



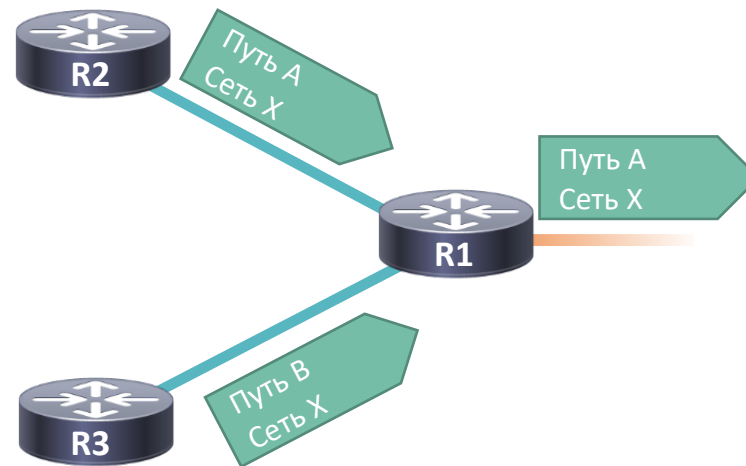
Networking
For everyone

Атрибуты BGP



Выбор лучшего маршрута

- BGP по умолчанию не использует балансировку
 - Его задача – найти одну лучшую точку выхода из автономной системы
 - Балансировать трафик одного приложения между разными AS – плохая идея
 - Если хочется, в пределах AS балансировку можно делать с помощью IGP
- Выбор лучшего маршрута (best path selection) основан на атрибутах

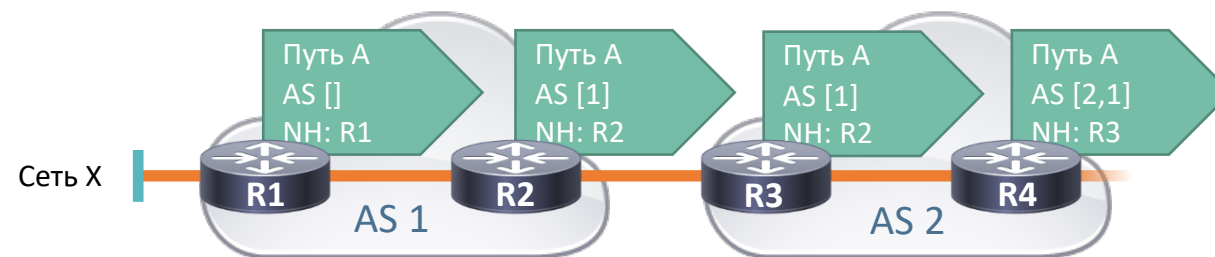
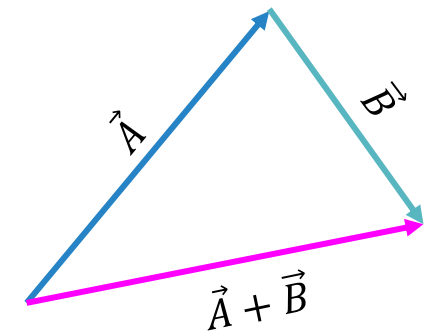




Основные атрибуты BGP

- Маршрутизаторы обмениваются векторами метрик для путей
- Вектор метрик в анонсе складывается с вектором до соседа

Атрибут	Описание
NEXT_HOP	IP-адрес Next Hop, меняется при отправке EBGP-соседу или политикой
AS_PATH	Список номеров AS для пути, дополняется при отправке EBGP-соседу
LOCAL_PREF	Приоритет маршрута в локальной AS, назначается политикой
MULTI_EXIT_DISC	Приоритет маршрута в соседской AS, назначается или наследуется
ORIGIN	Код происхождения маршрута, назначается или наследуется
Community	Список дополнительных меток, назначается политикой





