

Внешний протокол маршрутизации BGP-4

T. Li, Ed. S. Hares, Ed.

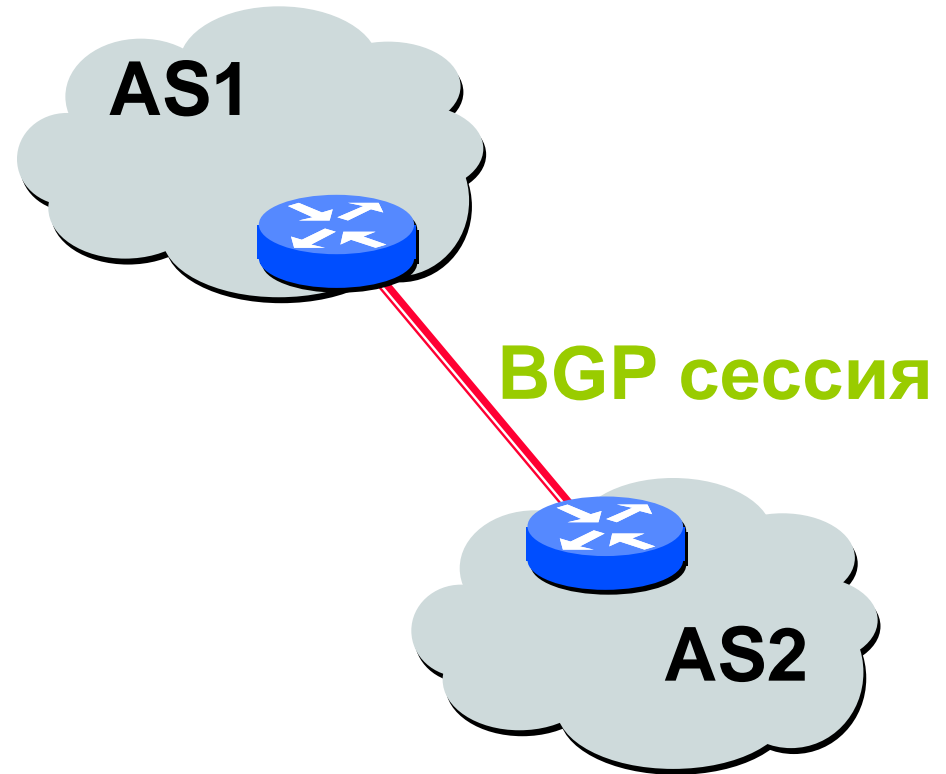
RFC 4271 January 2006

A Border Gateway Protocol 4 (BGP-4)

BGP-4

- BGP = Border Gateway Protocol
- Является протоколом маршрутизации на основе Политик
- Является фактически (de facto) EGP сегодняшнего глобального Internet
- **Относительно простой протокол, но конфигурация сложна и весь мир может видеть и сталкиваться с вашими ошибками**
 - 1989 : BGP-1 [RFC 1105]
 - Замена EGP (1984, RFC 904)
 - 1990 : BGP-2 [RFC 1163]
 - 1991 : BGP-3 [RFC 1267] – поддерживает только классовую маршрутизацию
 - 1995 : BGP-4 [RFC 1771] - поддерживает бесклассовую маршрутизацию (CIDR - Classless Interdomain Routing)

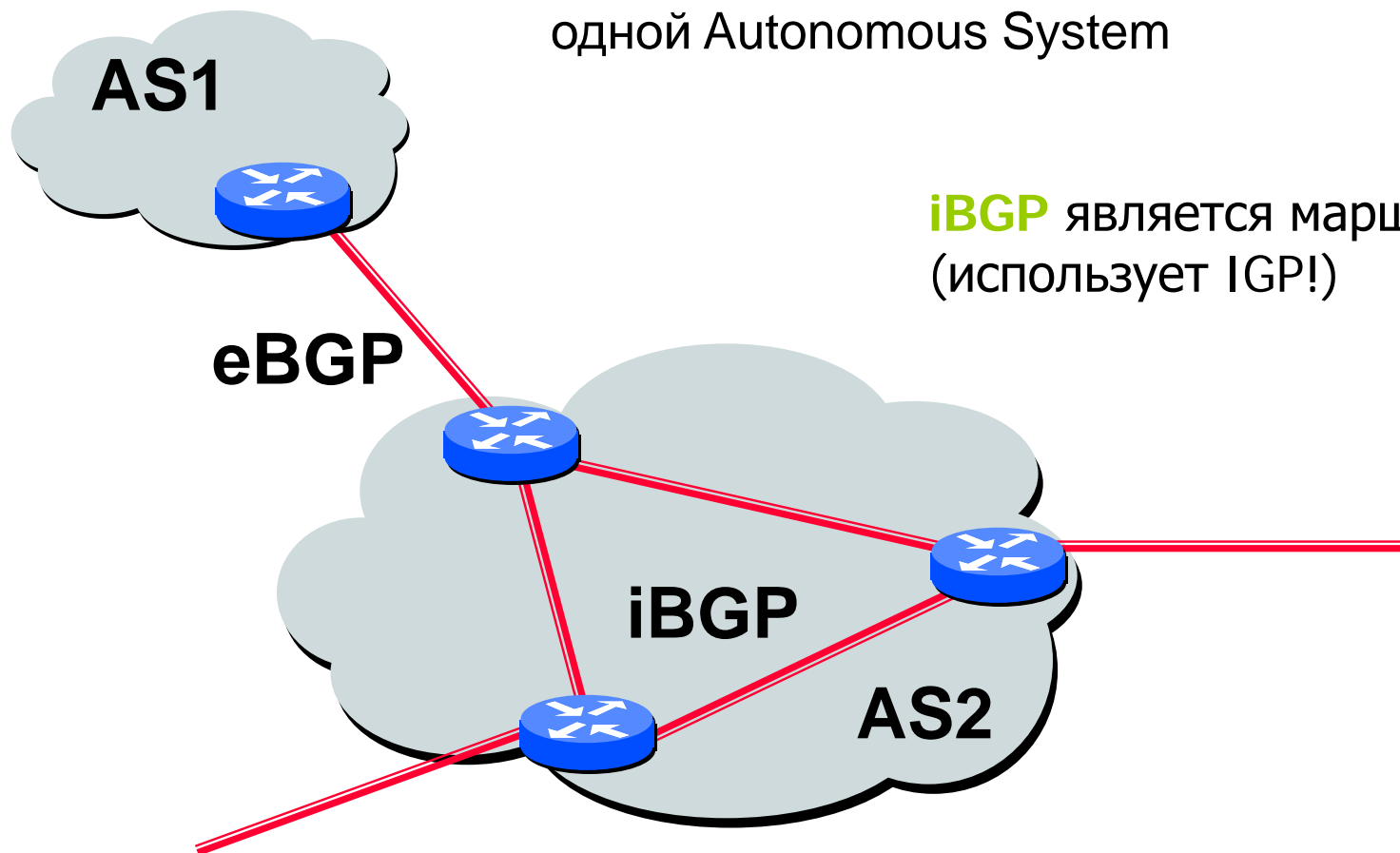
Операции BGP (упрощенные)



