

Лекция № 3. Автономные системы

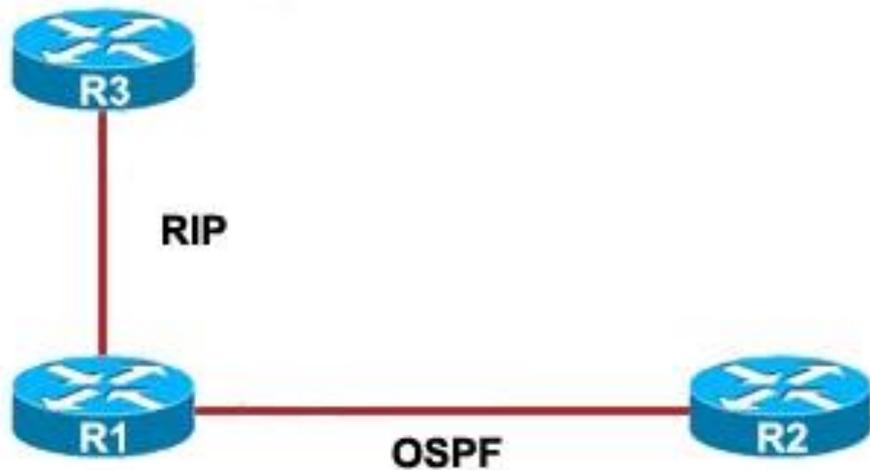
Интернет состоит не из хостов или сетей, а из автономных систем.

Протокол BGP выполняет междоменную маршрутизацию.

Управление трафиком в Итернете – вопрос не только технологический, но политический.

Редистрибуция маршрутов

Можно ли информацию, полученную средствами RIP, передавать процессу OSPF?



Да, такая возможность предусмотрена в программном обеспечении маршрутизаторов.

Мотивировка



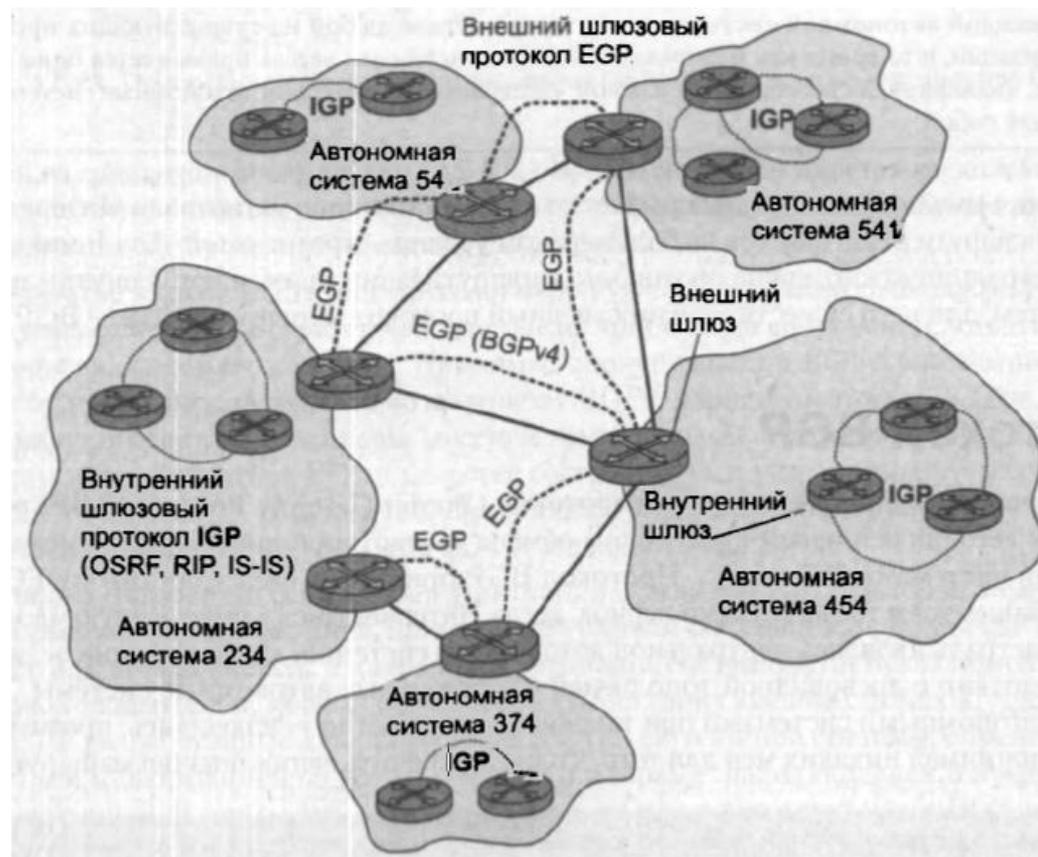
Можно ли использовать RIP или OSPF для решения задач маршрутизации в масштабах Интернета?

1. OSPF – это Link State протокол. Маршрутизатор знает топологию всей сети.
Из соображений безопасности не стоит показывать детали своей сети.
2. В Интернете существует более 450 тыс. маршрутов. Протоколы IGP не предназначены для обработки такого объема информации.

