



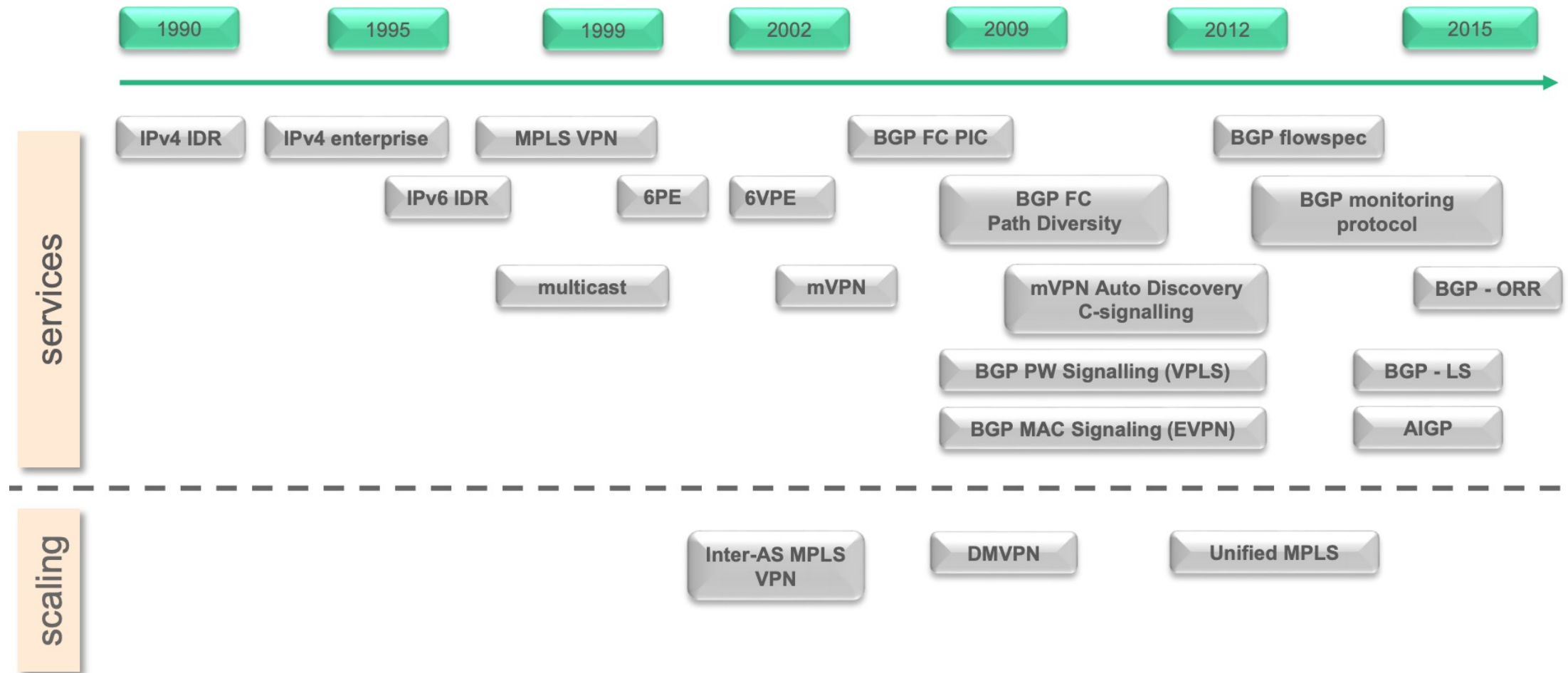
Networking
For everyone

Масштабируемость BGP

Слишком много BGP сервисов



Networking
For everyone



Утилизация памяти – проблемы и решения



Networking
For everyone

- Префиксы
 - Суммаризация
 - Фильтрация
 - Частичная инсталляция префиксов в RIB
- Пути
 - Уменьшение количество соседей
- Атрибуты
 - Фильтровать community



