

Unified MPLS

Александр Литвин

CSE

October 2012

Александр Высторопский

NCE

Почему Unified MPLS

Почему Unified MPLS

- На сегодняшний день пакетные транспортные технологии являются наиболее эффективным решением (технологии, цена, рост трафика...)
- Требование к транспортной инфраструктуре – максимальная утилизация оптической инфраструктуры, радио-релейной инфраструктуры – необходимость виртуализации транспортных сервисов
- Единая транспортная инфраструктура для mobile и wireline услуг
- Legacy TDM и ATM будет существовать на протяжении длительного периода времени
- MPLS выбран как транспортная технология – возможность виртуализации, поддержки TDM/ATM технологий, статистическое мультиплексирование
- Unified MPLS предоставляет масштабируемое решение

Тенденции в современных операторских сетях

- **Сети становятся больше**
 - Quad-play (Video, Voice, Data & Mobility)
 - Объединение сетей
 - Увеличение кол-ва узлов
- **Экспоненциальный рост потребления трафика**
 - Сервисы для бизнес абонентов
 - Мобильные сервисы
 - Услуги для домашних абонентов
- **MPLS на уровне сети доступа**
 - Seamless MPLS
 - MPLS-TP
- **Консолидация BGP ASN**

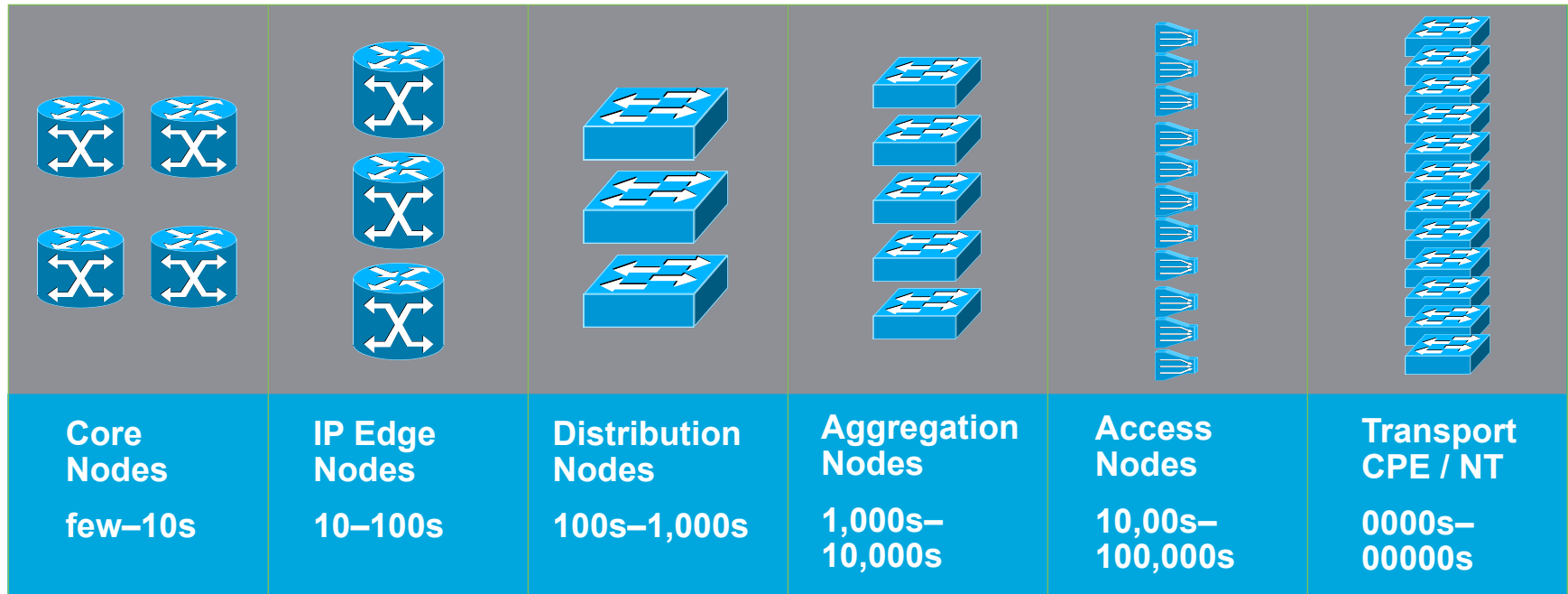
EXPERIENCE PROVIDER

Требования к современным сетям

- **Большая сетевая инфраструктура**
2000+ устройств в сети
- Возможность обеспечить подключение для обеспечения сервиса в любой точке сети
- **Наблюдаемость и управляемость End-to-End**
v4/v6 Uni/Multicast услуги
- Быстрая сходимость или **Восстановление**
Чем быстрее, тем лучше. 😊
- Масштабируемость и Производительность
- **Простота**

Any Play

Масштабируемость – от Ядра до Границы сети



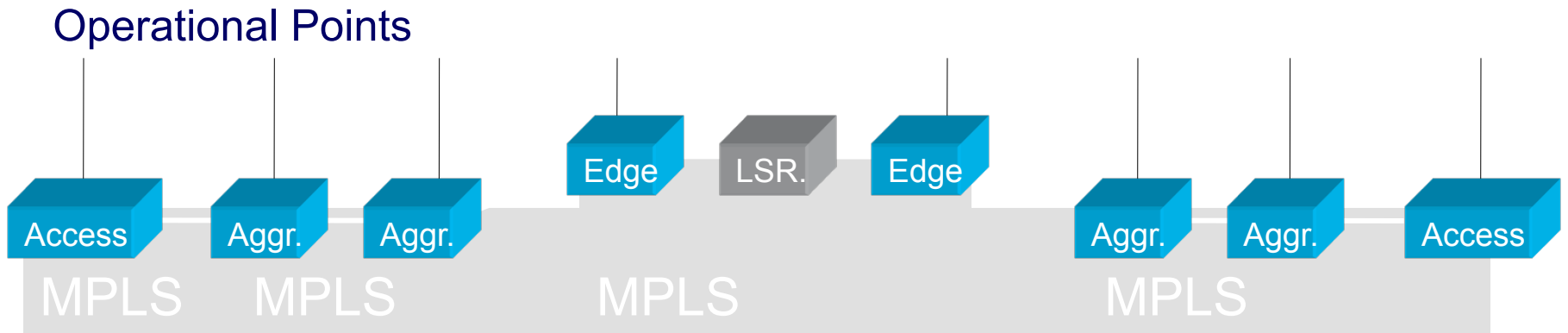
Продвижение транспортной технологии MPLS в сегмент агрегации и доступа

The Market Points to MPLS as the Right Choice

- **MPLS обеспечивает надежное транспортное решение**
 - Гибкость, статистическое мультиплексирование
 - Детерминированный маршрут - Traffic Engineering
 - Статические и динамические схемы построения защиты
 - Виртуализация услуг: L2VPN, L3VPN
 - Multipoint capabilities
- **74% предприятий применяют MPLS технологии в своих сетях и еще 10% планируют внедрение***
- **Снижение стоимости оборудования, поддерживающего MPLS.**
- **MPLS-TP**
 - OAM
 - Конфигурация - поддержка статического и динамического методов
 - Дальнейшее развитие в слиянии MPLS-TP с IP/MPLS
- **Unified MPLS от доступа к опорной инфраструктуре from access to core has the potential to:**
 - Единый control plane
 - Унификация в предоставлении услуг
 - Снижение операционных затрат

*Nemertes Research, [Press release](#), May 2008

Применение MPLS



- При внедрении сервиса все узлы вовлекаются в процесс конфигурирования. Большое кол-во точек изменения конфигурации. Система управления должна «знать» всю сетевую топологию.
- Задача состоит в минимизации точек изменения конфигурации
- Преимущество динамики в MPLS.
- Только интеграция всех MPLS сегментов позволяет минимизировать кол-во точек модификации конфигурации.

Обзор решения

Unicast Routing + MPLS

- Изоляция доменов IGP
- Подключение доменов IGP через BGP

Fast Restoration

- IP FRR (IGP LFA & BGP PIC)
- MPLS FRR (LDP FRR & TE FRR)

Маршрутизация + MPLS дизайн

Традиционный подход...

- Объявление инфраструктурных (e.g. PE) маршрутов в IGP
- Объявление инфраструктурных (e.g. PE) меток в LDP
- Сегментация IGP доменов (i.e. ISIS L1/L2 или OSPF Areas)



Маршрутизация + MPLS дизайн

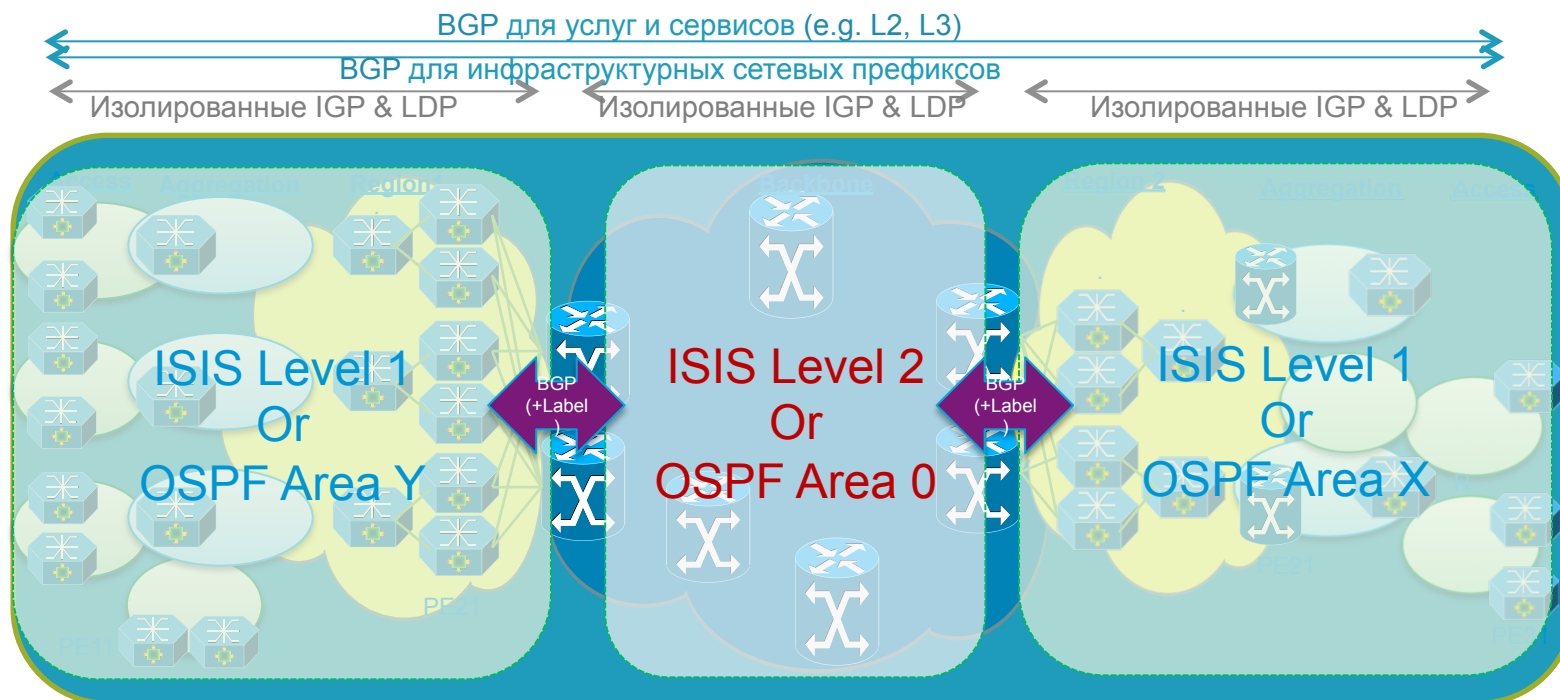
Традиционный подход не всегда является лучшим решением...

