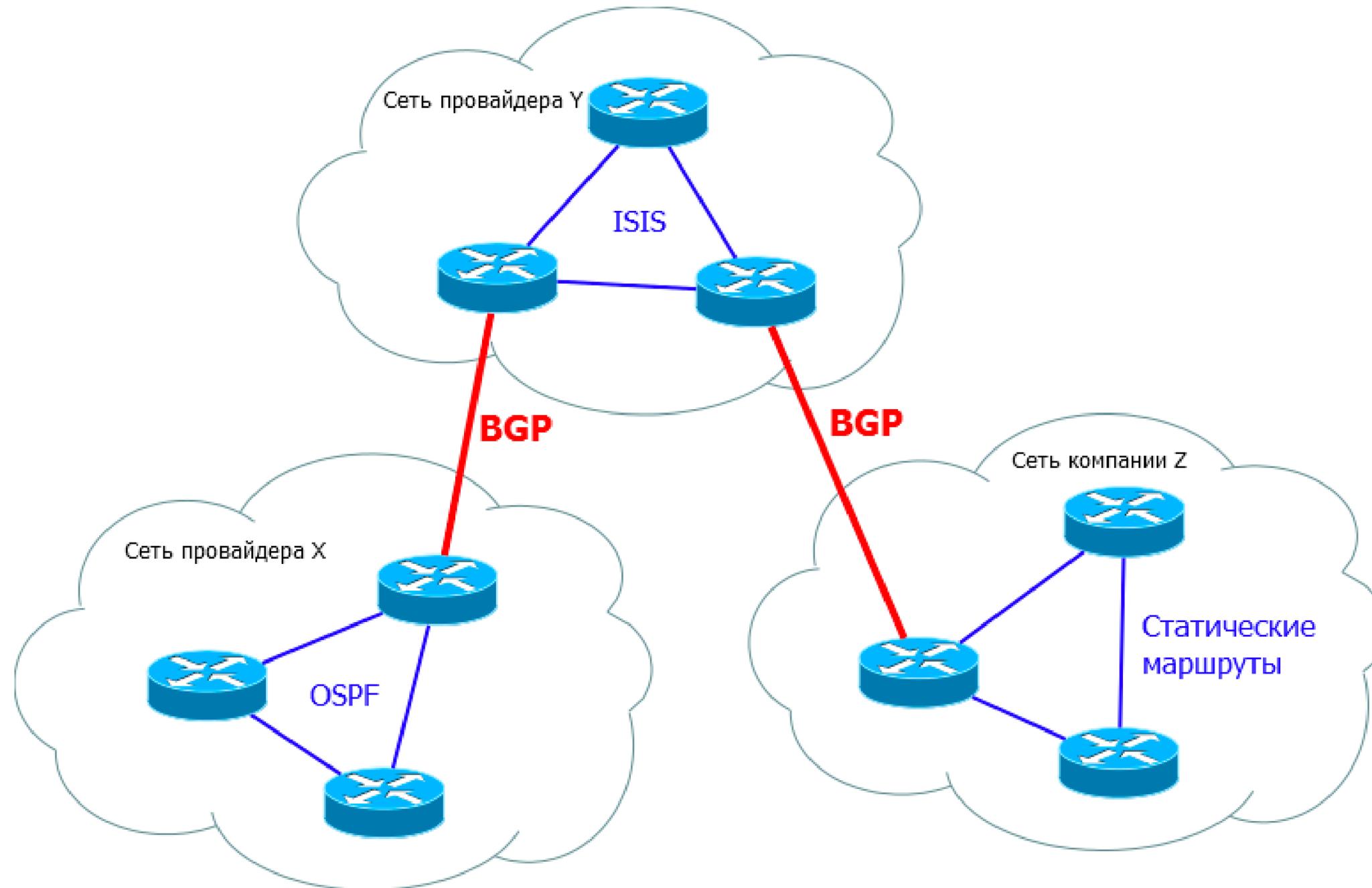


Принципы работы протокола BGP

Что такое BGP?

BGP — это протокол динамической маршрутизации, являющийся единственным EGP (External Gateway Protocol) протоколом. Данный протокол используется для построения маршрутов в интернете между отдельными автономными системами.