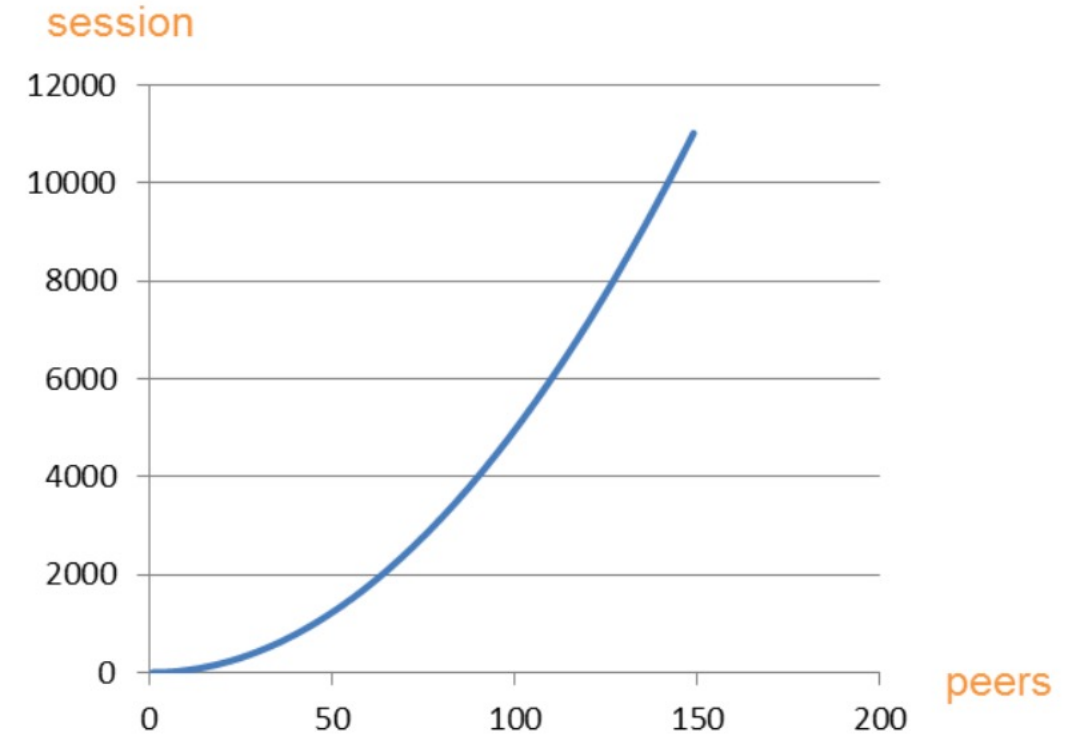
Networking
For everyone

Масштабируемость iBGP



iBGP масштабируем?

- Внутри одной AS необходима полная связность
- Полносвязная топология – $O(n^2)$
- Два основных решения
 - Отражатели маршрутов
 - конфедерации