- Большая база данных IGP
 - Проблемы быстрой сходимости
- Большой IGP домен
 - Устойчивость сетевой инфраструктуры
- Большая база данных LDP



Маршрутизация + MPLS дизайн 'Разделяй и властвуй' – новый подход

- Разделение и **изоляция доменов IGP**
Нет больше единой (end-to-end) IGP связности
- Использование **BGP** для инфраструктурных (i.e. PE) маршрутов
А также и для инфраструктурных (i.e. PE) меток (e.g. RFC3107)



Маршрутизация + MPLS дизайн

Разделяй и властвуй

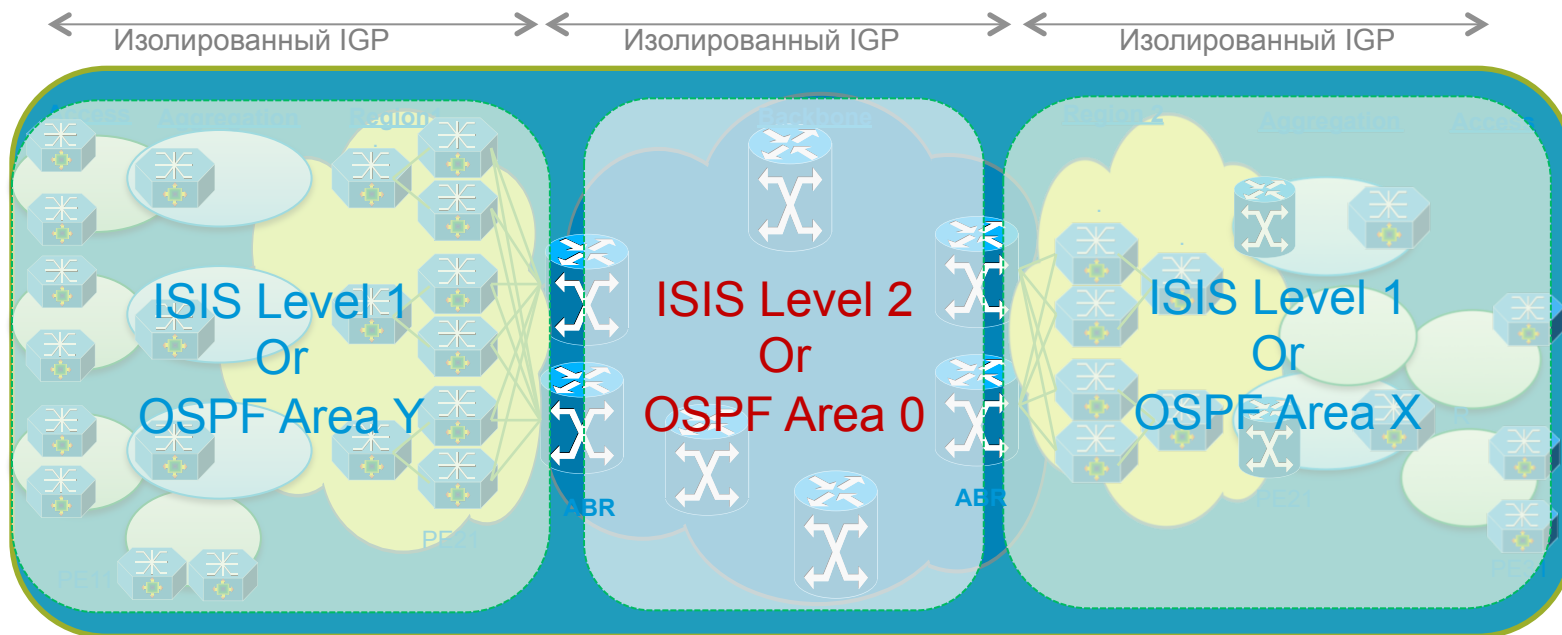
1. IGP используется только для распространения маршрутной информации только в рамках изолированного домена
 - a. Не нулевые или L1 area распространяют маршруты своего домена
 - b. Backbone распространяют маршруты только backbone зоны
2. BGP+Label (RFC3107) используется для обмена маршрутами между доменами
 - a. PE ре-дистрибуция своих loopback в IGP также как и в iBGP+Label
 - b. PE взаимодействуют со своими локальными ABRs используя iBGP
 - ABRs выполняют роль Route-reflectors
 - ABRs reflect `_only_` Infrastructure (i.e. PE) routes
 - c. ABR, в качестве RR, изменяет BGP Next-hop на своего Loopback
3. BGP, как и прежде, используется для УСЛУГ (L3, VPN)
 1. PEs имеют отдельные пиринговые отношения для обмена сервисными префиксами /NLRIs

Маршрутизация + MPLS Дизайн Unified MPLS

IGP используется только для распространения маршрутной информации
только в рамках изолированного домена

1

Не нулевые или L1 area распространяют маршруты своего домена
Backbone распространяют маршруты только backbone зоны



* ISIS L1->L2 (or L1->L1)
Redistribution can Be Avoided,
similar to that of OSPF Non-Zero-<->Zero Area Redistribution.

Маршрутизация + MPLS Дизайн Unified MPLS

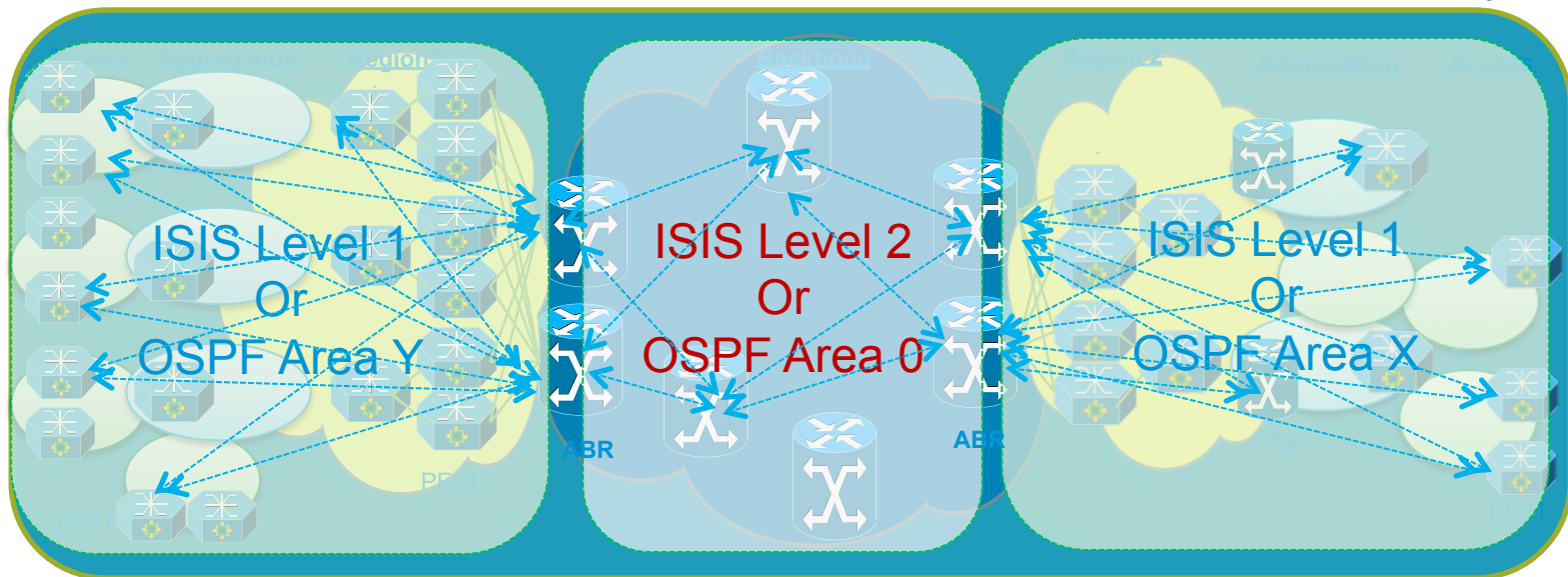
3 PE строит свои пириновые взаимодействия с локальными ABRs используя iBGP+label

ABRs выполняет роль Route-reflector-a

ABRs reflect `_only_` Infrastructure (i.e. PE) routes

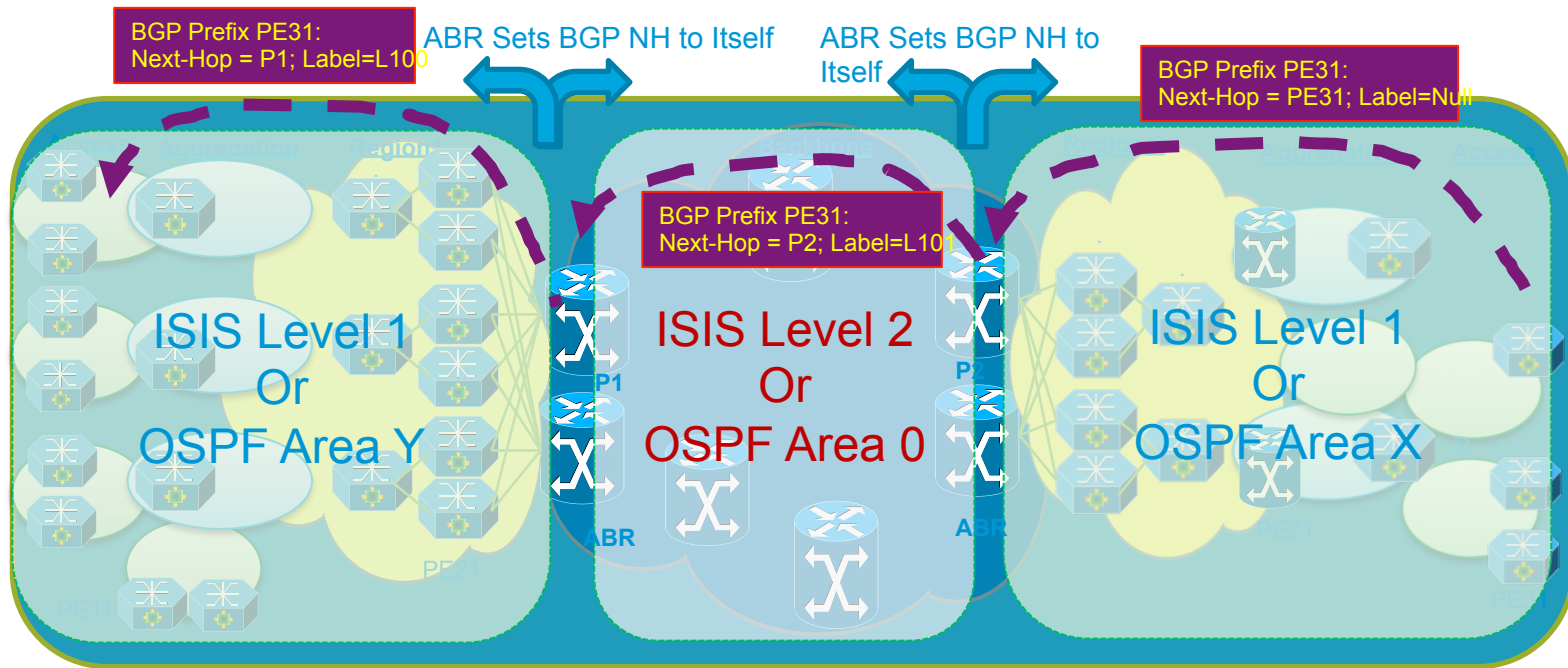
RRs также присутствует в опорной инфраструктуре

↔ iBGP+Label Peering



Маршрутизация + MPLS Дизайн Unified MPLS

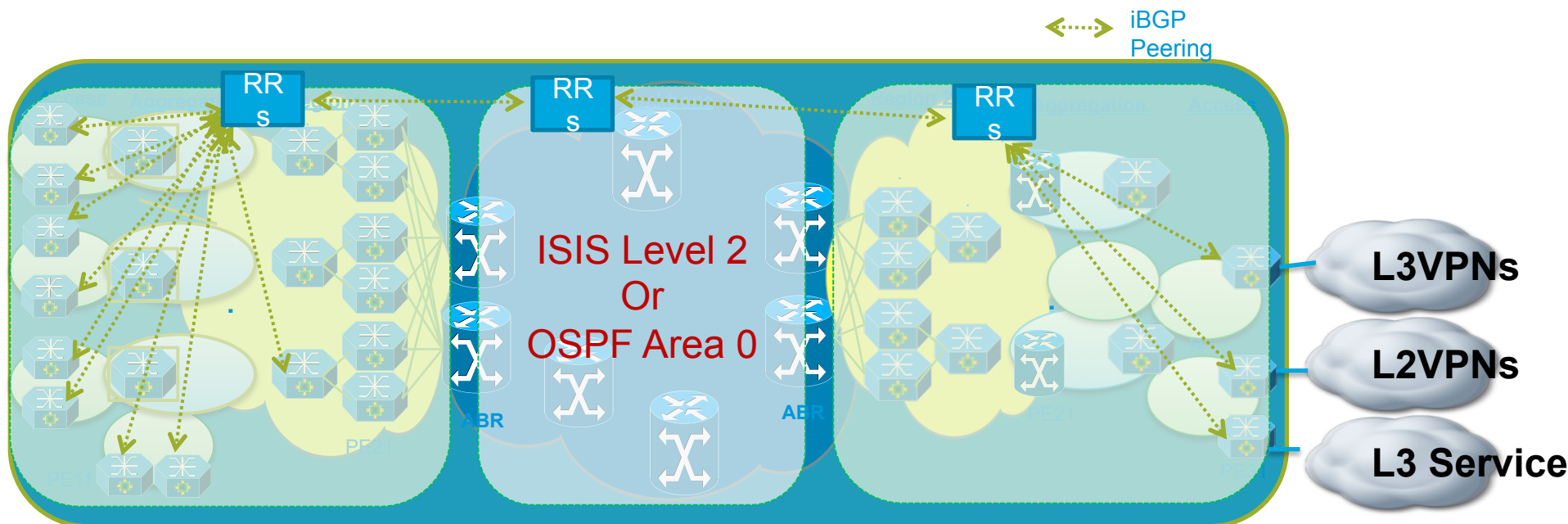
4. ABR, в качестве роут-рефлектора, изменяет BGP Next-hop на свой для каждого маршрута объявляемого через BGP



Маршрутизация + MPLS Дизайн Unified MPLS

- 5 PEs имеют отдельные iBGP взаимосвязи для **УСЛУГ**, одна или несколько AFI/SAFI (для IPv4/6, VPNv4/6, L2VPN, etc)

Для увеличения масштабируемости можно выделить RR для каких-то сервисов или групп сервисов.



