



Cisco Expo
2007

Построение IP и MPLS сетей высокой надежности



Антон Кириченко

ankirich@cisco.com

**Enable Your Network
Empower Your Business**

Что такое цена простоя ?

Потеря производства

- Кол-во пострадавших сотрудников
x
время простоя
x
ежечасное количество накладных расходов

Потеря производства

- Потеря прямой прибыли
- Выплата компенсаций
- Потеря будущих прибылей
- Потеря прибыльности инвестиций

**Цена
простоя**

Потеря производства

- Пользователи
- Поставщики
- Финансовые структуры
- Бизнес партнеры
- Сотрудники

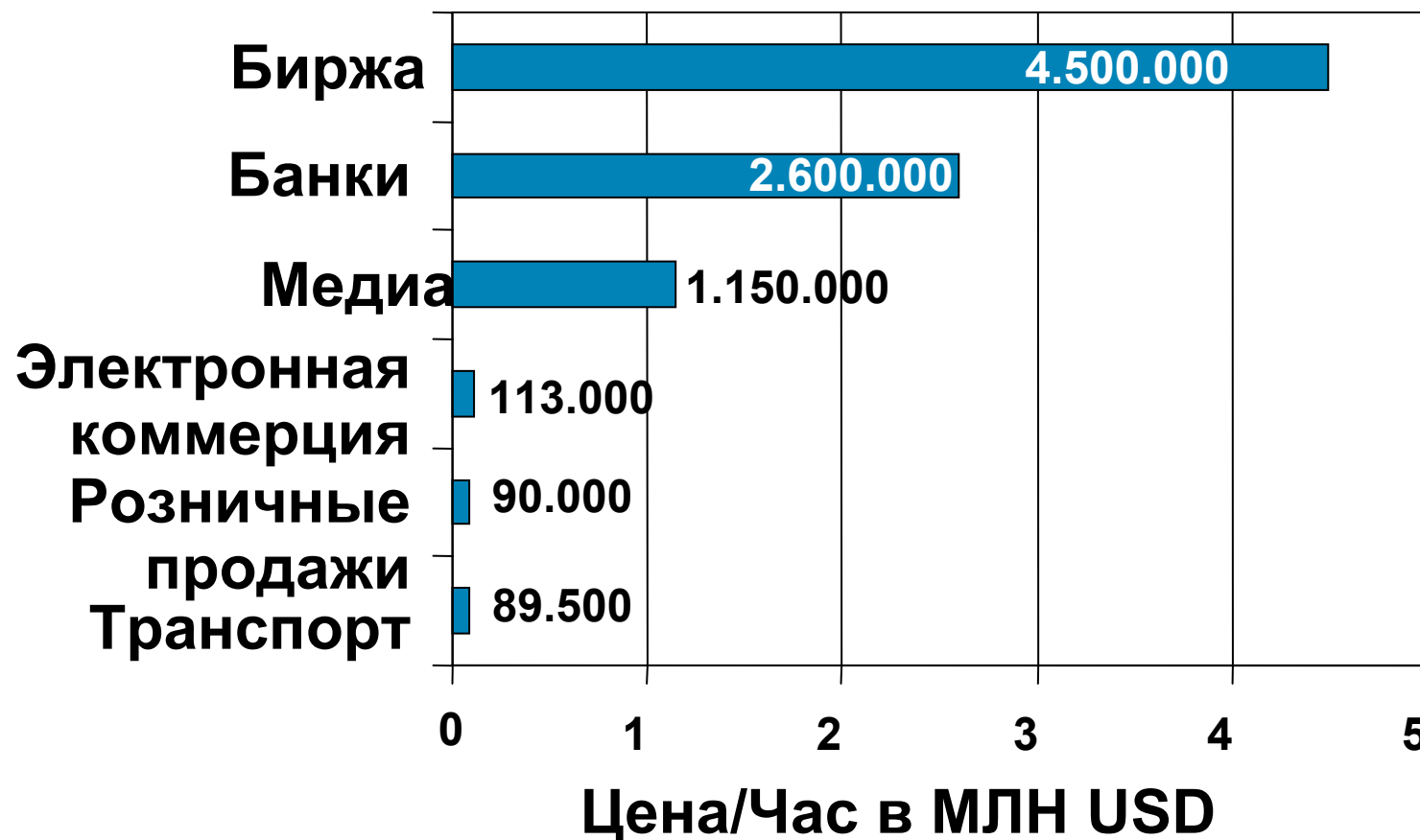
Потеря производства

- Потери денежных оборотов
- Гарантии погашения кредитных
- Кредитный рейтинг
- Реакция стокового рынка

Потеря производства

- Временный персонал
- Аренда оборудования
- Сверхурочные выплаты
- Затраты на командировки

Цена простоя сети в час



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

Содержание

- **Основы HA:**
 - MTBF, MTTR, Вычисленная и Измерения доступности
- **Надежность на уровне системы**
 - SSO, NSF, NSR, Warm Reload/Upgrade, ISSU
- **Надежность на уровне сети**
 - IP Event Dampening, BFD
 - Fast Convergence, Fast Rerouting
- **Ссылки**
- **Итоги**

Основы High Availability



Определение доступности

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

- **Mean Time Between Failure (MTBF)**
Как часто бывают сбои?
- **Mean Time To Repair (MTTR)**
Как долго их устранять?



Расчет надежности комплексных систем

- **MTBF -> Вычисляемое значение**

Вычислено Cisco на базе индустриальных стандартов

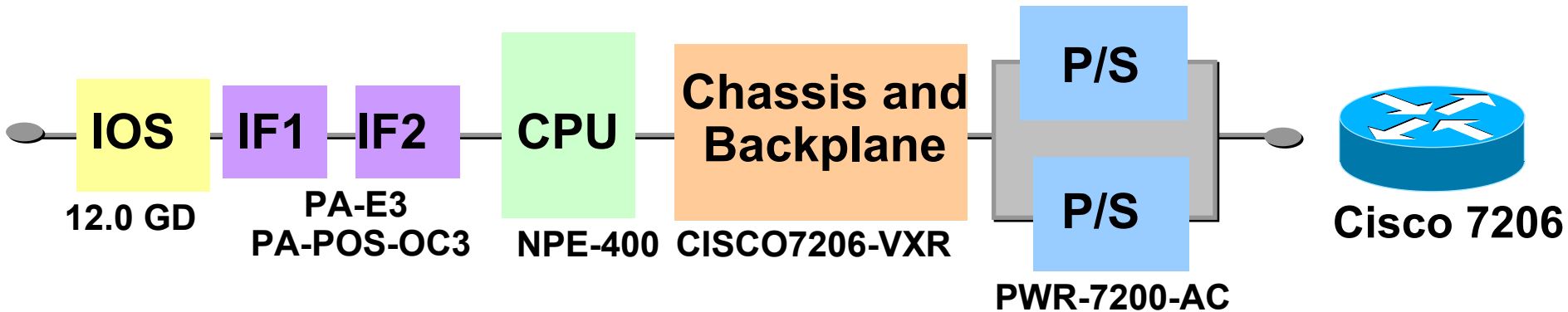
- **MTTR -> Ожидаемое значение**

Принимается как разумно ожидаемое (Reboot, exchange chassis/LC)

$$A_{\text{Series}} = \prod_{k=1}^N A_k = A_1 \times A_2 \times \dots \times A_N$$

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

Расчет надежности ус-ва: Cisco 7206



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

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

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

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

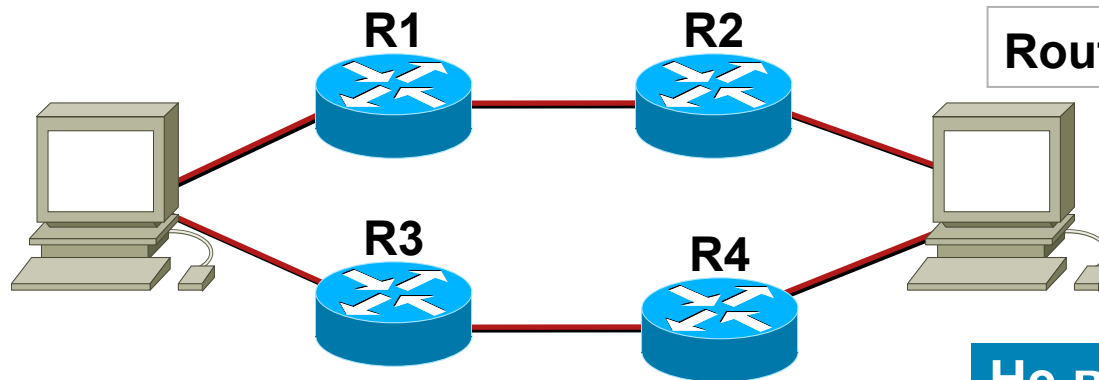
$$\text{BB} = \frac{460.000}{460.000 + 8} = 0.999983$$

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

$$\begin{aligned} \text{System Availability} &= 0.999997 * \dots * 0.999983 * (1 - (1 - 0.999995)^2) \\ &= 0.999961 = 99.9961\% \end{aligned}$$

Calculated MTBF Values from Cisco Database

Расчет надежности сети



Router R1,R2,R3, R4: 0.999961

1
Доступность R1 и R2 установленных последовательно
 $= (0.999961 * 0.999961) = 0.99992175$

Не взято в рассмотрение !!!:
-Links (WAN, LAN)
-Computer NICs
-Computer OS
-Computer Applications

2
Доступность параллельного участка
 $= 1 - ((1 - 0.999921)(1 - 0.999921)) = 0.999999994$

3
Сетевая надежность = 99.9999%
Рассчитано только на основе данных по надежности C7200 !!!

Фокус Cisco на HA: End-to-End

Надежность на уровне систем

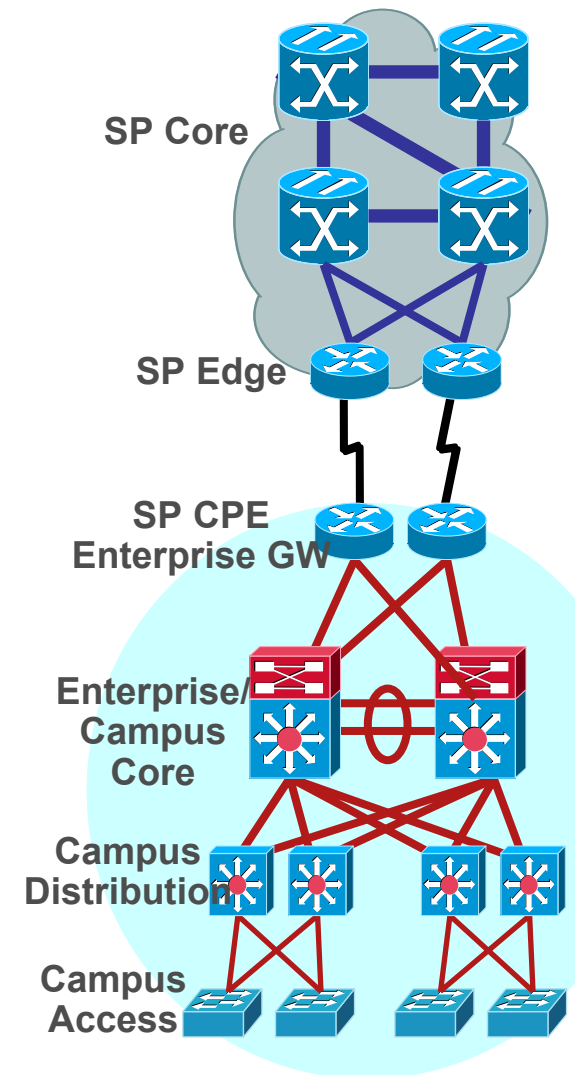
- увеличение MTBF
- уменьшение MTTR

Надежность на уровне сети

- **Fast rerouting** в ядре где присутствуют резервные маршруты
- Добавление новых ф-ий **fast network convergence, protection & restoration**