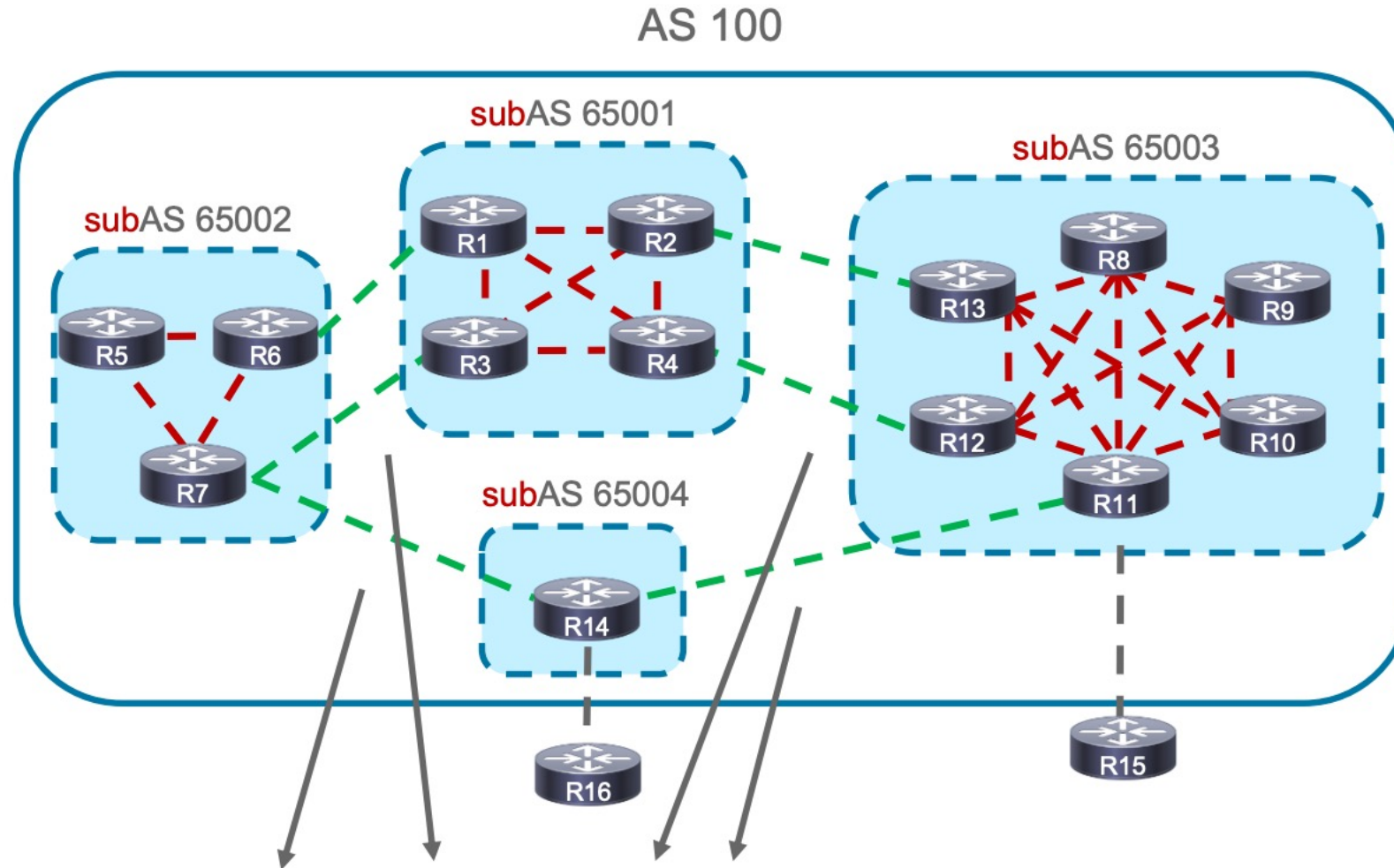
Конфедерации

- Суб-AS внутри родительской автономной системы
- Конфедерационные AS видны как обычная AS для внешних соседей
- Внутри суб-AS требуется полносвязный iBGP
- Между суб-AS – произвольная связность (eBGP!)
- Каждая суб-AS может иметь свой собственный IGP
- Есть маленькая 😊 проблема – нереально внедрить на уже работающей сети
 - Требуется полная перенастройка всей BGP системы

Конфедерации



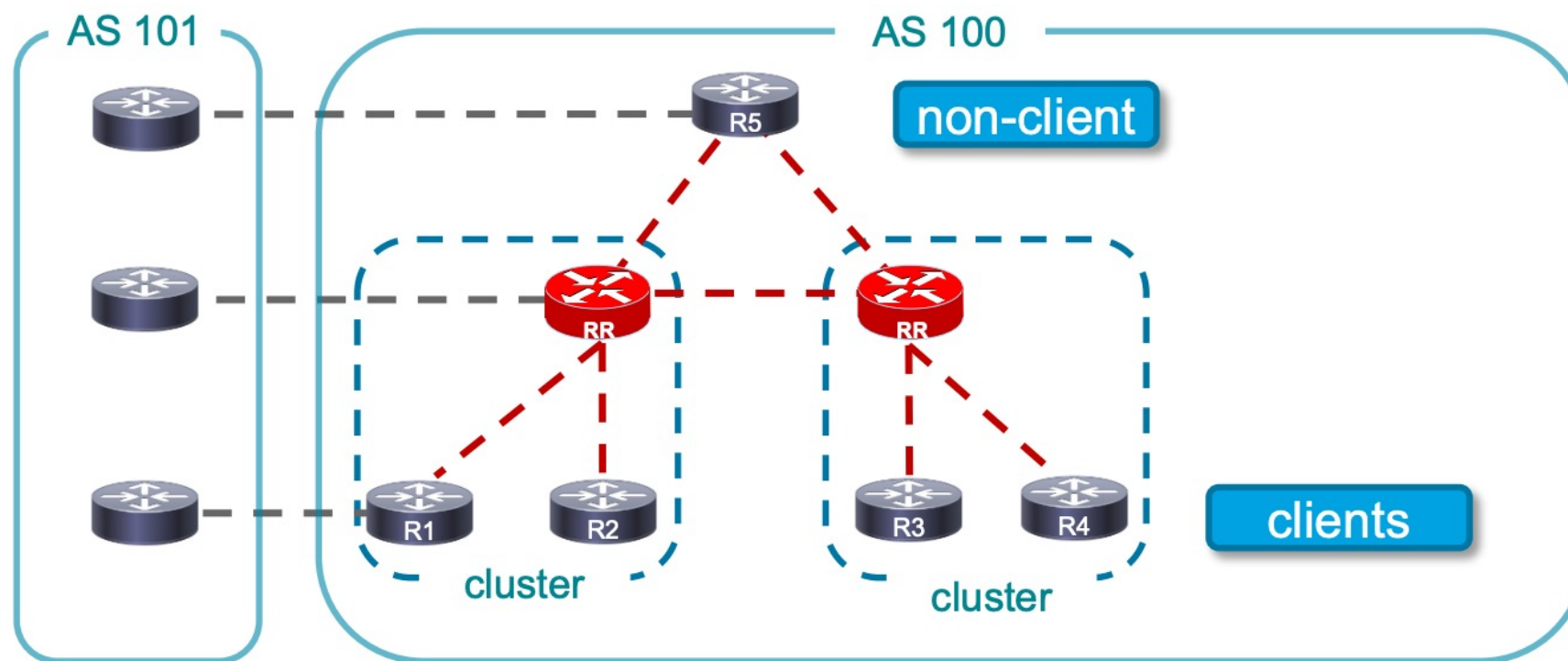
Networking
For everyone



Масштабируемость BGP

Отражатели маршруты (Route Reflectors, RR)