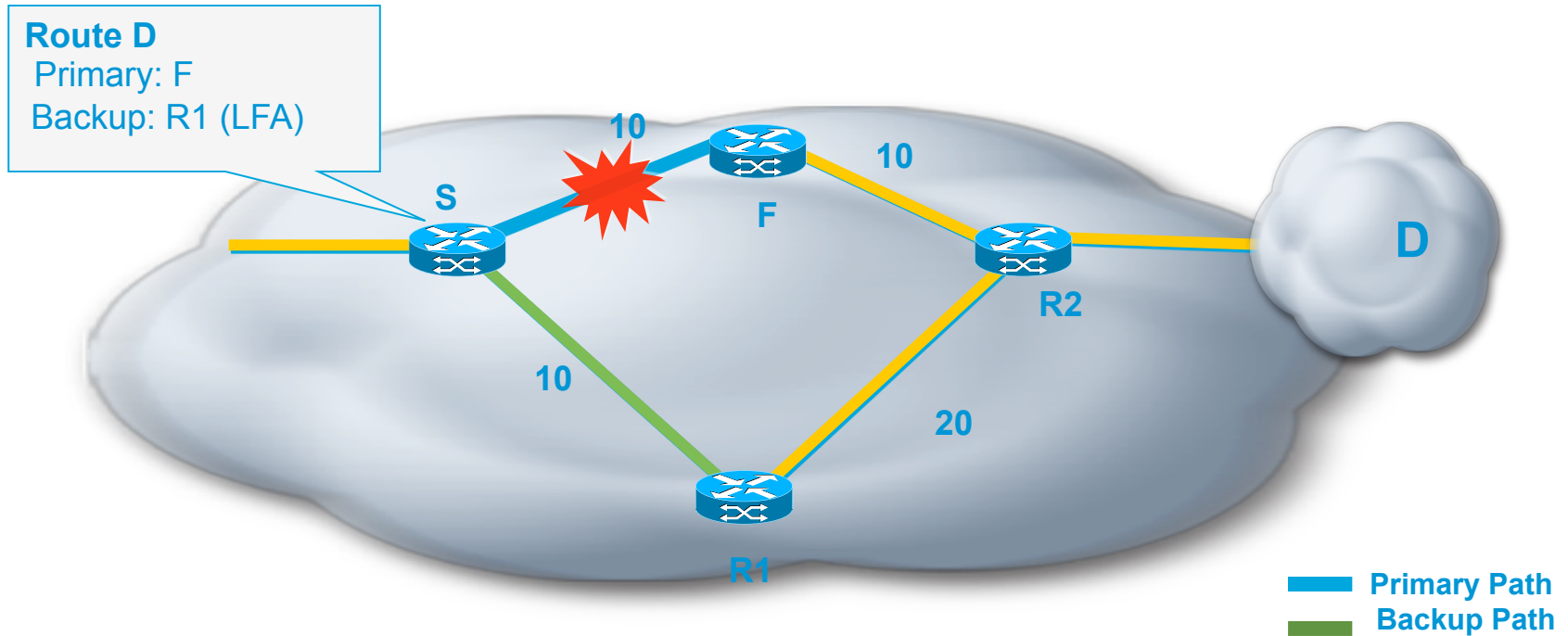
СХОДИМОСТЬ

- IGP fast convergence
 - Loop Free Alternate Fast (LFA FRR) внутри домена
 - Unified MPLS BGP PIC Core and Edge
 - Быстрое восстановление сервисов (BGP PIC Edge, PW FRR)
- 50ms for non recursive processes.
- 100ms convergence end-end.
- При этом модель практической реализации должна оставаться предельно простой

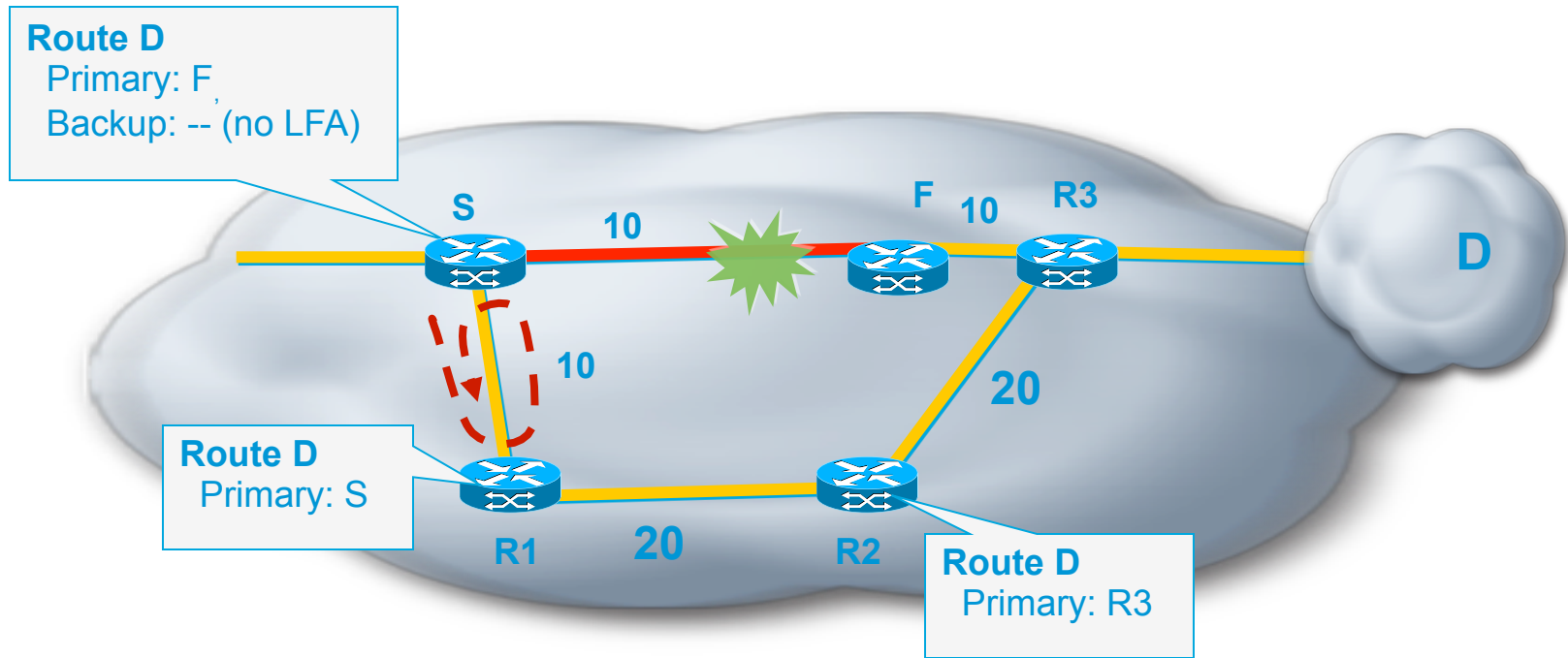
Что такое LFA FRR - обзор?

- Механизм быстрой сходимости, который позволяет достичь сходимости до 50 ms.
- Для достижения быстрой пере-маршрутизации трафика forwarding engine (CEF/FIB) специально программируется и хранит основной и запасной маршруты для каждого префикса. Каждый запасной маршрут предварительно рассчитан с использованием алгоритма LFA.
- Backup paths are computed AFTER the primary path and so do not delay normal convergence
- Для минимизации времени переключения forwarding engine (FIB) должен быть иерархическим.
- Необходима поддержка механизмов быстрого детектирования отказов для переключения на запасной маршрут (BFD, IPoDWDM proactive protection, ...)

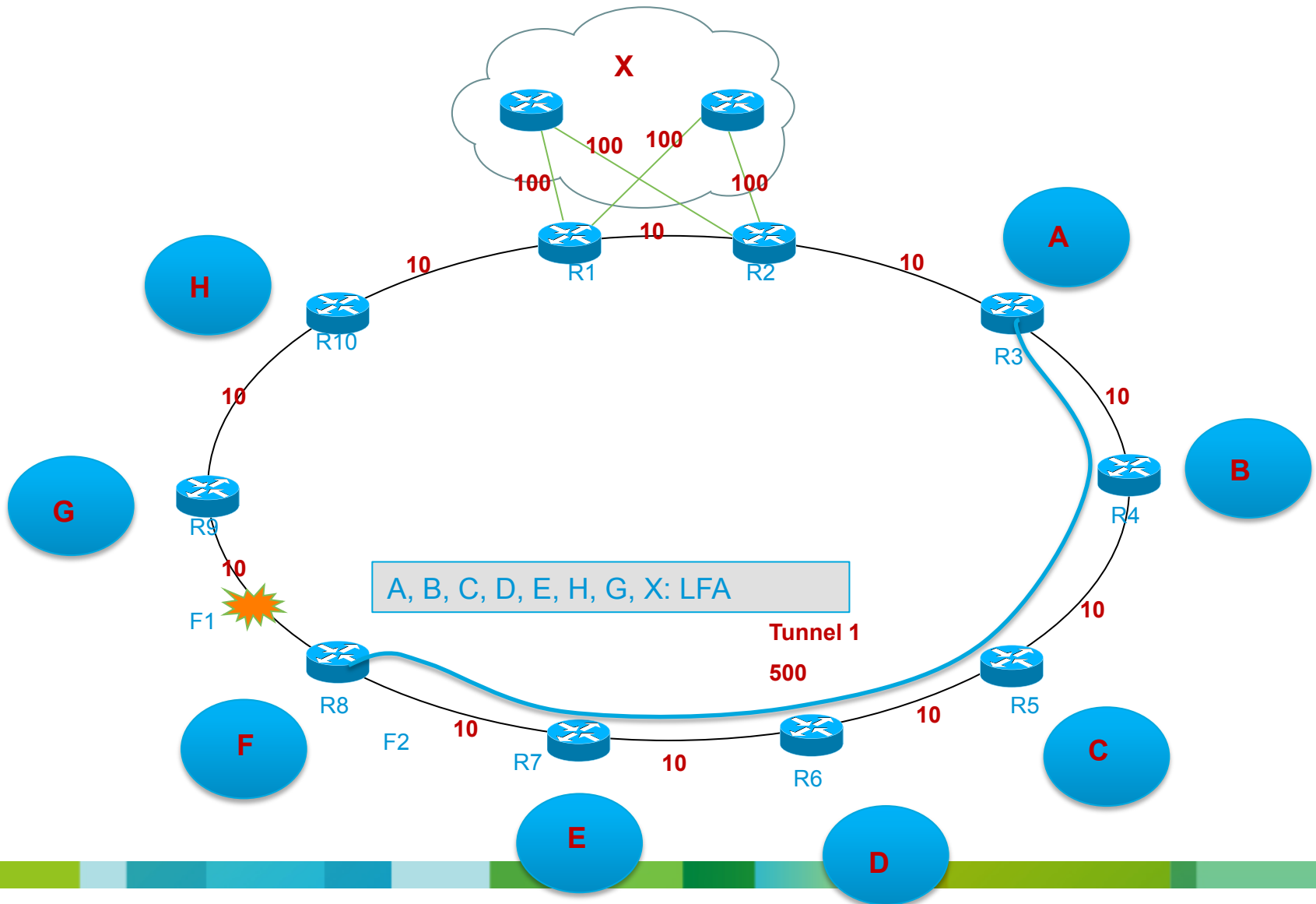
LFA найден



LFA не найден



Использование туннельных технологий для LFA FRR (Remote LFA FRR).

