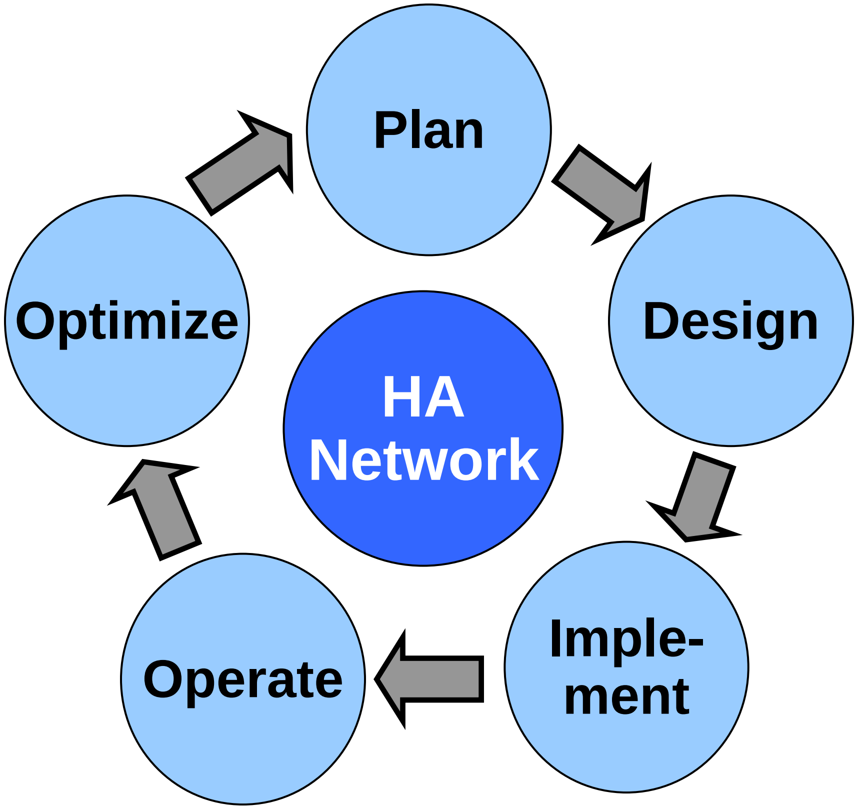
```
PODx#config t
Enter configuration commands, one per line.  End with CNTL/Z.
PODx(config)#event man applet one
PODx(config-applet)#event syslog pattern “COUNT”
PODx(config-applet)#action 1.0 syslog msg “applet one”
PODx(config-applet)#exit
PODx(config)#exit
00:04:01: %SYS-5-CONFIG_I: Configured from console by consol
PODx# clear counters
Clear "show interface" counters on all interfaces [confirm]y
00:04:14: %CLEAR-5-COUNTERS: Clear counter on all interfaces by
console
00:04:14: %HA_EM-6-LOG: one: applet one
PODx#
```

© 2007 Cisco Systems, Inc. All rights reserved.

HIGH AVAILABILITY BEST PRACTISES:

THE CULTURE OF AVAILABILITY

The Culture of Availability



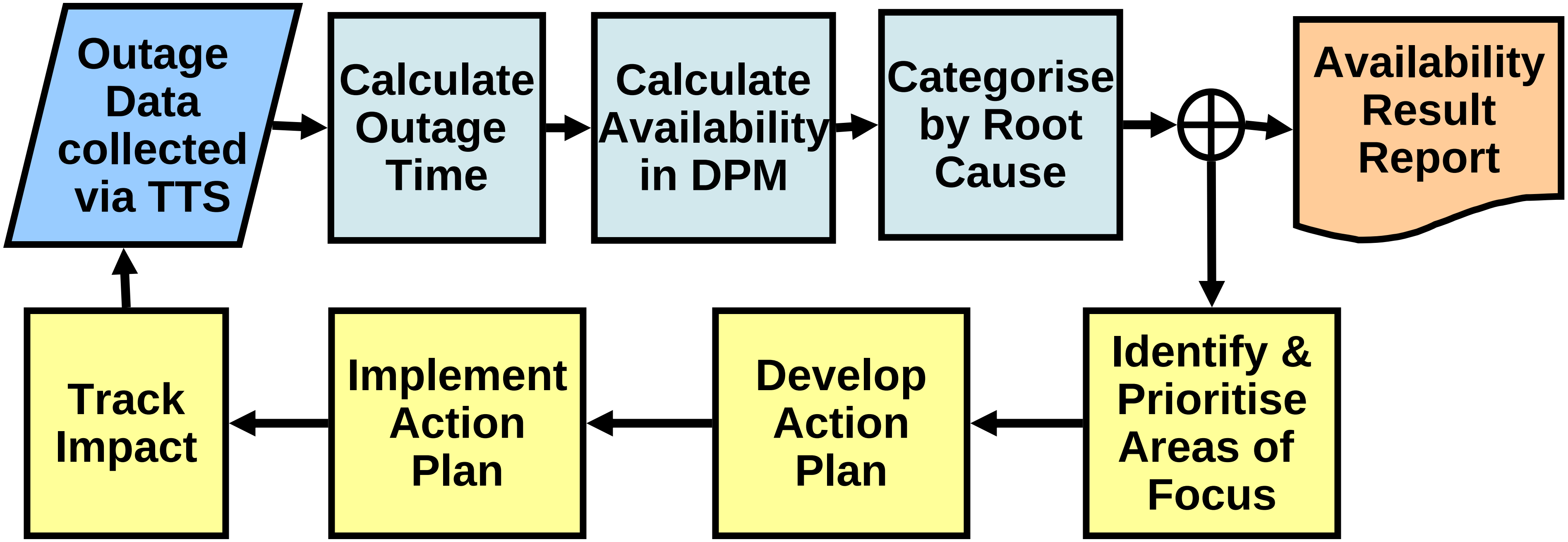
PDIOO
Framework

- Calculating, Measuring, and Improving Availability
- People, Process, and Tools for High Availability
- Configuration and Design

What Is Your Availability Level?