Когда соединение ALIVE обмениваться "route UPDATE" сообщениями

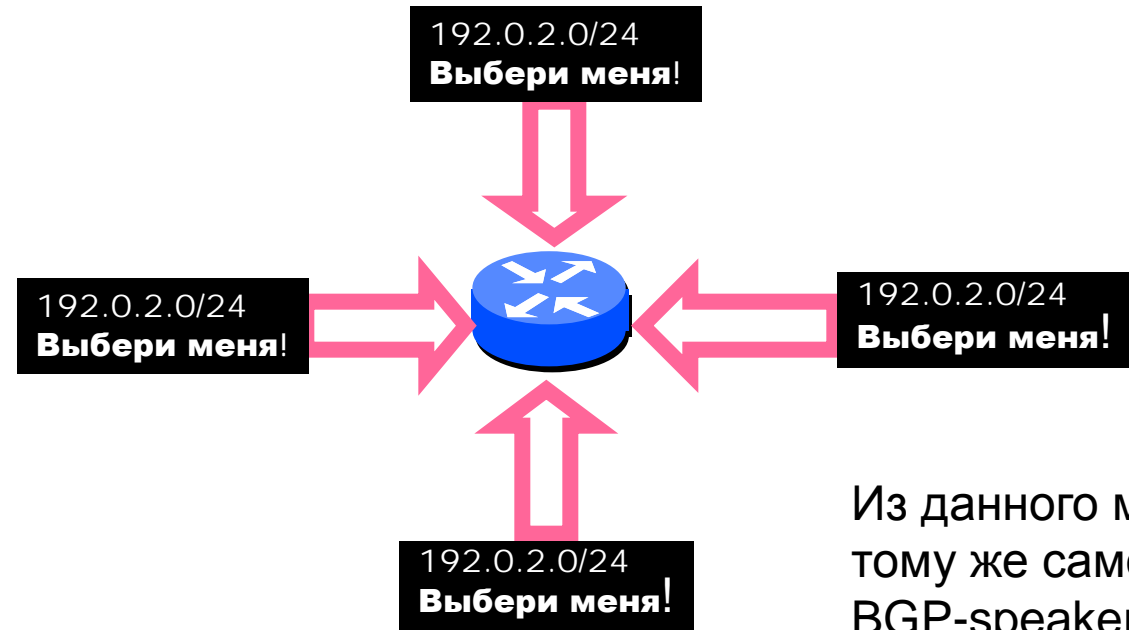
Существует два типа отношений между BGP-соседами

- Внешний сосед (External Neighbor) (eBGP) между различными Autonomous Systems
- Внутренний сосед (Internal Neighbor (iBGP) в одной Autonomous System



iBGP является маршрутизацией (использует IGP!)

Атрибуты, используемые для выбора наилучшего пути



Из данного множества маршрутов к одному и тому же самому IP-префиксу 192.0.2.0/24 BGP-speaker должен выбрать один наилучший маршрут

(Замечание: эдак можно и забраковать их все!
Или надо иметь произвольные предпочтения, основанное на атрибутах маршрута)

Список атрибутов на следующем слайде

BGP КОДЫ типов атрибутов

Смотрите IANA: <http://www.iana.org/assignments/bgp-parameters>

Номер Атрибута (Value)	Имя атрибута (Code)	категория/тип атрибута	RFC/проект док Reference
1	ORIGIN	общеизвестный обязательный	[RFC1771]
2	AS_PATH	общеизвестный обязательный	[RFC1771]
3	NEXT_HOP	общеизвестный обязательный	[RFC1771]
4	MULTI_EXIT_DISC	необязательный нетранзитивный	[RFC1771]
5	LOCAL_PREF	общеизвестный представленный на собственное усмотрение	[RFC1771]
6	ATOMIC_AGGREGATE	общеизвестный представленный на собственное усмотрение	[RFC1771]
7	AGGREGATOR	необязательный транзитивный	[RFC1771]
8	COMMUNITY	необязательный транзитивный	[RFC1997]
9	ORIGINATOR_ID	необязательный нетранзитивный	[RFC2796]
10	CLUSTER_LIST (список кластеров)	необязательный нетранзитивный	[RFC2796]
11	DPA (Destination Point Attribute)	атрибут точки назначения для BGP	[Chen]
12	ADVERTISER (объявитель маршрутов)	Сервер маршрутов BGP/IDRP	[RFC1863]
13	RCID_PATH / CLUSTER_ID	Сервер маршрутов BGP/IDRP	[RFC1863]
14	MP_REACH_NLRI	необязательный нетранзитивный	[RFC2283]
15	MP_UNREACH_NLRI	необязательный нетранзитивный	[RFC2283]
16	EXTENDED COMMUNITIES (дополнительный сообщества)		[Rosen]
...			
255	reserved for development		

Не все атрибуты должны присутствовать в каждом объявлении (анонсе)