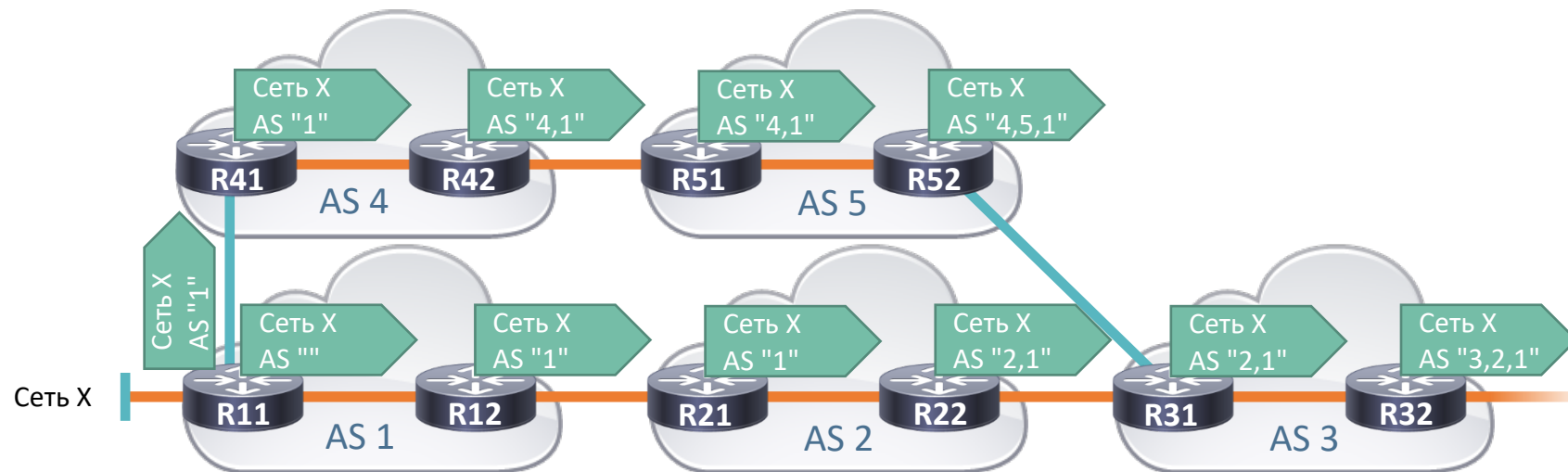
NEXT_HOP

- Адрес, который устанавливается в RIB в качестве шлюза для сети
 - Обычно не меняется при передаче по IBGP
 - Обязан резолвиться в RIB (хотя бы рекурсивно)
 - Может быть доступен с учетом балансировки
- Устанавливается в собственный адрес для:
 - Анонсов локально импортированных маршрутов
 - EBGP-анонсов



AS_PATH

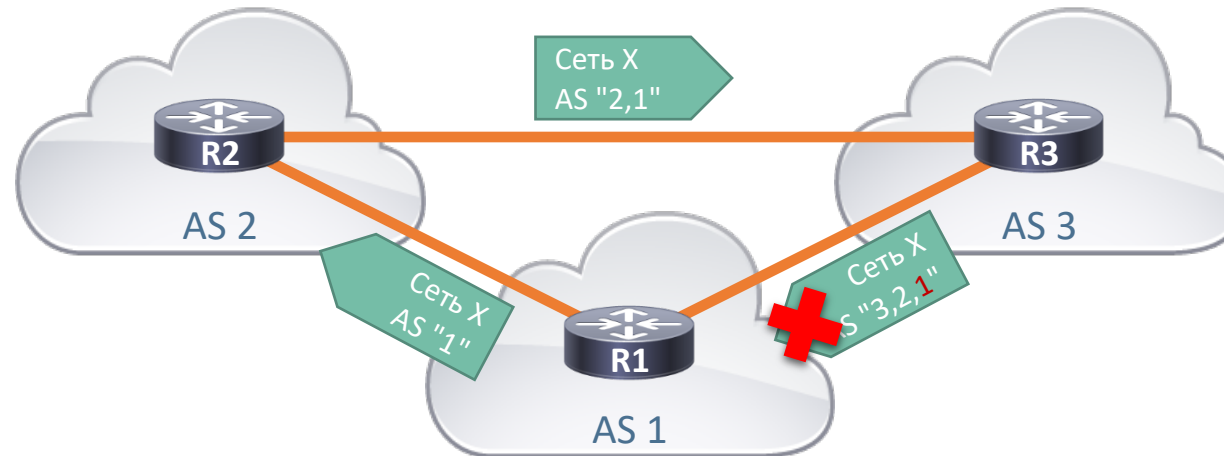
- Список автономных систем, через которые проложен маршрут
 - Часто записывается в виде строки с номерами АС через пробел или запятую
 - Такая строка также может содержать неожиданные знаки препинания
- У кого длиннее, тот лох





