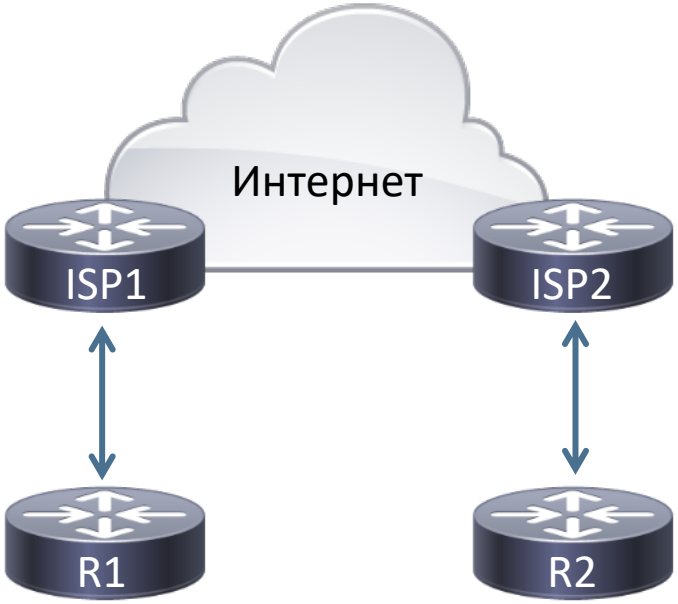
Networking
For everyone

Введение в BGP

BGP

Networking For everyone

- Border Gateway Protocol v4, RFC 4271 и много других
  - Протокол динамической маршрутизации класса EGP
    - Наследник протокола EGP (Exterior Gateway Protocol, RFC 827)
    - Не путайте EGP и EGP, это вам не это!
  - Работает между автономными системами
  - Multiprotocol Extensions (MP-BGP, RFC 4760)
- Использует TCP-подключение на порт 179
- Масштабируемый и безопасный
  - Создан для работы с миллионами маршрутов
    - Протокол-то выдержит, а вот железо…
  - Создан для работы с недоверенными пирами
  - Не создан для быстрой сходимости



Особенности IGP и EGP

Networking For everyone

- В своей AS все устройства находятся под общим контролем
  - Любые устройства – доверенные
  - Разумное (и предсказуемое) количество префиксов
  - Трафик отправляется по оптимальному маршруту
  - Скорость сходимости сети важна и подконтрольна
  - Если что-то не нравится, можно перенастроить
  - В случае возникновения проблемы понятны зоны ответственности
- При стыке с чужими AS все наоборот

Номера AS

Networking For everyone

- Важный параметр для работы BGP
  - Изначально определен как 16-битное число (0-65535)
  - В 2007 году расширен до 32 бит (поддержка обязательна с 2010 года)
    - Нотация ASPLAIN – одно десятичное число (например, 65550)
    - Нотация ASDOT – два 16-битных десятичных числа через точку (например, 1.14)
- Распределяются IANA через региональные регистраторы (RIR)
  - Зарезервированные IANA значения: 0, 65535 и 65535.65535
  - Значения для использования в документации: 64496–64511 и 1.1–1.15
  - Частные номера: 64512–65534 и 64086.59904 (4200000000)–65535.65534