Path Attributes: ORIGIN

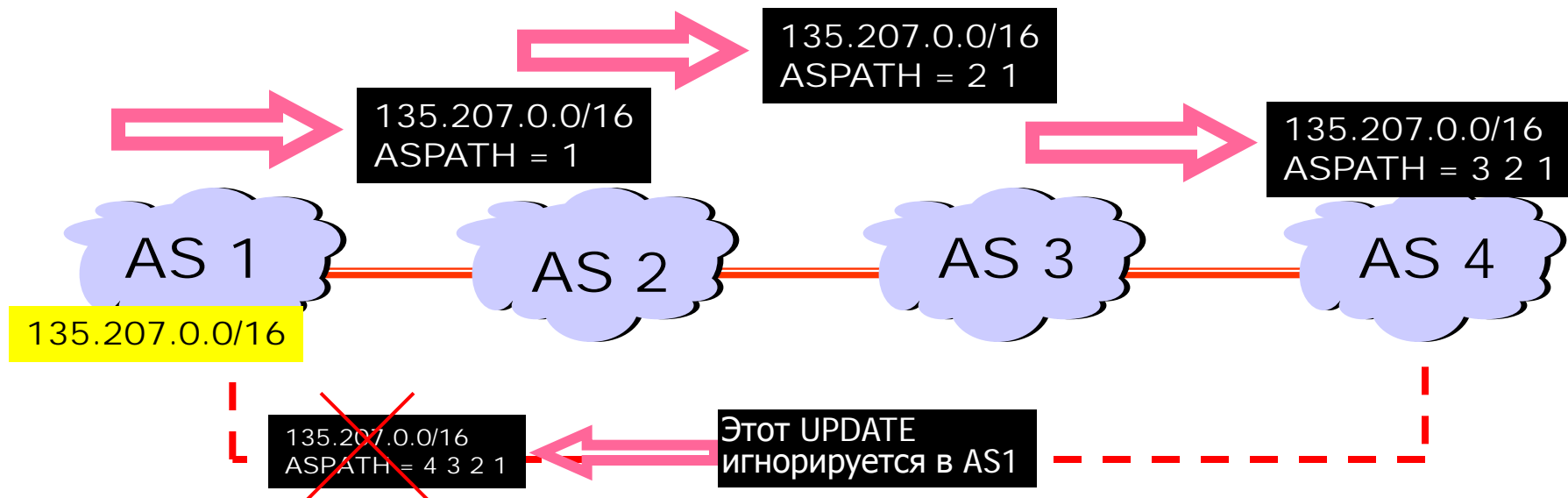
● ORIGIN (обязательный общеизвестный атрибут):

- Описывает источник обновления маршрута с учетом AS (происхождение маршрута: IGP, EGP, INCOMPLETE)
- Префикс, полученный от источника и *"введенный"* (*"injected"*) в BGP, может быть порожден:
 - ✓ прямо связанный с сетевым интерфейсом префикс:

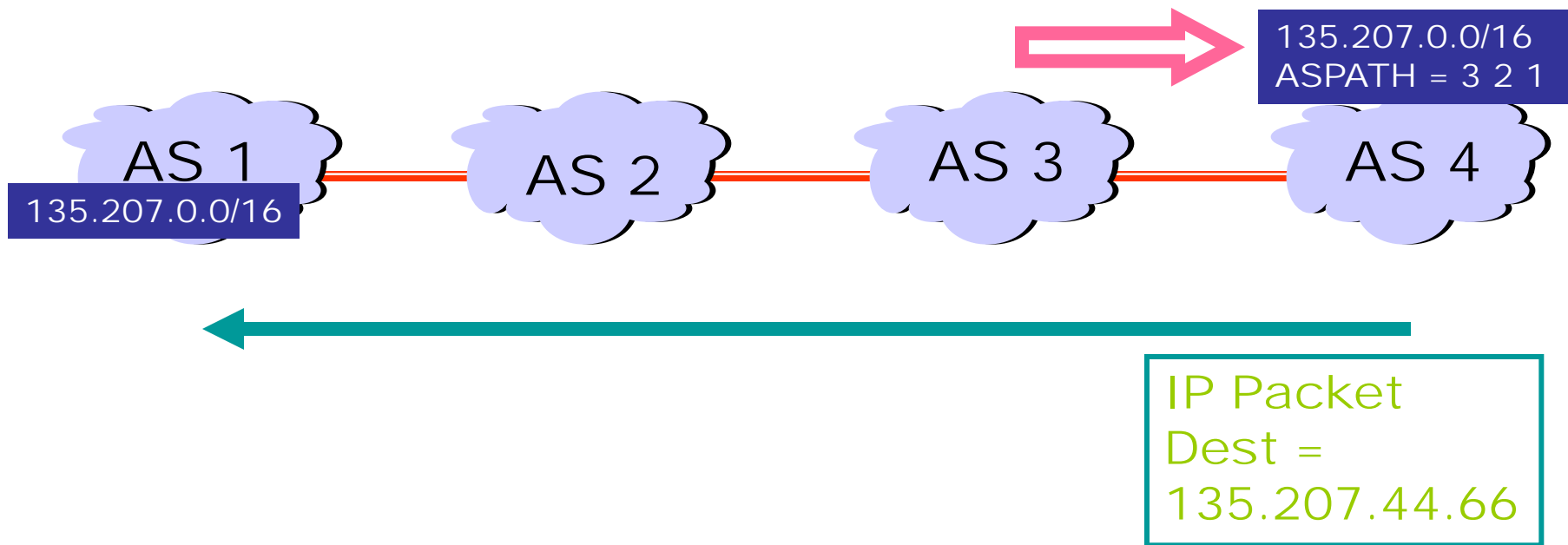
```
Interface Ethernet1/0
  ip address 172.16.65.1 255.255.255.192
```
 - ✓ вручную сконфигурированная статическая маршрутизация
 - ✓ динамически по IGP или EGP
- Порядок предпочитаемых значений: IGP / EGP / INCOMPLETE
 - ✓ IGP (EGP): префикс порожден IGP (EGP)
 - ✓ INCOMPLETE: другие средства сетевого уровня (статическая маршрутизация)