Свойства AS_PATH

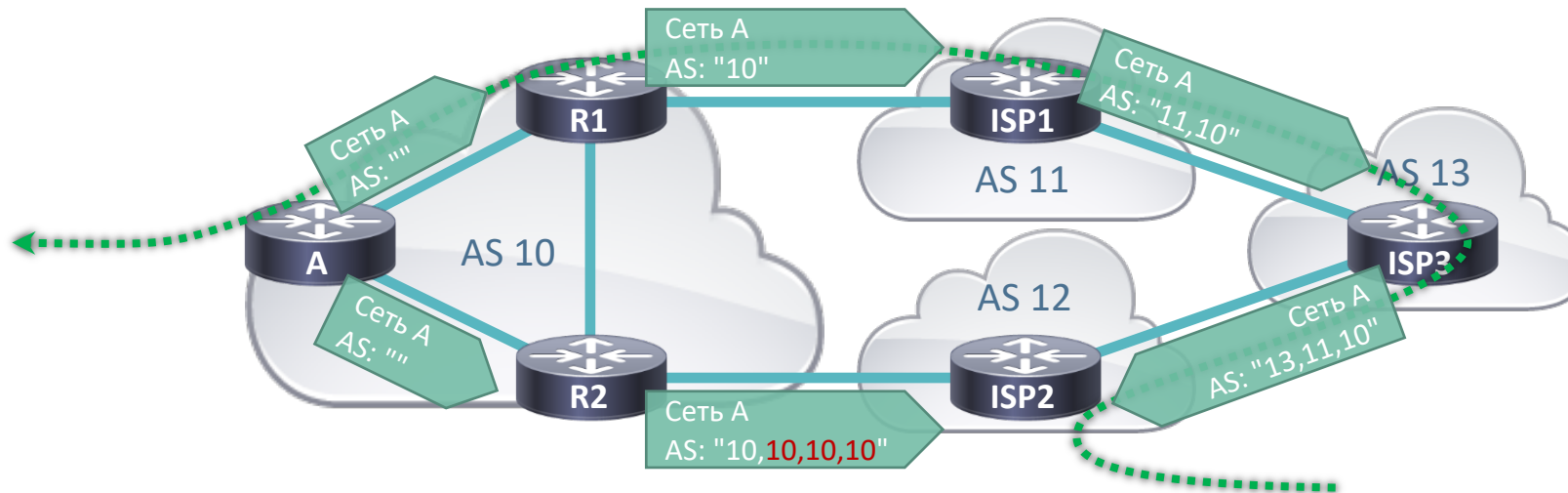
- На самом деле™ состоит из 4 субкомпонентов
 - AS_SEQUENCE, AS_SET, AS_CONFED_SEQ, AS_CONFED_SET
- При отправке маршрута EBGP-соседу AS_SEQUENCE дополняется "слева" номером собственной автономной системы
- Служит инструментом защиты от петель между AS
 - Маршрутизатор не примет от EBGP-соседа маршрут, если в AS_PATH указан номер его автономной системы