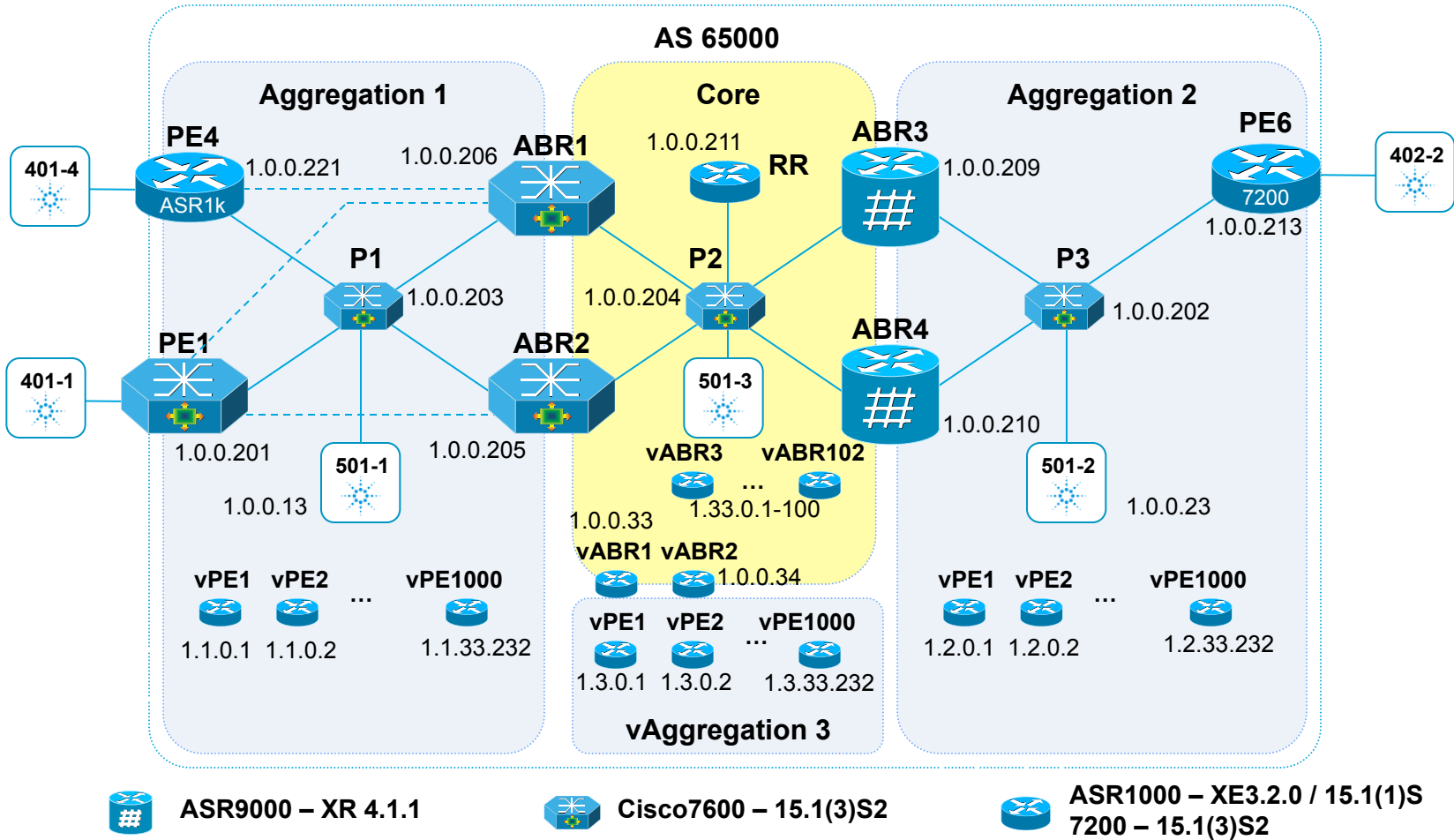


LFA FRR для кольцевых топологий

- LFA FRR with tunnel
- No PQ computation
- Manual placement
- RSVP-TE tunnel
- Remote LFA FRR
- PQ computation (find first node who is not sending traffic back)
- Auto placement of tunnel toward PQ node
- LDP tunnels

Параметры модели

50x Агрегационных доменов



Топология стенда

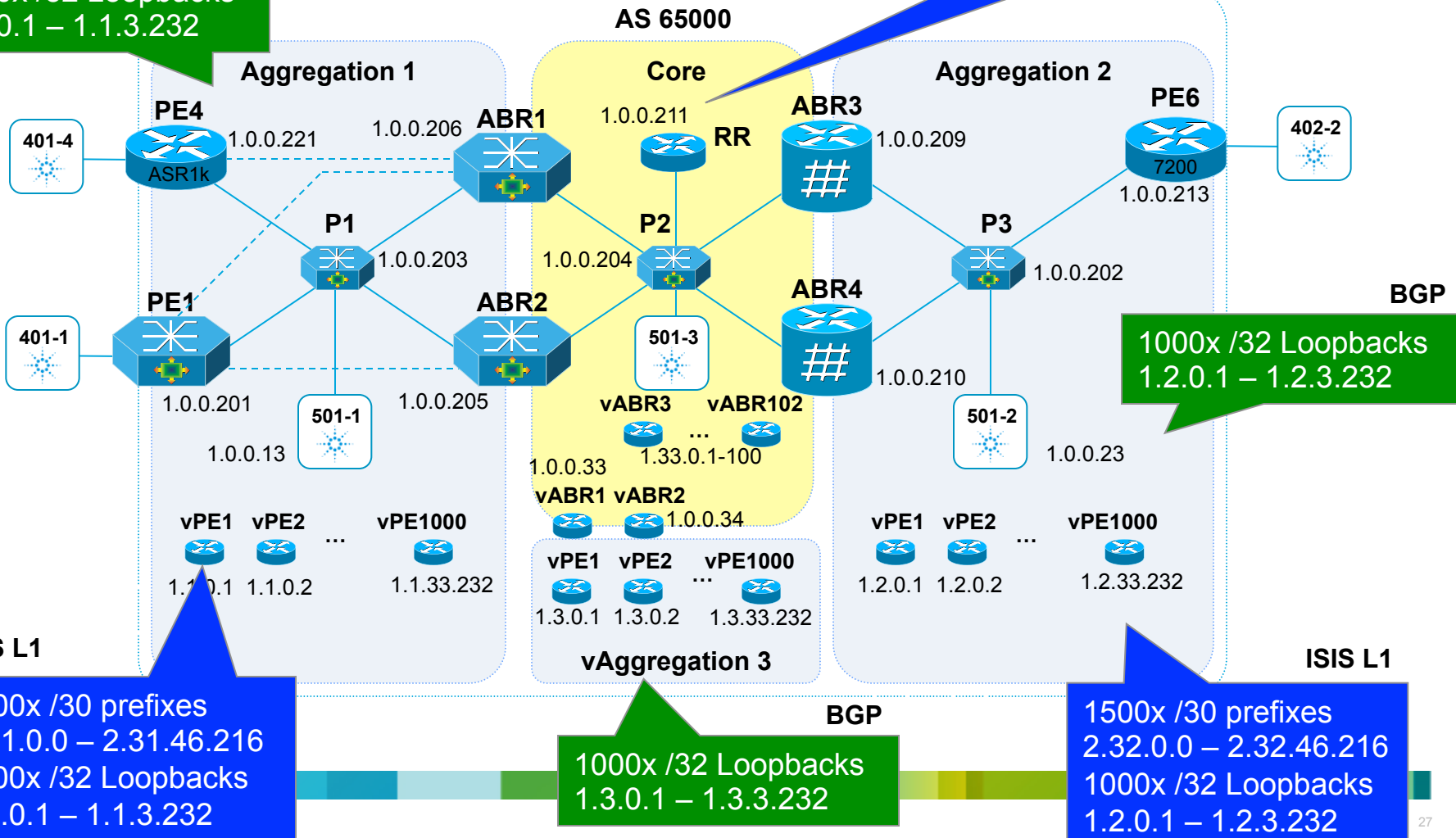
Симуляция сети размером 50,000 узлов

ISIS L2

1000x /30 prefixes
2.31.0.0 – 2.31.31.56
102x /32 Loopbacks
1.0.0.33, 1.0.0.34
1.3.0.1 – 1.3.3.232

BGP

1000x /32 Loopbacks
1.1.0.1 – 1.1.3.232



ISIS L1

1500x /30 prefixes
2.31.0.0 – 2.31.46.216
1000x /32 Loopbacks
1.1.0.1 – 1.1.3.232

BGP

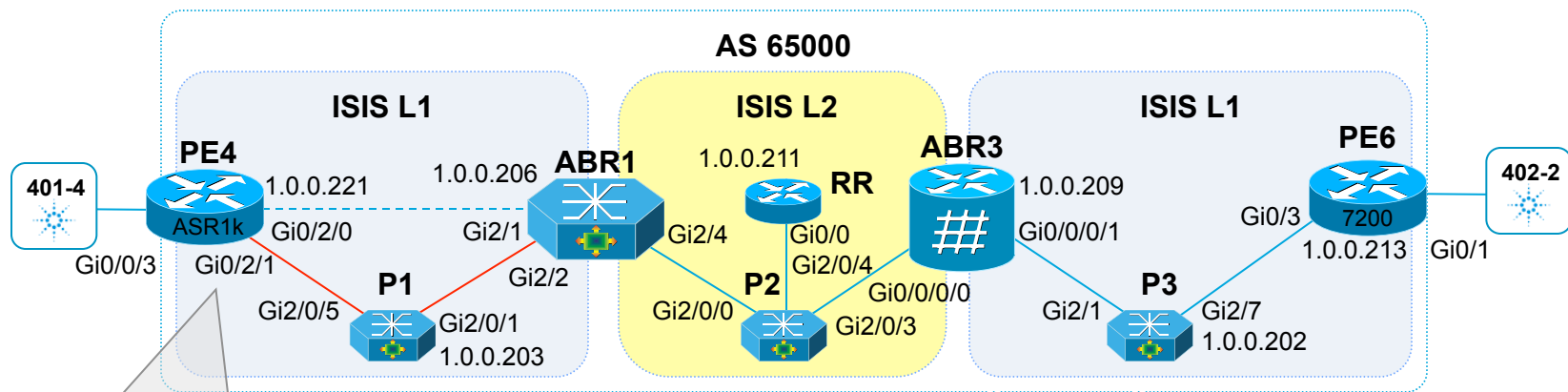
1000x /32 Loopbacks
1.3.0.1 – 1.3.3.232

ISIS L1

1500x /30 prefixes
2.32.0.0 – 2.32.46.216
1000x /32 Loopbacks
1.2.0.1 – 1.2.3.232

Конфигурация IGP

Узлы PE и P (IOS)



```
router isis
net 49.0001.0000.0000.0221.00
is-type level-1
ispf level-1
metric-style wide
fast-flood
ip route priority high tag 1000
max-lsp-lifetime 65535
lsp-refresh-interval 65000
spf-interval 5 20 50
prc-interval 5 20 50
lsp-gen-interval 5 20 50
no hello padding
mpls ldp sync
```

```
mpls ldp igp sync holddown 10000
mpls ldp router-id Loopback0
```

```
interface Loopback0
ip address 1.0.0.221 255.255.255.255
ip router isis
isis tag 1000
```

```
interface GigabitEthernet0/2/1
ip address 2.203.221.2
255.255.255.252
ip router isis
mpls ip
isis network point-to-point
```

```
interface GigabitEthernet0/2/0
ip address 2.206.221.2 255.255.255.252
ip router isis
mpls ip
isis network point-to-point
isis metric 50
```

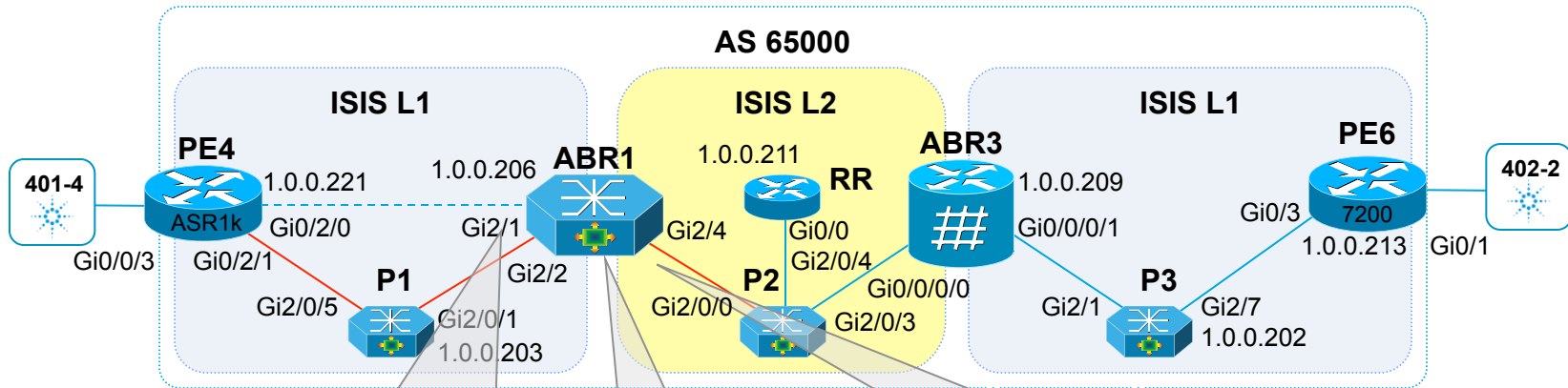
ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

Cisco7600 – 15.1(3)S2

ASR9000 – XR 4.1.1

Конфигурация IGP

Узлы ABR (IOS)



```
interface GigabitEthernet2/1
ip address 2.201.206.2 255.255.255.252
ip router isis
mpls ip
isis circuit-type level-1
isis network point-to-point
```

Aggregation Facing Interfaces

```
interface GigabitEthernet2/4
ip address 2.204.206.2 255.255.255.252
ip router isis
mpls ip
isis circuit-type level-2-only
isis network point-to-point
```