Атрибут AS-PATH

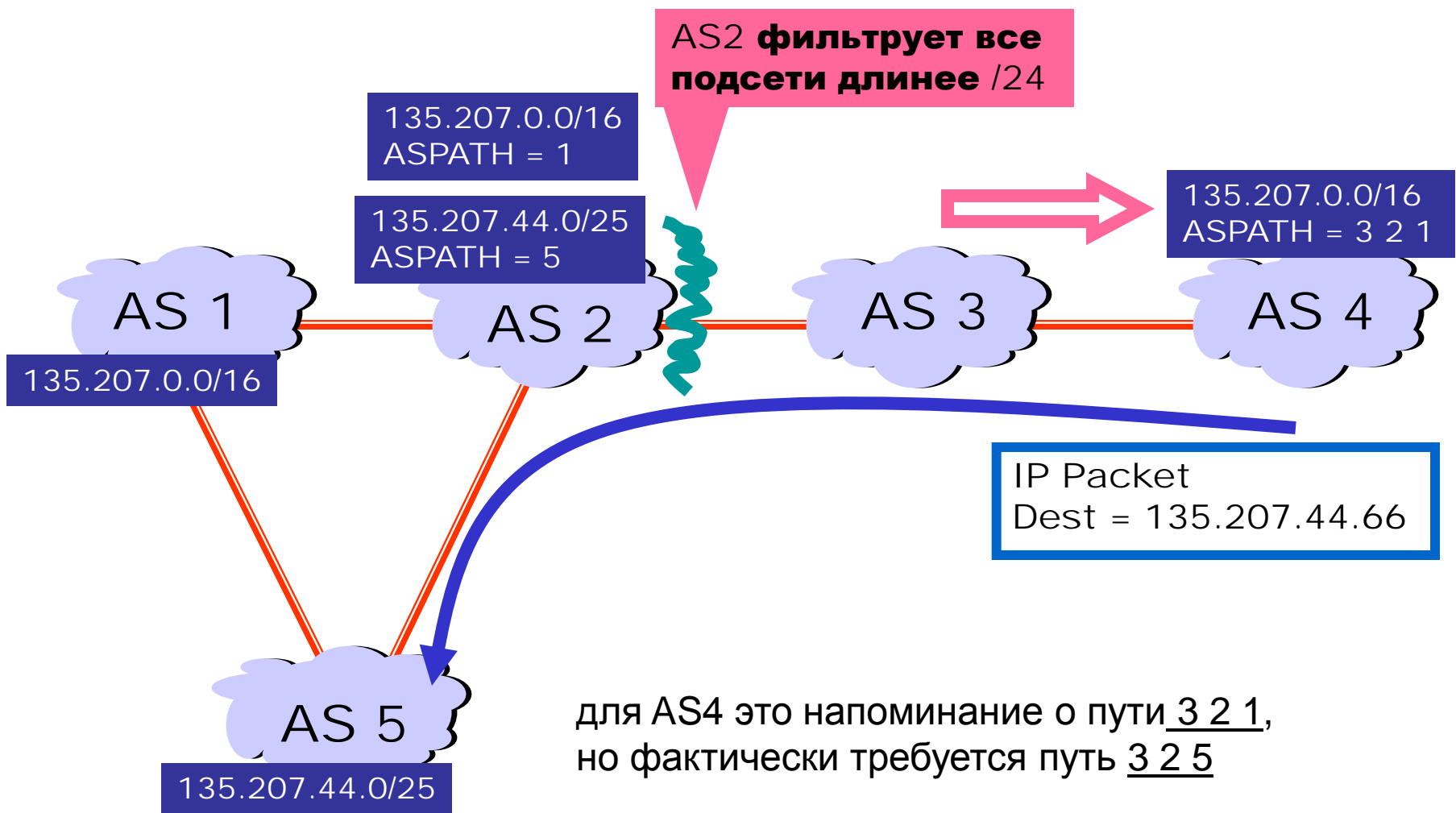
- **AS-PATH – это Список AS, через которые прошло префиксное объявление**
 - Синонимы: AS_PATH ⇔ AS_SEQUENCE ⇔ Список AS
 - Каждая AS по пути следования префикса добавляет в начало списка AS (prepending) свой номер AS
 - Номер AS, сгенерировавшей префикс, размещен в конце списка AS (перед кодом атрибута ORIGIN)
- **AS-PATH используется для**
 - обнаружения петель маршрутизации
 - ✓ если маршрут пришел в AS, номер которой уже имеется в AS_SEQUENCE, то UPDATE игнорируется
 - выбора пути
 - ✓ предпочитается более короткий список