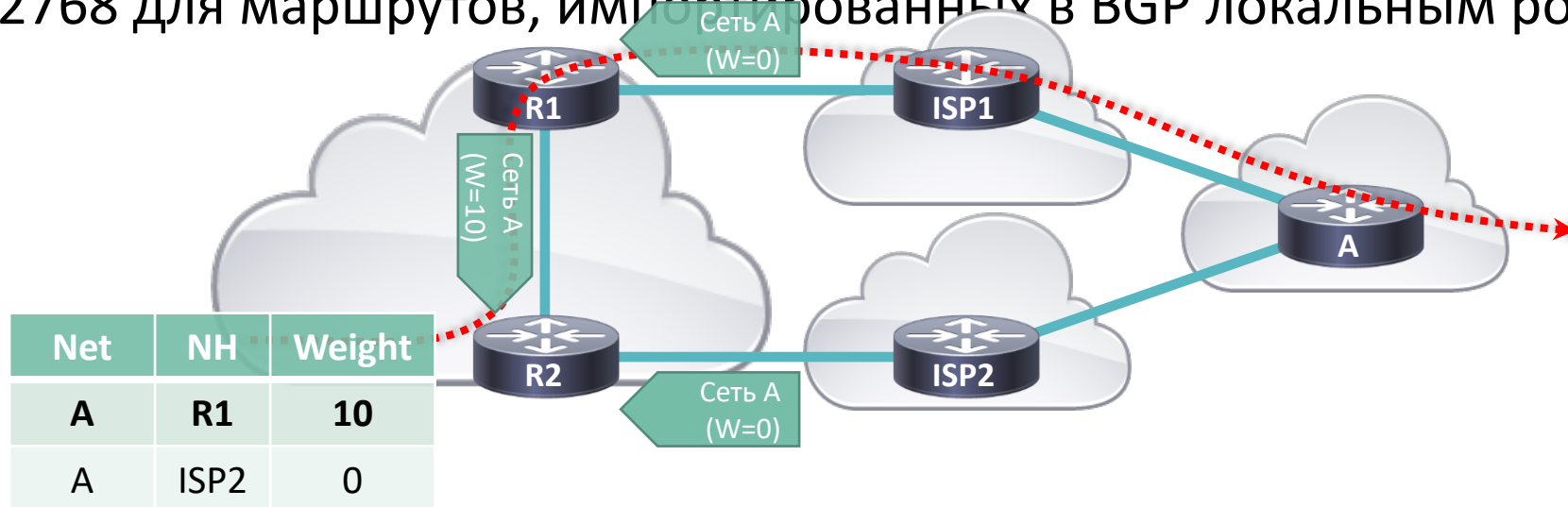
AS-Path Prepending

- Настройка требуется для управления исходящим трафиком
 - Настраивается через Route Map, обычно на EBGP-соседстве
 - Искусственно удлиняет AS-Path добавлением номера собственной AS
 - Также номер своей AS добавляется при отправке EBGP Update отдельно от Prepending



Weight

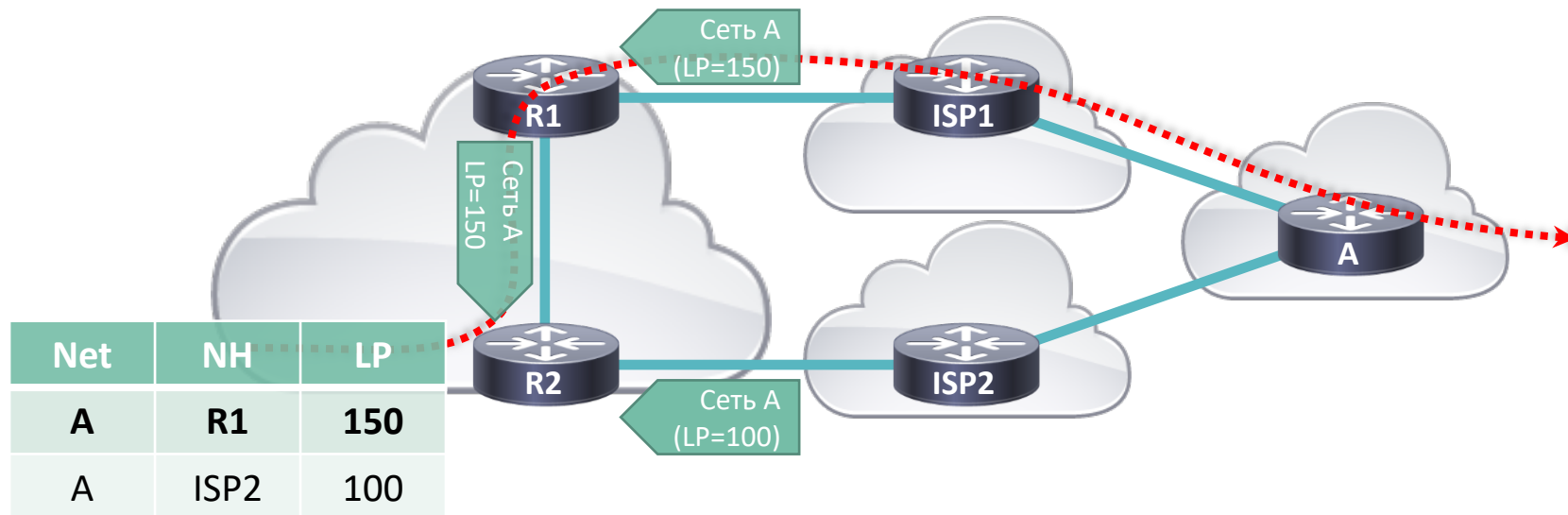
- Проприетарный атрибут Cisco
 - Локальный, назначается администратором, не передается в Update
 - 16-битное число, больше = лучше
 - инструмент выбора маршрута для исходящего трафика
- По умолчанию:
 - 0 для маршрутов, изученных от соседей
 - 32768 для маршрутов, импортированных в BGP локальным роутером