- **Analyze the Gaps: Reactive ~99%**
  - **Few, if any, identified processes (fix user reported problems)**
  - **Low tool utilization**
  - **Low level of consistency (HW, SW, config, design)**
  - **No quality-improvement processes**
- **Analyze the Gaps: Proactive ~99.9%**
  - **Good change management processes (what-if analysis, change validation)**
  - **Fault and configuration management tools**
  - **Improved consistency (HW, SW, config, design)**
  - **Typically no quality improvement process**
- **Analyze the Gaps: Predictive ~99.99+%**
  - **Consistent processes for fault, configuration, performance, and security**
  - **Fault, configuration, performance, and workflow process tools**
  - **Excellent consistency (HW, SW, config, design)**
  - **HA culture of quality improvement**

„Trouble Ticket Availability Measures“ Method


Outage Data collected via TTS
Calculate Outage Time
Calculate Availability in DPM
Categorise by Root Cause
Availability Result Report
Identify & Prioritise Areas of Focus
Develop Action Plan
Implement Action Plan
Track Impact

+ Easy to get started (no network overhead)

+ Assists Operational & Strategic Business Decisions

+ Better data quality (categorized outages, trend impact to network)

-- Outage may occur that are not included in Trouble Ticket System

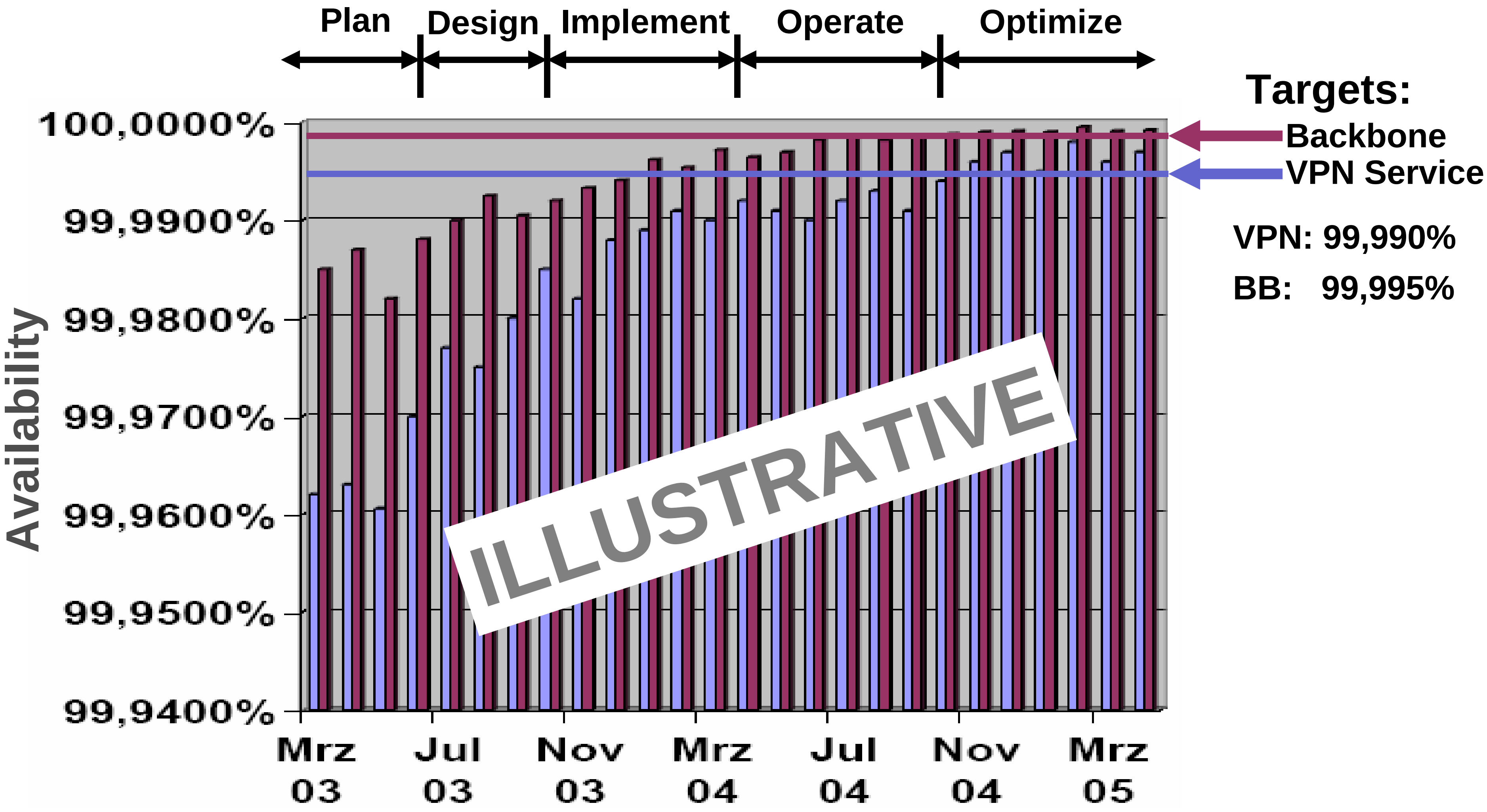
-- Internal consistency process issues

© 2007 Cisco Systems, Inc. All rights reserved.