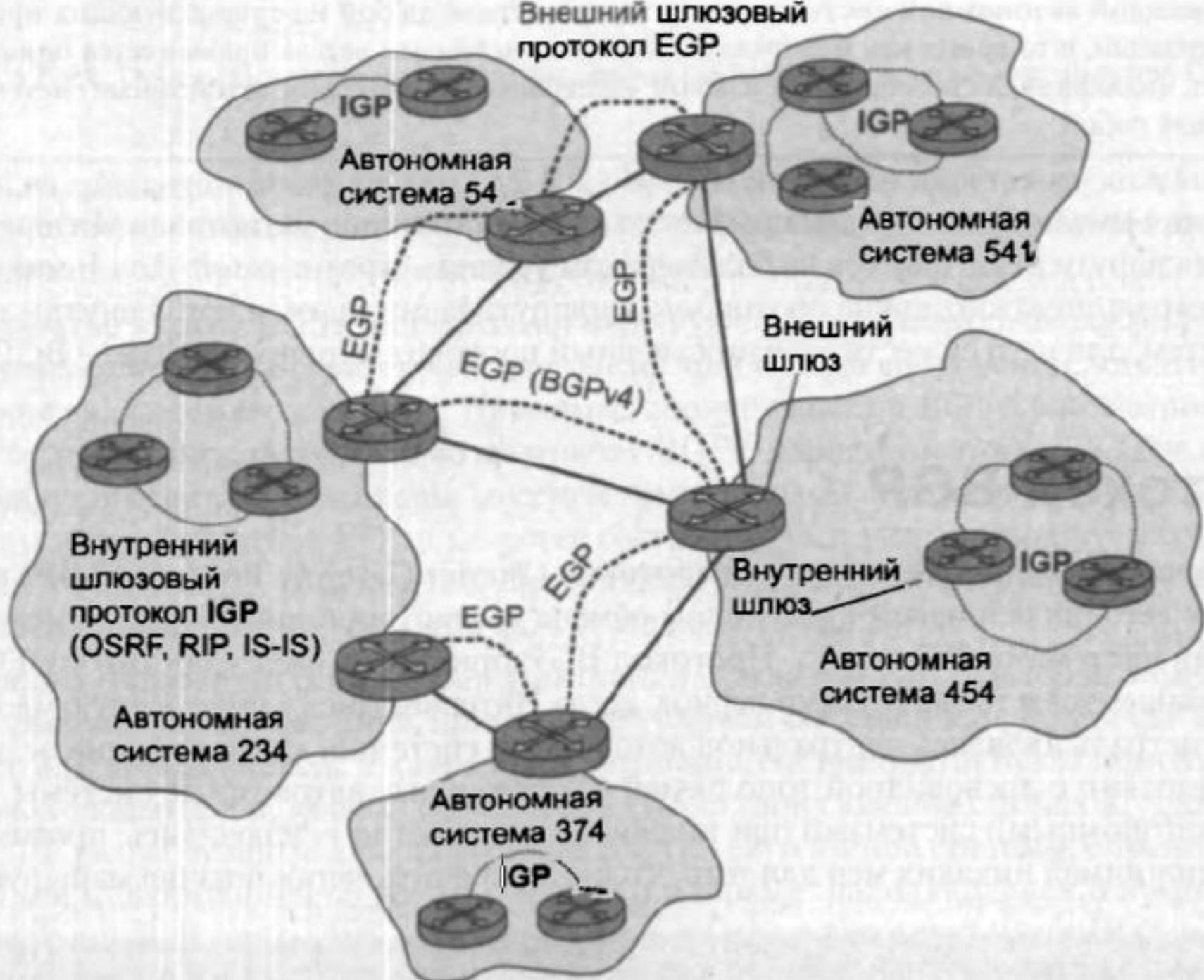
Автономные системы. Понятие

Автономная система (Autonomous System, AS) — это совокупность сетей под единым административным управлением, обеспечивающим общую для всех входящих в автономную систему маршрутизаторов политику маршрутизации.

RFC 1930 (март 1996)



Внешний шлюзовый протокол EGP.



Определение AS согласно RFC 4271

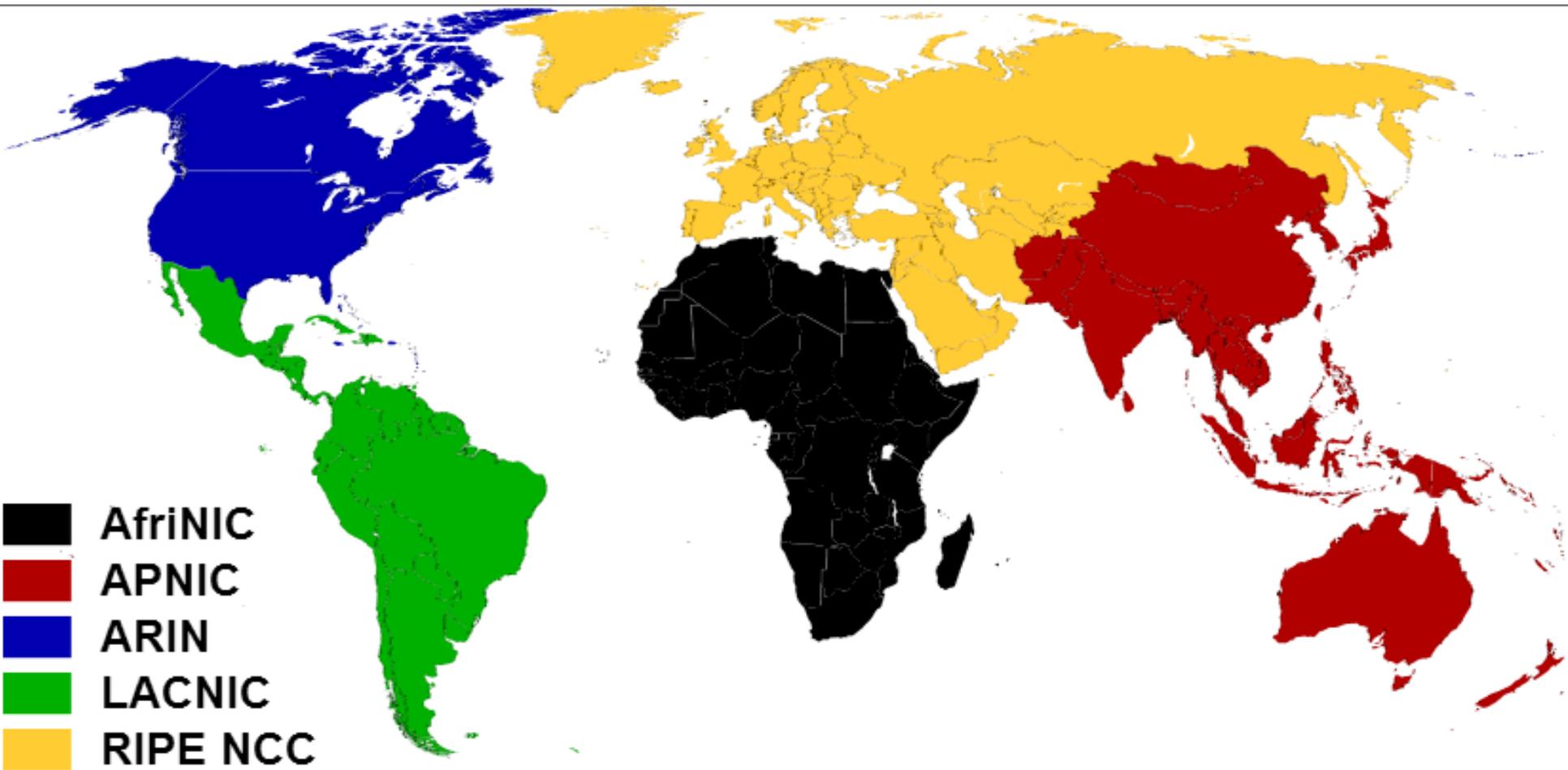
AS – элемент политики маршрутизации в современной среде внешней маршрутизации.

AS – множество маршрутизаторов с единым техническим администрированием, использующих один протокол внутренней маршрутизации (IGP) и единую метрику для маршрутизации пакетов внутри AS, а для передачи пакетов в другие автономные системы применяющих протокол внешней маршрутизации (EGP).

AS – группа из одного или нескольких префиксов IP, работающих у одного или нескольких сетевых операторов, которые имеют единую (SINGLE) и четко определенную (CLEARLY DEFINED) политику маршрутизации.

Политика маршрутизации (routing policy) в данном случае понимается как набор решений о пересылке, принимаемых в современной сети Internet.

Региональные интернет регистраторы



Классификация AS согласно RFC 1772

Тупиковая (stub) AS – автономная система, имеющая соединение только с одной AS.

Многодомная (multihomed) AS – автономная система, соединенная с несколькими AS, но не принимающая транзитный трафик.

Транзитная AS – автономная система, соединенная с множеством других AS и предназначенная (с некоторыми ограничениями на уровне политики) для поддержки как локального, так и транзитного трафика.

Протокол BGP не вносит ограничений в топологию соединений между AS.