Core Facing Interfaces

```
interface GigabitEthernet2/3
ip address 2.205.206.2 255.255.255.252
ip router isis
mpls ip
isis network point-to-point
```

Inter ABR Interface

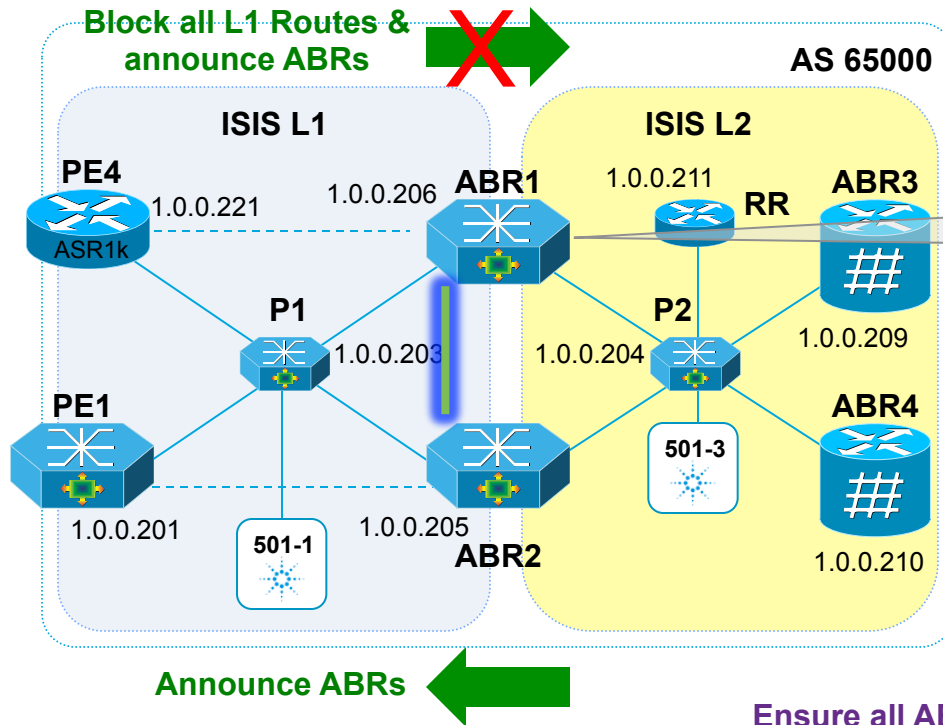
ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

Cisco7600 – 15.1(3)S2

ASR9000 – XR 4.1.1

Конфигурация IGP

Узлы ABR (IOS)



```
access-list 100 permit ip host 1.0.0.206 any
access-list 100 permit ip 2.205.206.0 0.0.0.3 any
access-list 100 deny ip any any
```

```
ip access-list standard ibgp_nexthops
permit 1.0.0.33
permit 1.0.0.34
permit 1.0.0.205
permit 1.0.0.206
permit 1.0.0.209
permit 1.0.0.211
permit 1.0.0.210
deny any
```

Ensure all ABRs are in IS-IS

```
router isis
 redistribute isis ip level-1 into level-2 distribute-list 100
 redistribute isis ip level-2 into level-1 route-map ibgp_nexthops
```

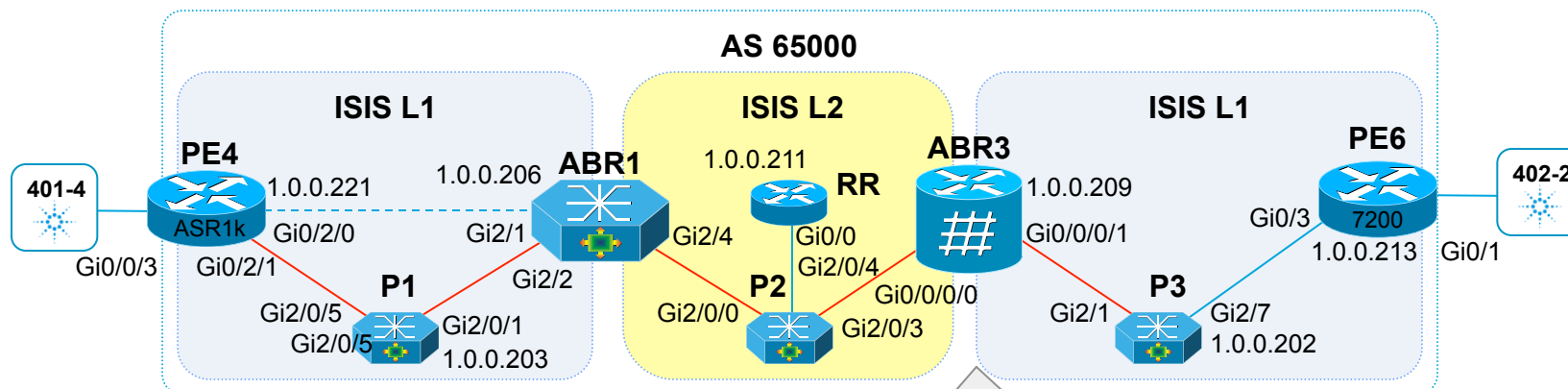
ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

Cisco7600 – 15.1(3)S2

ASR9000 – XR 4.1.1


Конфигурация IGP


Узлы ABR (IOS XR)



```

router isis core
net 49.001.000.000.209.00
address-family ipv4 unicast
metric-style wide
ispf
spf-interval maximum-wait 5000 initial-wait 20 secondary-wait 50
spf prefix-priority high tag 1000
!
interface Loopback0
point-to-point
address-family ipv4 unicast
tag 1000
  
```

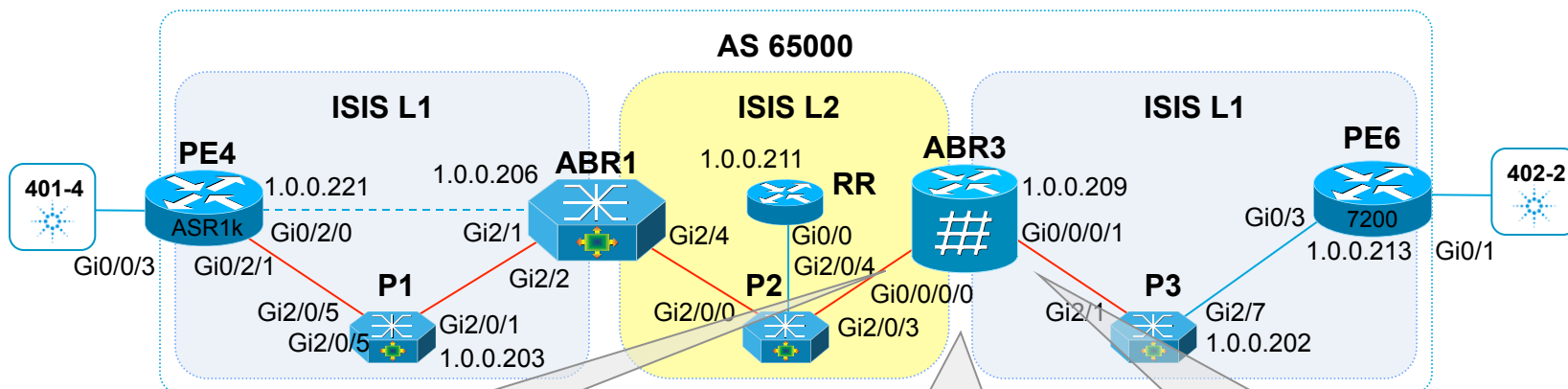
 ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

 Cisco7600 – 15.1(3)S2

 ASR9000 – XR 4.1.1

Конфигурация IGP

Узлы ABR (IOS XR)



```
router isis core
interface GigabitEthernet0/0/0/0
circuit-type level-2-only
point-to-point
address-family ipv4 unicast
mpls ldp sync
```

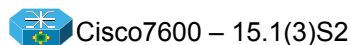
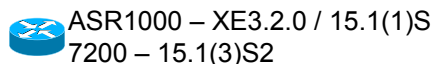
Aggregation Facing Interfaces

```
router isis core
interface GigabitEthernet0/0/0/4
point-to-point
address-family ipv4 unicast
mpls ldp sync
```

Inter ABR Interface

```
router isis core
interface GigabitEthernet0/0/0/1
circuit-type level-1
point-to-point
address-family ipv4 unicast
mpls ldp sync
```

Core Facing Interfaces



Конфигурация IGP

Узлы ABR (IOS XR)

Basic BGP
Configuration

Redundancy and
<1s Convergence

Unified MPLS
specific

```
prefix-set ibgp_nexthops
```

```
1.0.0.209,  
1.0.0.210,  
1.0.0.205,  
1.0.0.206,  
1.0.0.211,  
1.0.0.33,  
1.0.0.34  
end-set
```

```
!  
route-policy ibgp_nexthops  
  if destination in ibgp_nexthops then  
    pass  
  endif  
end-policy
```

```
prefix-set inter_area
```

```
1.0.0.209/32  
end-set
```

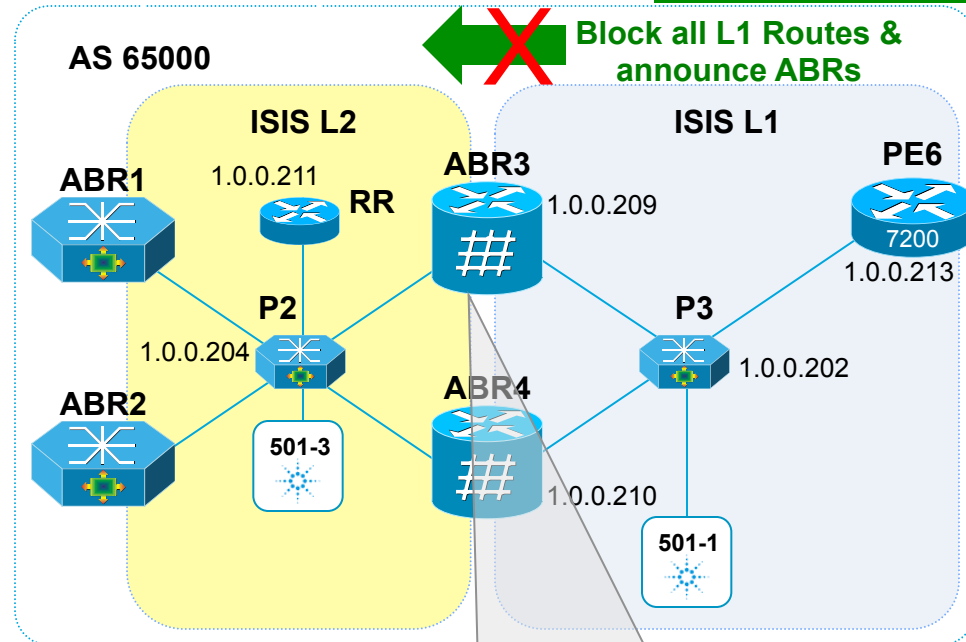
```
!  
route-policy inter_area  
  if destination in inter_area then  
    pass  
  endif  
end-policy
```

```
router isis core
```

```
  address-family ipv4 unicast
```

```
    propagate level 1 into level 2 route-policy inter_area
```

```
    propagate level 2 into level 1 route-policy ibgp_nexthops
```



Ensure all ABRs are in ISIS

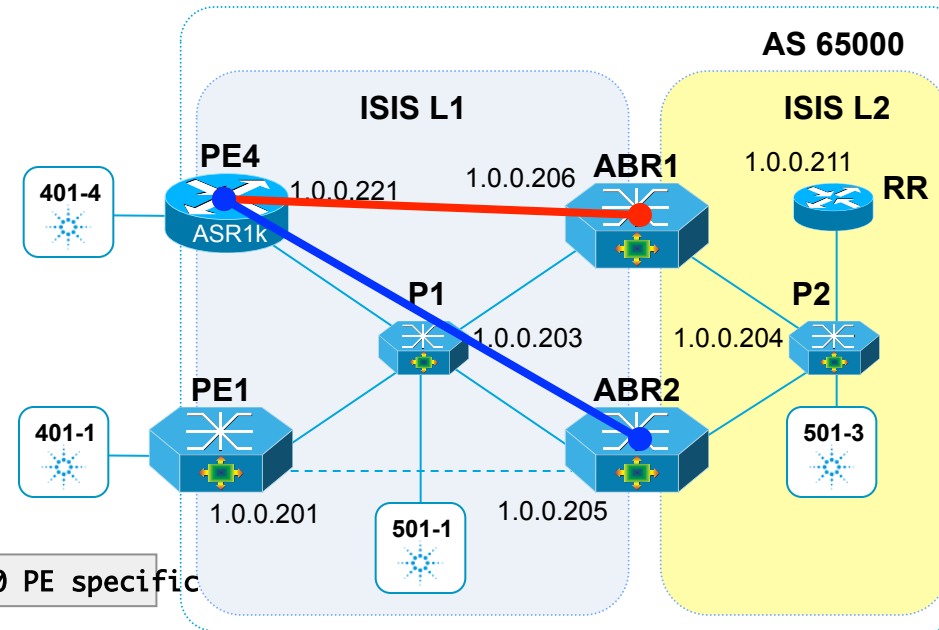
Конфигурация BGP

Узлы PE (IOS)

```

router bgp 65000
  bgp log-neighbor-changes
  neighbor 1.0.0.205 remote-as 65000
  neighbor 1.0.0.205 update-source Loopback0
  neighbor 1.0.0.206 remote-as 65000
  neighbor 1.0.0.206 update-source Loopback0
  !
  address-family ipv4
    bgp additional-paths install
    bgp nexthop trigger delay 1
    bgp mpls-local-label
    redistribute connected route-map inter_area
    neighbor 1.0.0.205 activate
    neighbor 1.0.0.205 send-label
    neighbor 1.0.0.206 activate
    neighbor 1.0.0.206 route-map primary_weight in
    neighbor 1.0.0.206 send-label
    no neighbor 1.0.0.211 activate
  exit-address-family
  