Cisco Advanced Services NAIS: Network Availability Improvement Service

- Service based on „Trouble Ticket Measures“ method
- Customized Result Packages
  - Overall Network Availability (Baseline, Trends)
  - Downtime Analysis:
    - Planned vs unplanned
    - root cause
    - Resolution
    - Equipment Tpye
  - MTTR Analysis
- 100+ Networks worldwide used service to improve availability
- Contact you local Cisco sales team for further information
- http://www.cisco.com/en/US/partner/netsol/ns206/networking_solutions_white_paper09186a008015829c.shtml

Availability Measurement and Improvement

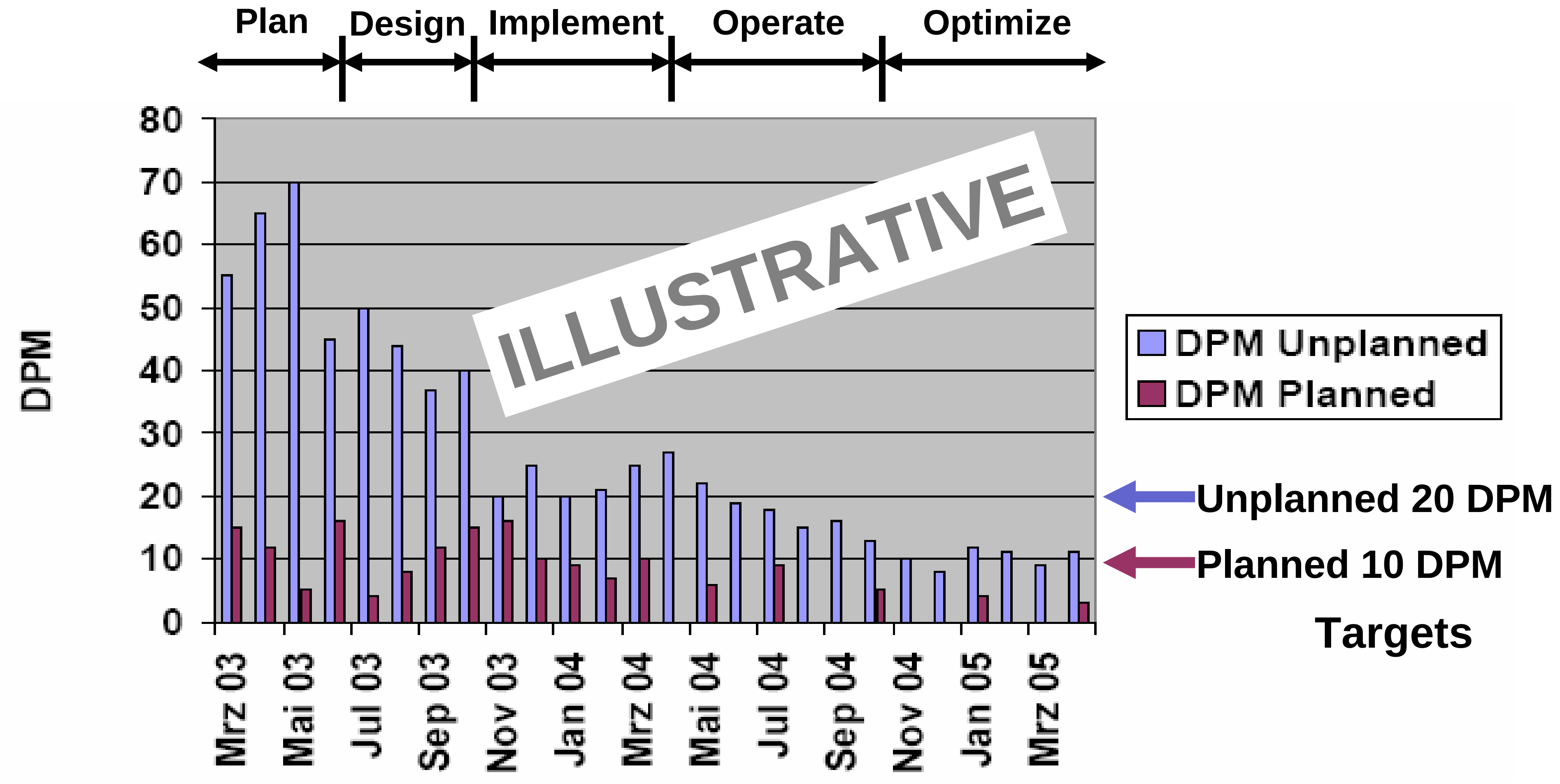


VPN: 99,990%

BB: 99,995%

© 2007 Cisco Systems, Inc. All rights reserved.