Общие характеристики BGP4 по RFC 4271

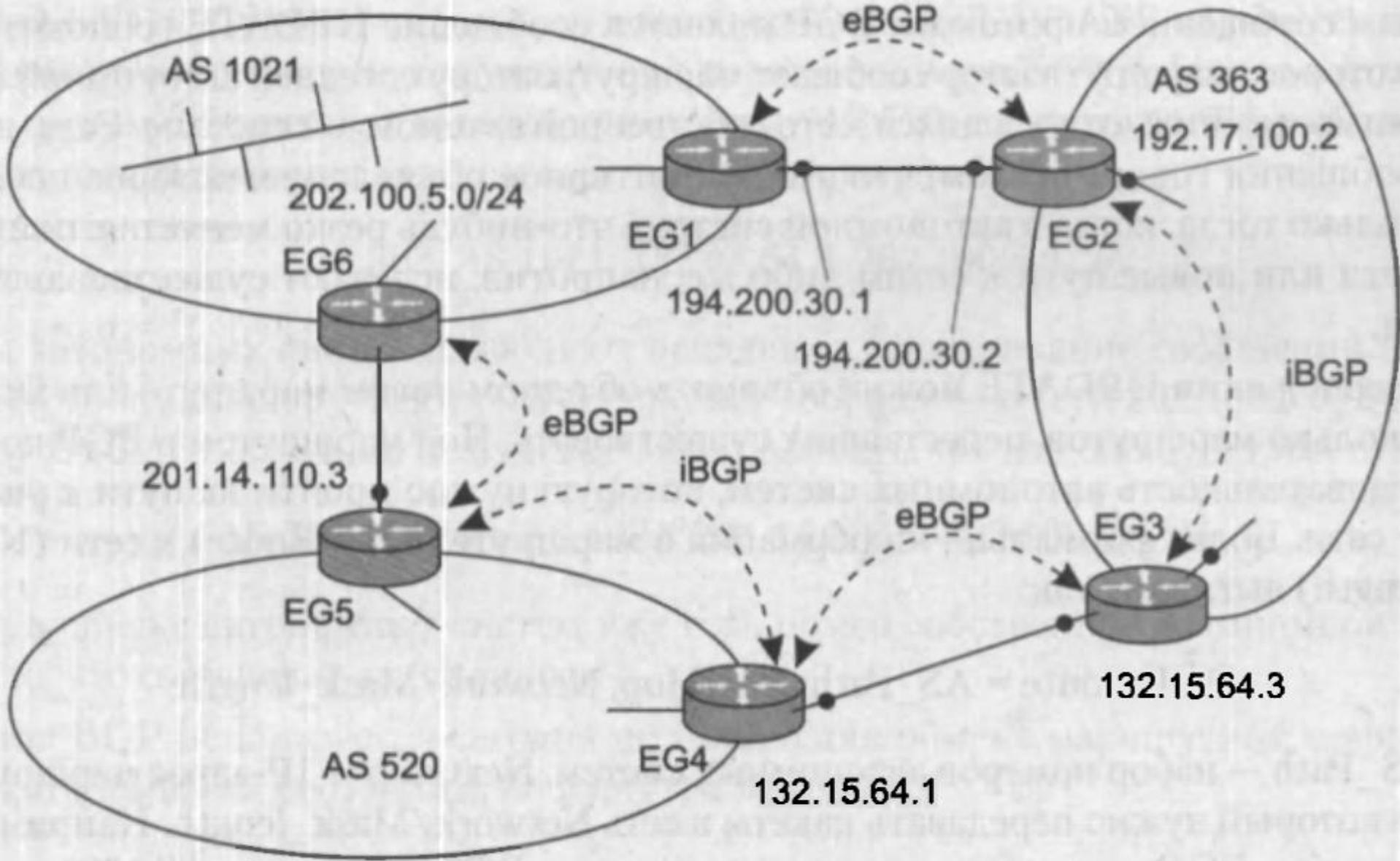


Основной функцией поддерживающей протокол BGP системы является обмен информацией о доступности сетей с другими системами BGP.

Поддержка бесклассовой междоменной маршрутизации.
Поддержка объединения маршрутов.

BGP поддерживает только парадигму пересылки на основе адреса получателя. Однако есть еще политики

BGP использует протокол TCP/179.



BGP_Route = AS_PATH; NEXT_HOP; IP-prefix

AS 1021; 194.200.30.1; 202.100.5.0/24

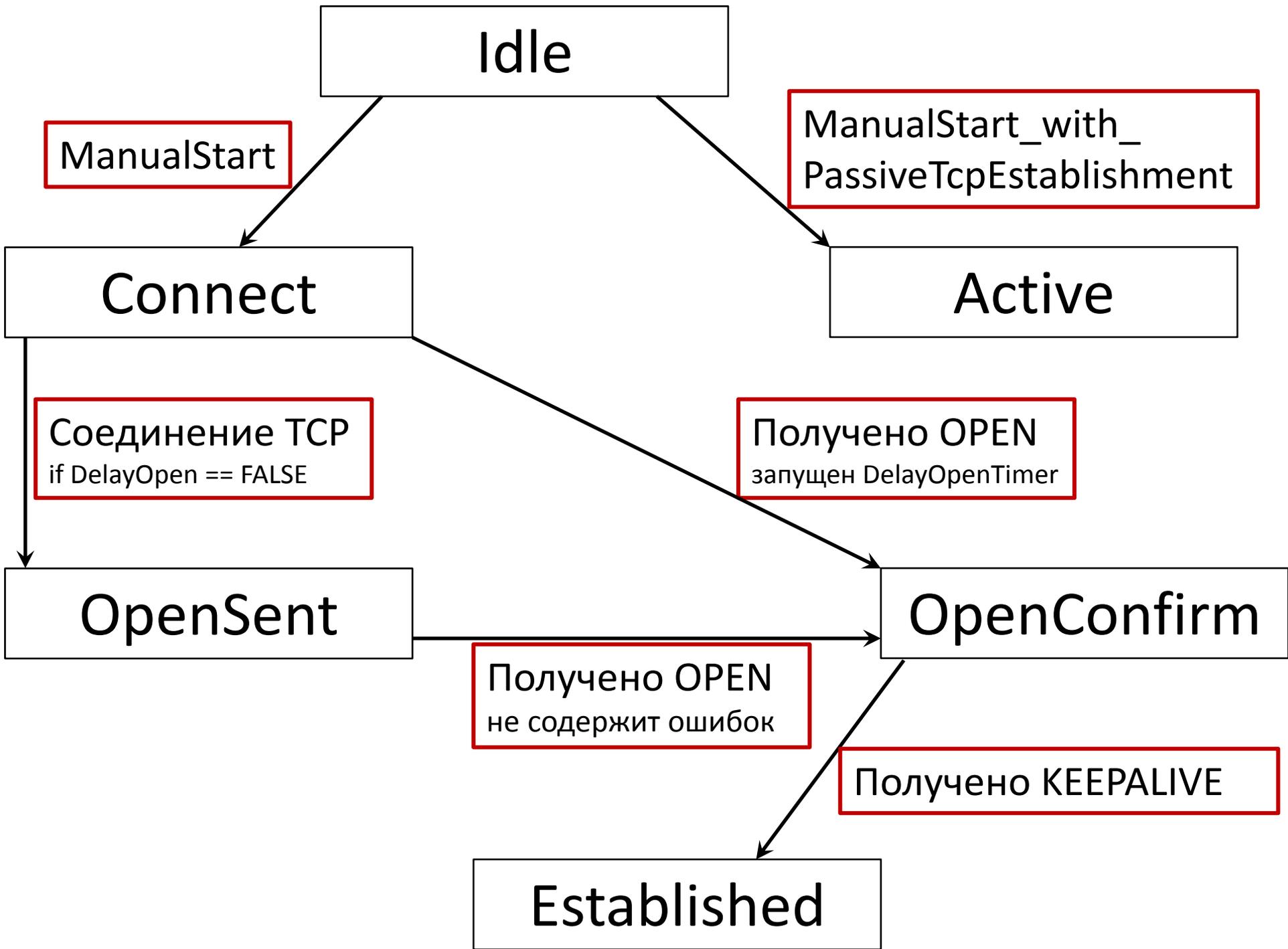
База маршрутной информации BGP4

База маршрутной информации (Routing Information Base)

Adj-RIBs-In – маршрутные данные, полученные из входящих сообщений UPDATE, которые были приняты от других узлов BGP.

Loc-RIB – локальная маршрутная информация узла BGP, выбранная путем применения локальной политики к маршрутам, содержащимся в Adj-RIBs-In.

Adj-RIBs-Out – информация локального узла BGP, выбранная им для анонсирования своим партнерам.



Заголовок BGP

1	2	3	4
Marker, 16 байт			
Length, 2 байт		Type, 1 байт	

Marker содержит 1 в каждом бите; включено для совместимости с предыдущими версиями BGP.

Length – целое число без знака показывает общий размер сообщения (включая заголовок) в октетах.
 $19 \leq \text{Length} \leq 4096$.