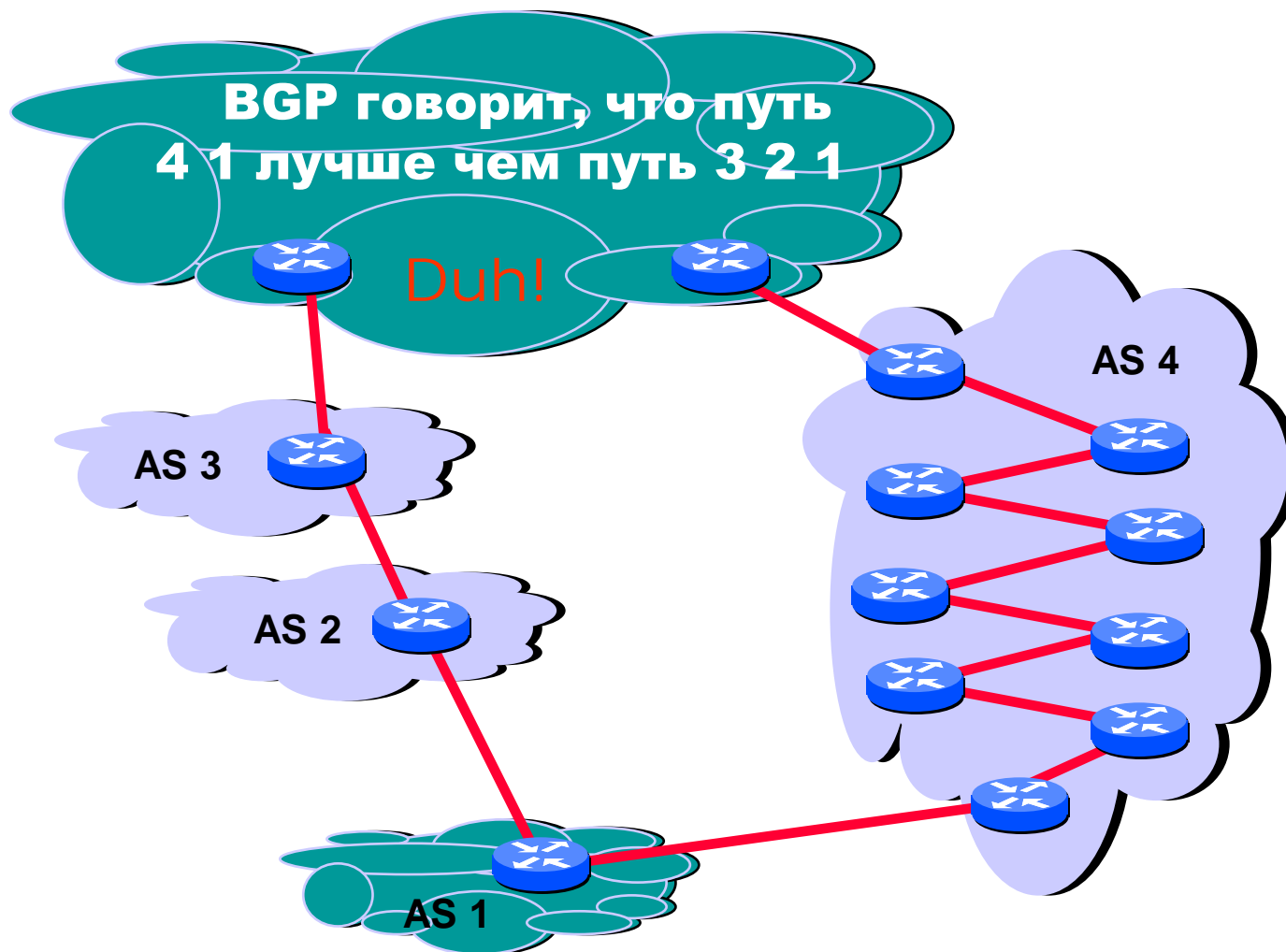
Трафик часто следует согласно AS-PATH



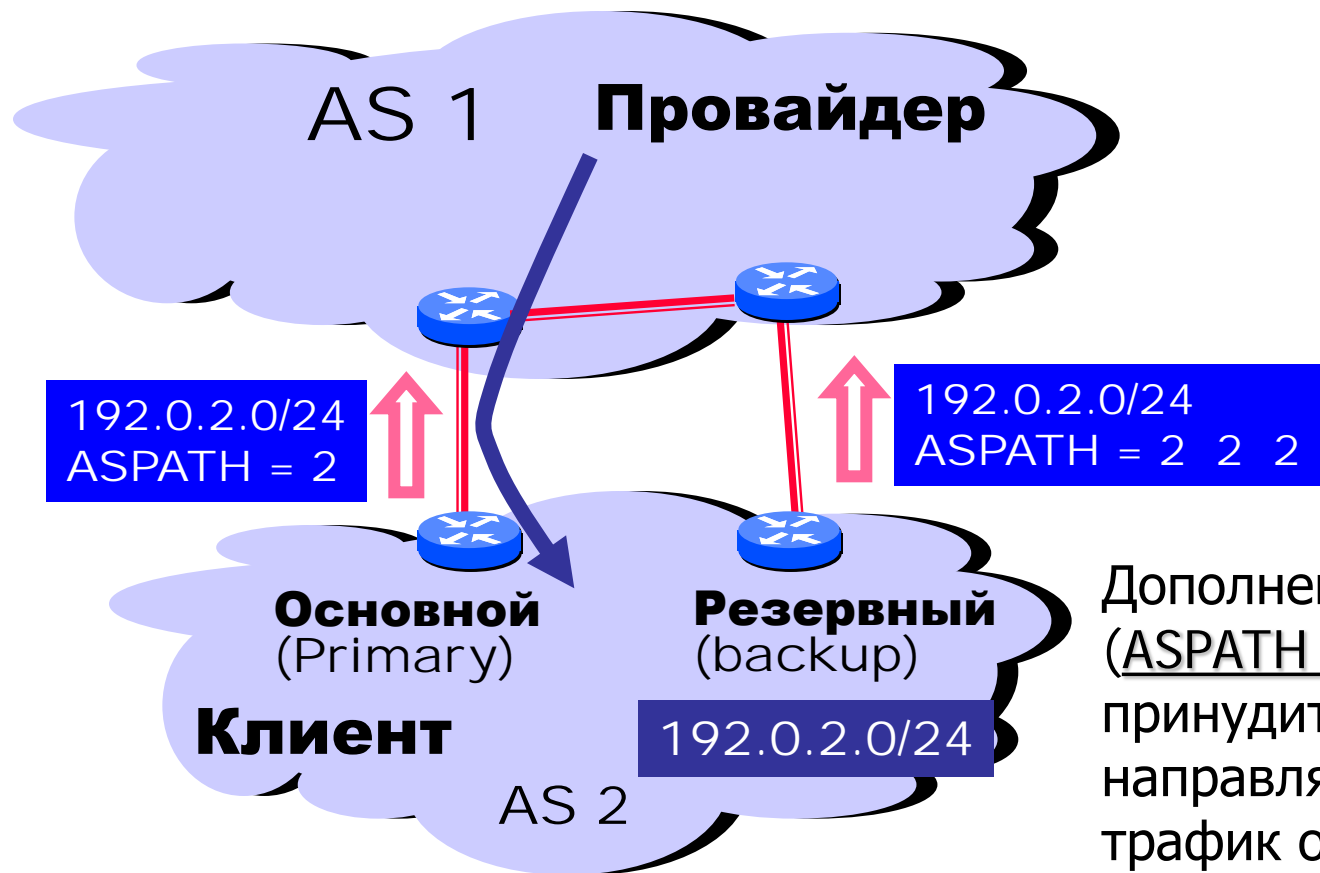
... но этого может и не быть



Более короткий AS-path не означает меньшее число скачков



AS-PATH Padding: Ограничение входящего трафика



Дополнение (ASPATH Padding) будет принудительно направлять входящий трафик от AS1 по основному (primary) каналу связи