Reduce Unplanned Outages

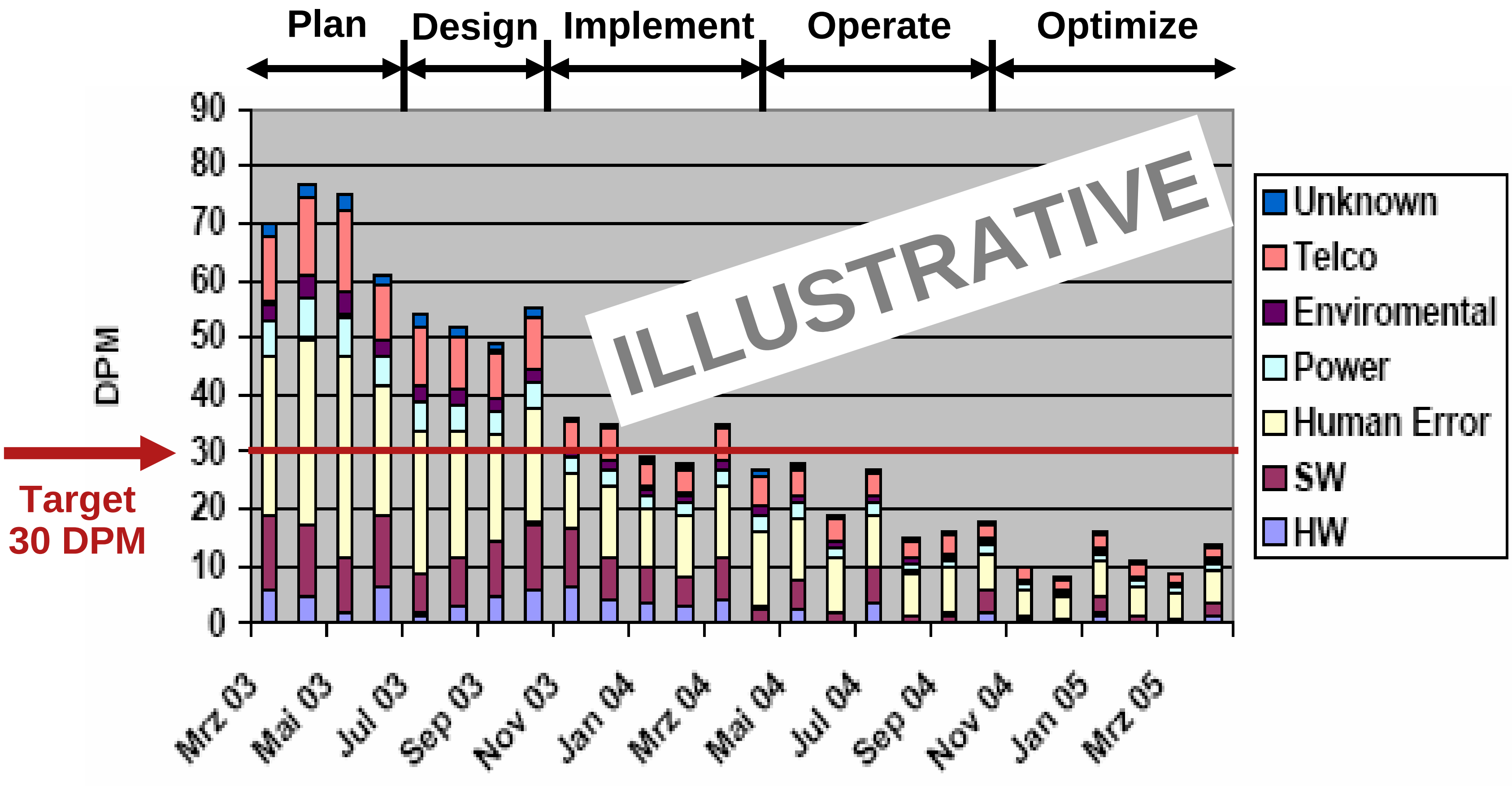


Unplanned 20 DPM

Planned 10 DPM

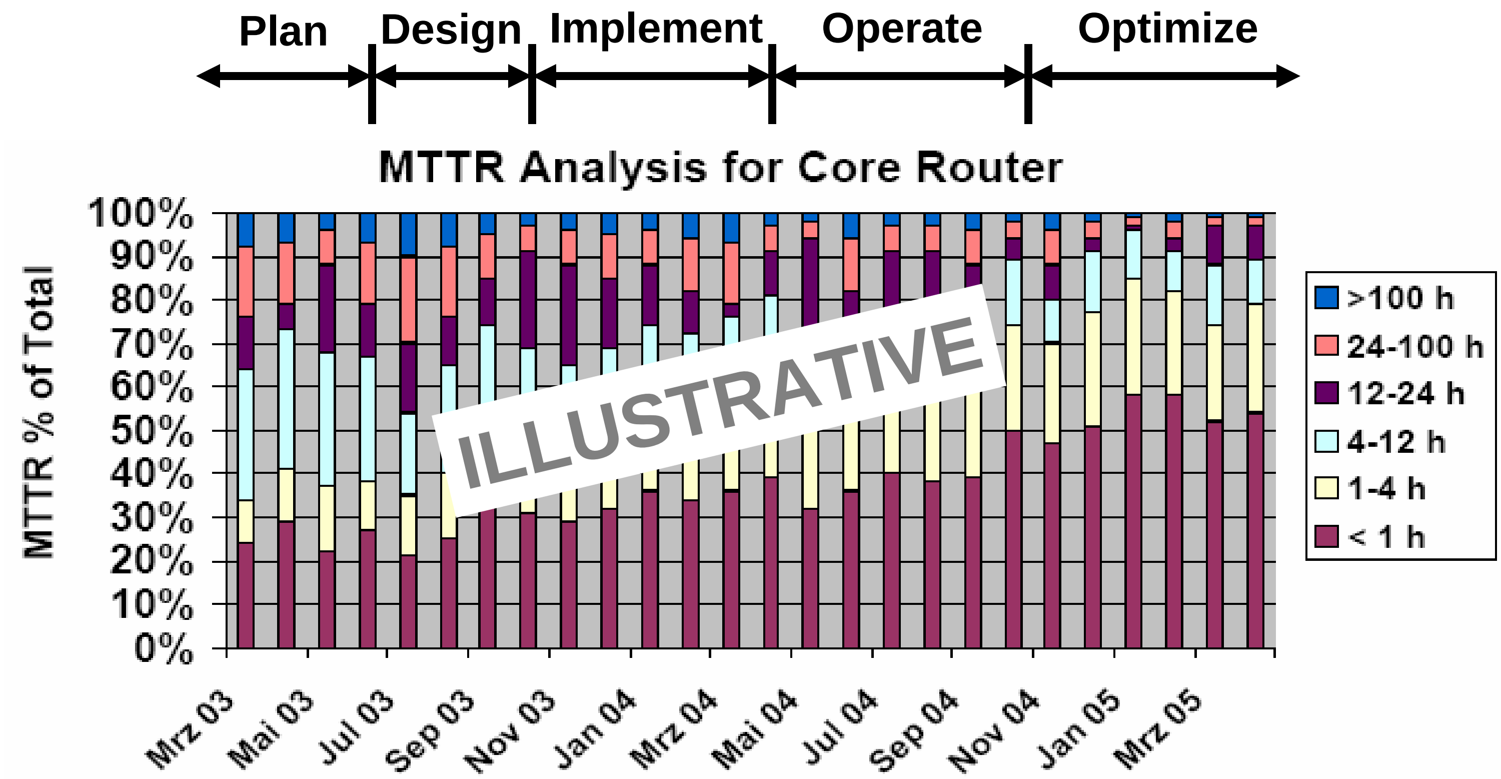
Targets

Analyse DPM by Cause



© 2007 Cisco Systems, Inc. All rights reserved.

MTTR Analysis – Hardware Faults



# of MTTR hours decreased for Core Router HW

> 50% of faults resolved in < 1 h

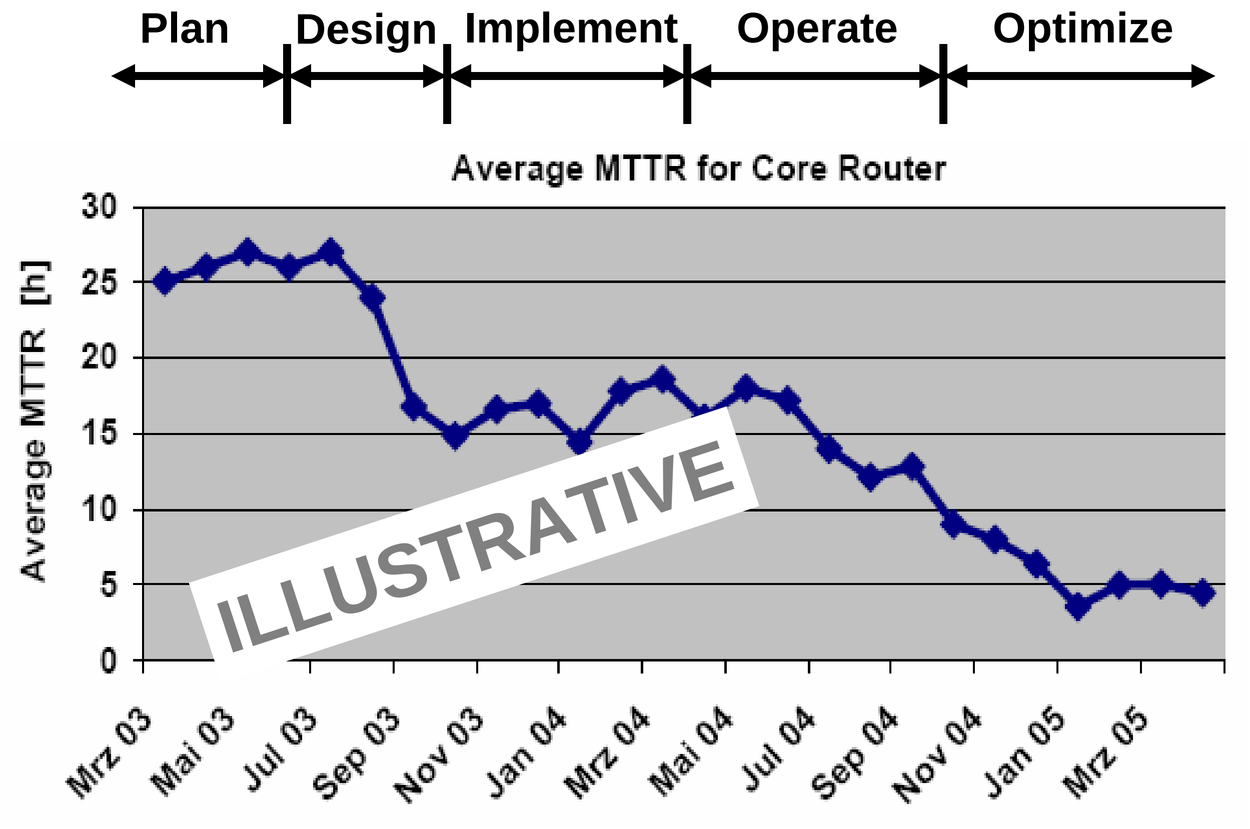
> 70% of faults resolved in < 4 h

< 5% of fault not resolve in <24 h

Produced for each Fault Type

© 2007 Cisco Systems, Inc. All rights reserved.

MTTR Analysis – Hardware Faults



# of MTTR hours decreased for Core Router HW reduced average MTTR from 25 h down to 5 h

Produced for each Fault Type

© 2007 Cisco Systems, Inc. All rights reserved.

"The real value of the RT59 program was to drive tangible availability improvements from assessment through to implementation. Cisco enabled us to prioritize our efforts to bring about the greatest improvements in the shortest amount of time. In addition we have seen productivity benefits and found an opportunity to sell better services to our customers."