LOCAL_PREF

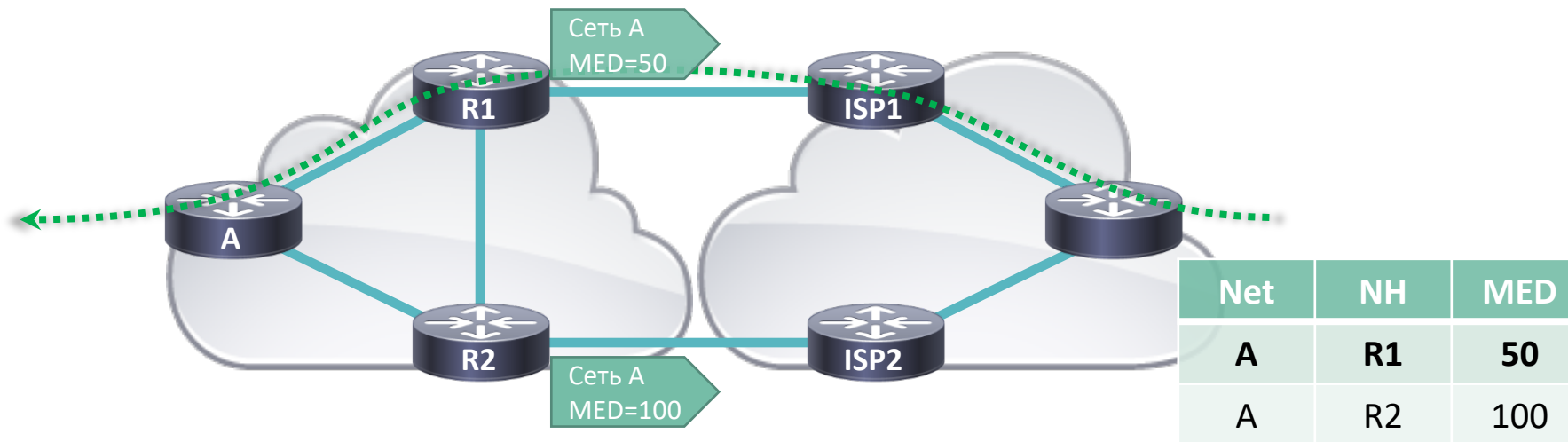
- Local Preference
 - Удобный инструмент выбора маршрута для исходящего трафика
- Приоритет маршрута в локальной автономной системе
 - Не отправляется соседям EBGP (маршруты получают LP по умолчанию)
- 32-битное число
 - Больше – лучше





MULTI_EXIT_DISC

- MED, Multi-Exit Discriminator
 - Инструмент указания предпочтения маршрута для соседней AS
 - Предназначен для контроля входящего в нашу AS трафика
 - По умолчанию MED маршрутов из разных AS не сравниваются
 - 32-битное число (меньше – лучше)
 - Часто наследуется от метрики IGP





Особенности MED

- Атрибут MED не очень удобен в использовании
 - Не передается EBGP-соседям, если изучен по IBGP
 - Обычно проставляется на пограничных роутерах
 - Распространяется только по соседней автономной системе
 - Сравнение имеет смысл у EBGP-маршрутов из одной AS
 - Основная польза – при dual-homed подключении
- Часто для контроля входящего маршрута вместо MED используют:
 - Атрибут Community
 - Технику AS Prepending (искусственное "удлинение" AS-PATH)