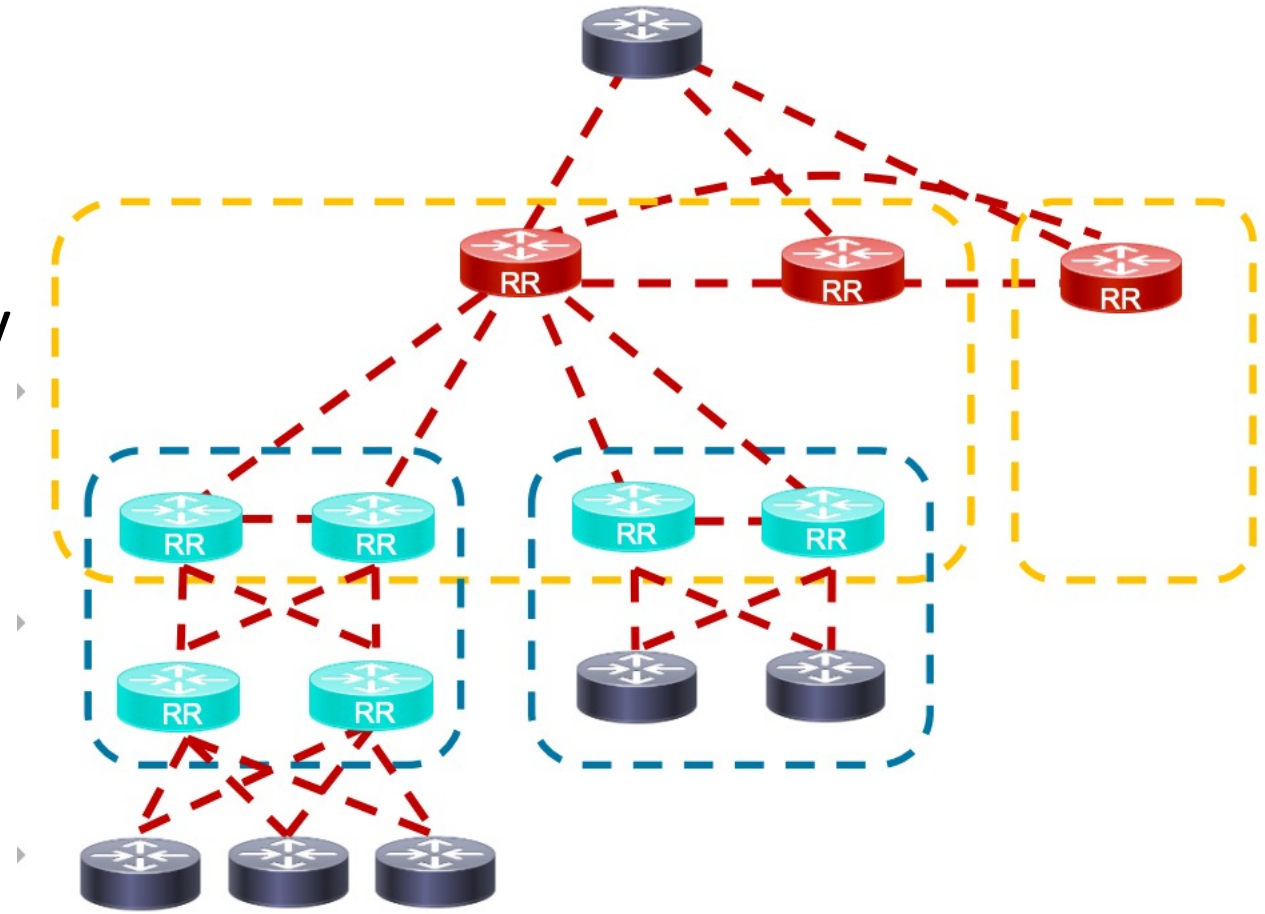
- Встречаются гораздо чаще, чем конфедерации
- Полносвязная топология преобразуется в Hub-Spoke
 - RR выступают в роли Hub
- Клиенты ничего не знают про RR





А если сеть очень большая?