Manager of Operations of a Service Provider
about NAIS Program, 2003

Getting to 4 Nines

Roadblocks to 4 Nines (99.99%)

- Single point of failure
  (edge card, edge router, single trunk)
- Outage required for hardware and software upgrades
- Long recovery time for reboot or switchover
- No tested hardware spares available on site
- Long repair times due to a lack of troubleshooting guides and process
- Inappropriate environmental conditions

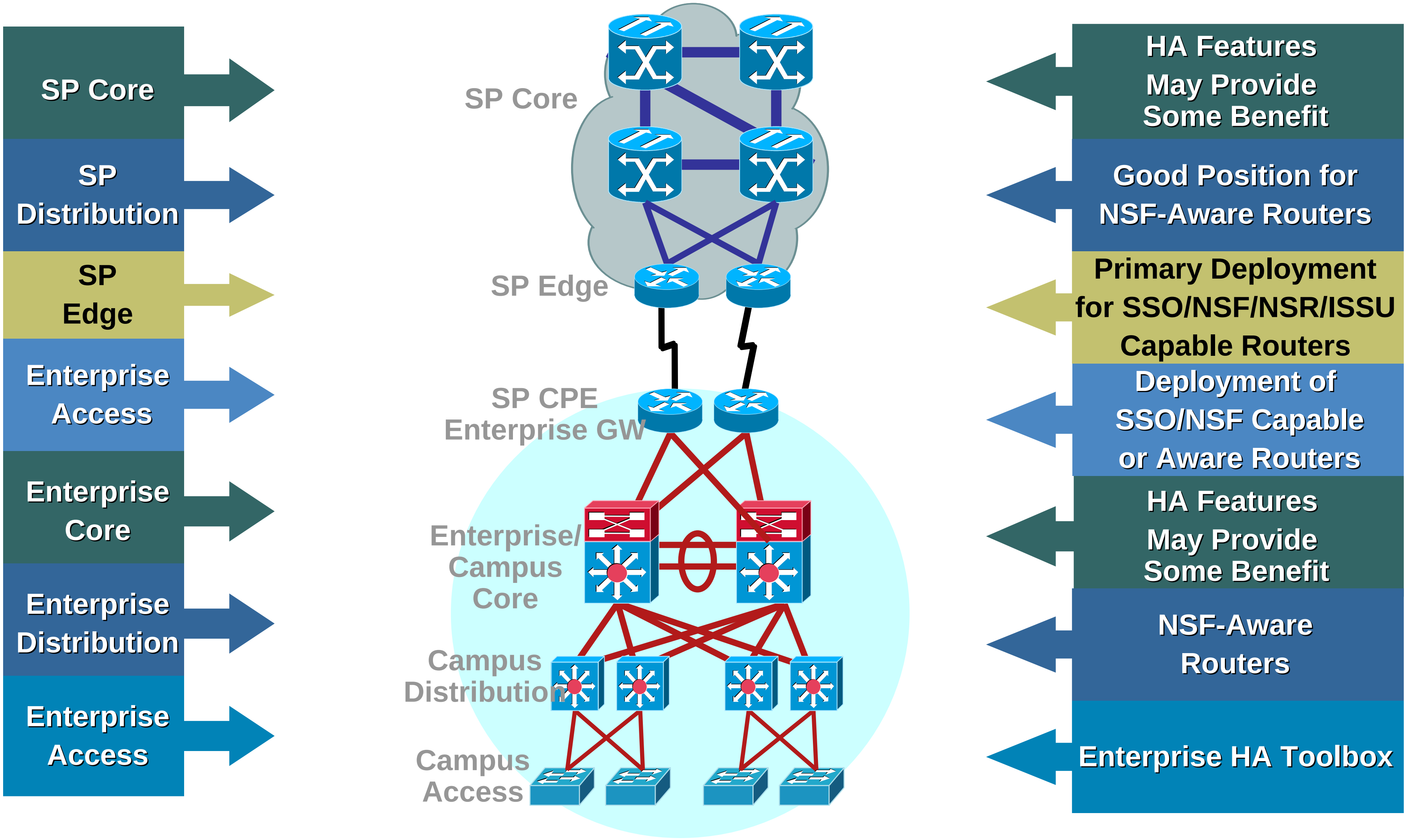
© 2007 Cisco Systems, Inc. All rights reserved.

Getting to 5 Nines

Roadblocks to 5 Nines (99.999%)

- High probability of redundancy failure (failure not detected— redundancy not implemented)
- High probability of double failures
- Long convergence time for rerouting traffic around a failed trunk or router in the core
- Rely on manual operations

NSF/SSO: Deployment Strategies



SP Core → HA Features May Provide Some Benefit

SP Distribution → Good Position for NSF-Aware Routers

SP Edge → Primary Deployment for SSO/NSF/NSR/ISSU Capable Routers

Enterprise Access → Deployment of SSO/NSF Capable or Aware Routers

Enterprise Core → HA Features May Provide Some Benefit

Enterprise Distribution → NSF-Aware Routers

Enterprise Access → Enterprise HA Toolbox

REFERENCES

© 2007 Cisco Systems, Inc. All rights reserved.