Особенности BGP

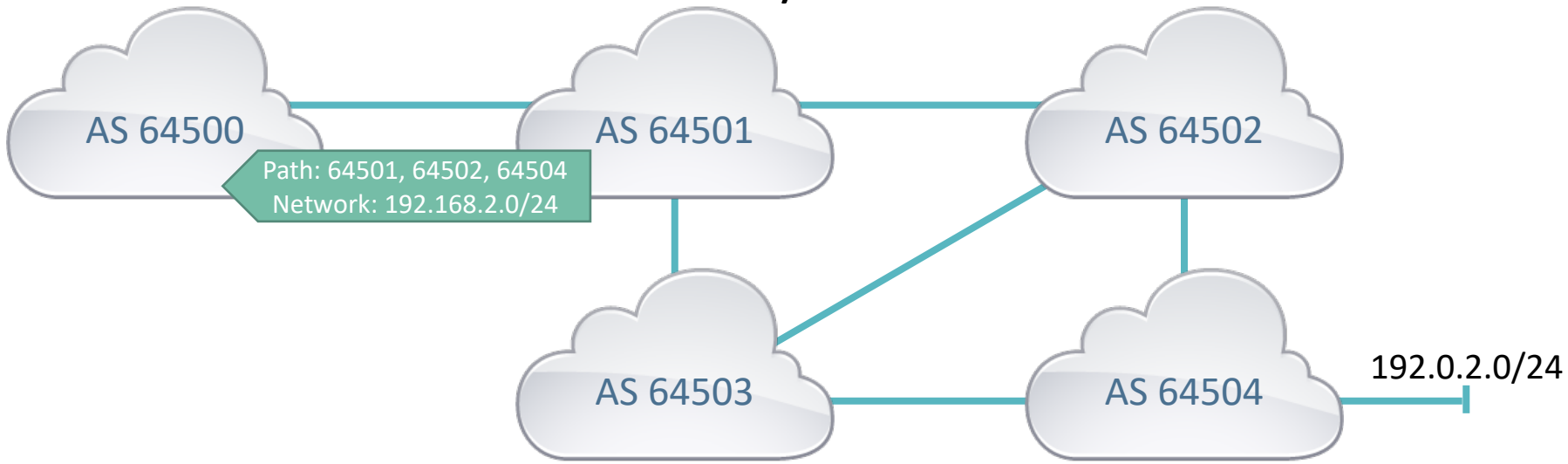
Networking For everyone

- BGP принимает "чужие" маршруты
- BGP выбирает один лучший маршрут до сети (best path)
  - Выбор основывается на атрибутах пути, которые не отражают физические свойства маршрута; лучший – не обязательно самый быстрый
  - Исходящий трафик можно контролировать, выбирая лучший маршрут
- BGP анонсирует "свои" префиксы
  - Пользуются анонсируемыми нами маршрутами наши пиры
  - Контролировать выбор маршрута пирами в общем случае нельзя
- Самое главное в BGP – политики
  - Политики приема маршрутов от пира
  - Политики выбора маршрутов для отправки пиру
  - Политики модификации маршрутов и выбора лучшего маршрута

Принцип работы BGP

Networking For everyone

- Path Vector Algorithm
  - Дистанционно-векторный алгоритм дополнен информацией о пути трафика
  - Роутеры получают часть информации о топологии (но не саму топологию)
- BGP отправляет в анонсах лучшие пути трафика и их атрибуты
  - Префиксы, которые доступны по этому пути
  - Список автономных систем на пути



Характеристики BGP

Networking For everyone

- Использует TCP, порт 179
  - Требует ручного прописывания соседей
  - Keepalive-сообщения для проверки соседства
- Огромный вектор метрик (атрибутов)
  - Список автономных систем по пути маршрута (AS-Path)
  - Несколько вариантов приоритета
    - Приоритет в собственной AS (Local Preference)
    - Рекомендуемый приоритет в соседской AS (Multi-Exit Discriminator)
  - Next-hop
  - Происхождение маршрута (Origin)
  - …
- IP-адреса и номера AS соседей нужно прописывать вручную

Сообщения BGP

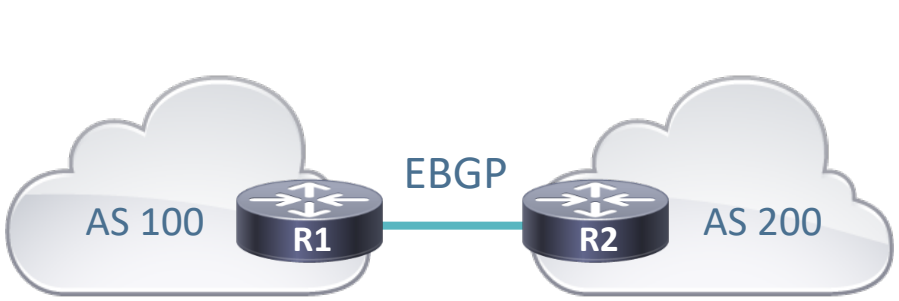
Networking For everyone

- В TCP-сессии BGP обменивается следующими сообщениями:
  - **Open**: согласование параметров соседства
  - **Keepalive**: проверка соседства
  - **Update**: передача маршрутной информации
  - **Notification**: завершение сессии

EBGP и IBGP

Networking For everyone

- BGP, как протокол EGP, рассчитан на работу в определенных условиях
  - Много маршрутов, в том числе и от неподконтрольных соседей
- Некоторые BGP-пиры могут быть из своей автономной системы
  - EBGP: сосед из чужой AS, режим тотального недоверия
  - IBGP: сосед из своей AS, режим частичного доверия
  - Протокол один и тот же, разные именно режимы работы