Type содержит код типа сообщения:

1 – OPEN, 2 – UPDATE, 3 – NOTIFICATION, 4 – KEEPALIVE.

Type=5 используется для сообщений Route-Refresh, добавленных в RFC 2918.

Сообщение OPEN

1	2	3	4
			Version
My Autonomous System		Hold Time	
BGP Identifier			
Opt Par Len	Optional Parameters (переменный размер)		

Version – номер версии протокола.

My AS – номер AS отправителя сообщения.

Hold Time – число секунд, которое отправитель предлагает установить для таймера удержания.

BGP Identifier – идентификатор узла BGP (IP адрес, присвоенный этому узлу BGP).

Сообщение OPEN

1	2	3	4
			Version
My Autonomous System		Hold Time	
BGP Identifier			
Opt Par Len	Optional Parameters (переменный размер)		

Opt Par Len – общий размер поля Optional Parameters в октетах.

Optional Parameters – список дополнительных параметров, представленных в формате <Param Type, Param Length, Param Value>.

В [RFC3392] определен дополнительный параметр Capabilities.

Сообщение UPDATE

Withdrawn Routes Length, 2 байта

Withdrawn Routes (переменный размер)

Total Path Attribute Length, 2 байта

Path Attributes (переменный размер)

Network Layer Reachability Information (переменный размер)

Withdrawn Routes – список префиксов IP-адресов для отзываемых маршрутов.

Префикс = <length, prefix>

22	202.100.44
00010110	11001010.01100100.001011 00

Сообщение UPDATE

Withdrawn Routes Length, 2 байта

Withdrawn Routes (переменный размер)

Total Path Attribute Length, 2 байта

Path Attributes (переменный размер)

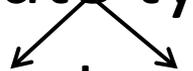
Network Layer Reachability Information (переменный размер)

Path Attributes – последовательность переменной длины с атрибутами пути.

Атрибут пути =

<attribute type, attribute length, attribute value>

| attr. Flags | attr. Type Code |



Сообщение UPDATE

Withdrawn Routes Length, 2 байта

Withdrawn Routes (переменный размер)

Total Path Attribute Length, 2 байта

Path Attributes (переменный размер)

Network Layer Reachability Information (переменный размер)

NLRI – последовательность переменной длины с адресными префиксами IP.

Префикс = <length, prefix>

Сообщение UPDATE может анонсировать не более одного набора атрибутов пути, но этому пути может соответствовать множество адресатов, путь к которым описывается общим набором атрибутов.

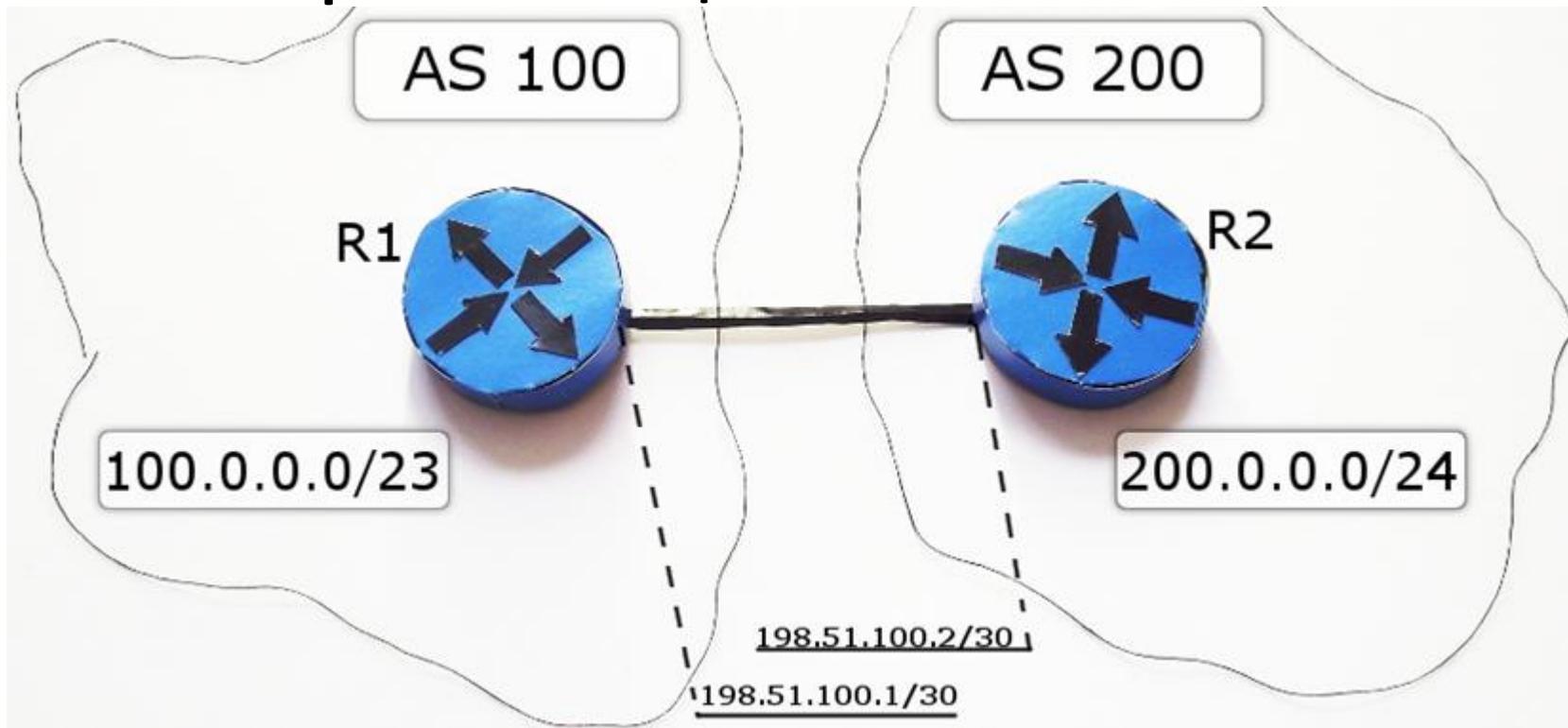
Сообщение KEEPALIVE и NOTIFICATION

BGP не использует на уровне TCP каких-либо механизмов для проверки доступности других узлов.

Вместо этого используются сообщения KEEPALIVE.

Сообщения NOTIFICATION передаются в случаях обнаружения ошибок. Соединение BGP незамедлительно закрывается после передачи такого сообщения.

Организация BGP сессии



```
R1(config-router)# do show ip bgp neighbor
BGP neighbor is 198.51.100.2, remote AS is 20, external link
BGP version 4, remote router ID 0.0.0.0
BGP state = Idel
```

Организация BGP сессии: TCP SYN

11	18:30:57.802710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [SYN] Seq=
12	18:30:57.822710	198.51.100.2	198.51.100.1	TCP	60 bgp > 33887 [SYN, ACK]
13	18:30:57.832710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [ACK] Seq=
14	18:30:57.852710	198.51.100.1	198.51.100.2	BGP	99 OPEN Message
15	18:30:57.872710	198.51.100.2	198.51.100.1	BGP	118 OPEN Message, KEEPALIVE
16	18:30:57.892710	198.51.100.1	198.51.100.2	BGP	73 KEEPALIVE Message
17	18:30:57.902710	198.51.100.1	198.51.100.2	BGP	92 KEEPALIVE Message, KEE
18	18:30:57.922710	198.51.100.2	198.51.100.1	BGP	92 KEEPALIVE Message, KEE
19	18:30:58.132710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [ACK] Seq=