Протокол граничного шлюза



Один из основных атрибутов, который передается с информацией о маршруте — это список автономных систем, через которые прошла эта информация. Эта информация позволяет BGP определять где находится сеть относительно автономных систем, исключать петли маршрутизации, а также может быть использована при настройке политик.

Маршрутизация осуществляется пошагово от одной автономной системы к другой. Все политики BGP настраиваются, в основном, по отношению к внешним/соседним автономным системам. То есть, описываются правила взаимодействия с ними.

Терминология протокола

- **Внутренний протокол маршрутизации (interior gateway protocol)** – протокол, который используется для передачи информации о маршрутах внутри автономной системы.
- **Внешний протокол маршрутизации (exterior gateway protocol)** – протокол, который используется для передачи информации о маршрутах между автономными системами.
- **Автономная система (autonomous system, AS)** — набор маршрутизаторов, имеющих единые правила маршрутизации, управляемых одной технической администрацией и работающих на одном из протоколов IGP (для внутренней маршрутизации AS может использовать и несколько IGP).
- **Транзитная автономная система (transit AS)** — автономная система, через которую передается трафик других автономных систем.
- **Путь (path)** — последовательность состоящая из номеров автономных систем через которые нужно пройти для достижения сети назначения.
- **Атрибуты пути (path attributes, PA)** — характеристики пути, которые позволяют выбрать лучший путь.
- **BGP speaker** — маршрутизатор, на котором работает протокол BGP.
- **Соседи (neighbor, peer)** — любые два маршрутизатора, между которыми открыто TCP-соединение для обмена информацией о маршрутизации.
- **Информация сетевого уровня о доступности сети (Network Layer Reachability Information, NLRI)** — IP-префикс и длина префикса.

Информация о маршруте в BGP состоит из:

- Автономной системы (одна или несколько IP сетей)
- Вектора пути (список промежуточных автономных систем)
- Следующий маршрутизатор

Политика маршрутизации:

- можно выбирать, какие автономные системы анонсировать, а какие нет

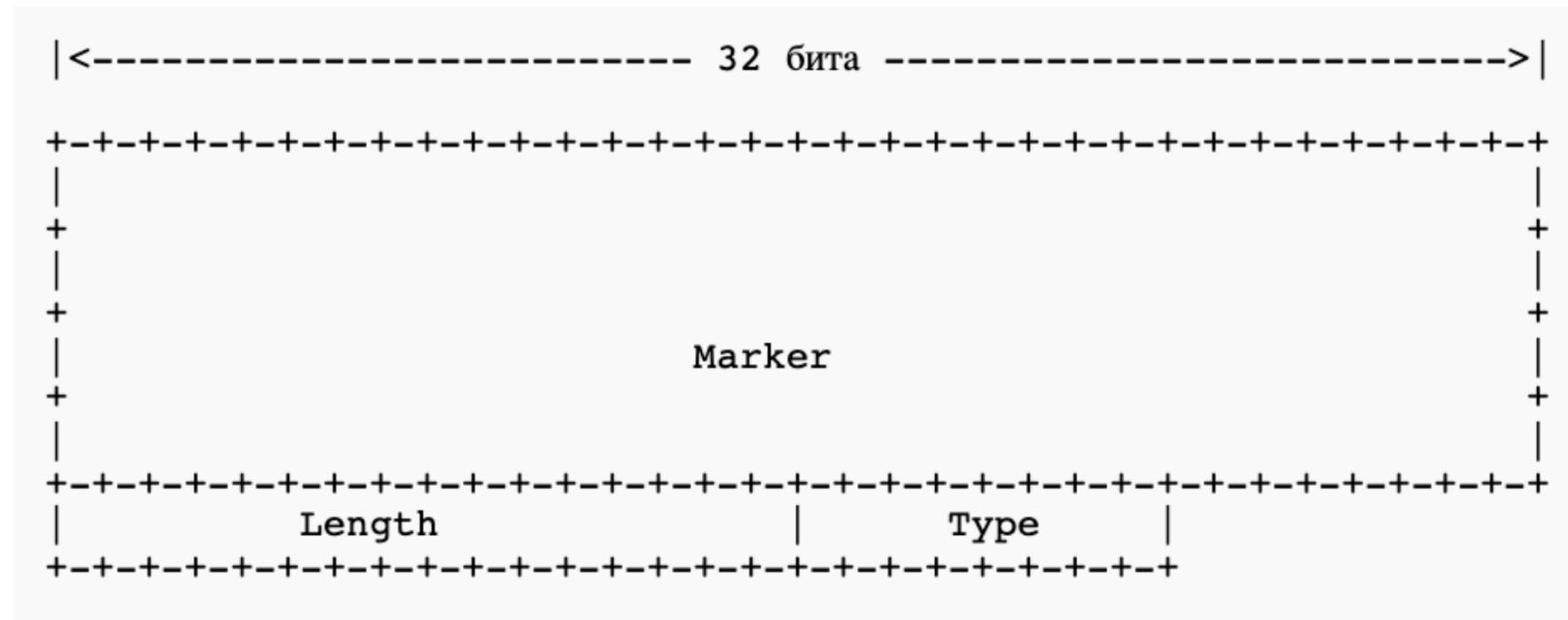
Общие характеристики

- Использует TCP для передачи данных, это обеспечивает надежную доставку обновлений протокола (порт 179)
- Отправляет обновления только после изменений в сети (нет периодических обновлений)
- Периодически отправляет keepalive-сообщения для проверки TCP-соединения
- Метрика протокола называется path vector или атрибуты (attributes)

Внутренний BGP (Internal BGP) и Внешний BGP (External BGP)

- Внутренний BGP (Internal BGP, iBGP) — BGP работающий внутри автономной системы. iBGP-соседи не обязательно должны быть непосредственно соединены. Другими словами это можно назвать внутренними пирингами, которые следуют внутренним правилам BGP. Используются при формировании пирингов внутри автономной системы.
- Внешний BGP (External BGP, eBGP) — BGP работающий между автономными системами. По умолчанию, eBGP-соседи должны быть непосредственно соединены. Другими словами - пиринг между автономными системами.

Типы сообщений BGP



- Marker — поле, которое включено в заголовок для совместимости. Размер поля — 16 байт, все байты должны быть 1.
- Length — длина всего сообщения в октетах, включая заголовок. Поле может принимать значения от 19 до 4096.
- Type — тип передаваемого сообщения:
 - 1 — OPEN
 - 2 — UPDATE
 - 3 — NOTIFICATION
 - 4 — KEEPALIVE

Типы передаваемого сообщения

- **Open** — используется для установки отношений соседства и обмена базовыми параметрами. Отправляется сразу после установки TCP-соединения.
- **Update** — используется для обмена информацией маршрутизации.
- **Notification** — используется когда возникают ошибки BGP. После отправки сообщения сессия с соседом разрывается.
- **Keepalive** — используется для поддержания отношений соседства, для обнаружения неактивных соседей.

Состояние соседства

Состояние подключения - это состояние, в котором устройство BGP ожидает завершения TCP- соединения с соседним устройством.

Состояние	Ожидание TCP?	Инициация TCP?	Установлено TCP?	Отправлено Open?	Получено Open?	Сосед Up?
Idle	Нет					
Connect	Да					
Active	Да	Да				
Open sent	Да	Да	Да	Да		
Open confirm	Да	Да	Да	Да	Да	
Established	Да	Да	Да	Да	Да	Да

Атрибуты пути

- **Well-known mandatory** — все маршрутизаторы, работающие по протоколу BGP, должны распознавать эти атрибуты. Должны присутствовать во всех обновлениях (update).
- **Well-known discretionary** — все маршрутизаторы, работающие по протоколу BGP, должны распознавать эти атрибуты. Могут присутствовать в обновлениях (update), но их присутствие не обязательно.
- **Optional transitive** — могут не распознаваться всеми реализациями BGP. Если маршрутизатор не распознал атрибут, он помечает обновление как частичное (partial) и отправляет его дальше соседям, сохраняя не распознанный атрибут.
- **Optional non-transitive** — могут не распознаваться всеми реализациями BGP. Если маршрутизатор не распознал атрибут, то атрибут игнорируется и при передаче соседям отбрасывается.

Выбор пути

Характеристики процедуры выбора пути протоколом BGP:

- В таблице BGP хранятся все известные пути, а в таблице маршрутизации — лучшие.
- Пути выбираются на основании политик.
- Пути не выбираются на основании пропускной способности.

Сначала проверяется:

- Доступен ли next-hop (Route Resolvability Condition)
- Для того чтобы next-hop считался доступным (accessible), необходимо чтобы в таблице маршрутизации был IGP-маршрут, который ведет к нему.