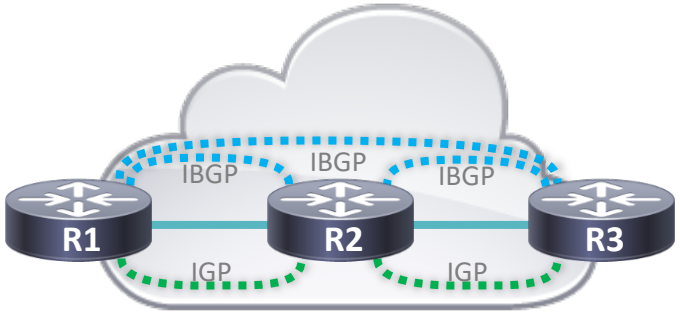


Характеристики EBGP и IBGP соседств

Networking For everyone

- Exterior BGP neighborship
  - Соседство между пирами разных AS
  - Должны быть в одном канале
  - Атрибуты модифицируются отправителем
    - В первую очередь - AS-Path и Next-Hop

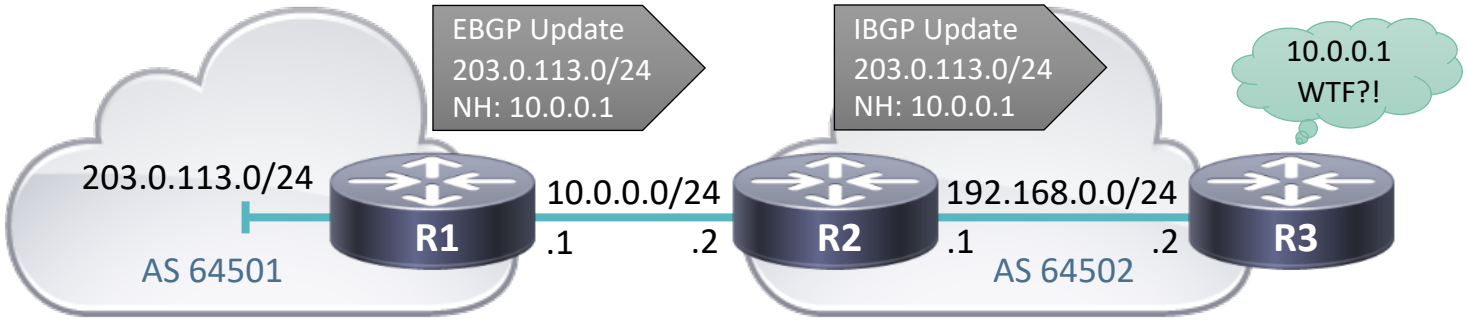
- Interior BGP neighborship
  - Соседство между пирами одной AS
  - Могут соединяться через транзитные роутеры
    - Требует IGP для нахождения пути до соседей
    - Требует полносвязность соседств
  - Атрибуты не модифицируются при отправке



IBGP и Next Hop

Networking For everyone

- При отправке маршрутов IBGP-соседу атрибуты не меняются
  - Как следствие, атрибут Next Hop для внешних маршрутов будет сохраняться
  - Обычно у IBGP-соседей нет маршрутов до этих адресов