- ⊕ Frame 11: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
- ⊕ Ethernet II, Src: c0:00:14:74:00:00 (c0:00:14:74:00:00), Dst: c0:01:14:74:00:00 (c0:01:14:74:00:00)
- ⊕ Internet Protocol Version 4, Src: 198.51.100.1 (198.51.100.1), Dst: 198.51.100.2 (198.51.100.2)
- ⊖ Transmission Control Protocol, Src Port: 33887 (33887), Dst Port: bgp (179), Seq: 0, Len: 0
 - Source port: 33887 (33887)
 - Destination port: bgp (179)
 - [Stream index: 3]
 - Sequence number: 0 (relative sequence number)
 - Header length: 24 bytes
- ⊕ Flags: 0x002 (SYN)
 - window size value: 16384
 - [Calculated window size: 16384]
- ⊕ Checksum: 0x2656 [validation disabled]
- ⊕ Options: (4 bytes), Maximum segment size

Организация BGP сессии: OPEN message

11	18:30:57.802710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [SYN] Seq=0
12	18:30:57.822710	198.51.100.2	198.51.100.1	TCP	60 bgp > 33887 [SYN, ACK]
13	18:30:57.832710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [ACK] Seq=1
14	18:30:57.852710	198.51.100.1	198.51.100.2	BGP	99 OPEN Message
15	18:30:57.872710	198.51.100.2	198.51.100.1	BGP	118 OPEN Message, KEEPALIVE
16	18:30:57.892710	198.51.100.1	198.51.100.2	BGP	73 KEEPALIVE Message
17	18:30:57.902710	198.51.100.1	198.51.100.2	BGP	92 KEEPALIVE Message, KEEP
18	18:30:57.922710	198.51.100.2	198.51.100.1	BGP	92 KEEPALIVE Message, KEEP
19	18:30:58.132710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [ACK] Seq=10

- ⊕ Frame 14: 99 bytes on wire (792 bits), 99 bytes captured (792 bits)
- ⊕ Ethernet II, Src: c0:00:14:74:00:00 (c0:00:14:74:00:00), Dst: c0:01:14:74:00:00 (c0:01:14:74:00:00)
- ⊕ Internet Protocol Version 4, Src: 198.51.100.1 (198.51.100.1), Dst: 198.51.100.2 (198.51.100.2)
- ⊕ Transmission Control Protocol, Src Port: 33887 (33887), Dst Port: bgp (179), Seq: 1, Ack: 1, Len: 45
- ⊖ Border Gateway Protocol - OPEN Message
 - Marker: ffffffffffffffffffffffffffffffff
 - Length: 45
 - Type: OPEN Message (1)
 - Version: 4
 - My AS: 100
 - Hold Time: 180
 - BGP Identifier: 198.51.100.1 (198.51.100.1)
 - Optional Parameters Length: 16
 - ⊖ Optional Parameters
 - ⊖ Optional Parameter: Capability
 - Parameter Type: Capability (2)
 - Parameter Length: 6
 - ⊕ Capability: Multiprotocol extensions capability
 - ⊖ Optional Parameter: Capability
 - Parameter Type: Capability (2)
 - Parameter Length: 2
 - ⊕ Capability: Route refresh capability
 - ⊕ Optional Parameter: Capability

Организация BGP сессии: KEEPALIVE

11	18:30:57.802710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [SYN] Seq=0
12	18:30:57.822710	198.51.100.2	198.51.100.1	TCP	60 bgp > 33887 [SYN, ACK]
13	18:30:57.832710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [ACK] Seq=1
14	18:30:57.852710	198.51.100.1	198.51.100.2	BGP	99 OPEN Message
15	18:30:57.872710	198.51.100.2	198.51.100.1	BGP	118 OPEN Message, KEEPALIVE
16	18:30:57.892710	198.51.100.1	198.51.100.2	BGP	73 KEEPALIVE Message
17	18:30:57.902710	198.51.100.1	198.51.100.2	BGP	92 KEEPALIVE Message, KEEPALIVE
18	18:30:57.922710	198.51.100.2	198.51.100.1	BGP	92 KEEPALIVE Message, KEEPALIVE
19	18:30:58.132710	198.51.100.1	198.51.100.2	TCP	60 33887 > bgp [ACK] Seq=10

- ⊕ Frame 15: 118 bytes on wire (944 bits), 118 bytes captured (944 bits)
- ⊕ Ethernet II, Src: c0:01:14:74:00:00 (c0:01:14:74:00:00), Dst: c0:00:14:74:00:00 (c0:00:14:74:00:00)
- ⊕ Internet Protocol Version 4, Src: 198.51.100.2 (198.51.100.2), Dst: 198.51.100.1 (198.51.100.1)
- ⊕ Transmission Control Protocol, Src Port: bgp (179), Dst Port: 33887 (33887), Seq: 1, Ack: 46, Len: 64
- ⊖ Border Gateway Protocol - OPEN Message
 - Marker: ffffffffffffffffffffffffffffffff
 - Length: 45
 - Type: OPEN Message (1)
 - Version: 4
 - My AS: 200
 - Hold Time: 180
 - BGP Identifier: 198.51.100.2 (198.51.100.2)
 - Optional Parameters Length: 16
 - ⊕ Optional Parameters
- ⊖ Border Gateway Protocol - KEEPALIVE Message
 - Marker: ffffffffffffffffffffffffffffffff
 - Length: 19
 - Type: KEEPALIVE Message (4)

Обмен маршрутной информацией: UPDATE

272	18:44:45.729710	198.51.100.1	198.51.100.2	BGP	106	UPDATE	Me
273	18:44:45.759710	198.51.100.2	198.51.100.1	BGP	106	UPDATE	Me
274	18:44:45.799710	198.51.100.1	198.51.100.2	BGP	92	KEEPALIVE	
275	18:44:45.819710	198.51.100.2	198.51.100.1	BGP	92	KEEPALIVE	
276	18:44:46.029710	198.51.100.1	198.51.100.2	TCP	60	60882 > b	

- ⊕ Frame 272: 106 bytes on wire (848 bits), 106 bytes captured (848 bits)
- ⊕ Ethernet II, Src: c0:00:14:74:00:00 (c0:00:14:74:00:00), Dst: c0:01:14:74:00:00 (c0:01:14:74:00:00)
- ⊕ Internet Protocol version 4, Src: 198.51.100.1 (198.51.100.1), Dst: 198.51.100.2 (198.51.100.2)
- ⊕ Transmission Control Protocol, Src Port: 60882 (60882), Dst Port: bgp (179), Seq: 65, A
- ⊖ Border Gateway Protocol - UPDATE Message
 - Marker: ffffffffffffffffffffffffffffffffff
 - Length: 52
 - Type: UPDATE Message (2)
 - Unfeasible routes length: 0 bytes
 - Total path attribute length: 25 bytes
 - ⊖ Path attributes
 - ⊕ ORIGIN: IGP (4 bytes)
 - ⊕ AS_PATH: 100 (7 bytes)
 - ⊕ NEXT_HOP: 198.51.100.1 (7 bytes)
 - ⊕ MULTI_EXIT_DISC: 0 (7 bytes)
 - ⊖ Network layer reachability information: 4 bytes
 - ⊖ 100.0.0.0/23
 - NLRI prefix length: 23
 - NLRI prefix: 100.0.0.0 (100.0.0.0)

Атрибуты пути

Информация о новых
или удалённых маршрутах

Выбор маршрута

- Максимальное значение Weight
- Максимальное значение Local Preference
- Предпочесть локальный маршрут маршрутизатора
(next hop = 0.0.0.0)
- Кратчайший путь через автономные системы
(короткий AS_PATH)
- Минимальное значение Origin Code
- Минимальное значение MED
- Путь eBGP лучше чем путь iBGP
- и т. д.