Встроенное управление и автоматизация процессов мониторинга

- **embedded management**
- **intelligent event management**



Набор инструментов Cisco HA

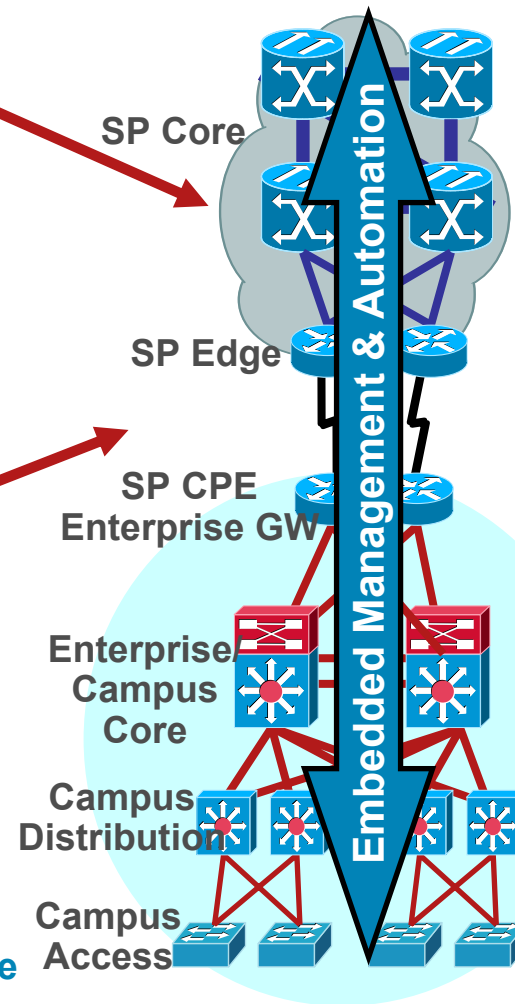
Надежность на уровне сети

- Поддержка NSF
- IP Event Dampening
- Bi-Directional Forwarding Detection (BFD)
- **Fast Convergence**
 - BGP Convergence Optimization
 - iSPF Optimization (OSPF, IS-IS)
 - Multicast Subsecond Convergence

Fast Rerouting (IP and MPLS)

Надежность на уровне устройства

- Надежность control/data plane:
 - HSA, RPR, RPR+, Stateful NAT/IPSec/FW,**
 - NSF /w SSO including MPLS**
 - Control Plane Policing, GLBP, HSRP,
 - Warm Reload**
- Процессов замены ПО : **ISSU, Warm Upgrade**
- Надежность линков:
 - Line Card Redundancy with Y-Cable
 - Link Bundling (Etherchannel/POS-Channel)



Embedded Management & Automation:

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

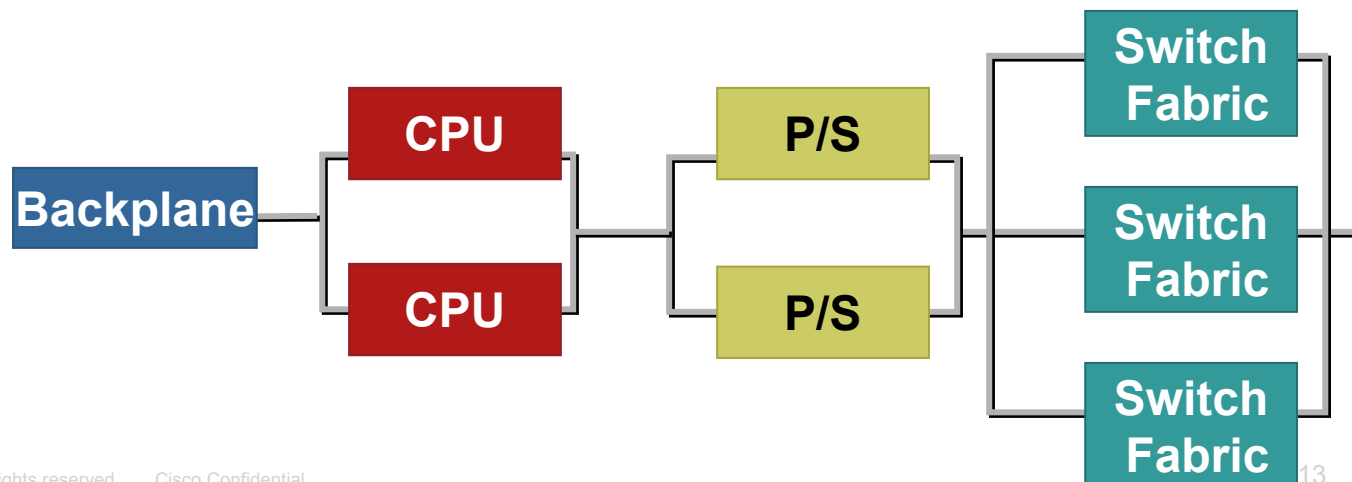
**Надежность на
системном
уровне:**

**Stateful
Switchover (SSO)**

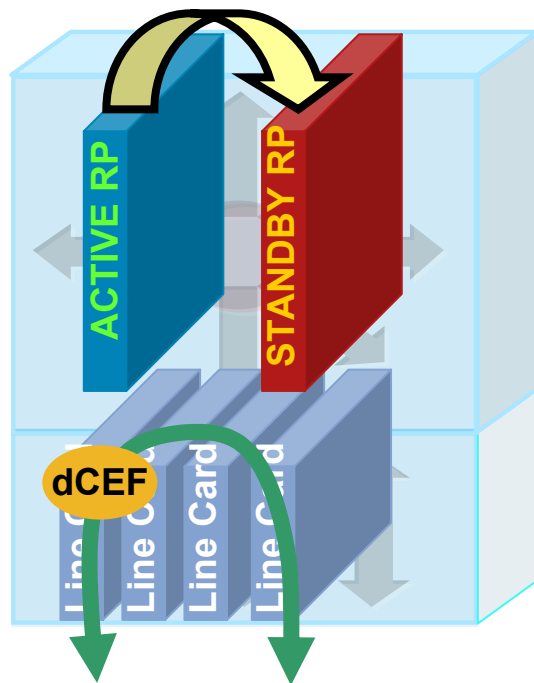


Компоненты аппаратной надежности

- Надежная аппаратная составляющая
- Распределение нагрузки
- **Active/standby резервирование** (processor, power, fans, line-cards)
- **Active/standby определение сбоя**
- MTBF карт (100,000 hrs)
- **Разделение control и forwarding plane**
- **Горячая замена (OIR)**



Восстановление систем с двумя RP



- **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)

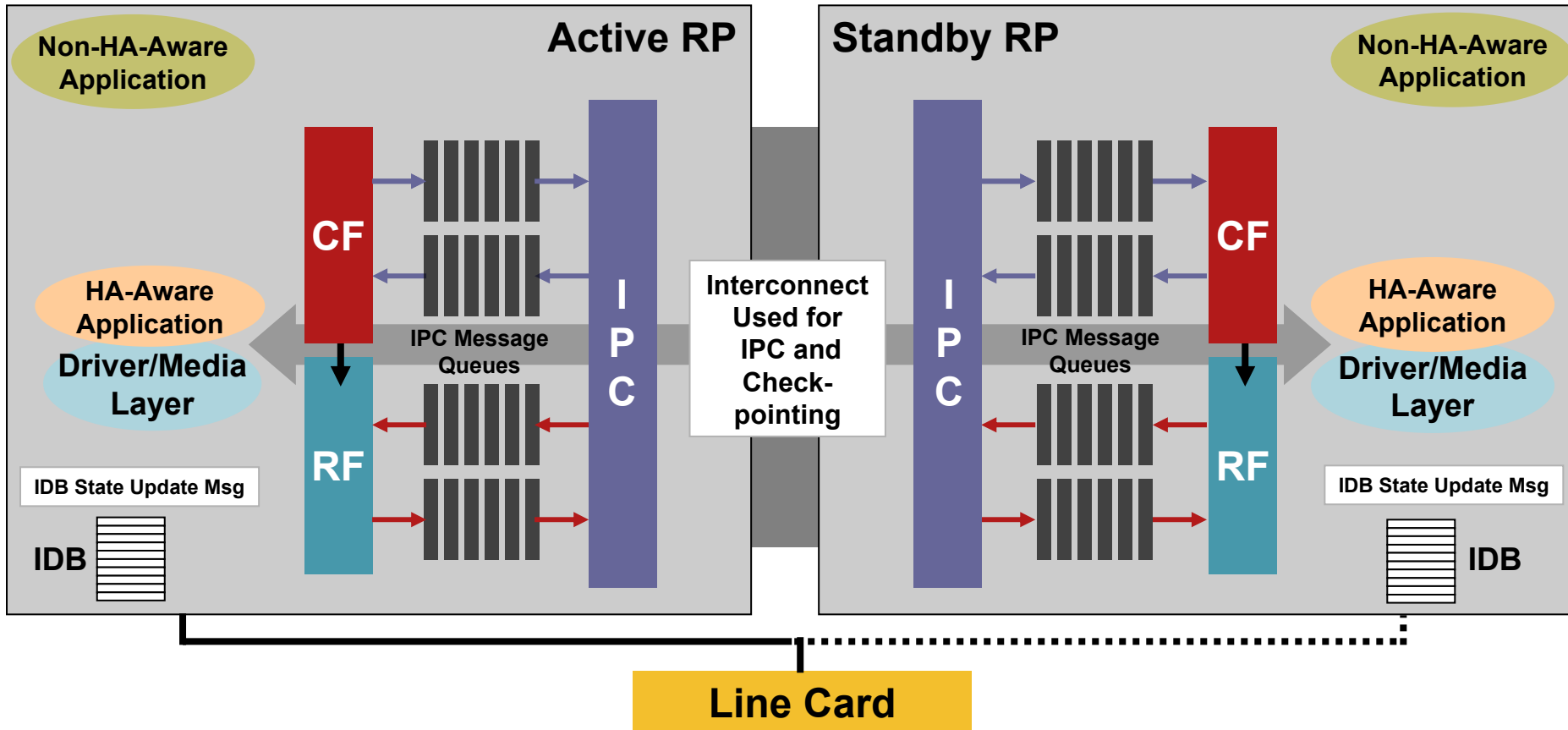
- **Hot Redundancy (2004)**

 - **GR / NSF / NSR w/ SSO** Graceful Restart / Non-Stop Forwarding/Routing with Stateful Switchover

- **In-Service-Software-Upgrade: ISSU (2005)**

- **Standby RP** получает контроль после программного или аппаратного сбоя основного RP
- **SSO** позволяет резервному RP сразу получать состояние и статус контрольных протоколов
- **NSF** возможность продолжать передачу пакетов до тех пор пока не сойдется таблица маршрутизации, требуется взаимодействие с другими маршрутизаторами
- **NSR** работает в режиме SSO для синхронизации роутинговой информации между активным и резервным процессором
- **GR** (graceful restart) переустанавливает роутинговую информацию без перестроения в сети
- **Цели** : добиться **0% потери пакетов**