Reference Materials

CCO: http://www.cisco.com/go/availability

- High Availability White Papers

  http://www.cisco.com/en/US/partner/tech/tk869/tk769/tech_white_papers_list.html

Good

- Cisco Non-stop Forwarding with Stateful Switchover Deployment Guide:

  http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper0900aecd801dc5e2.shtml

- MPLS High Availability:

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122s/122snwft/release/122s25/fshaov.htm

- IP Event Dampening:

  http://www.cisco.com/en/US/products/sw/iosswrel/ps1829/products_feature_guide09186a00806994c7.html

- Bidirectional Forwarding Detection:

  http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122limit/122sx/12218sxe/fs_bfd.htm

- In Service Software Upgrade (ISSU)

  http://www.cisco.com/en/US/products/ps7149/products_ios_protocol_group_home.html

High Availability Reports

–Yankee Report, 2/2004: „The Road to a Five-Nines Network“

–Gartner Report 2001: „Survive in a 24 hours world“

–Infonetics Report 2/2003: „The Costs of Enterprise Downtime“

–Meta Group Report 4/2004: „ Comprehensive View of HA Data Center Networking”

© 2007 Cisco Systems, Inc. All rights reserved.

Associated Sessions (1/2)

IP Routing

- Advances in BGP (BRKIPM-3005: Wedney 08:30, Thursday 15:30)
- Advances in OSPF (BRKIPM-3006: Thursday 08:30, Friday 13:30)
- Advances in EIGRP (BRKIPM-3008: Thursday 08:30)
- IGP, BGP and PIM Fast Convergence (BRKIPM-3004: Wednesday 15:30)
- IP Fast ReRoute Technologies (BRKIPM-3017: Friday 08:30)
- IP Routing Design and Deployment Techtorial (TECIPM-3003)

MPLS Technology

- MPLS Techtorial (TECIPM-300x)
- MPLS Architectures for Enterprise Networks (BRKIPM-2013: Wednesday 15:30)
- MPLS Security in Service Provider Networks (BRKIPM-3012: Thursday 13:30)
- Advanced MPLS Deployment in Enterprise Networks (BRKIPM-3014: Friday 13:30)
- Layer 2 VPNs and Pseudo Wire (BRKIPM-3002: Thursday 08:30, Friday 08:30)
- Advanced Topics and Future Directions in MPLS (BRKIPM-3003: Thursday 15:30, Friday 08:30)

© 2007 Cisco Systems, Inc. All rights reserved.

Associated Sessions (2/2)

QoS Technology

- QoS Decomposed: The Components of the QoS Toolkit **(BRKIPM-2010: Thursday 1330)**
- End-to-end QOS Design: Deploying IP and MPLS QoS for Multiservice Networks **(BRKIPM-3009: Wednesday 08:30, Friday 13:30)**

Security:

- Network Core Infrastructure Protection **(BRKSEC-2013)**
- Network-Based Intrusion Prevention Systems **(BRKSEC-2009)**
- Detecting and Mitigating Denial of Service Attack **(DRKSEC-2014)**
- Detecting Router Abuse **(BRKSEC-2015)**

Network Management

- Operating MPLS Networks and Services **(TECNMS-2001)**
- Protecting the SP network against attacks **(TECNMS-2003)**
- Advanced Network Performance measurements /w IP SLA **(BRKNMS-3004)**

Recommended Reading

- **High Availability Network Fundamentals**
  By Chris Oggerino (2001)
  ISBN: 1587130173