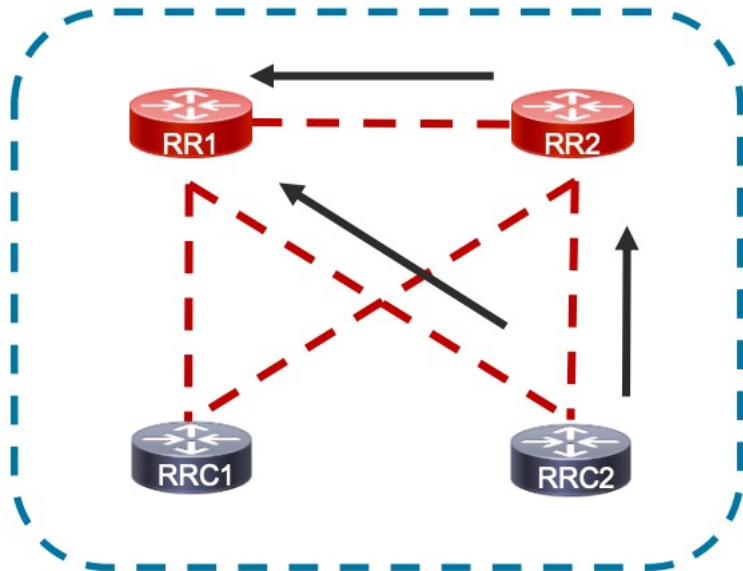
- Можно сделать RR клиентом другого RR
 - иерархия
- Выбирать RR лучше всего с учётом физической топологии
- Нет ограничений по количеству уровней