- LOCAL-PREF: **исходящий трафик, локальное предпочтение**
- MED: **входящий трафик, обычно от того же ISP (link-level knob)**

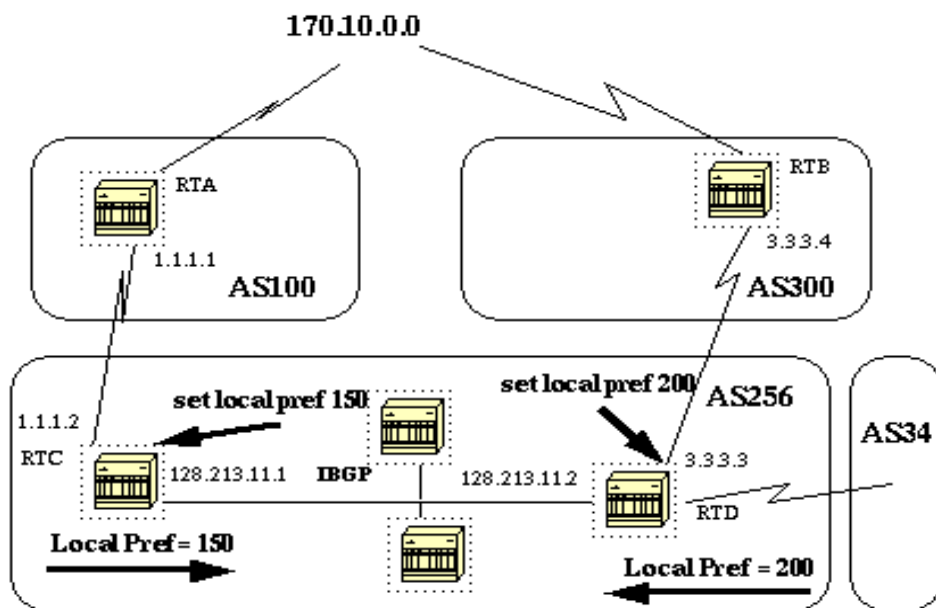


Атрибут: LOCAL-PREF (локальные предпочтения)

Локально конфигурируемый

- передается между спикерами только внутри AS
- действителен для всех узлов внутри одной AS
- предпочитается путь с наибольшим LOCAL-PREF, если существует несколько маршрутов к одному и тому же префиксу
 - ✓ По умолчанию LOCAL-PREF = 100
- влияет на исходящий трафик, не влияет на входящий трафик

Используется для установки точки выхода из AS для достижения определенного префикса



- **Пример:** AS256 имеет доступ в Интернет через AS100 и AS300
- **Установим в AS256:**
 - в RTD - LOCAL-PREF = 200
 - в RTC - LOCAL-PREF = 150
- **Тогда:**
 - RTC присваивает маршрутам, объявленным AS100 LOCAL-PREF = 150
 - RTD присваивает маршрутам, объявленным AS300 LOCAL-PREF = 200
 - RTC и RTD обмениваются по iBGP UPDATE сообщениями о маршрутах
- **И в AS256 выберутся маршруты к префиксу 170.10.0.0 через AS300**

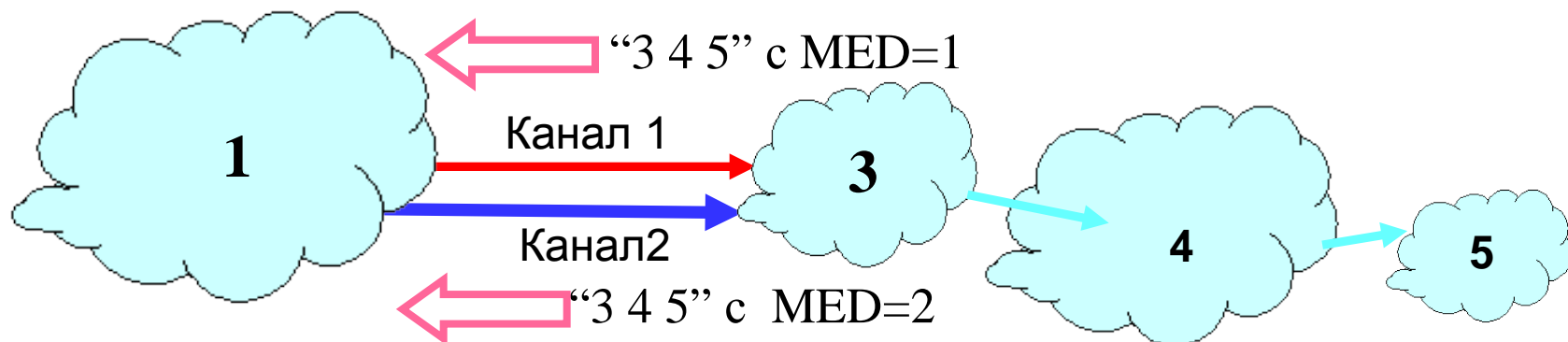
“Скажите вашему соседу, что Вы хотите ???”