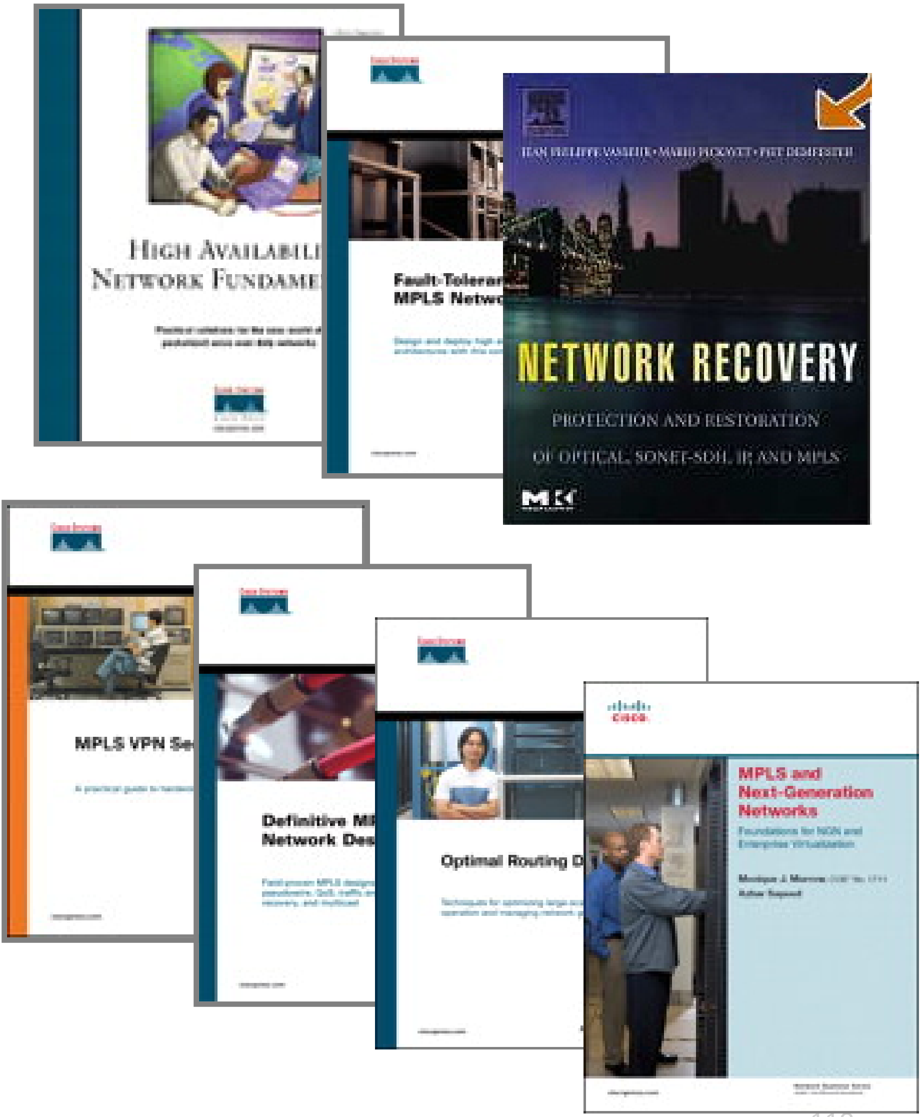
Fundamentals

- **Network Recovery**
  By Vasseur, Picavet, Demeester (2004)
  ISBN:012715051X
- **Fault-Tolerant IP and MPLS Networks**
  by Iftekhar Hussain (2005)
  ISBN: 1587051265

Overview

- **MPLS VPN Security**
  by M. Behringer/M. Morrow (2005)
  ISBN: 1587051834
- **Definitive MPLS Network Designs**
  by J.Guichard, F. LeFaucheur, J.-P. Vasseur (2005)
  ISBN: 1587051869
- **Optimal Routing Design**
  by Russ White, Alvaro Retana, Don Slice (2005)
  ISBN: 1587051877
- **MPLS and Next-Generation Networks: Foundations for NGN and Enterprise Virtualization**
  By A. Sayeed, M. Morrow (2006)
  ISBN-10: 1-58720-120-8

New


High Availability Network Fundame
Fault-Toleran MPLS Netwo
NETWORK RECOVERY
PROTECTION AND RESTORATION OF OPTICAL, SONET-SDH, IP, AND MPLS
MPLS VPN Se
Definitive MP Network Des
Optimal Routing D
MPLS and Next-Generation Networks
Foundations for NGN and Enterprise Virtualization

© 2007 Cisco Systems, Inc. All rights reserved.

SUMMARY

Summary: Building Highly Available IP and MPLS Networks

- **High Availability Fundamentals**

  how to calculate and estimate System/Network Availability

- **System Level Resiliency**

  Stateful Switchover (SSO) basic infrastructure for NSF, NSR, ISSU
  Non-Stop Forwarding (NSF) with Zero Packet Loss
  Non-Stop Routing (NSR) for PE-CE Deployments
  Warm Reload / Warm Upgrade for Single Processor Systems
  In-Service Software Upgrade (procedure)

- **Network Level Resiliency**

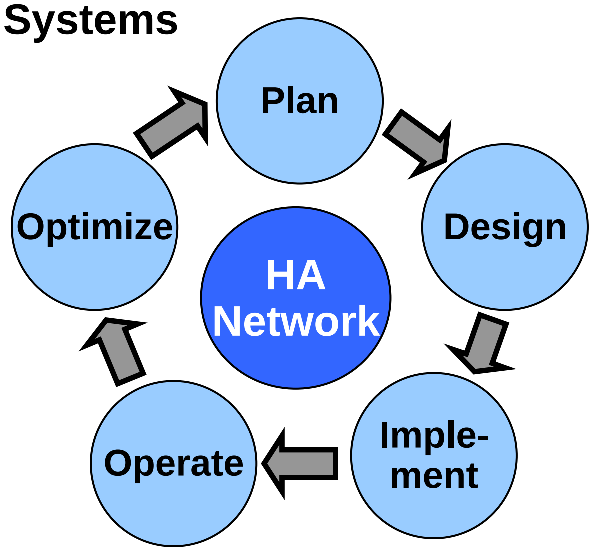
  IP Event Dampening to isolate unstable links
  BFD as single mechanism for liveness detection
  Fast Convergence and Fast Rerouting

- **Embedded Management**

  MPLS Diagnostics Expert (MPLS OAM)
  Generic Online Diagnostics (GOLD)
  Embedded Event Manager (EEM)

- **High Availability Best Practises**

  High Availability Network Life Cycle (PDIOO)
  Trouble-Ticket Availability Measures Method (Cisco NAIS Service)
  Roadblocks to 4 and 5 Nines



HA = Strategy

Recommended Reading

BRKIPM - 3011

- Continue your Networkers learning experience with further reading from Cisco Press.
- Visit the on-site Cisco company store, where the full range of Cisco Press books is available for you to browse.

Meet the Experts
IP and MPLS Infrastructure Evolution

- Andy Kessler
  Technical Leader

- Beau Williamson
  Consulting Engineer

- Benoit Lourdelet
  IP services Product manager

- Bertrand Duvivier
  Consulting Systems Engineer

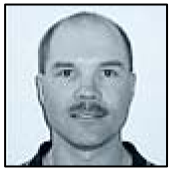
- Bruce Davie
  Cisco Fellow

- Bruce Pinsky
  Distinguished Support Engineer

© 2007 Cisco Systems, Inc. All rights reserved.

Meet the Experts
IP and MPLS Infrastructure Evolution

- Gunter Van de Velde
  Technical Leader
- John Evans
  Distinguished Systems Engineer
- Oliver Boehmer
  Network Consulting Engineer
- Patrice Bellagamba
  Consulting Engineer
- Shannon McFarland
  Technical Leader



© 2007 Cisco Systems, Inc. All rights reserved.

Meet the Experts
IP and MPLS Infrastructure Evolution

- Andres Gasson
  Consulting Systems Engineer

- Steve Simlo
  Consulting Engineer

- Toerless Eckert
  Technical Leader

- Dino Farinacci
  Cisco Fellow & Senior Software Engineer

CISCO