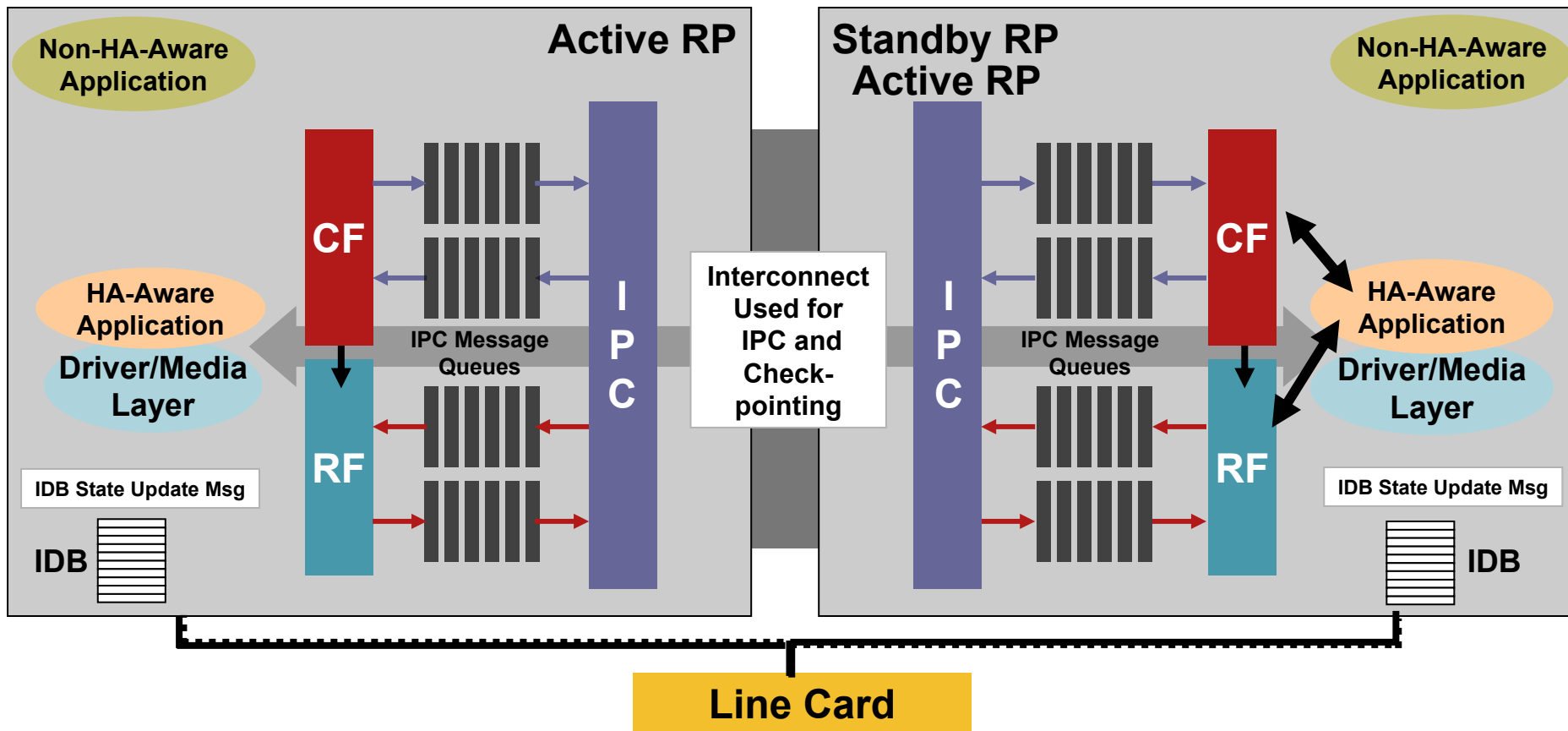
Подсистемы SSO



- **RF (Redundancy Facility):** мониторинг и отчет о состоянии RP
- **CF (Checkpointing Facility):** позволяет клиентам-приложениям пересылать апдейты с Активного на Резервный RP
- **IPC (Inter-Process Communication):** используется для передачи CF, RF и Config Sync сообщений
- **Driver/Media layer:** код для поддержки состояния IDB
- **Config Sync:** поддержка одинаковых конфигураций на обоих RP

Архитектура SSO для Stateful L2 протоколов (PPP, FR, HDLC ...)

animated



- Сбой на активном RP через RF сообщения инициализирует переключение
- L2 информация поддерживается в процессе переключения используя сообщения
- Line Cards переключаются на новый RP
- Соседние устройства “не видят” сбоев на интерфейсах интерфейсов

Протоколы и приложения поддерживающие SSO

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

Для полного списка рекомендуется
ознакомиться в RN для версий ПО и платформ

Активация и мониторинг SSO

```
RouterA(config)# redundancy
```

```
RouterA(red-config)# mode ?
```

```
rpr      Route Processor Redundancy
```

```
rpr-plus Route Processor Redundancy Plus
```

```
ssu     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
```

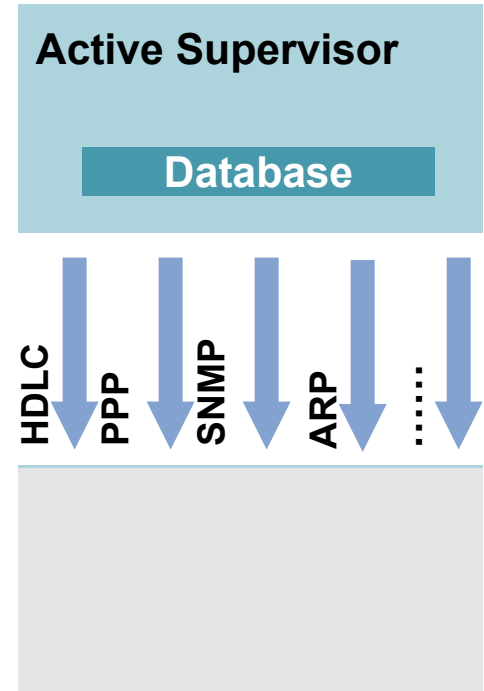
```
Redundancy Mode = SSO
```

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**

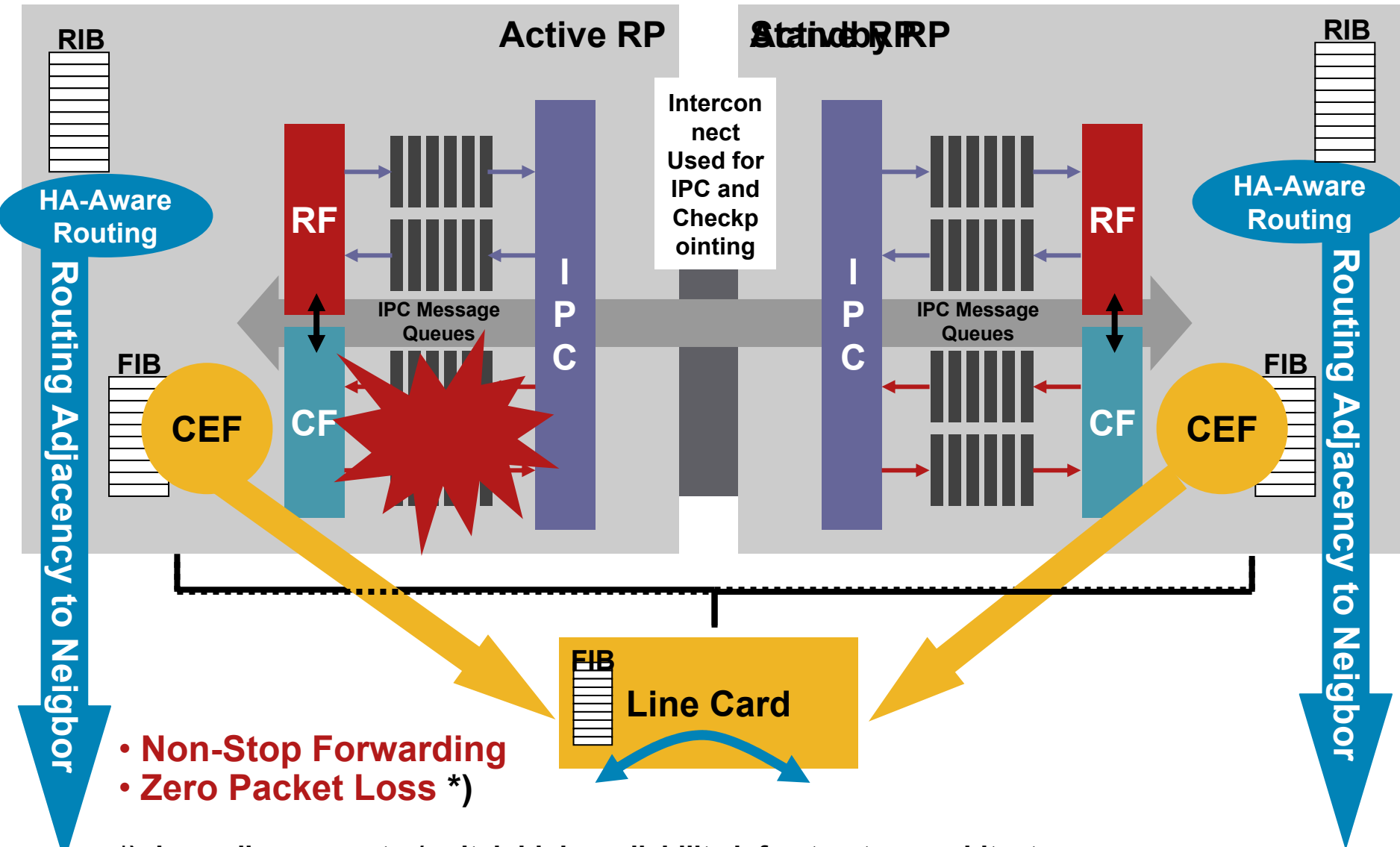
Index	Prev Active	Curr Active	Swact Reason	Swact Time
1	1	0	unsupported	8:03:52 UTC Thu Nov 29 2003
2	0	1	unsupported	08:07:00 UTC Thu Nov 29 2003

**Надежность на
системном
уровне:**

**Non-Stop
Forwarding (NSF)**



Архитектура NSF с HA-aware маршрутизацией



- Non-Stop Forwarding
- Zero Packet Loss *)

*) depending on router/switch high availability infrastructure architecture

Требования и расширения для NSF-aware роутинговых протоколов

Требования:

- **Переключение ДОЛЖНО быть закончено до срабатывания dead/hold таймеров**
 - Иначе будет перемаршрутизация
- **FIB ДОЛЖЕН быть неизменным в процессе переключения**
 - Текущие маршруты маркируются как “stale” в процессе переключения
 - “Cleaned” после переключения
 - Routing loop или black hole могут появиться, если в течении переключения изменилась топология сети

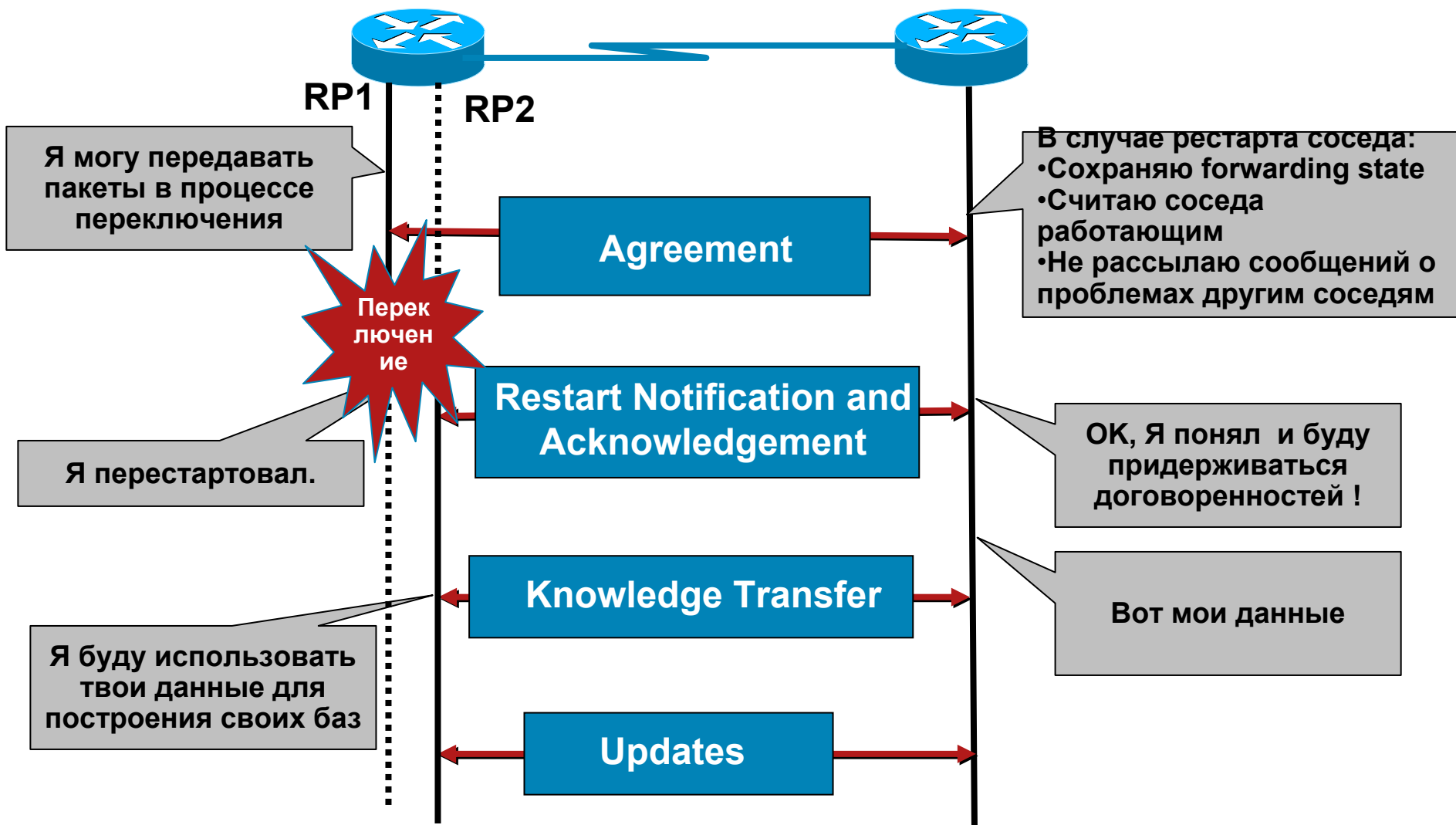
Расширение:

- Соседние маршрутизаторы должны знать что NSF маршрутизатор может продолжать посылать пакеты– NSF aware
- Расширения к ISIS, OSPF, EIGRP, BGP and LDP созданы для предотвращения route flapping

Установка соединения NSF-aware протоколов маршрутизации

NSF/SSO Capable Router

NSF Aware Peer



NSF расширения для IGP протоколов

Включение NSF для IGP:

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

Стандарты и драфты:

- Механизм для обеспечения бесперебойной передачи пакетов в процессе переключения **полностью не стандартизирован !**
- Cisco поддерживает два типа для OSPF: Cisco's **OSFP NSF** и IETF **Graceful OSPF Restart**
- Cisco's NSF реализация **ISIS (ietf option)** в соответствии с **RFC 3847 Restart Signaling**

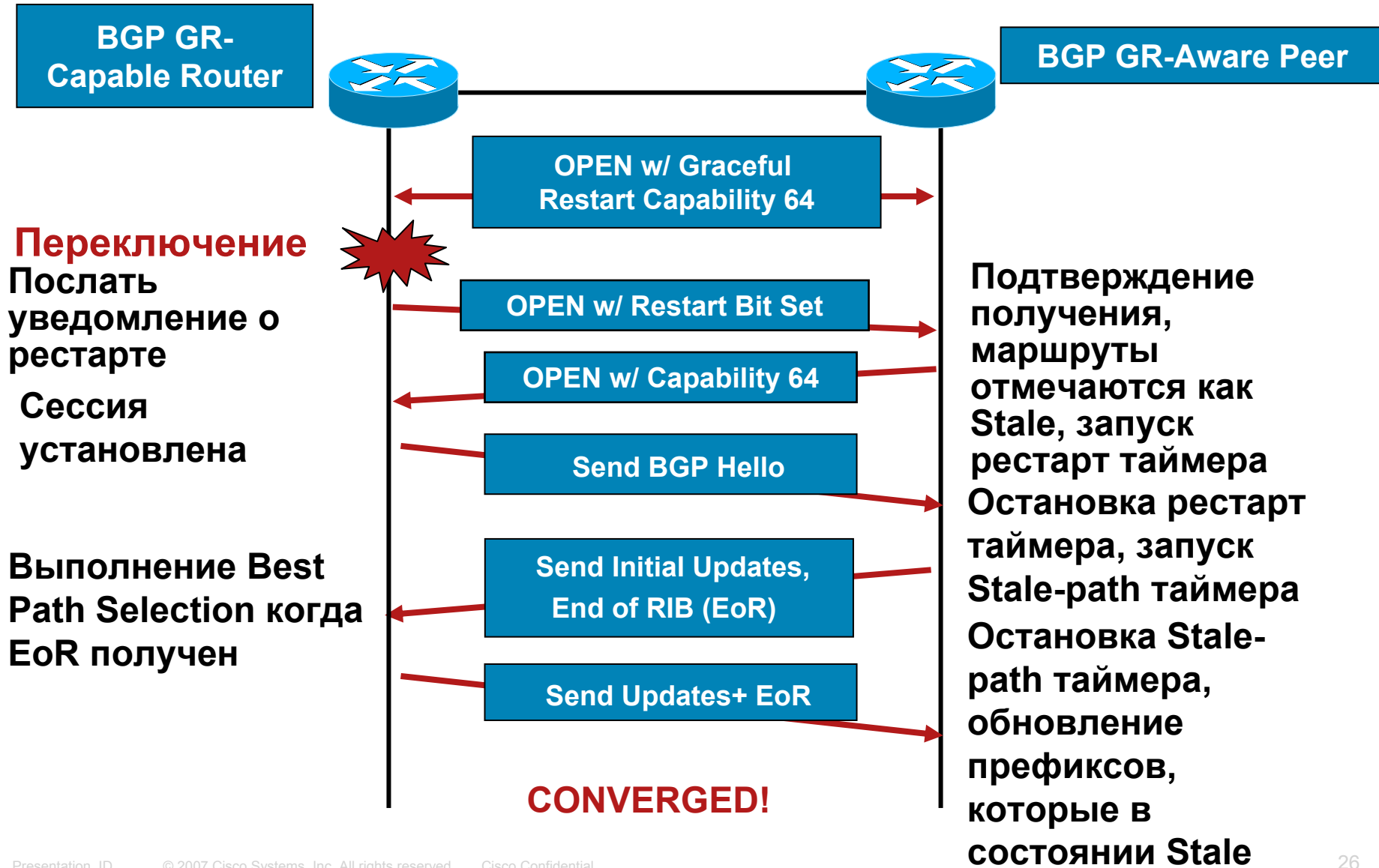
Подробности в NSF Deployment Guide:

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

BGP Graceful Restart

- IETF: **draft-ietf-idr-restart-13.txt**
 - submitted to IESG as proposed standard (expires 1/2007)
- Обеспечивает **graceful recovery механизм** для рестарта BGP процесса
- Реализация в Cisco IOS/XR:
 - 12.0S, 12.2T, 12.2S, XR 2.0
- Требуют GR-aware соседние маршрутизаторы
- GR поддерживаются в 7300, 7500, 7600, 12000, 10000, CRS-1

Graceful Restart BGP



Команда BGP Graceful Restart

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

- **Restart timers:** макс время на восстановление BGP сессии (def. 120s, adv. to peer)
- **Stalepath timers:** как долго можно использовать маршрут в состоянии Stall (def. 360s, used internally)
- **BGP hold time:** 180 seconds (3 x 60 sec keep-alive)

show ip bgp neighbor

```
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

Эволюция HA RP

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

MPLS and IP NSF/SSO-Coexistence

- **IP HA ускоряет восстановление MPLS**

 - No waiting for the route processor

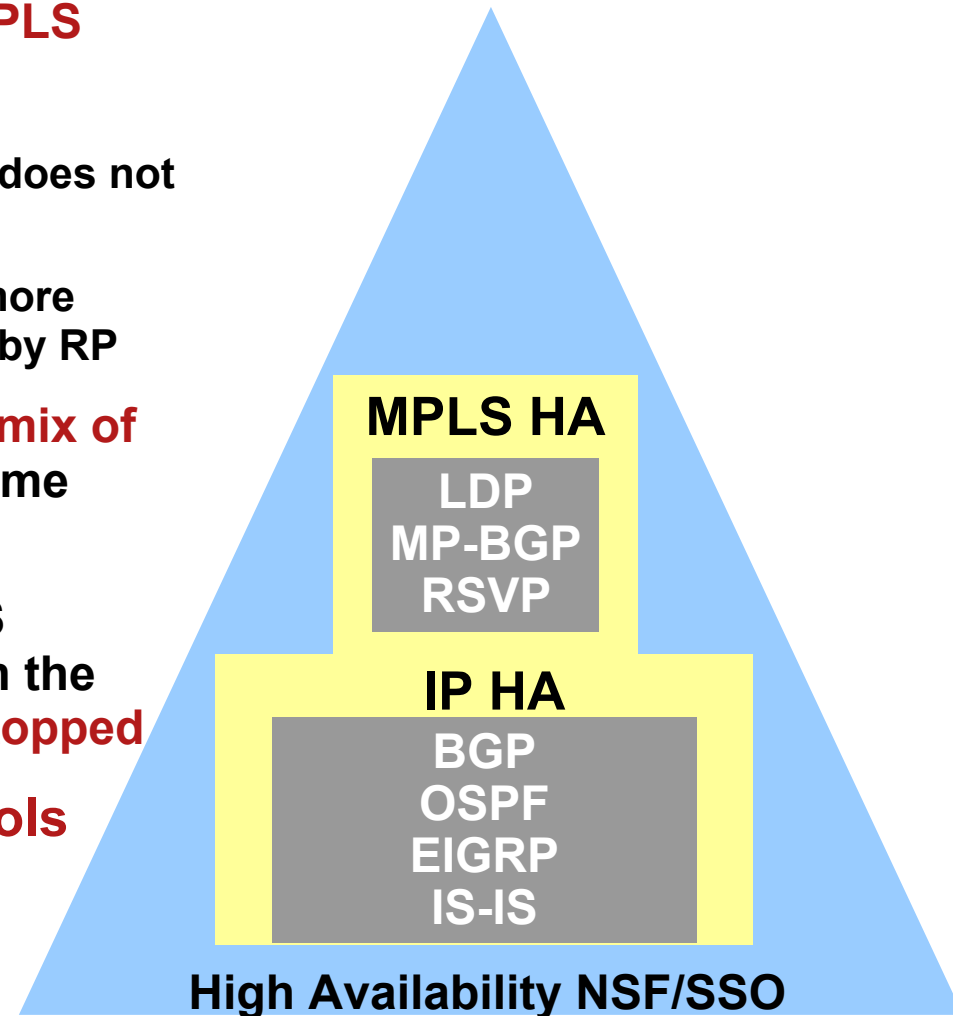
 - No loss of Layer 2 connectivity, so it does not need to be re-established

 - 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



Компоненты и ключевые элементы MPLS HA

- **MPLS HA—LDP NSF/SSO**

1. **Контроль передачи local label bindings** на резервный RP

Где: на устройстве с двумя процессорами

2. **LDP graceful restart**

Где: PE, RR и P роутеры

- **MPLS HA—BGP VPNv4 NSF/SSO**

Подробности:

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

Конфигурирование MPLS VPN HA