● MED указывает предпочитаемый канал входящего трафика

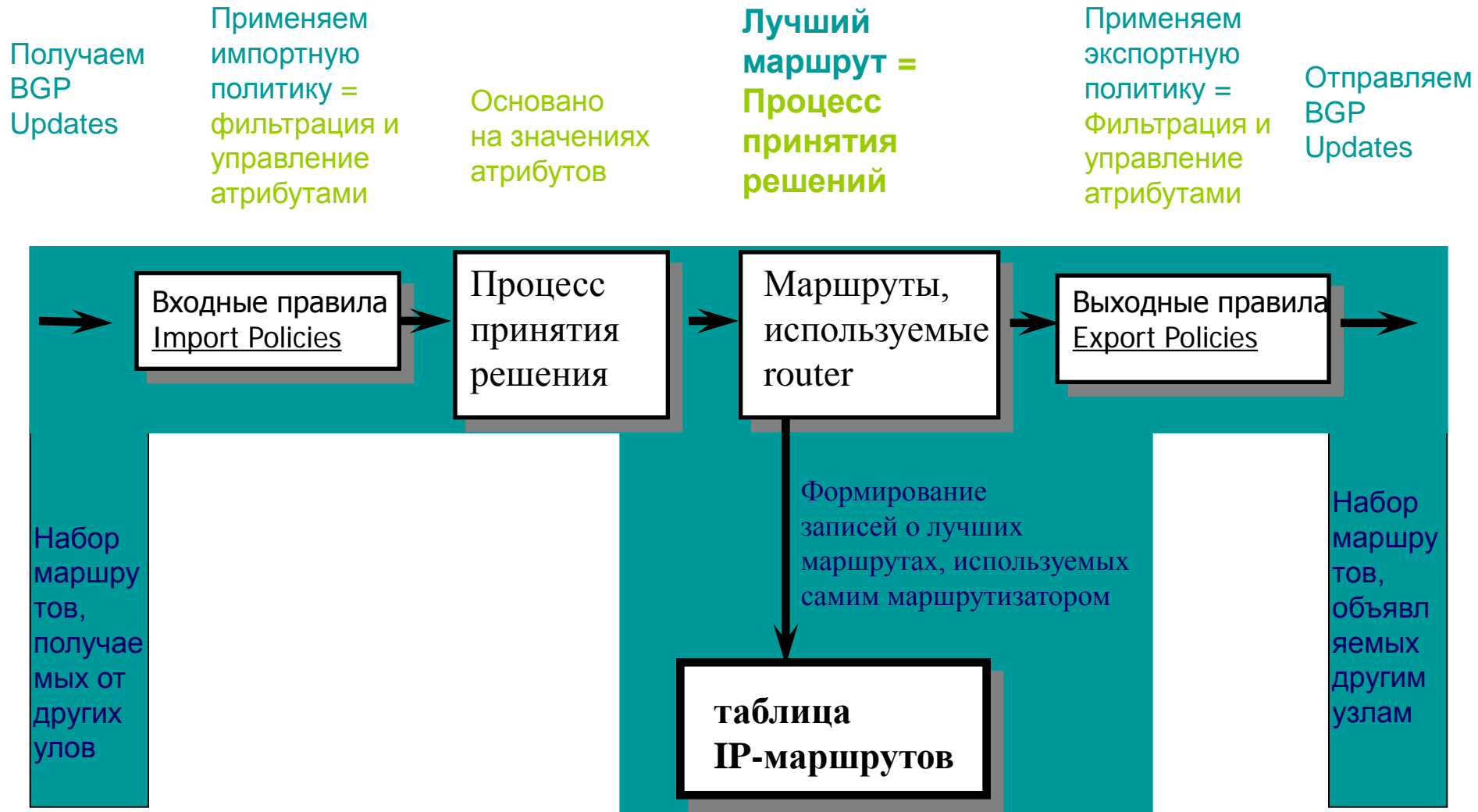
- выбирается маршрут с наименьшей MED
- подсказка внешним соседним узлам о предпочтительном маршруте в AS, имеющей несколько точек входа
 - ✓ Необходимо заручится поддержкой neighbor, чтобы он принял MED
- MED известен также как “внешняя метрика маршрута”
- допускается обмен MED между различными AS (в отличие от LOCAL_PREF)
- принятый извне UPDATE с атрибутом MED не должен покидать пределов AS

● Если маршрут формируется в AS, то чаще всего значение MED соответствует метрике IGP-маршрута

- Метрика IGP в этом случае показывает расстояние от префикса до точки входа в AS



Упрощенная схема процесса BGP маршрутизации



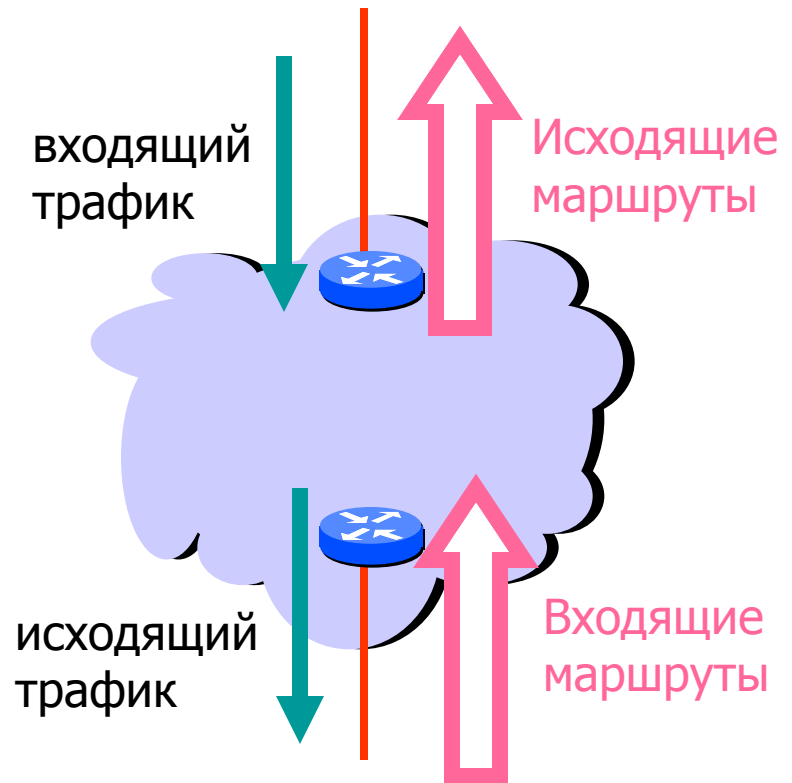
Импортная и Экспортная политика

• Для входящего трафика

- Фильтрация (отсеивание) исходящих маршрутов (экспортная политика)
- Задание списка атрибутов на исходящих маршрутах с надеждой на выбор лучшего маршрута у соседних узлов