```

Cisco 7600 PE specific



Prefer Path via ABR1

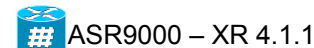
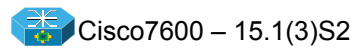
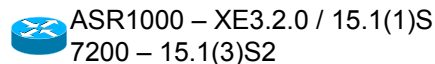
```

ip access-list standard all
  permit any
  !
  route-map primary_weight permit 10
  match ip address all
  set weight 10
  
```

Ensure PE is reachable from other Aggregation Areas

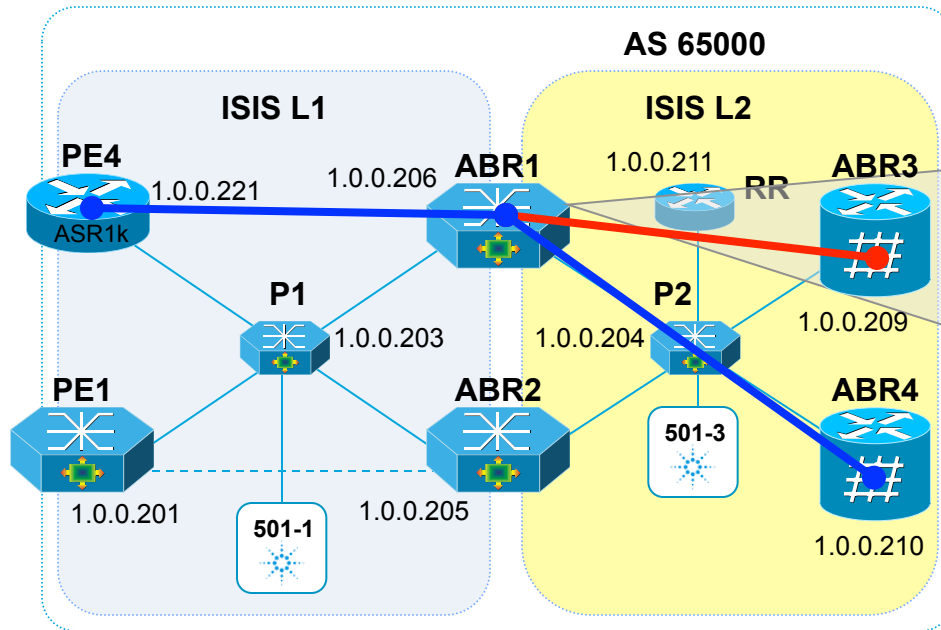
```

ip access-list standard inter_area
  permit 1.0.0.221
  deny any
  !
  route-map inter_area permit 10
  match ip address inter_area
  
```



Конфигурация BGP

Узлы ABR (IOS)



```

router bgp 65000
  bgp cluster-id 10
  bgp log-neighbor-changes
  neighbor 1.0.0.209 remote-as 65000
  neighbor 1.0.0.209 update-source Loopback0
  neighbor 1.0.0.210 remote-as 65000
  neighbor 1.0.0.210 update-source Loopback0
  neighbor 1.0.0.221 remote-as 65000
  neighbor 1.0.0.221 update-source Loopback0
  !
  address-family ipv4
    bgp nexthop trigger delay 1
    neighbor 1.0.0.209 activate
    neighbor 1.0.0.209 route-reflector-client
    neighbor 1.0.0.209 next-hop-self all
    neighbor 1.0.0.209 route-map primary_weight in
    neighbor 1.0.0.209 send-label
    neighbor 1.0.0.210 activate
    neighbor 1.0.0.210 route-reflector-client
    neighbor 1.0.0.210 next-hop-self all
    neighbor 1.0.0.210 send-label
    neighbor 1.0.0.221 activate
    neighbor 1.0.0.221 route-reflector-client
    neighbor 1.0.0.221 send-label
  
```

Prefer Path via ABR3

```

ip access-list standard all
  permit any
  !
  route-map primary_weight permit 10
  match ip address all
  set weight 10
  
```

ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

Cisco7600 – 15.1(3)S2

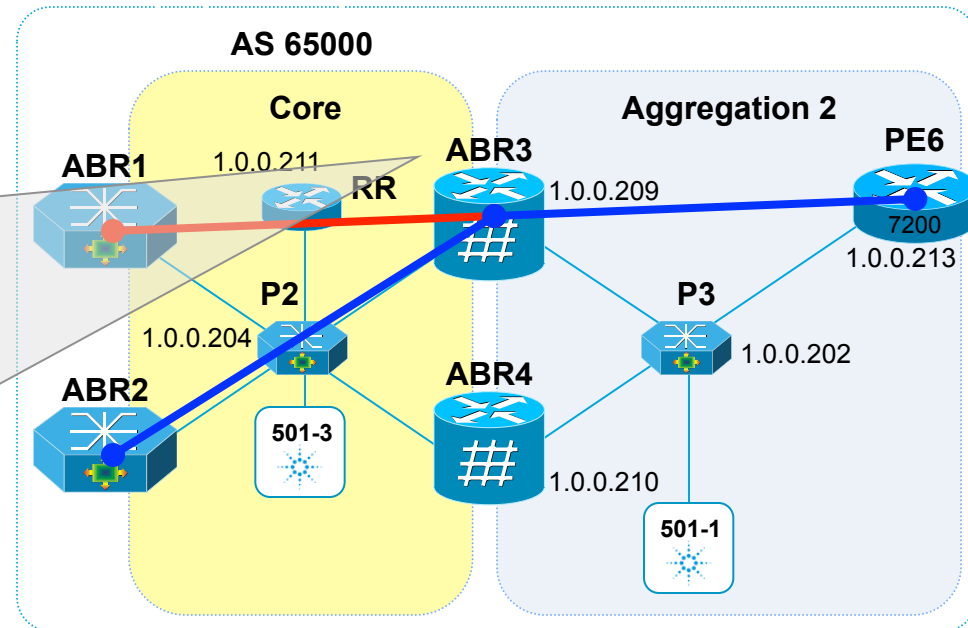
ASR9000 – XR 4.1.1

Конфигурация BGP

Узлы ABR (IOS XR)

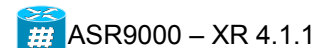
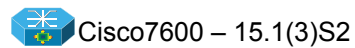
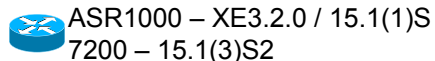
```

router bgp 65000
  bgp cluster-id 20
  ibgp policy out enforce-modifications
  address-family ipv4 unicast
    nexthop trigger-delay critical 1000
    allocate-label all
  !
  neighbor 1.0.0.205
    remote-as 65000
    update-source Loopback0
    address-family ipv4 labeled-unicast
    route-reflector-client
    next-hop-self
  !
  neighbor 1.0.0.206
    remote-as 65000
    update-source Loopback0
    address-family ipv4 labeled-unicast
    route-policy primary_weight
    route-reflector-client
    next-hop-self
  !
  neighbor 1.0.0.213
    remote-as 65000
    update-source Loopback0
    address-family ipv4 labeled-unicast
    route-reflector-client
  
```



```

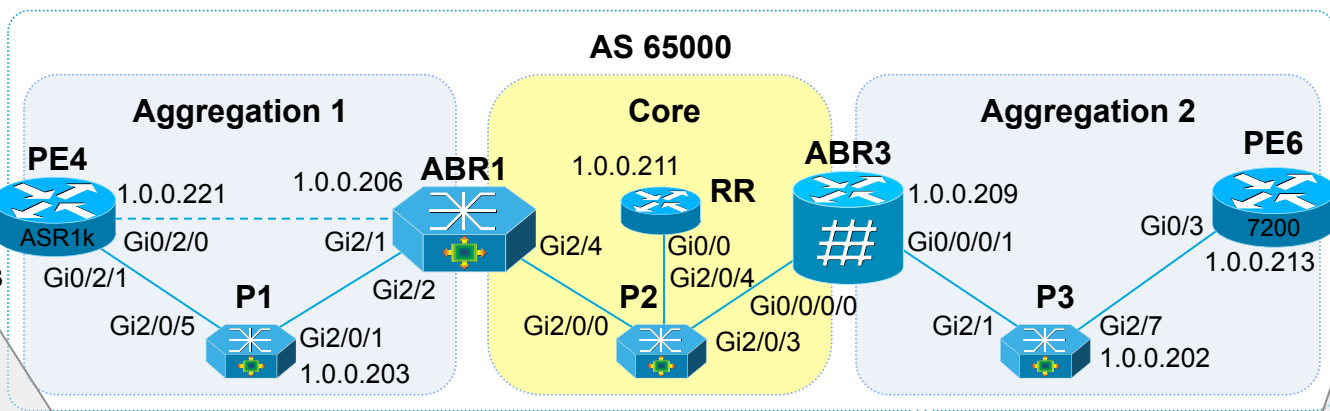
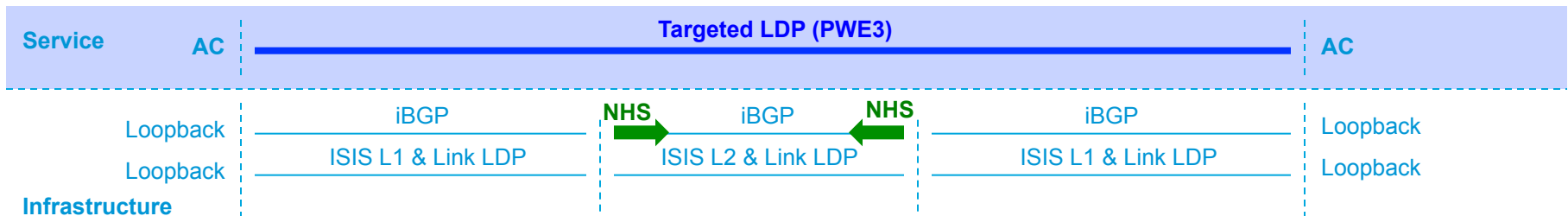
route-policy primary_weight
  set weight 10
end-policy
  
```



Конфигурация сервиса

Пример построения EoMPLS

Only 2 Touch-points to bring up the Service between any two Nodes !!



```
interface GigabitEthernet0/0/3.146
encapsulation dot1Q 146
xconnect 1.0.0.213 213221146 pw-class mpls
```

Access Circuit (AC) Configuration

```
interface GigabitEthernet0/1.146
encapsulation dot1Q 146
xconnect 1.0.0.221 213221146 pw-class mpls
```