```
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
```

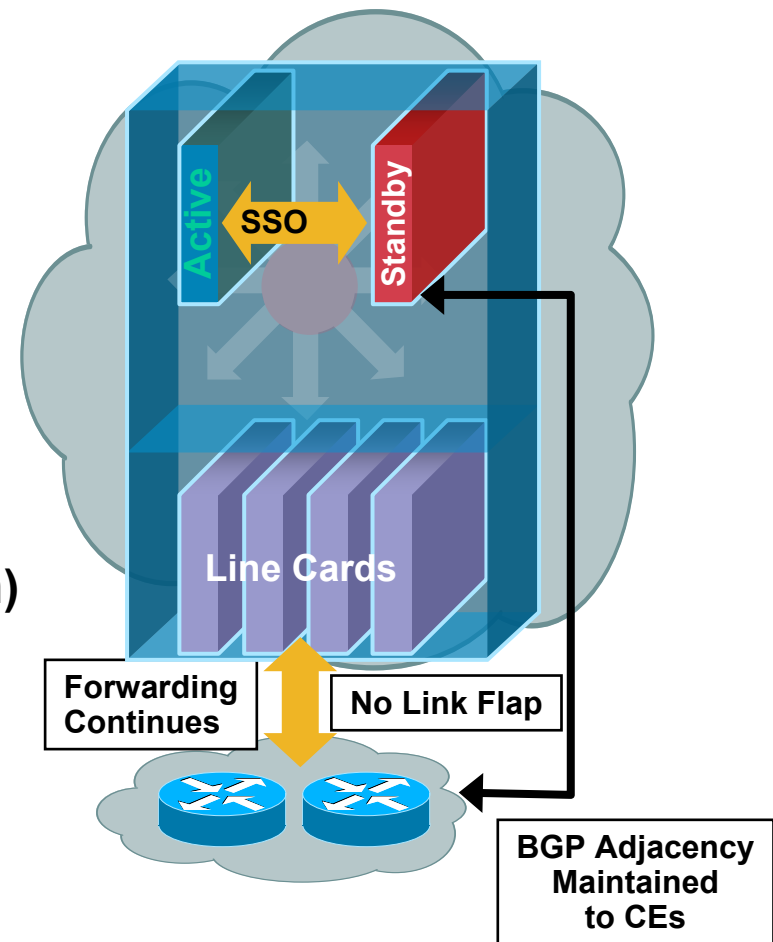
**Надежность на
системном
уровне:**

**Non-Stop Routing
(NSR)**



Non-Stop Routing (NSR)

- **NSR** и **NSF** это не одно и то же
- **NSR** вкратце
 - Передача трафика и сохранение роутинг в процессе переключения процессоров как dNSF
 - Не требует протокольных расширений
 - Не требует взаимодействия с соседними маршрутизаторами
- **IOS Support for NSR:**
 - **ISIS NSR** (stateful NSF!, Cisco Version)
 - **BGP NSR** (начиная с 12.2(28)SB for Cisco 10000 PRE2)
 - **LDP NSR** (IOS-XR в 2007)



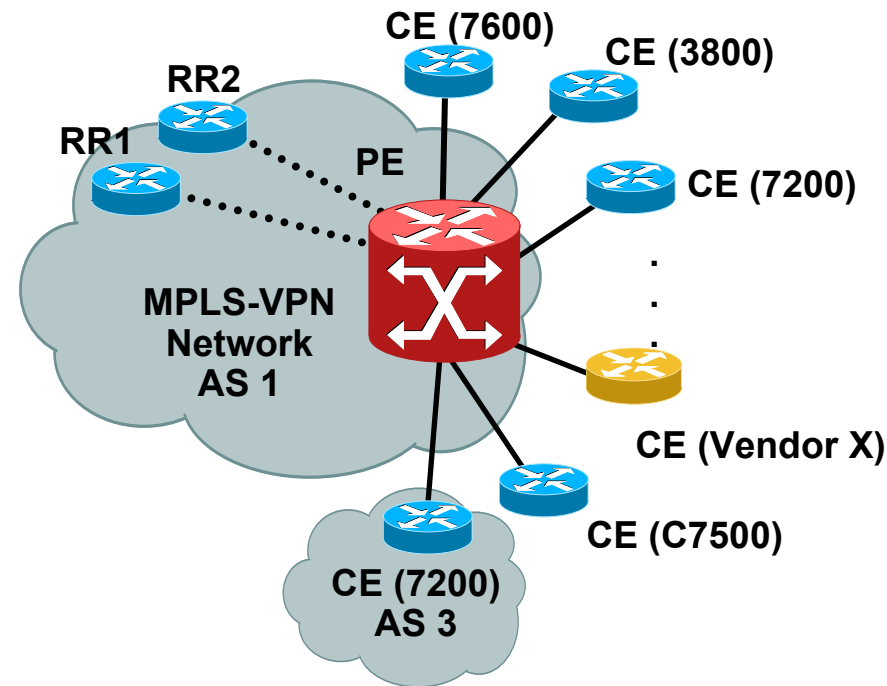
NSR – конфигурация PE

Configuration:

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

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
  !
```

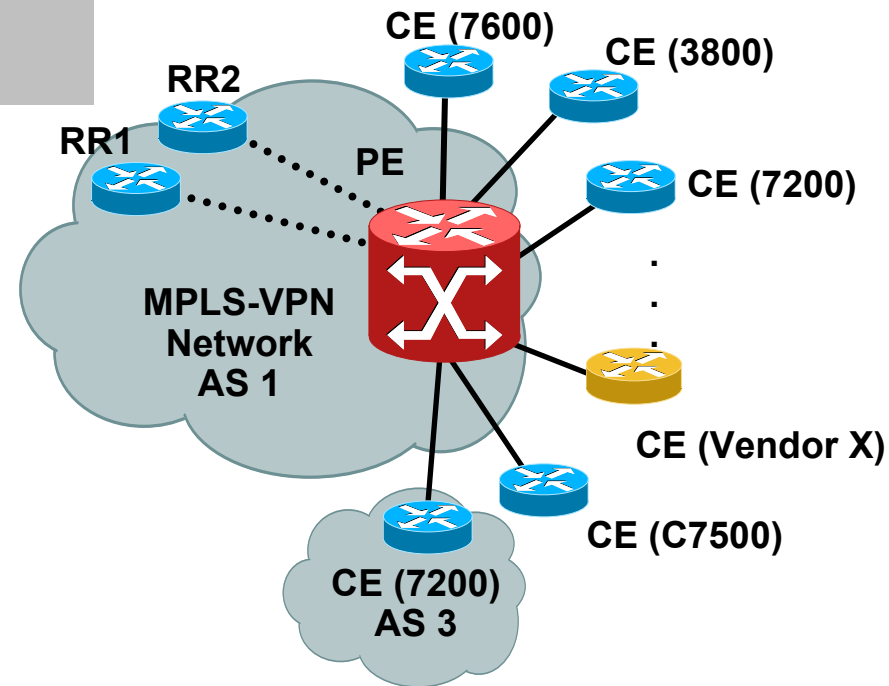


10.1.1.1 RR—Route Reflector
10.2.2.2 PE—Provider Edge
10.3.3.3 CE—Customer Edge

NSR – конфигурация CE

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

NOTE: Никакой поддержки расширений на стороне CE не надо!



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

Проверка поддержки BGP для NSR с SSO

```
Router# show ip bgp vpnv4 all sso summary  
  
stateful switchover support enabled for 40 neighbors
```

Показывает количество BGP соседей с поддержкой SSO

```
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)
```

показывает информацию BGP о VPN подтверждающую что SSO активирован или нет

Детали - BGP NSR Feature Guide:

http://www.cisco.com/en/US/products/ps6566/products_feature_guide09186a008067a607.html

Troubleshooting BGP NSR с SSO

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

Показывает **BGP-related SSO events**

Показывает отладочную информация о взаимодействии VFP между активным и резервным процессором

Используется для мониторинга и траблшуринга BGP в процессе переключения процессоров или ISSU

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

Показывает **TCP HA события** или отладочную информацию о взаимодействии TCP стека между активным и резервным RP

```
show tcp ha connections
```

Показывает **детали TCP соединений** которые поддерживают BGP NSR с SSO

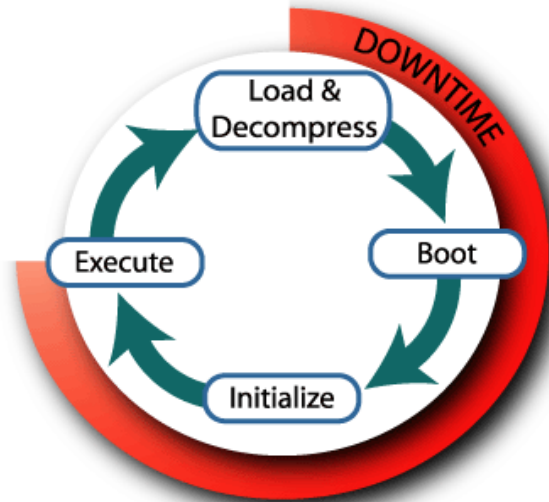
**Надежность на
системном
уровне:**

**Warm Reload /
Warm Upgrade**

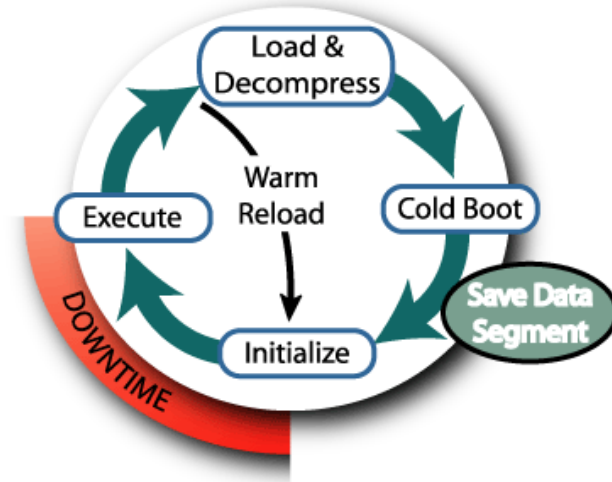


Cisco IOS Warm Reload

Загрузочный процесс
до Warm Reload



Warm Reload
Process



Значительно уменьшается время загрузки -> уменьшение
MTTR для сбоев ПО

Доступно и для однопроцессорных систем

Детали Warm Reload

- Заранее загруженный и декомпрессированный имидж
- Нужна дополнительная память для хранения компрессированных переменных для инициализации – примерно 1-2 МВ
- Полезно в случае сбоев ПО:
 - Критический сбой системы
 - Необходимость перезагрузки для восстановления работы
- Аппаратный сбой может привести к холодной перезагрузке
- Если маршрутизатор перегрузился в течении 5 минут с той же причиной, то будет холодная перезагрузка

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

count – максимальное количество warm перезагрузок (**default 5, value 1-50**)

uptime – минимально время между инициализацией системы и сбоем перед тем как будет предпринята попытка warm перезагрузки (**default value is 5 minutes**)

Улучшение показателей доступности систем с одним процессором

- NPE-G1 Setup

Обычная перегрузка: 3:43
минуты

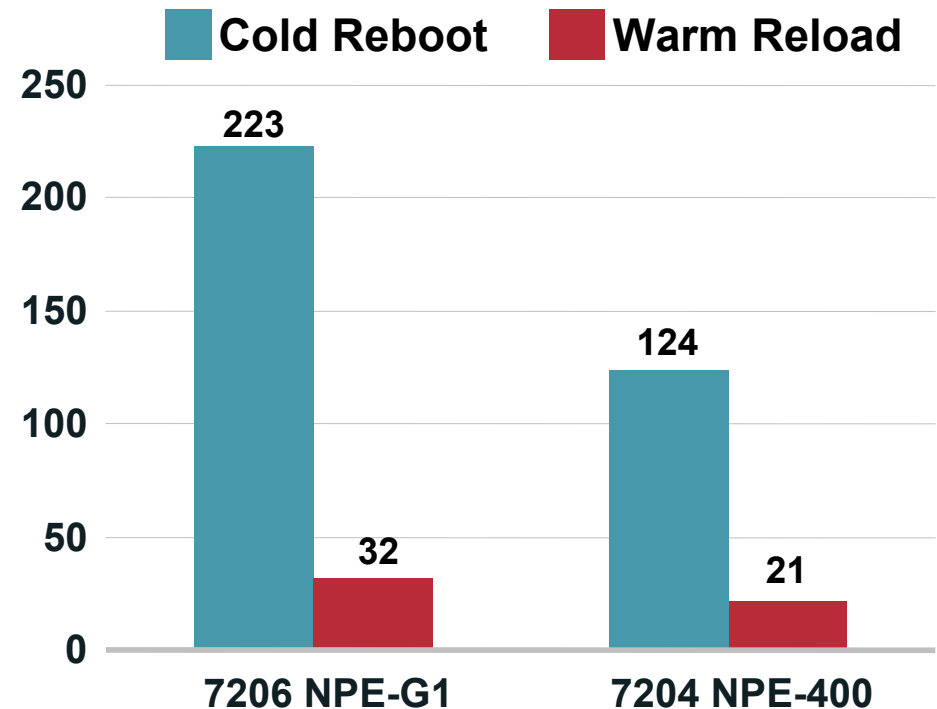
Warm reload: 0:32 минуты

Улучшение - 86%

- NPE-400 Setup

Обычная перегрузка: 2:04
минуты

Warm reload: 0:21 minutes

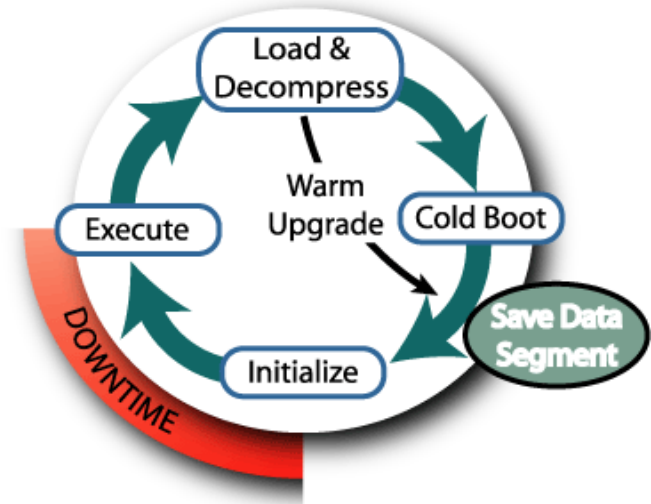


http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper0900aecd801778e8.shtm