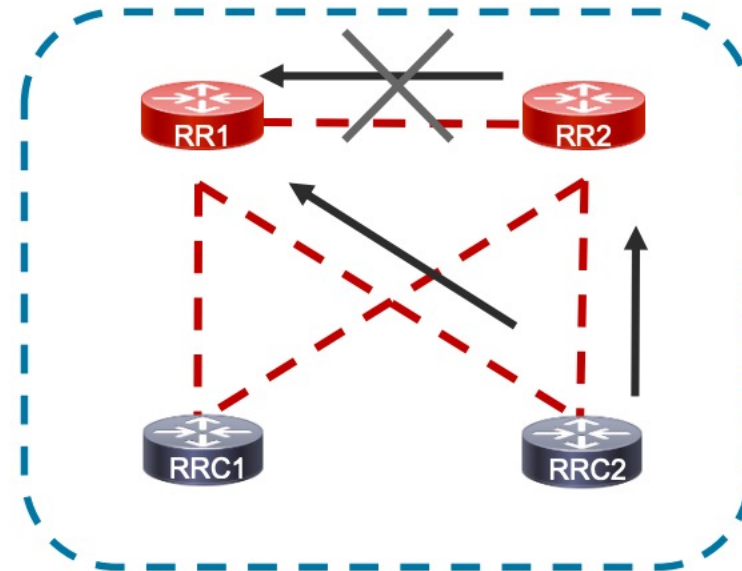


Cluster-id

- Разные
 - Требуется больше памяти



- Одинаковые
 - Могут «теряться» пути при отказах



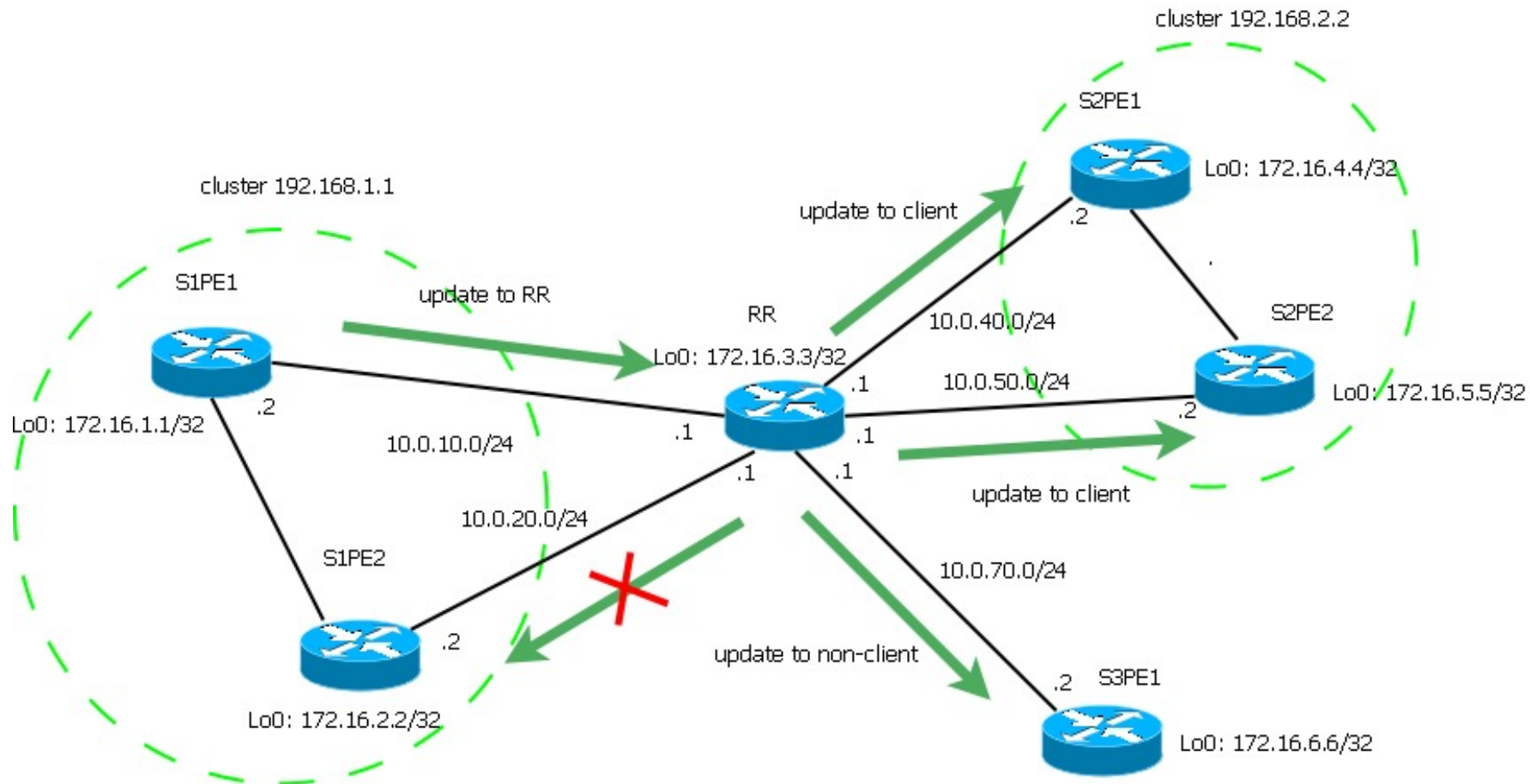


Выборочная установка маршрутов в RIB

- Если RR не стоит на пути Data Plane, зачем ему иметь префиксы в RIB/FIB?
- Позволяет сохранить ресурсы ЦП и TCAM
- Реализуется с помощью фильтра **table-map**
- Необходимо только для IPv4/IPv6
 - Нет необходимости в применении для VPNv4/VPNv6

Multi Cluster

- Один RR может относиться к нескольким кластерам
- Можно отключить передачу маршрутов client-client
 - *no bgp client-to-client reflection intra-cluster cluster-id 192.168.1.1*





Networking
For everyone

Масштабируемость eBGP



BGP Route Server

- Аналог RR для eBGP
- Часто используется в IX
- Увеличивается прозрачность
- Не изменяется AS-PATH
- Не изменяется Next-Hop
- Остаётся возможность применения политик/фильтрации

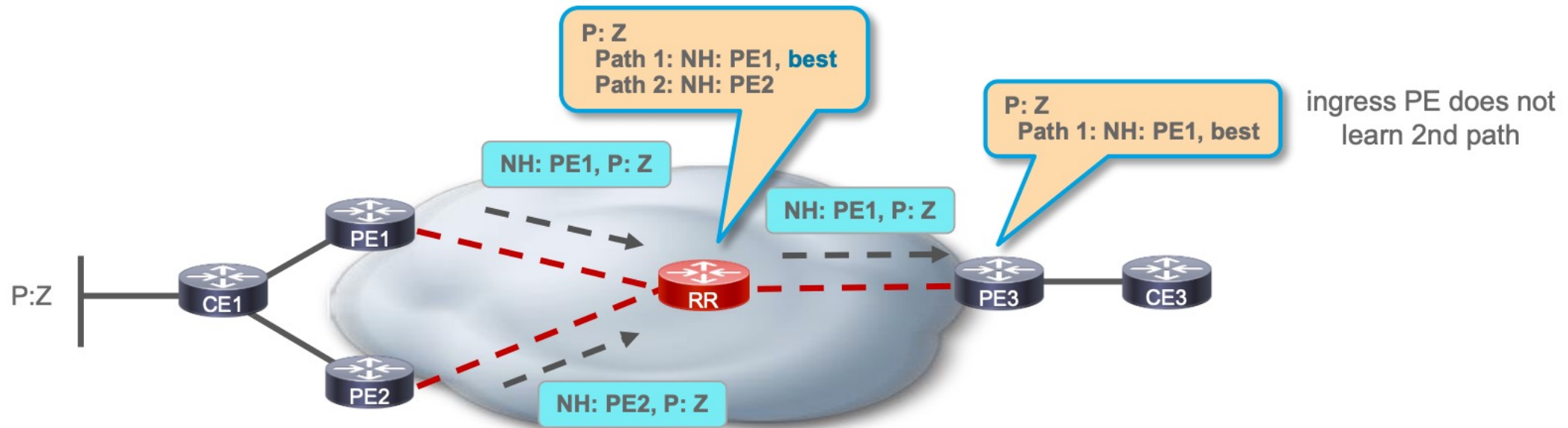


Networking
For everyone

Проблемы при
использовании RR и их
решения

Выбор наилучшего пути

- RR выбирает один лучший путь и отражает его своим клиентам
- В случае multihoming подключений это может быть проблемой





Зачем иметь несколько путей?