Атрибут Community

- Расширение функциональности BGP с помощью добавления меток
 - Неупорядоченный список меток, проставленных транзитными роутерами
 - На самом деле™ атрибутов три, каждый для разных форматов меток:
 - Communities (RFC 1997): записывается как <ASN>:<Value>, 16 бит ASN, 16 бит Value
 - Extended Communities (RFC 4360): 16 бит ASN, несколько типов меток по 6 байт
 - Large Communities (RFC 8092): 32 бита ASN, записывается <ASN>:<Value1>:<Value2>
- Метка может обозначать все, что угодно, нужны договоренности
 - Пример договоренностей: <https://kb.msk-ix.ru/ix/services/route-server/#bgp>
- Некоторые известные "стандартные" метки:
 - NO_EXPORT (65535:65281, 0xFFFFFFFF01) – не анонсировать EBGP-соседям
 - NO_ADVERTISE (65535:65281, 0xFFFFFFFF02) – не анонсировать никому
 - BLACKHOLE (65535:666, 0xFFFF029A) – установить Next Hop в Null0



ORIGIN

- Код происхождения маршрута в BGP (меньше = лучше)

ORIGIN	Обозначение	Смысл
0	IGP	NLRI принадлежит исходной AS
1	EGP	NLRI в исходной AS изучена по протоколу EGP (RFC 904)
2	INCOMPLETE	NLRI в исходной AS появилась каким-то иным способом

- Когда-то ORIGIN позволял сосуществовать в Интернете BGP и EGP
 - EGP давно устарел, в XXI веке атрибут практического смысла не имеет
 - Правила хорошего тона предписывают анонсировать NLRI с ORIGIN=IGP



Свойства атрибутов

- **Well-known mandatory** – атрибут обязан присутствовать в UPDATE
- **Well-known discretionary** – атрибут может не передаваться в UPDATE, но обязан поддерживаться в любой реализации
- **Optional transitive** – атрибут может не поддерживаться в конкретной реализации, но в любом случае должен передаваться соседям
- **Optional non-transitive** – если атрибут не поддерживается в конкретной реализации, он игнорируется и соседям не передается

Атрибут	Свойства
Next Hop	Well-known mandatory
AS_PATH	Well-known mandatory
LOCAL_PREF	Well-known discretionary (mandatory в IBGP, отсутствует в EBGP)
MULTI_EXIT_DISC	Optional non-transitive (часто отсутствует, может быть удален)
ORIGIN	Well-known mandatory
Community	Optional transitive (часто присутствует, сохраняется при передаче)



Best Path Selection

- При выборе маршрута RFC 4271 предписывает следовать алгоритму:
 1. Сравниваются только доступные маршруты до одного и того же префикса
 - NEXT_HOP должен быть доступен (в RIB)
 - AS_PATH не должен содержать номер собственной AS
 2. Предпочитаются маршруты с **большим** Local Preference
 3. Предпочитаются маршруты с меньшим количеством AS в AS_PATH
 4. Предпочитаются маршруты с меньшим Origin (IGP<EGP<INCOMPLETE)
 5. Предпочитаются маршруты с меньшим MED
 - MED сравнивается только для маршрутов, полученных из одной и той же AS
 - Если MED отсутствует, полагать его значение наименьшим возможным
 6. Предпочитаются маршруты, полученные по EBGP
 7. Предпочитаются маршруты с меньшей IGP-метрикой до Next Hop
 8. Предпочитаются маршруты, полученные от соседа с **большим** Router-ID
 9. Предпочитается маршрут, полученный от соседа с **большим** IP-адресом

NWLLA OMNI



Networking
For everyone

- При выборе маршрута Cisco IOS использует алгоритм:

Код	Мнемоника	Описание
N	n ext-hop	маршруты с доступным NEXT_HOP
W	w eight	маршруты с бóльшим Weight
L	l ocal-pref	маршруты с бóльшим Local Preference
L	l ocal-origin	маршруты с нулевым Next Hop (вброшены в BGP локально)
A	a s-path	маршруты с меньшим количеством AS в AS_PATH
O	o origin	маршруты с меньшим Origin (IGP<EGP<INCOMPLETE)
M	m ed	маршруты с меньшим MED
N	n eighbor	маршруты, полученные от EBGП (hot potato routing)
I	i gp-metric	маршруты с меньшей IGP-метрикой до Next Hop
non-ECMP	ebgp-age	только для EBGП: маршрут, полученный первым
	neighbor-id	маршруты, полученные от соседа с бóльшим Router-ID
	neighbor-ip	маршрут, полученный от соседа с бóльшим IP-адресом



Networking
For everyone