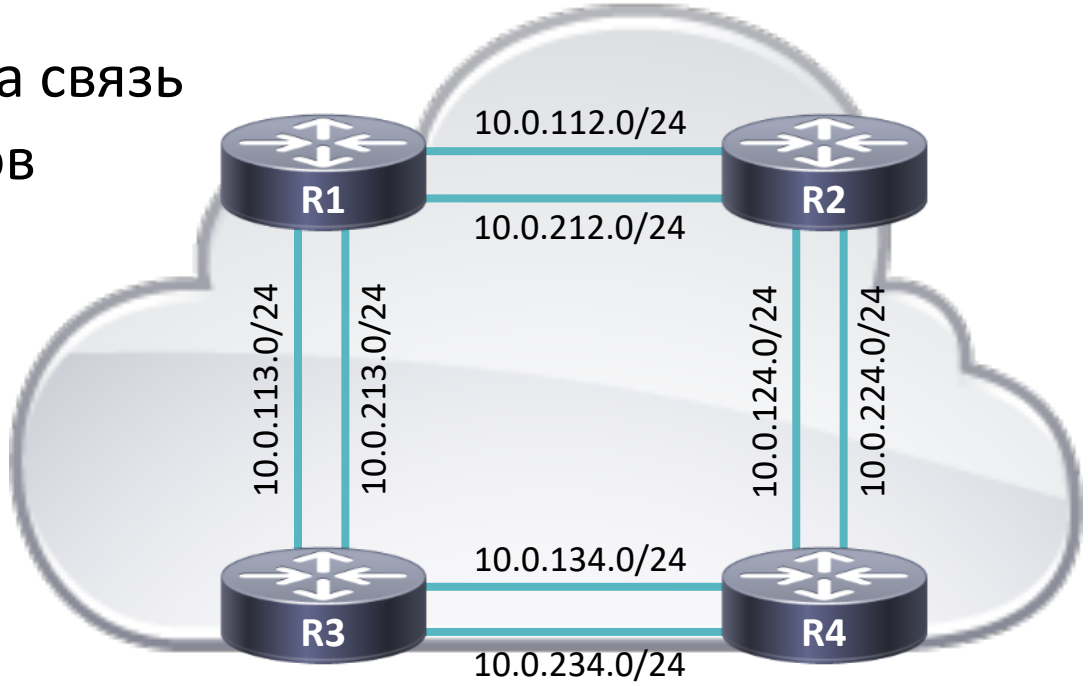


- Два варианта решения проблемы
  - Импортировать адреса EBGP-соседей в IGP
  - Поменять адрес Next Hop перед отправкой маршрутов IBGP-соседям

Дизайн соседства IBGP

Networking For everyone

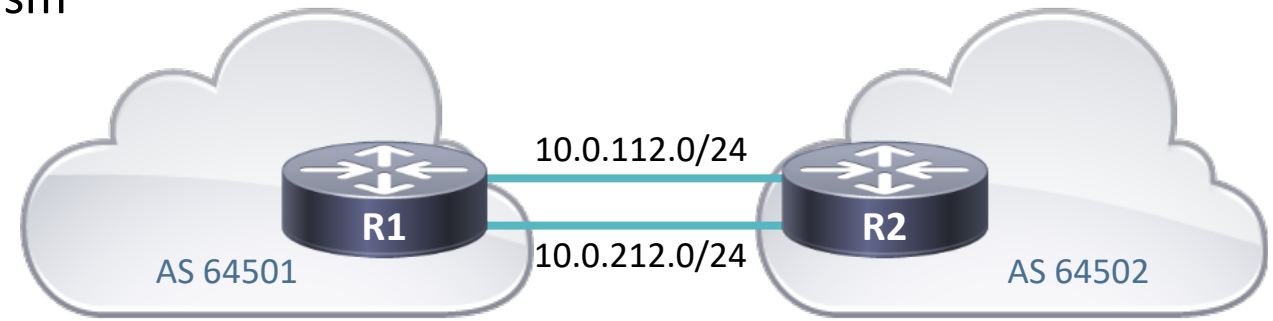
- Маршруты от IBGP-соседей не передаются другим IBGP-соседям
  - Между любыми двумя роутерами AS требуется соседство
  - На каждом роутере может быть много адресов
- Сессии ставятся с адресов виртуальных интерфейсов (Loopback)
  - Адреса анонсируются в IGP
  - Отказ отдельного канала не влияет на связь
  - Одна сессия на каждую пару роутеров



Дизайн соседства EBGP

Networking For everyone

- По умолчанию IP-пакеты до EBGP-соседа отправляются с TTL=1
  - А принимаются – с любым
  - Иногда это неудобно (например, если хочется поднять сессию с Loopback)
  - Иногда это небезопасно (можно устроить SYN DoS атаку)
- Проблему можно решить двумя способами:
  - **EBGP Multihop** предписывает отправлять пакеты с заданным TTL
  - **BGP TTL Security** предписывает отправлять пакеты всегда с TTL=255
    - А вот принимать – только с указанным, это легко проверить и тяжело подделать
    - Generalized TTL Security Mechanism
    - Нужно включать с обеих сторон



Состояния соседства BGP

Networking For everyone

- Соседство BGP проходит через состояния:
  - **Idle**: роутер не пытается установить соседство
    - Нет маршрута до соседа, наложенные санкции, и т.п.
  - **Active**: отправлен TCP SYN
  - **Connect**: получен TCP SYN (в том числе при входящей сессии)
  - **OpenSent**: TCP-сессия установлена, отправлено Open
  - **OpenConfirm**: получено корректное Open, ждем Keepalive
  - **Established**: все в порядке, соседство установлено

Проблемы с установлением соседства

Networking For everyone

- Неправильный IP-адрес соседа
  - Также недоизученный в IGP адрес IBGP-соседа
- Неправильный IP-адрес источника
  - Также недопрописанное соседство на соседе
- Ошибки в указании номеров AS
  - Open будет принято только от правильного IP-адреса и с правильной AS

Networking For everyone