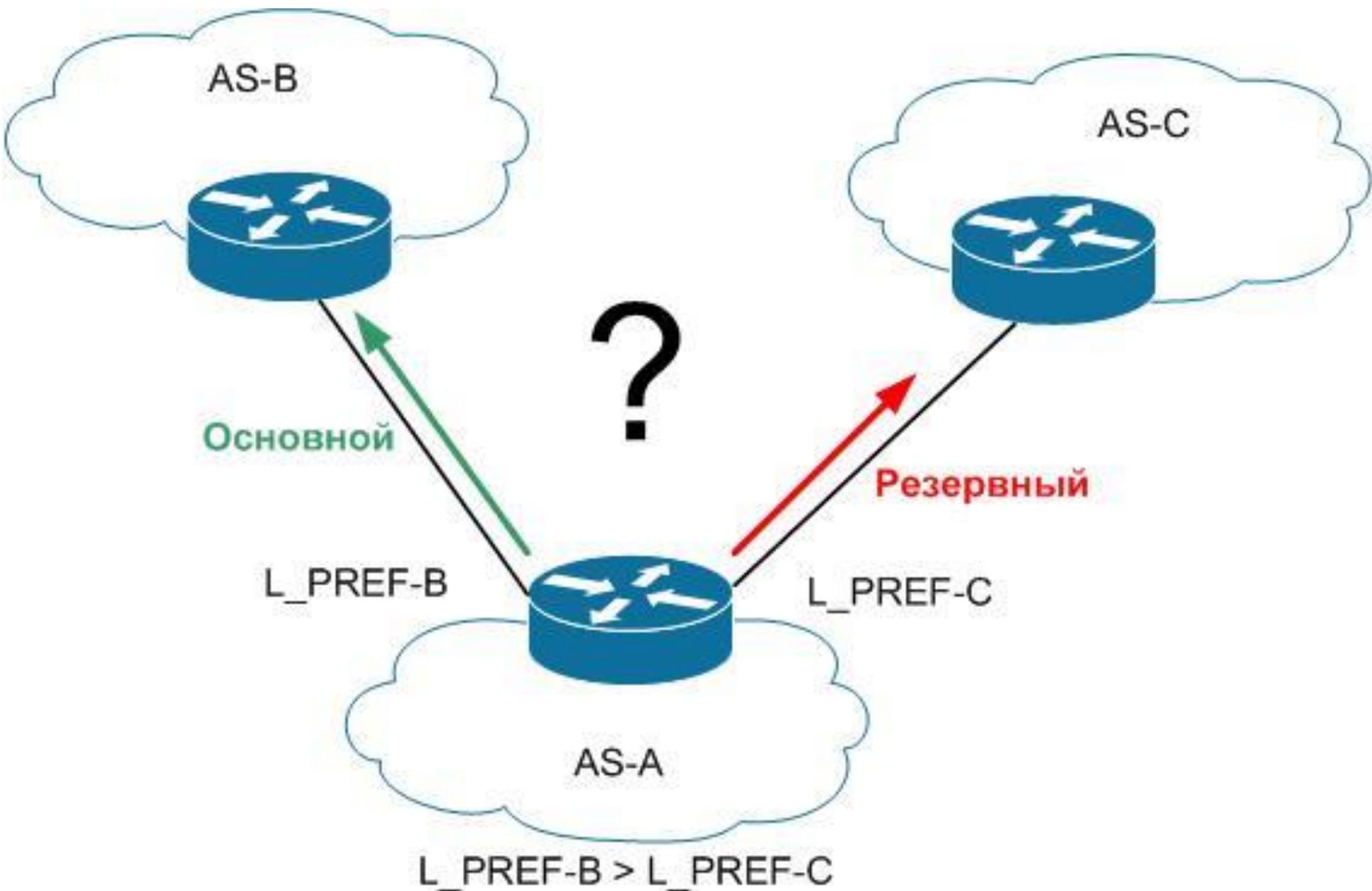
Инструменты управления трафиком

AS-Path ACL

Prefix-list

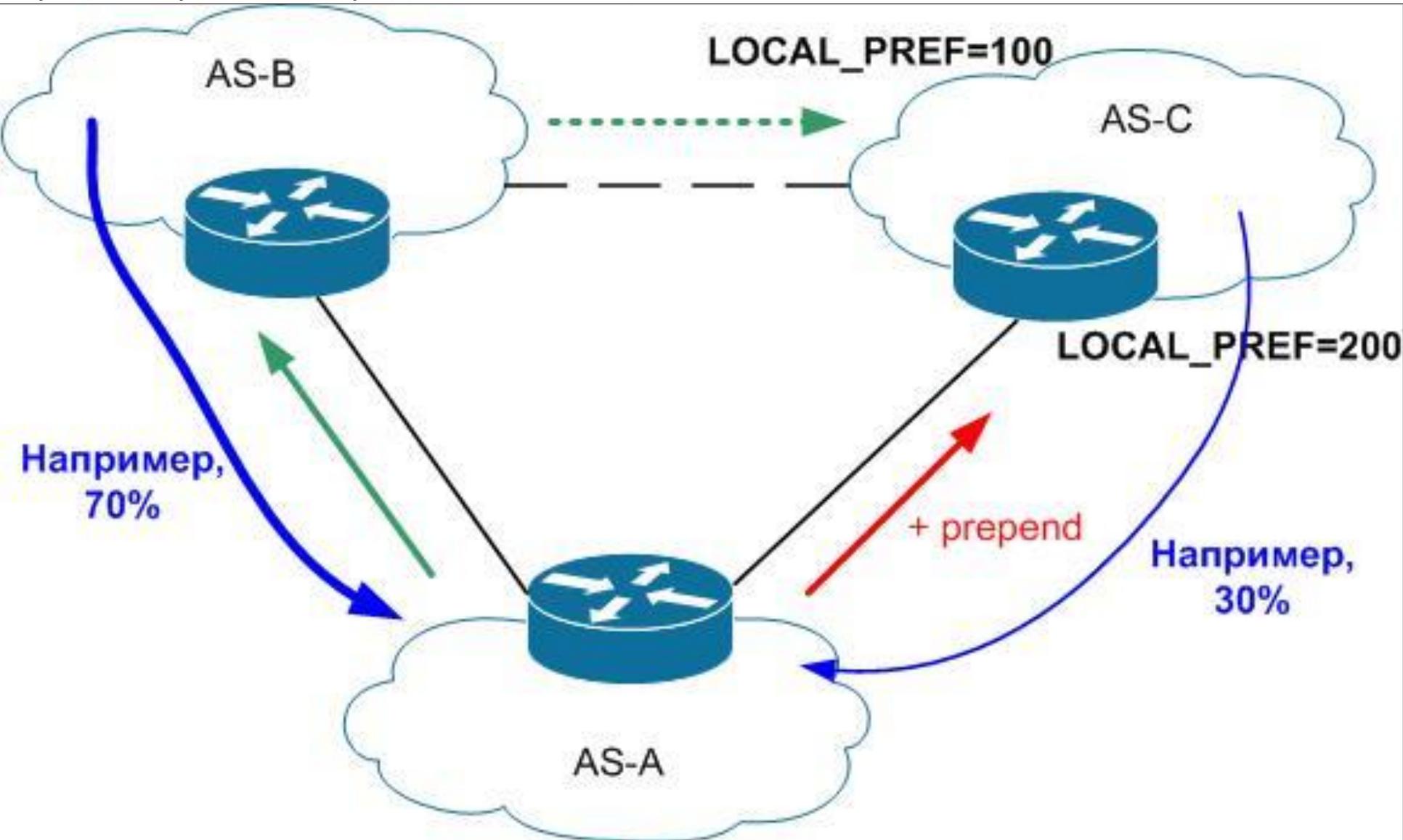
Входящий	Исходящий
AS_Path prepend	Weight
MED	Local Preference
Community	BGP multipath
Анонс разных префиксов через разных ISP	