Access Circuit (AC) Configuration

ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

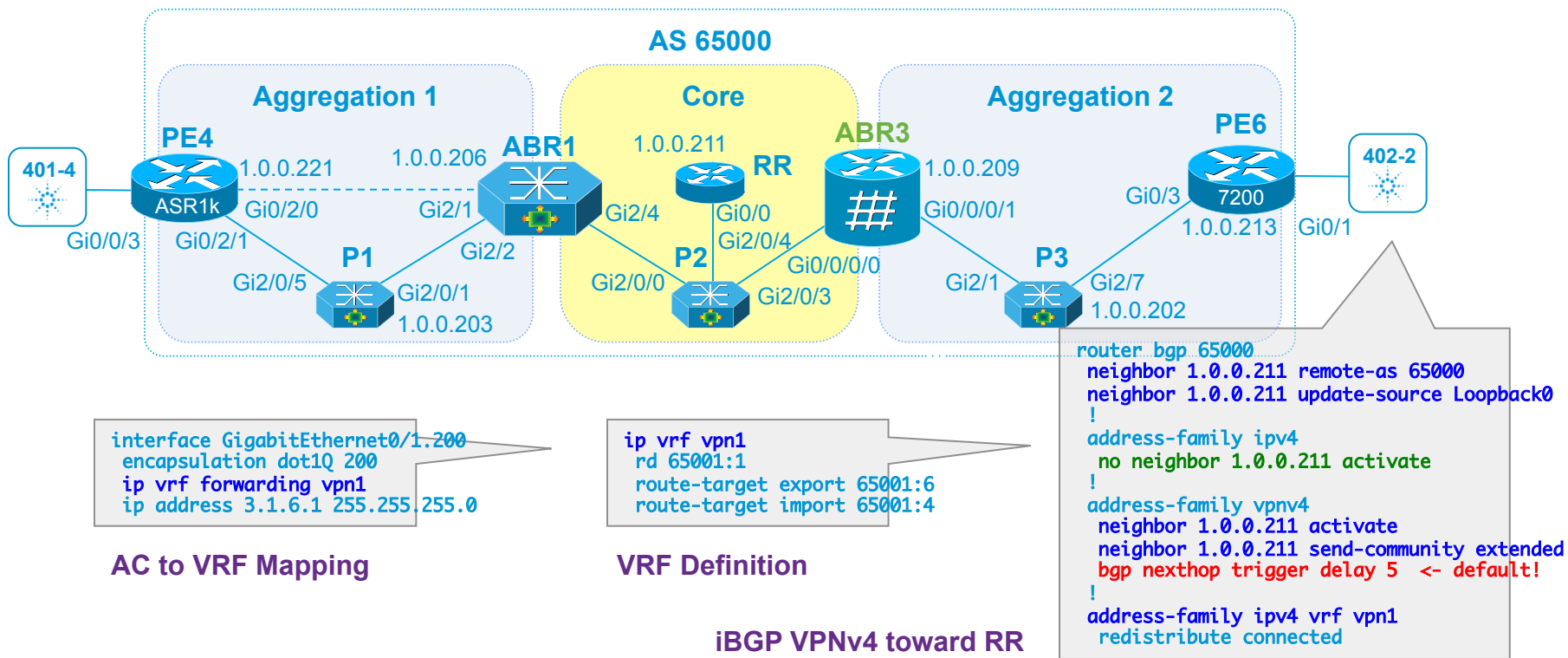
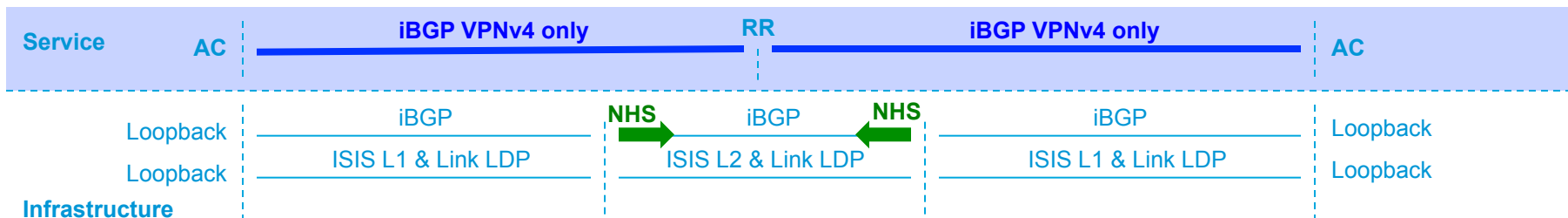
Cisco7600 – 15.1(3)S2

ASR9000 – XR 4.1.1

Конфигурация сервиса

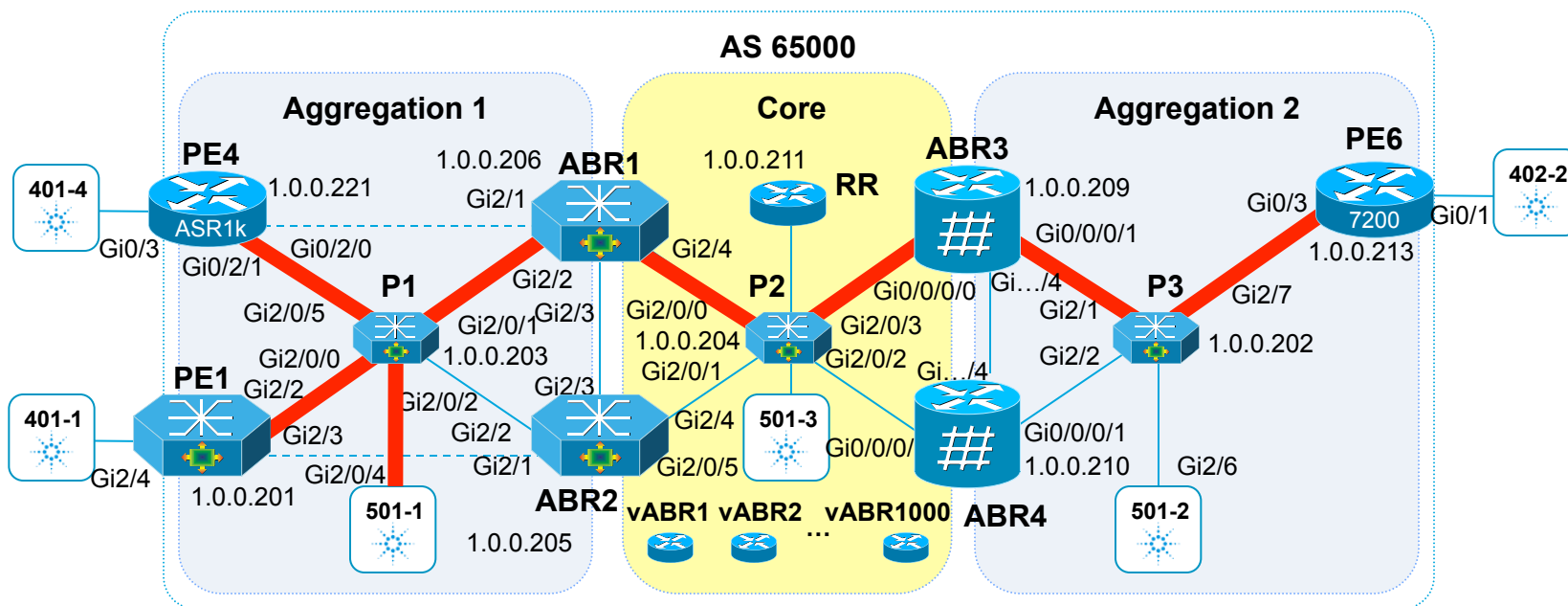
Пример построения L3VPN

VPNv4 Layer is independent of Infrastructure BGP Layer !!



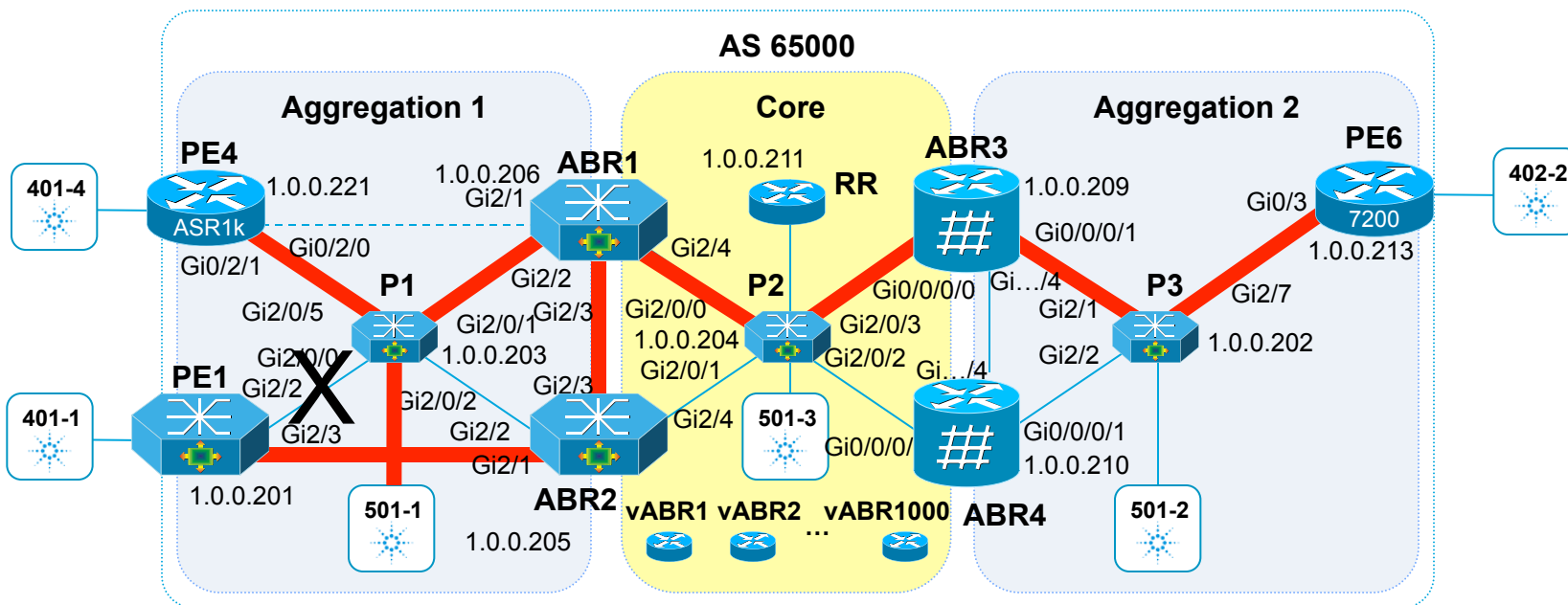
Анализ отказоустойчивости

Отказ Primary Path



Анализ отказоустойчивости

Авария на канале PE1 ↔ P1



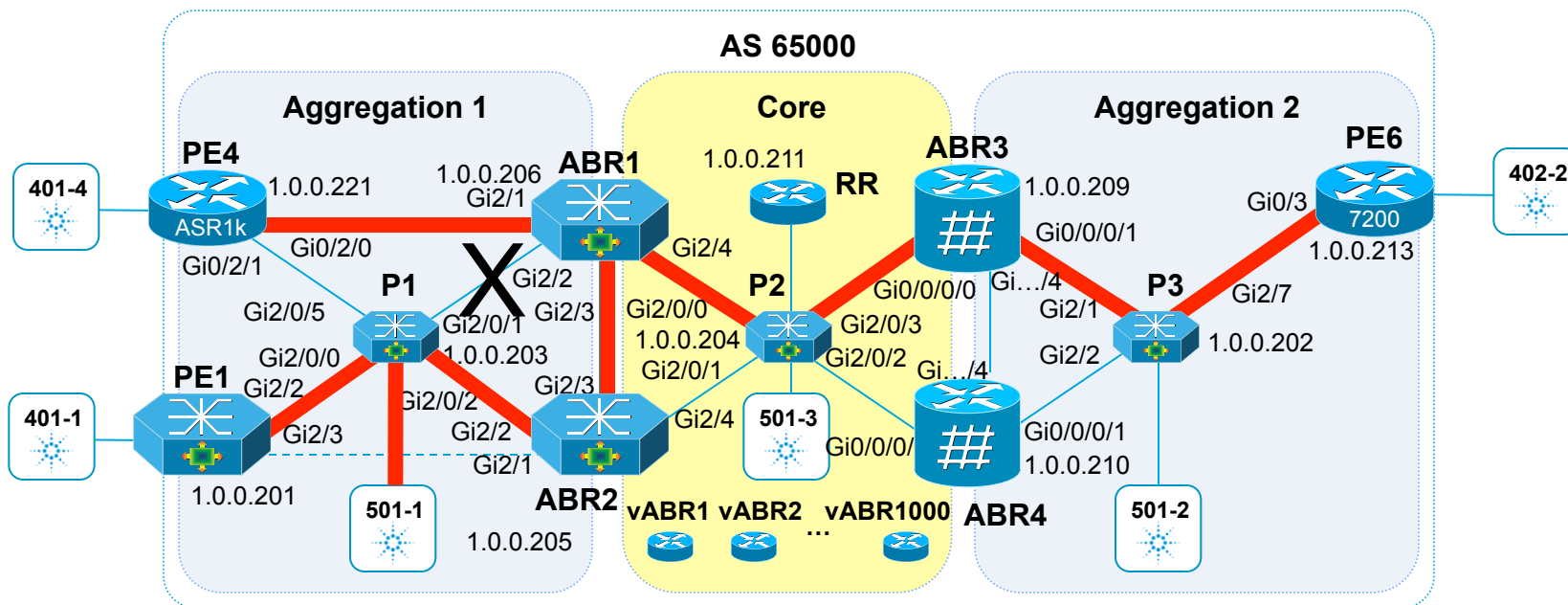
ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