- Скорость конвергенции
 - BGP Fast Convergence
 - Несколько путей в BGP таблице
 - BGP PIC Edge
 - Несколько путей в FIB
- Балансировка
 - ECMP
- Hot potato маршрутизация
 - Оптимальная маршрутизация с т.з. PE

Как распространить несколько путей?

- BGP shadow RR
- BGP shadow session
- BGP Add-Path
- BGP Optimal Route Reflection



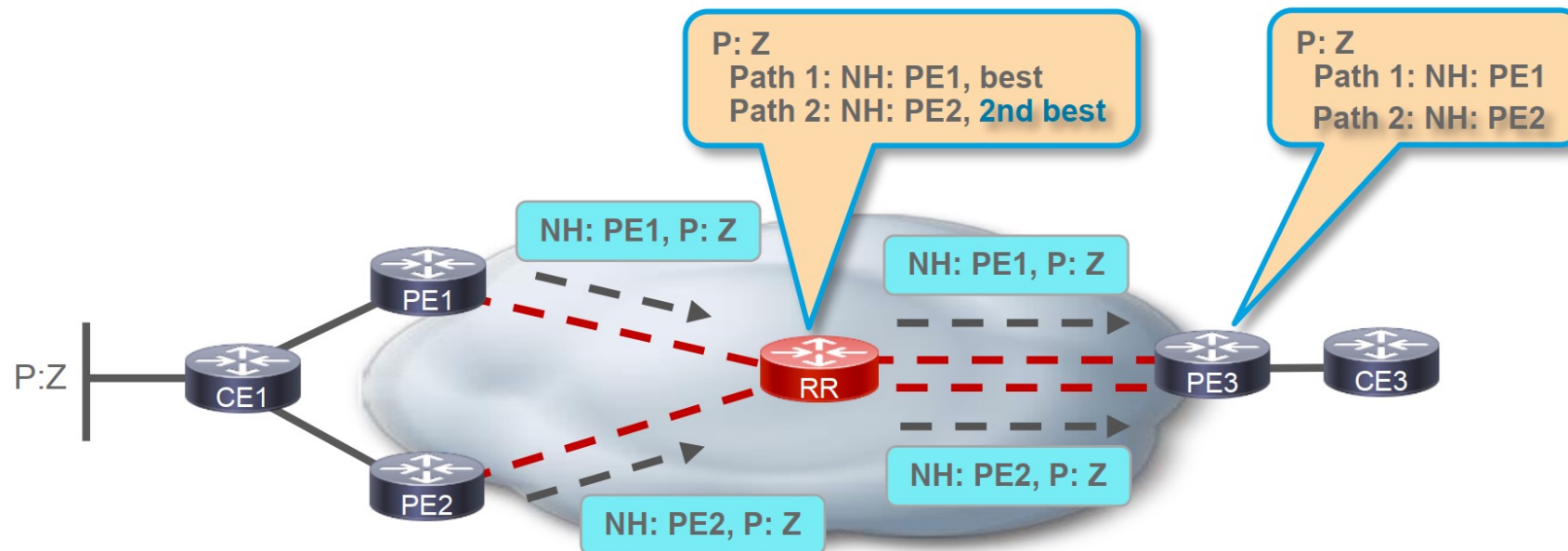
BGP Shadow RR

- Основной RR работает как обычно – отдаёт лучший маршрут
- Дополнительный RR отдаёт 2-ой лучший маршрут
 - Чаще всего он будет отличаться от основного маршрута атрибутом Next-Hop
- В результате на PE есть два маршрута «на выбор»
- Есть вопросы как размещать основной RR и shadow – на одном устройстве или нет



BGP shadow session

- Строится несколько сессий от RR до PE
- В рамках второй сессии отдаётся альтернативный маршрут
 - По сути это функционал Shadow RR, перенесённый на multisession
- Требуется поднятие дополнительных Loopback интерфейсов для настройки 2-ой сессии





Add-path

- Возможность передавать несколько путей для одного префикса
- Используется дополнительное capability
 - Все PE должны понимать это Capability
- Для отслеживания путей используется новый идентификатор path-identifier