• Для исходящего трафика

- Фильтрация входящих маршрутов (импортная политика)
- Задание списка атрибутов на входящих маршрутах, чтобы влиять на выбор лучшего маршрута у соседних узлов



Вообще-то, AS имеет больше возможностей управления исходящим трафиком

Входная и выходная фильтрация

С точки зрения обмена маршрутной информацией:

● **Входная фильтрация** управляет исходящим трафиком

- отсеивание маршрутной информации (filters route updates), поступающей от других peers
- фильтрация, основанная на IP-префиксах, AS_PATH, community

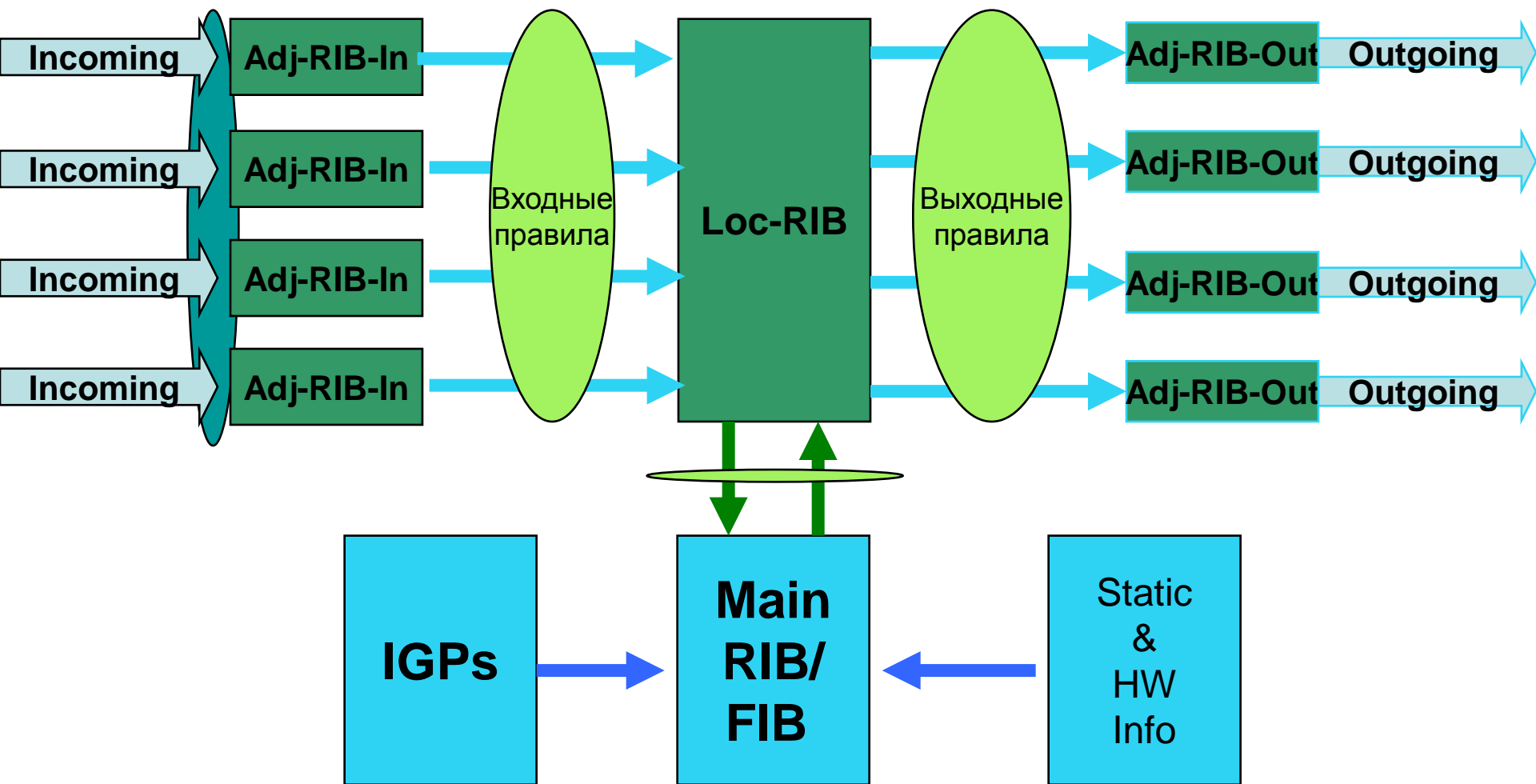
● **Выходная фильтрация** управляет входящим трафиком

- объявление маршрута другим AS означает, что можно достичь IP-префикс через нас
- не объявление маршрута другим AS означает, что для достижения префикса должен использоваться другой маршрутизатор (другая AS)

● **Манипулирование Атрибутами**

- Импортная политика: LOCAL_PREF (манипулируют доверием)
- Экспортная политика: AS_PATH и MEDs

Выполняемый поток политик (правил)



Концептуальная модель BGP операций

● RIB: Routing Information Base

- информационная база о маршрутах

● Adj-RIB-In:

- Префиксы, полученные от соседей
- сколько пэров - столько Adj-RIB-In

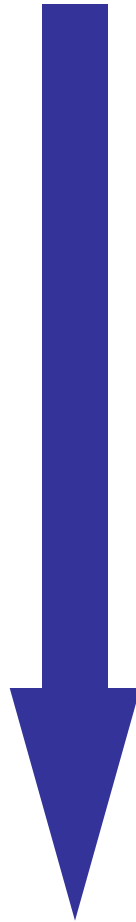
● Loc-RIB:

- Префиксы, выбранные для локального использования после анализа Adj-RIB-Ins. Эта маршрутная база данных, рекламируется внутри AS

● Adj-RIB-Out:

- Префиксы, рекламируемые соседям
- Столько много Adj-RIB-Out сколько соседей

Резюме: выбор маршрута



Наибольший **Local Preference**

Предписанная зависимость

Короткий **AS-PATH**

Низкий **MED**

iBGP < eBGP

Низкая IGP стоимость
до выходного BGP

Проектирование трафика

Самый **низкий ID**
(идентификатор) **маршрутизатора**

Throw up hands and
break ties \ (руки вверх
и разорвать связь)