Улучшение процесса апгрейда с использованием Warm reload

- Построен на базе **Warm Reload**
- Уменьшенное время для процессов апгрейда-даунгрейда
- Позволяет маршрутизатору загрузить и декомпрессировать ПО и затем передать управление, в то время как будет продолжаться передача трафика
- Если апгрейд неудачен, то текущая версия IOS будет работать далее

Warm Upgrade process

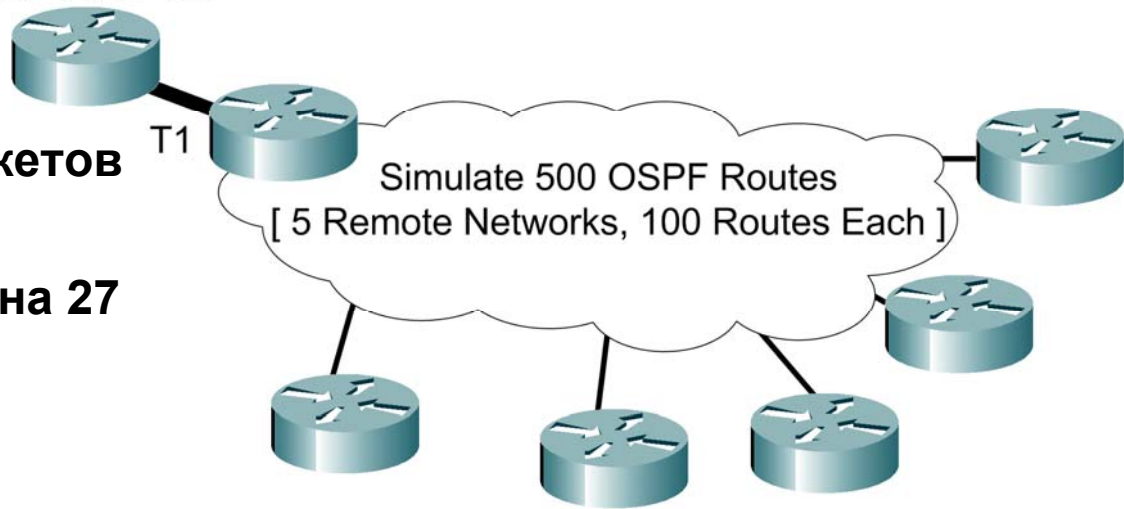


```
Router# reload warm file disk2:c7200-js-mz.122-18.S3
```

Warm Upgrade уменьшает время отсутствия сервиса

- **Меншее время потерь пакетов**
- **Время недоступности уменьшено как минимум на 27 секунд**
- **68 % времени меньше отсутствие сервиса**

Router Under Test
7206 - NPE-400



	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

**Надежность на
системном
уровне:**

**In-Service
Software Upgrade
(ISSU)**



Шаги Cisco IOS Full Image ISSU

animated



1. Подготовка к ISSU:
Копирование новой версии
Cisco IOS на оба RP

2. Загрузка
резервного RP
Standby RP
перезагружается
Теперь запущен с
новым IOS, в режиме
SSO

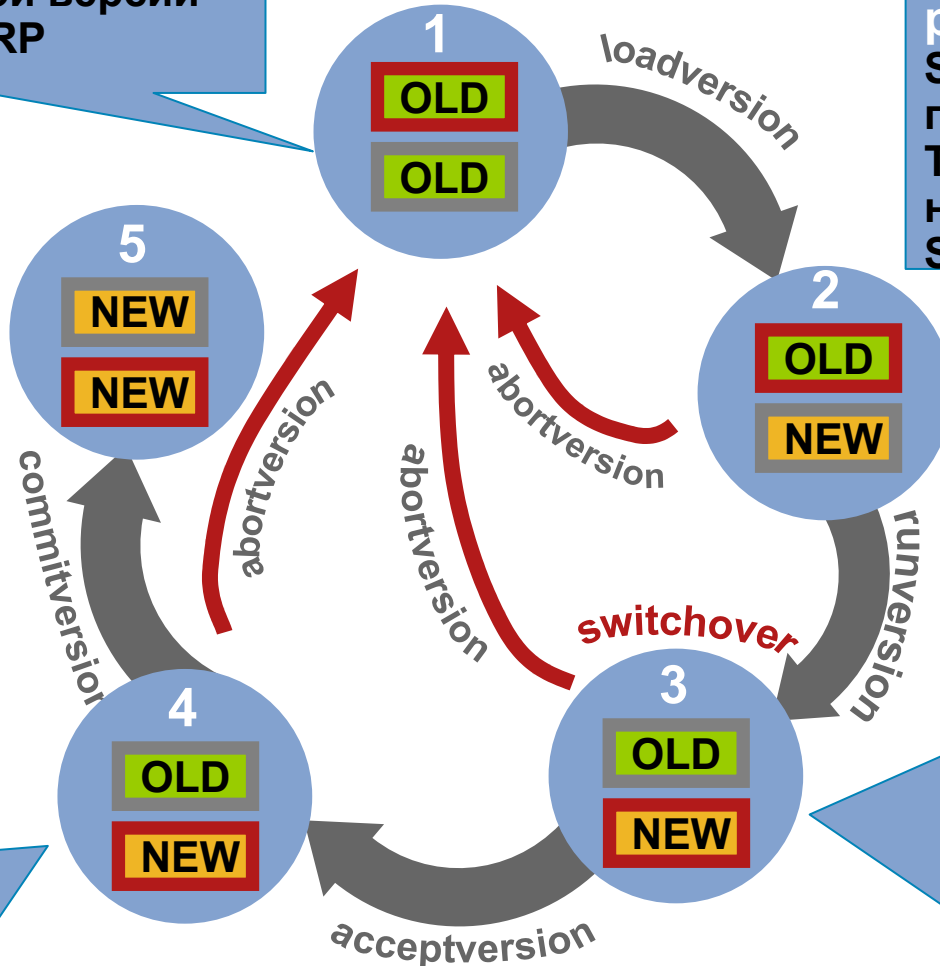
5. Завершение
процесса

Подтверждение
новой версии,
резервный
процессор
перезагружается
на новую версию

3. Переключение
RP и запуск
новой версии

Старый RP
перегружается и
становится
резервным со
старой версией

По прежнему в
SSO режиме



Команды ISSU для Full Software Upgrade

- **issu loadversion**

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

Optional
Parameter

- **issu runversion**

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

- **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
```

- **issu abortversion**

Show ISSU State Detail

After “issu runversion”

```
router#sh issu state det
```

```
Slot = B  
RP State = Active
```

Slot B Is Active

Bootvar Adjusted

```
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
```

New Version
“2”

```
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
```

Old Version
“1”