Исходящий трафик

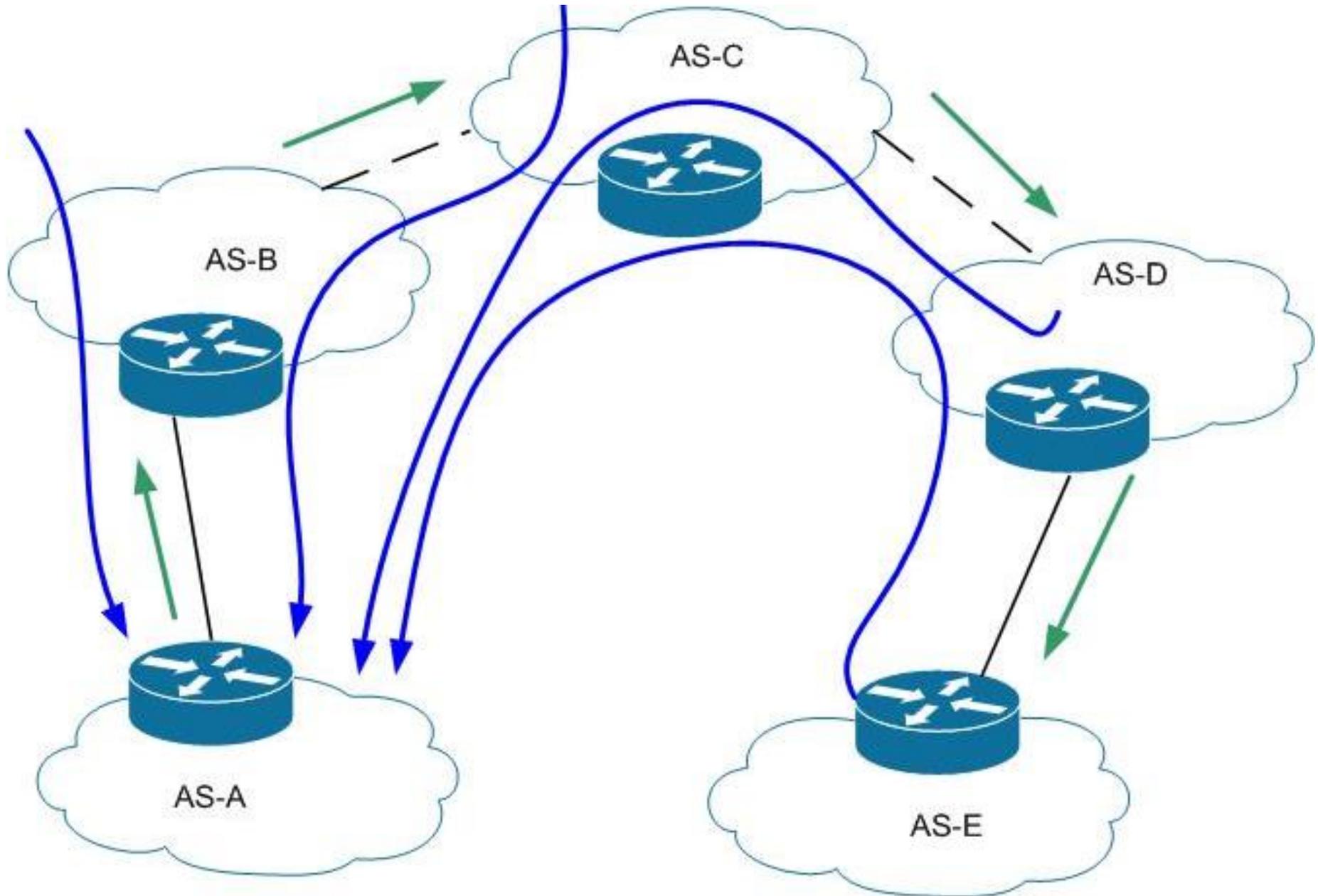


Анонсируем свои сети обоим провайдерам, но в сторону резервного мы специально удлиняем AS_PATH.

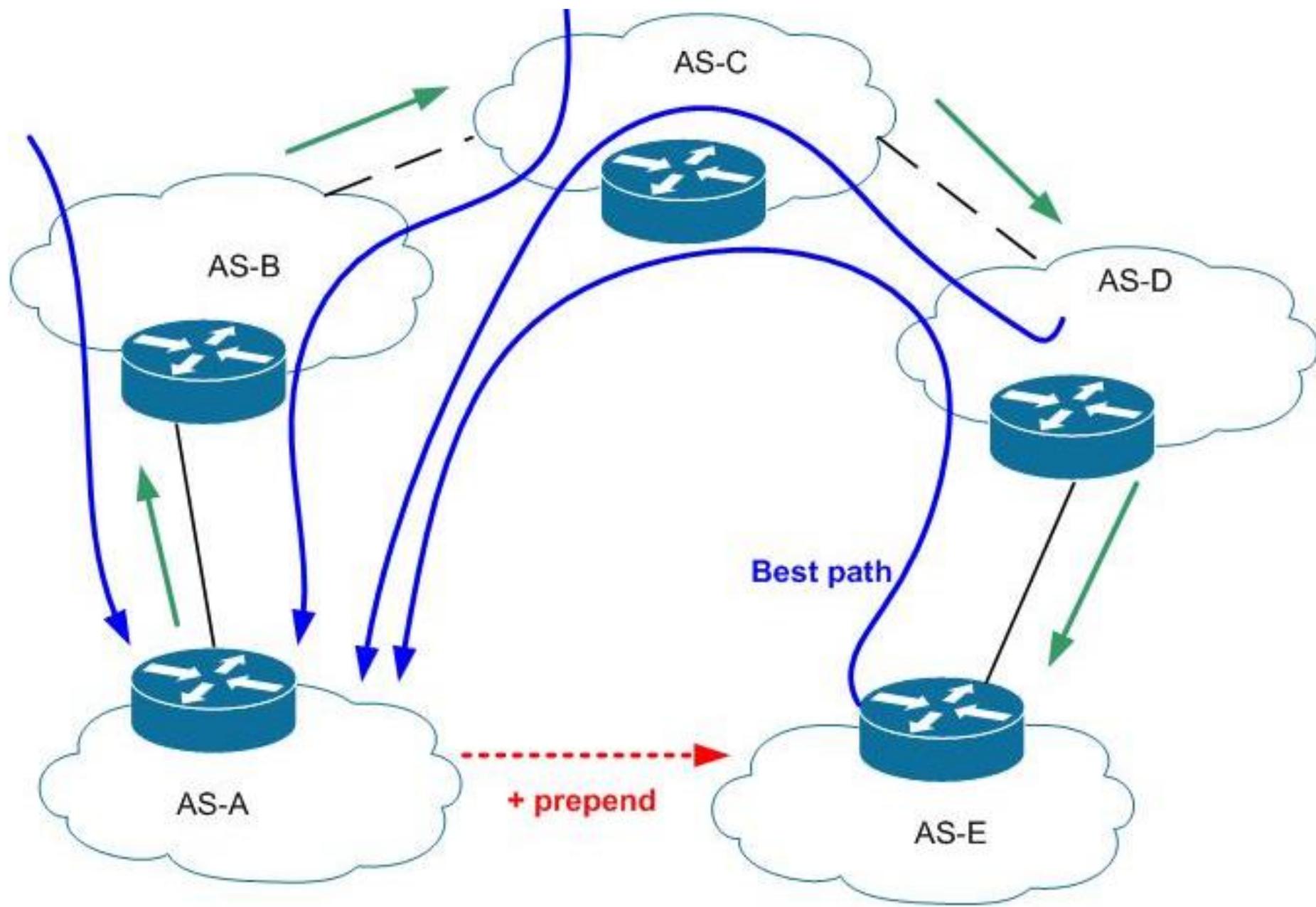
Однако провайдер С считает более приоритетным путем в клиентскую сеть тот путь, который непосредственно соединяет его с клиентом.