Возможности Add-Path

add-all-path

- RR делает расчёт лучшего пути и отправляет все пути к соседям
- Используется для:
 - ESMР
 - обеспечение hot potato

add-n-path

- Соседям отправляются N путей
- Используется для:
 - N+1 резервирование
 - N = 3 для IOS
 - N = 2 для IOS-XR

multipath

- Соседям отправляются мультипасы
- Используется для:
 - балансировка
 - доступно только на IOS-XR



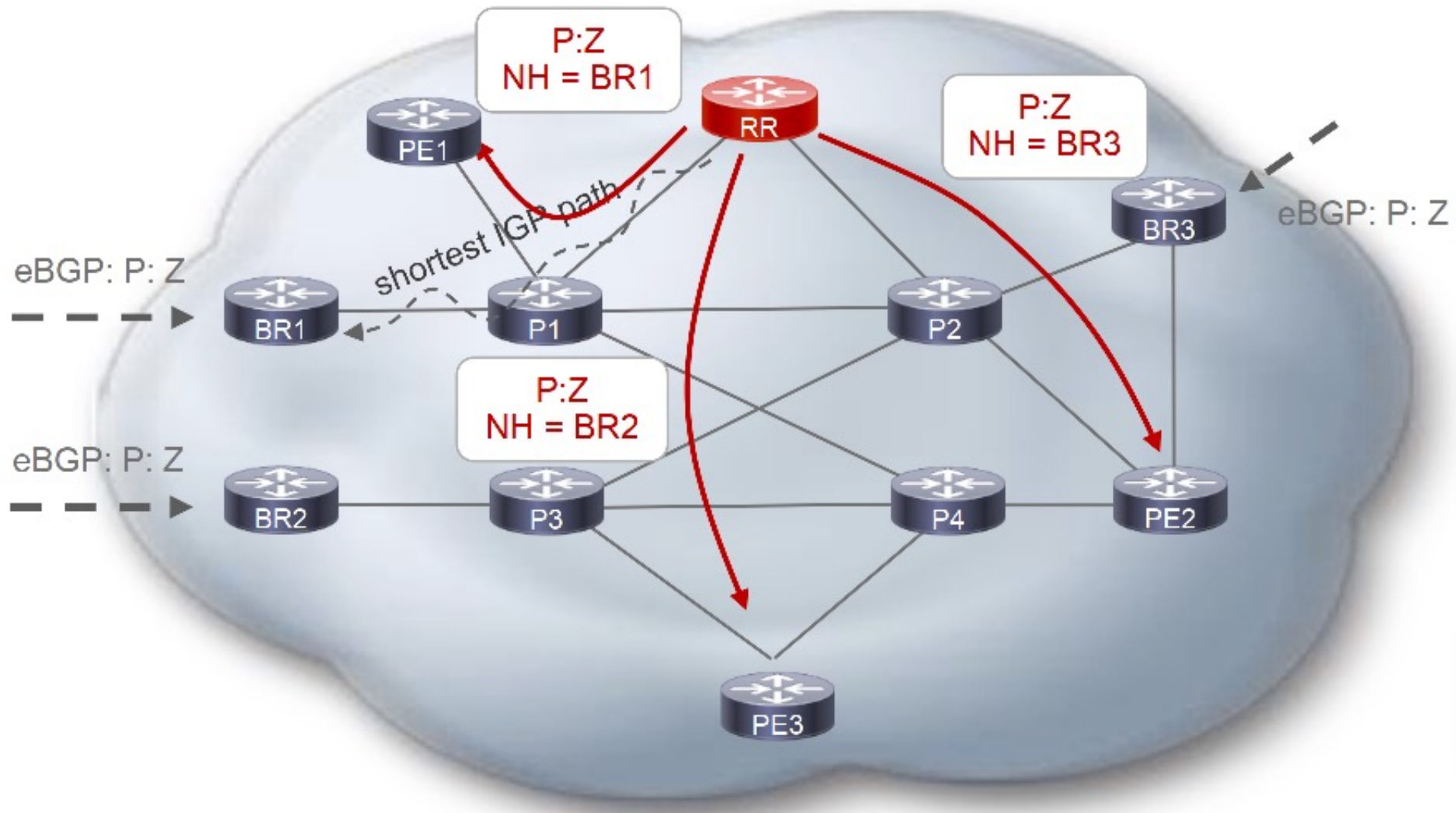
BGP Optimal Route Reflection

- У RR появляется возможность отдавать разные пути разным клиентам
- Работает только, если лучший путь был выбран на основе параметра **IGP cost**
- RR запускает алгоритм поиска лучшего пути, смотря на него глазами клиента
 - как rSPF в LFA
- Обязательно наличие link-state протокола маршрутизации
- От клиента ничего не требуется
- RR должен быть виртуальным с огромным количеством CPU и RAM

BGP Optimal Route Reflection



Networking
For everyone





Networking
For everyone