Cisco7600 – 15.1(3)S2

ASR9000 – XR 4.1.1

Анализ отказоустойчивости

Авария на канале P1 ↔ ABR1



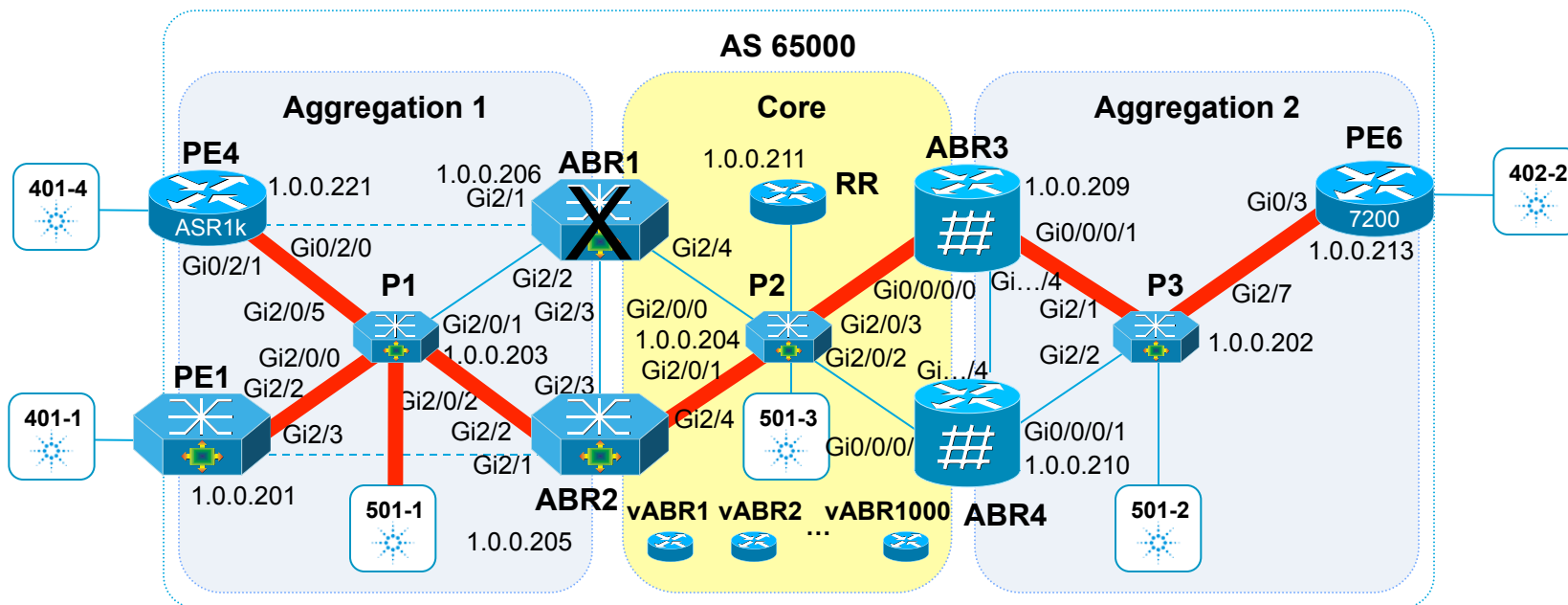
ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2

Cisco7600 – 15.1(3)S2

ASR9000 – XR 4.1.1

Анализ отказоустойчивости

Отказ узла ABR1



ASR1000 – XE3.2.0 / 15.1(1)S
7200 – 15.1(3)S2



Cisco7600 – 15.1(3)S2



ASR9000 – XR 4.1.1

Сигнализация Multi-Path

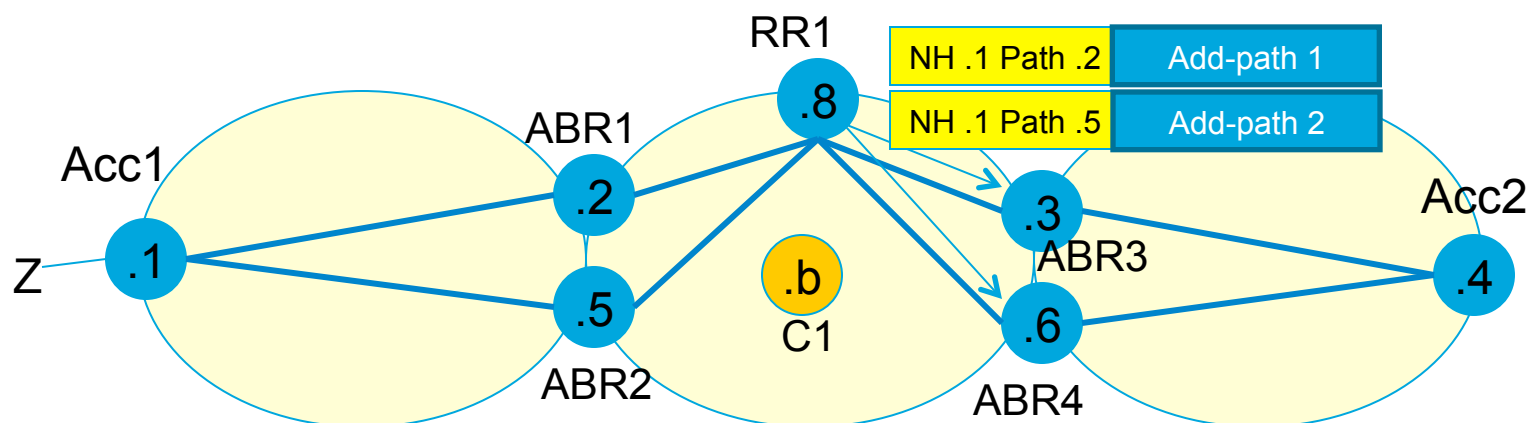
Оптимизация

- BGP Add-Path
[draft-ietf-idr-add-paths-06](#)
- BGP Accumulated IGP
[draft-ietf-idr-aigp-08.txt](#)

Сигнализация Multi-Path

XR 4.0.0
IOS XE 3.7.0S

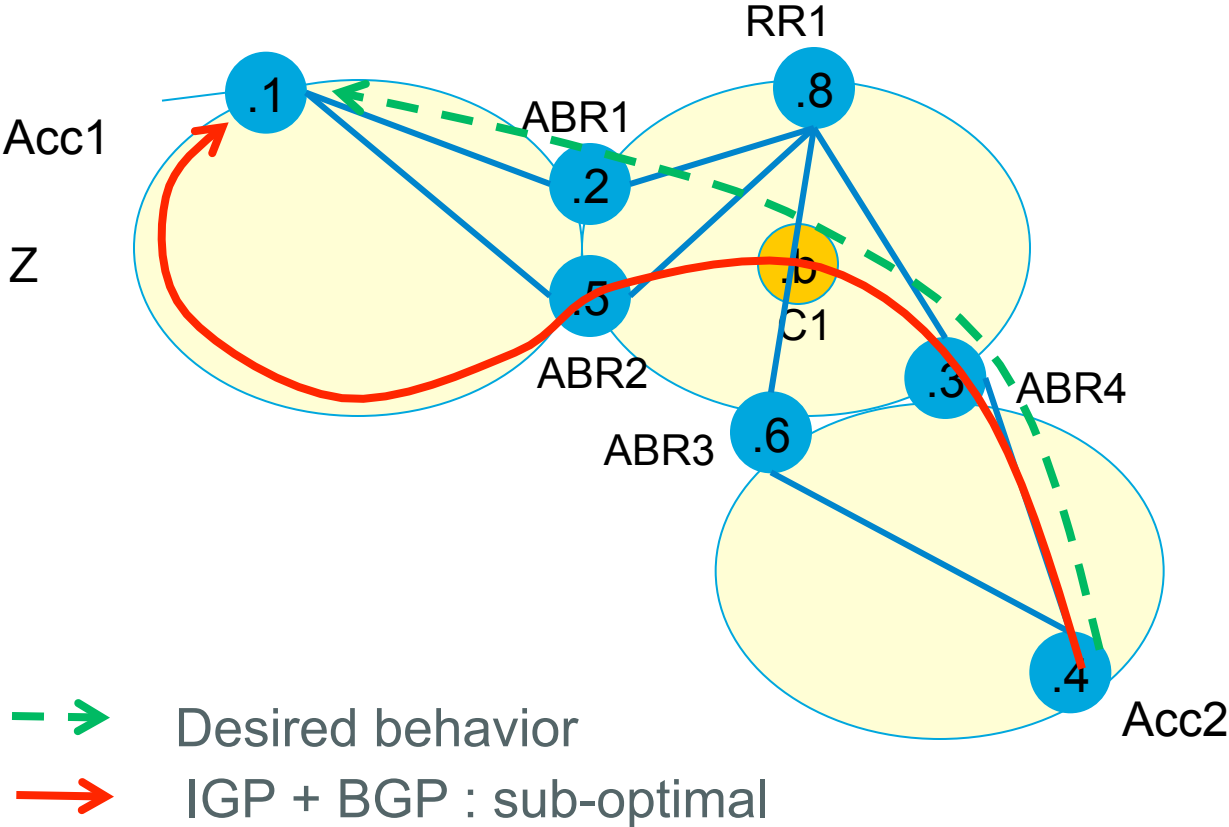
BGP Add-Path



- BGP Add path :
 - add-all-path:
 - RR1 forward all path without best path selection for backup path
 - add-1-path:
 - RR do backup best path selection and send one backup path.

Сигнализация Multi-Path

Unified MPLS sub-optimal path

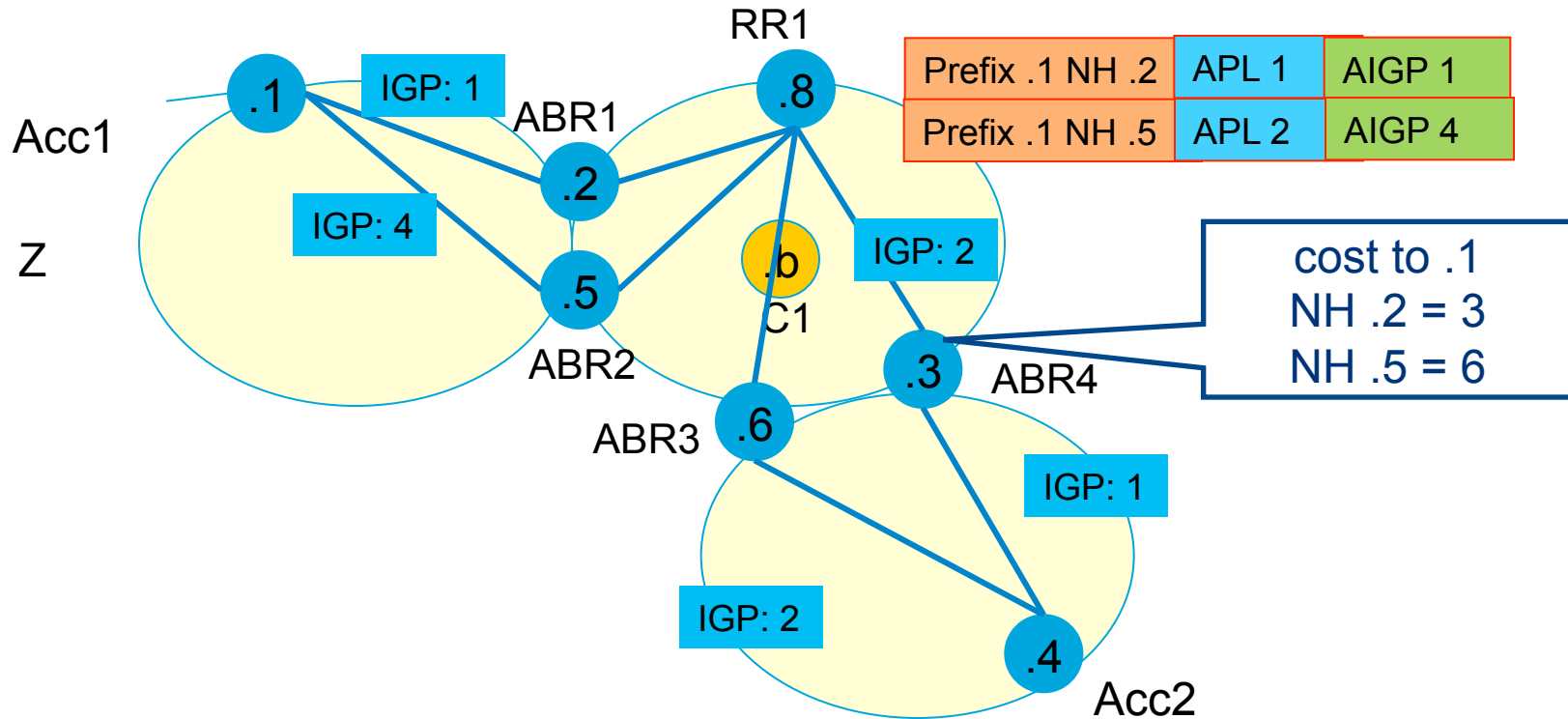


Reason: BGP cost is unchanged.

Сигнализация Multi-Path

Today XR
Radar IOS

Accumulated IGP Attribute for BGP



BGP accumulated IGP allows to carry and accumulate IGP cost across different intra-MPLS domains, providing end-end path optimization.

Поддержка функциональности

Cisco IOS/IOS XR

Next Hop Self on Route Reflector for iBGP + Label

- IOS XR – 4.1.1 для CRS and ASR9k

Unified MPLS iBGP Client

- IOS – 15.2(4)S для 7600 Sup720/MSFC3 RSP720/MSFC4

IPv4 BGP Label Distribution

- IOS – 12.0S, 12.2S, 12.3 и выше
- IOS XE – 2.2 и выше
- IOS XR – 3.5 и выше

Thank you.