```
router# show issu rollback-timer
```

```
Rollback Process State = In progress
```

```
Configured Rollback Time = 45:00
```

```
Automatic Rollback Time = 29:03
```

Сетевая доступность



Сетевой уровень надежности

IP Event Dampening

Bi-Directional Forwarding Detection

Fast Convergence

iSPF Optimization (OSPF, IS-IS)

BGP Optimization

FC and NSF/SSO Coexistence

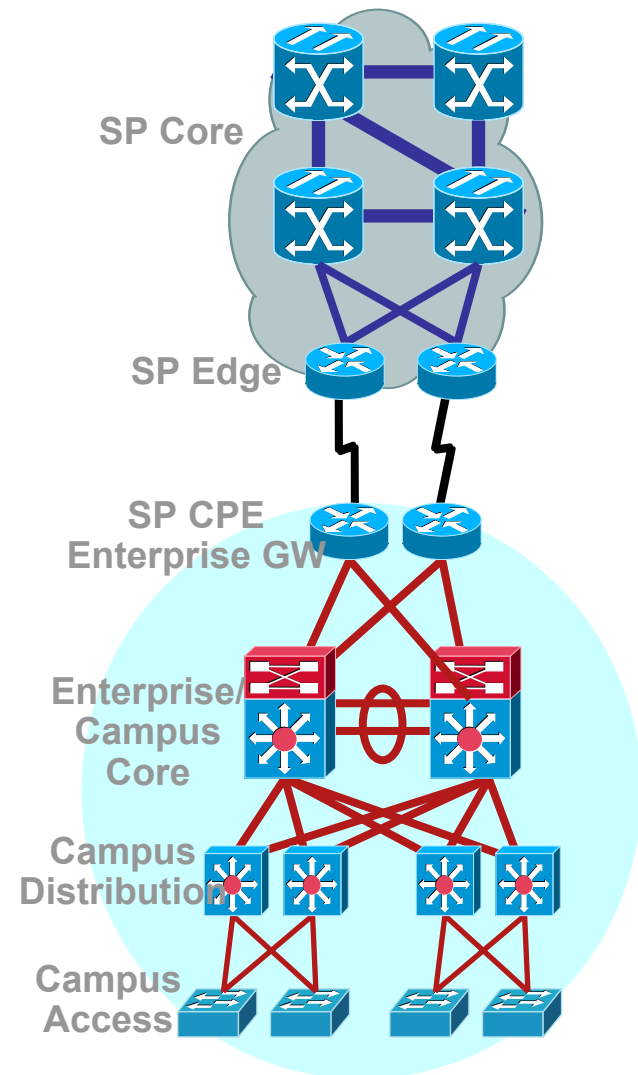
GR Shutdown

Fast ReRoute (FRR)

MPLS FRR

IP FRR

.....



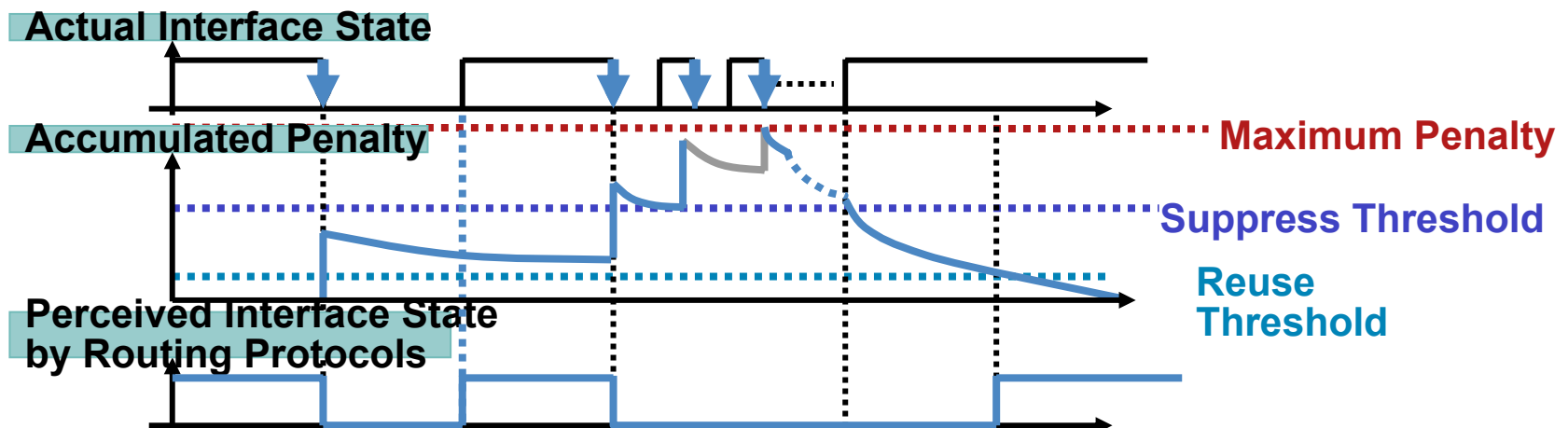
**Сетевая
доступность**

**IP EVENT
DAMPENING**

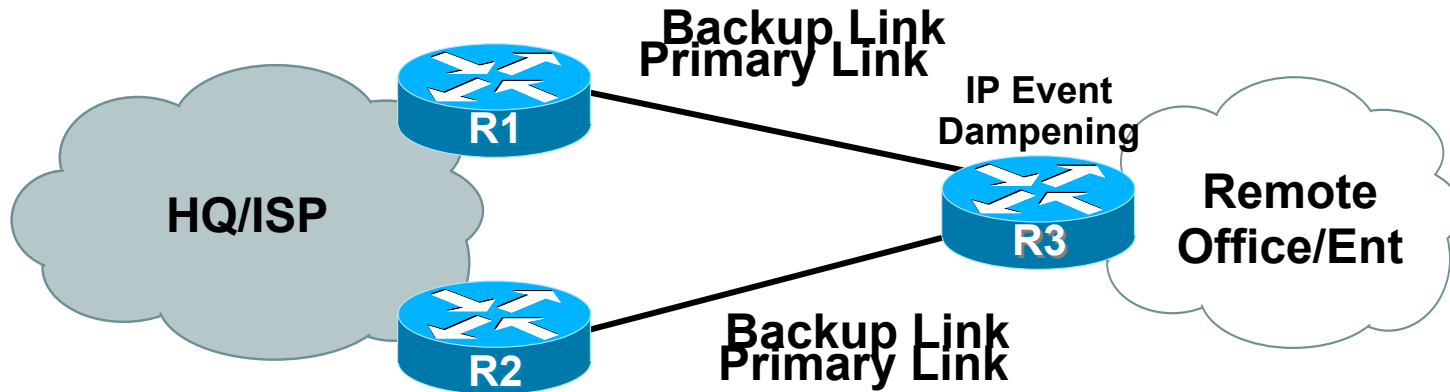


Основы IP Event Dampening

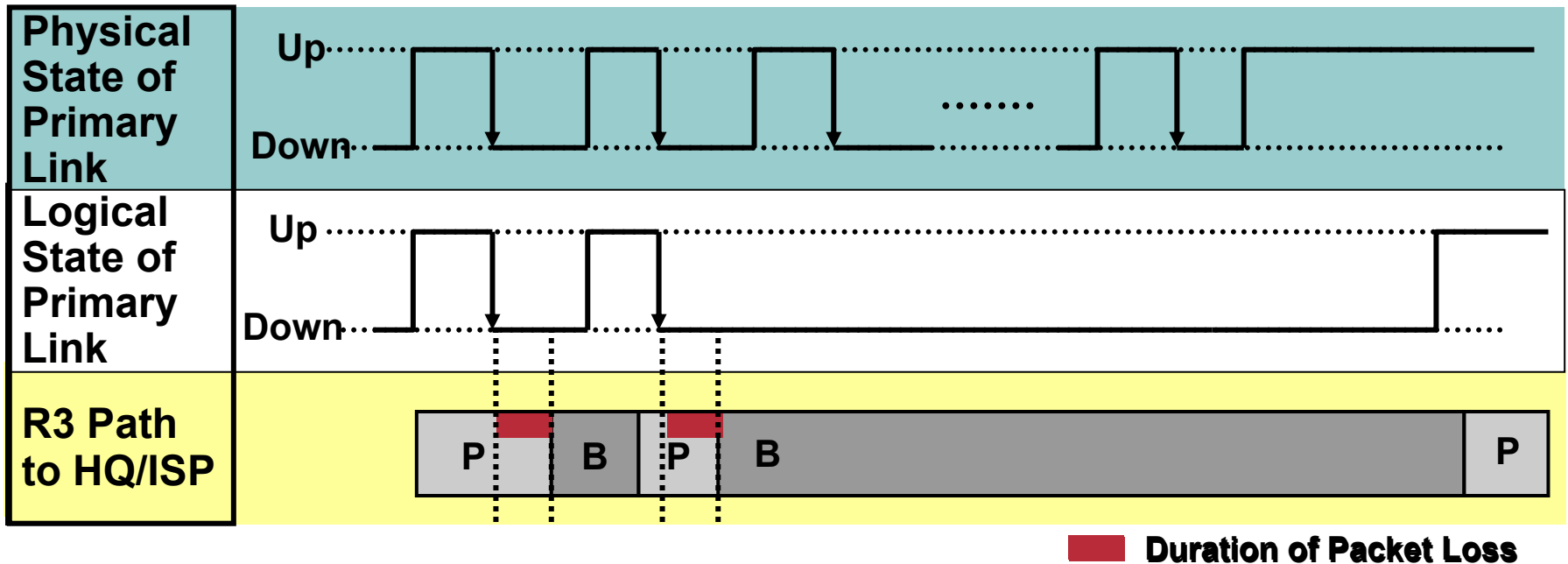
- IP Event Dampening **логически изолирует нестабильные линки**:
 - уменьшение потерь пакетов и нагрузки на CPU
 - Уменьшение повторяющихся сбоев в сети
 - Изоляция нестабильных элементов сети
- Общая концепция с **BGP route-flap dampening**
- Слежение за нестабильным интерфейсом и добавление penalty
- Перевод интерфейса в состояние **“down”** с точки зрения **протоколов маршрутизации при превышении критического значения penalty**
- Использование экспоненциального алгоритма для уменьшения penalty и перевода интерфейса в состояние **“up”**



Применение IP Event Dampening



IP Event Dampening скрывает Link-Flapping эффеты от протоков роутинга



Конфигурирование IP Event Dampening

```
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.**

**Сетевая
доступность**

**Bi-Directional
Forwarding
Detection (BFD)**



Проблемы сходимости

Процесс сходимости

- **Определение сбоя**, Принятие решения, Восстановление

Определение сбоя наиболее проблематично:

- **Различающиеся методы определения потери соседнего роутера с точки зрения роутингового протокола**
- **Медленные таймеры по умолчанию**
- **Определение сбоя на layer2 зависит от среды и типа интерфейса**

Требуется единый стандартный метод для определения сбоя линка, устройства или протокола

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

Задачи:

Быстрое определение сбоев в обоих направлениях между интерфейсами и протоколами

Независимость

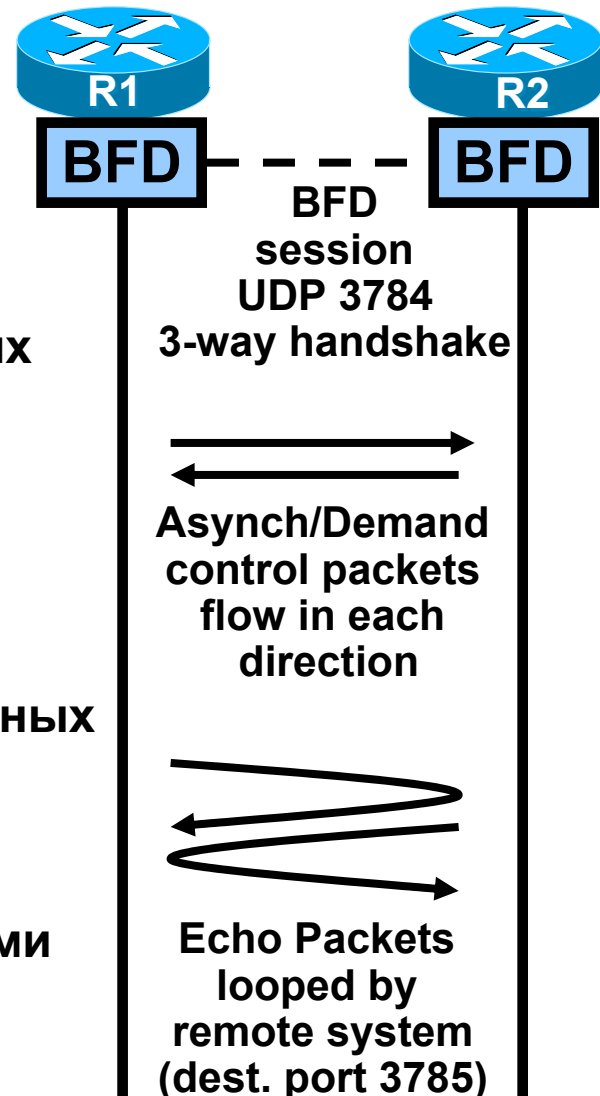
“Легкость” для CPU

Различные режимы

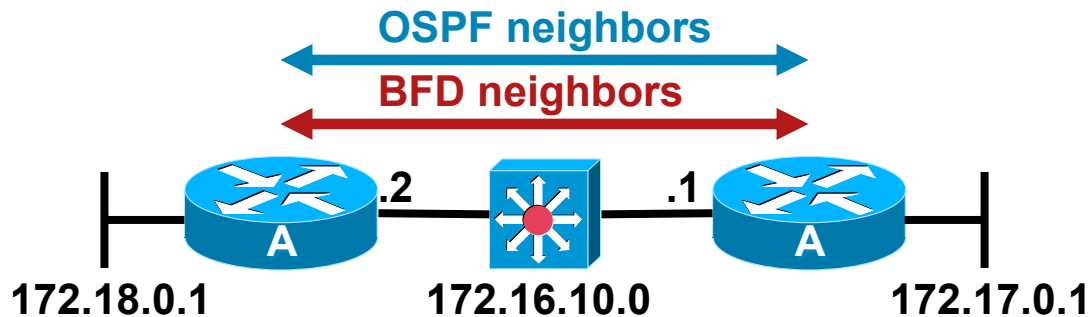
Асинхронный: периодическая передача контрольных BFD пакетов

По запросу: после установки BFD сессии контрольные пакеты посылаются по запросу. Используется с более простыми маршрутизаторами

Эхо функция: возврат эхо пакета через путь прохождения пакета (аппаратная реализация)



BFD и OSPF



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 по прежнему нужны для control plane проверки, обнаружения и т.д.
- OSPF процесс регистрирует соседние маршрутизаторы и интерфейсах с BFD
- BFD отслеживает работу системы передачи трафика (forwarding plane)
- Быстро извещает OSPF о падении BFD сессии

BFD: Поддержка, масштабирование и производительность

IOS/XR:

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

IOS-XR: 3.2

CPU платформы:

+2% CPU @ 100 BFD session

Distributed platforms: Cisco 12000

Нет нагрузки на CPU

+2% LC-CPU @ 100 session / LC
со 150ms задержкой

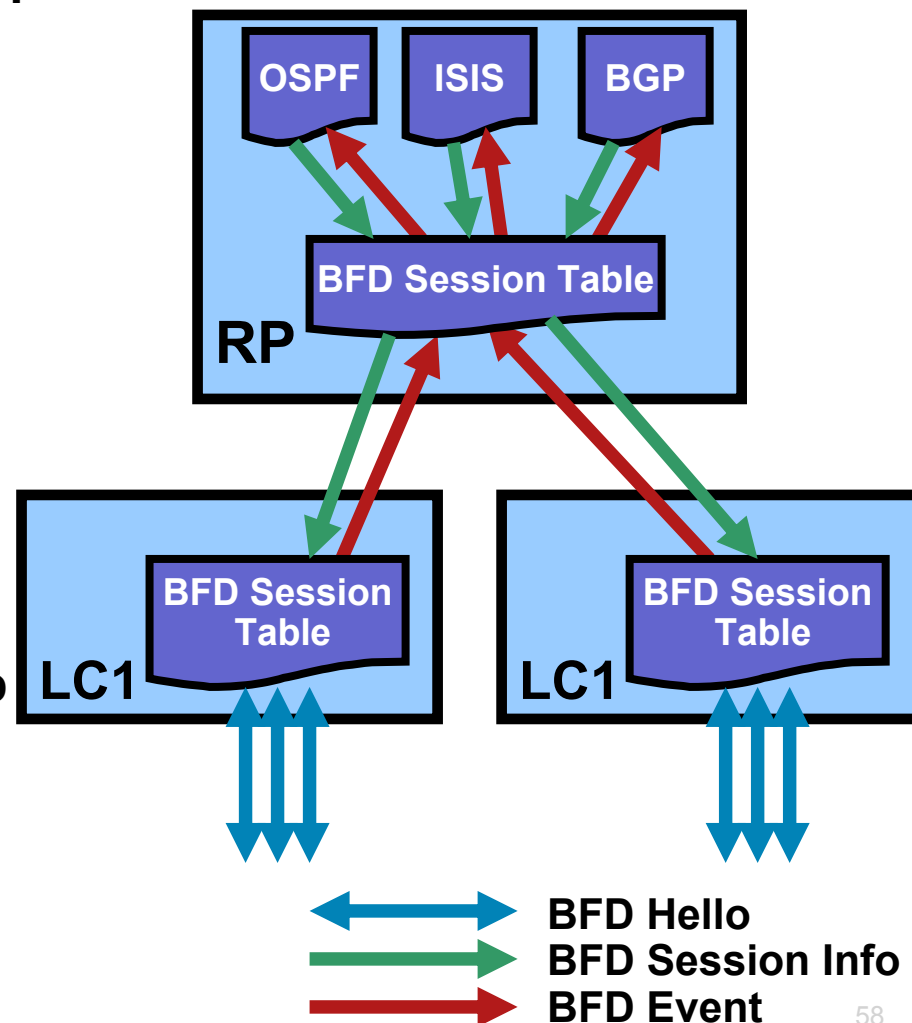
Поддержка протоколов:

ISIS, OSPF, EIGRP, BGP-SingleHop
(BGP-Multihop, LDP, IPv6)

Взаимодействие с другими средствами

IP Event Dampening

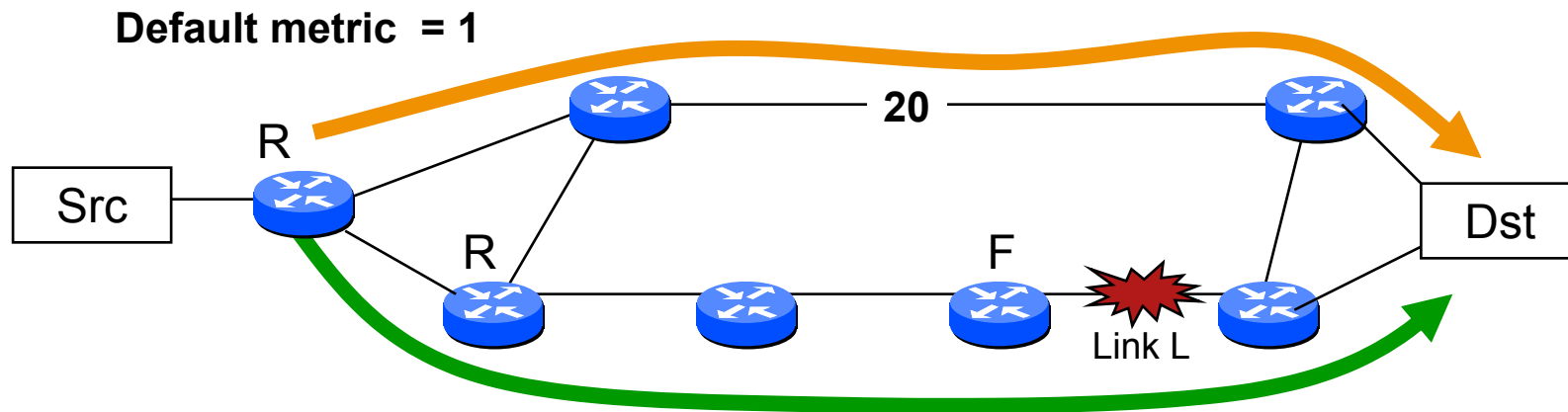
Cisco 12000 BFD Architecture



Сетевая доступность
FAST CONVERGENCE
and
FAST REROUTING



Задачи Fast Convergence



- Время потери коннективности: $T2 - T1$, называется “сходимость”
- Каковы должны быть значения сходимости?:
 - **Sub-Second**: требования для большинства сетей
 - **Sub-200ms**: отдельные приложения с требованием LoC $\leq 200ms$
 - **Sub-50ms**: Бизнес требования некоторых приложений

Fast Convergence в целом

- **NSF /w SSO**: сохраняет передачу трафика

Роутинговая информация восстанавливается независимо

- **Fast Convergence**: быстро перемаршрутизирует трафик на резервные пути

Скорейший путь обнаружения сбоя интерфейса: сигнализация POS
IS-IS < 10 msec

Более быстрое извещение о сбое в сети: opt. flooding

Ускоренный расчет новой топологии: **iSPF**

Исключение нестабильности: **exponential backoff timers**

- **Сценарий сбоя:**

GR/NSF покрывает только проблемы резервного RP

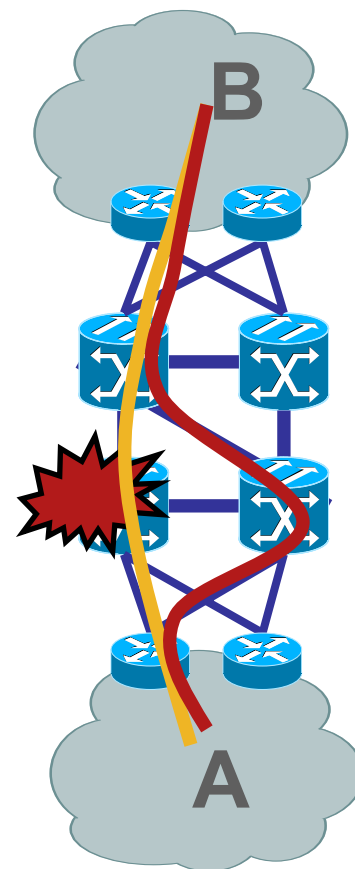
FC покрывает все сбои – узла, линка и т.д.

Могут ли NSF и IGP Fast Hello работать вместе?

- **NSF/SSO и FC имеют противоположные задачи:**
 - NSF: обеспечить передачу трафика через аварийный маршрутизатор
 - FC: быстрое переключение в обход сбойного узла
- **Сценарии реализации могут быть различными:**
 - NSF: SP edge
 - FC: IGP FC фокусировано на Core
- **NSF/SSO протестированные с различными IGP таймерами**
 - Стенд с 3 маршрутизаторами: Cisco 12000, 10000, 12000 / 12.0(22)S
 - ISIS с 5000 маршрутами: hello: 1 sec (multiplier: 3)
 - OSPF с 5000 маршрутами: hello: 2 sec (dead: 8 sec)
 - NSF/SSO продолжает работать нормально с такими таймерами
 - Первый hello посылается ~2 sec после переключения
 - Более консервативные настройки требуют >4sec

Детали:

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



Traffic flow with
— NSF/SSO
— Fast Converg.

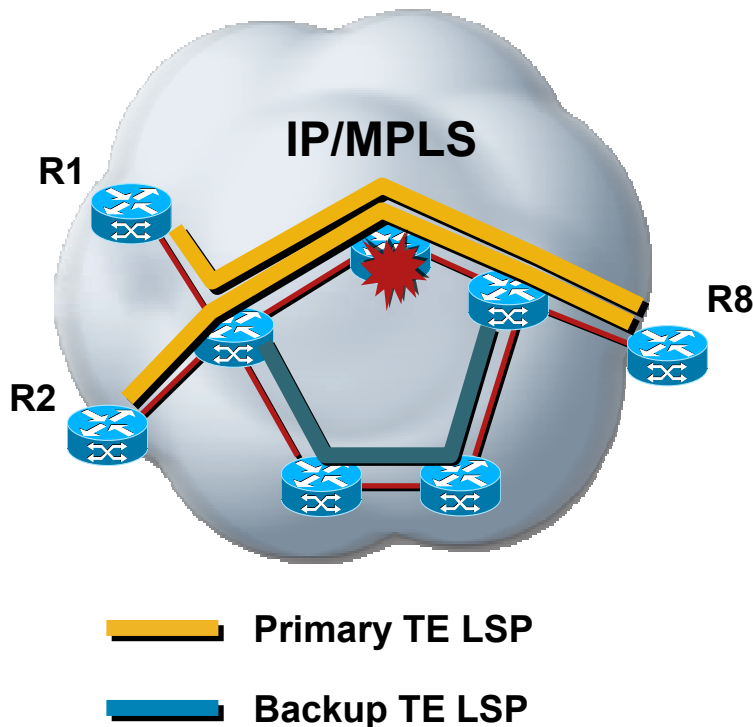
Текущий статус IGP FC

- **Link/node down event :**
 - Failure Detection (POS today, BFD emerging) < ~ 20ms
 - Origination < ~ 10ms
 - Queueing, Serialization, Propagation < 30ms
- **Распространение изменений в сети**
 - Flooding < $5 * 2ms = 10ms$
- **Перерасчет SPF**
 - SPF < $n * 40us$
- **Перезапись маршрутов в FIB**
 - FIB update: $p * 100us$
 - FIB Distribution Delay: 50ms
 - $\sim 100ms + p * 0.1 ms$
- **500 наиважнейших префиксов: ~ 150ms**
- **Sub-50ms impossible today -> need Fast Reroute**

MPLS Fast Re-Route (FRR)

Ключевые элементы:

- **Предварительно рассчитанный путь**
- **Локальное действие** (для избегания реакции по все сети)
- **Туннелирование** (для избегания реакции по все сети)



MPLS FRR:

- **fast recovery** для node/link сбоев
- Масштабируемое решение
- Более эффективно в плане стоимости чем защита на уровне оптики
- **Возможность защиты по полосе пропускания**

Ссылки



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

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

 http://www.cisco.com/en/US/tech/tk869/tk769/technologies_white_paper0900aec801dc5e2.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)**

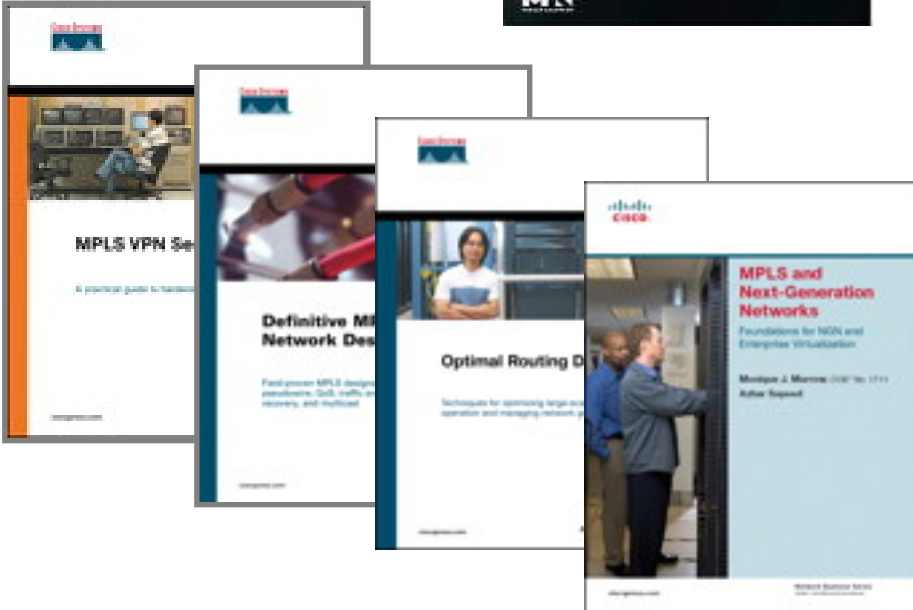
http://www.cisco.com/en/US/products/ps7149/products_ios_protocol_group_home.html

- High Availability Network Fundamentals**
 By Chris Oggerino (2001)
 ISBN: 1587130173
- Network Recovery**
 By Vasseur, Picavet, Demeester (2004)
 ISBN:012715051X
- Fault-Tolerant IP and MPLS Networks**
 by Iftekhar Hussain (2005)
 ISBN: 1587051265
- 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

Fundamentals

Overview

New



Вопросы ?