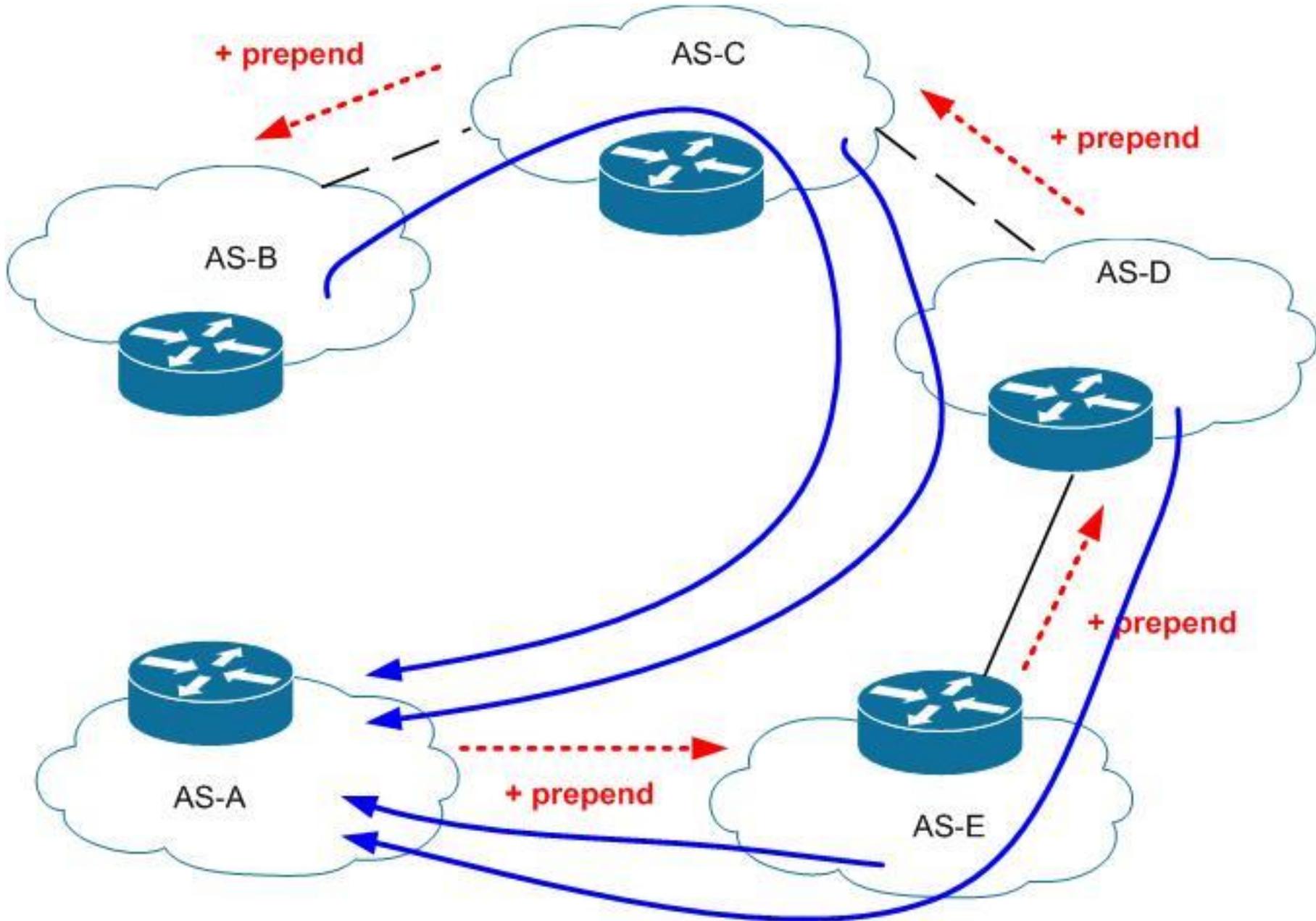
Влияние порядка установки сеанса



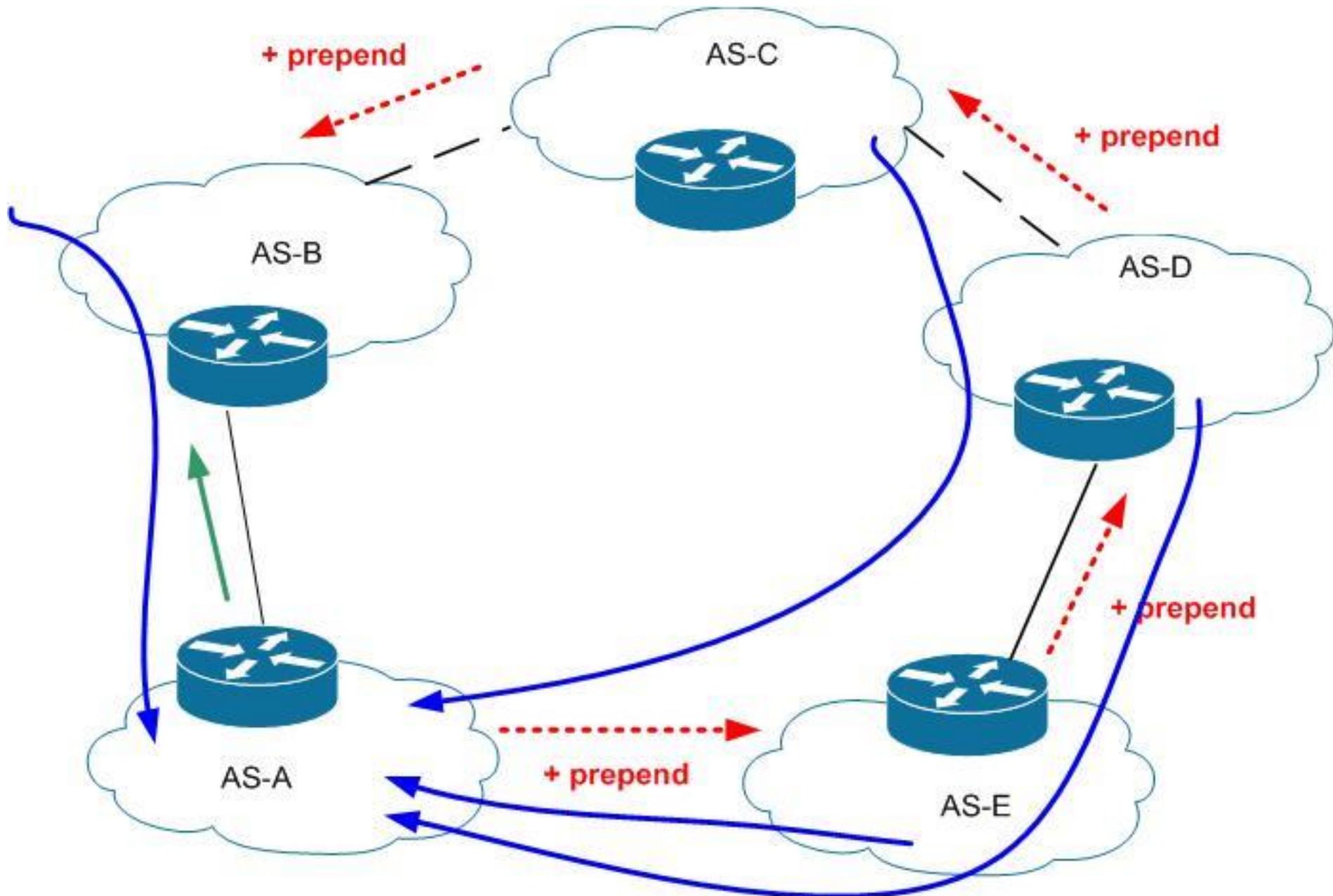
Установили связь партнером



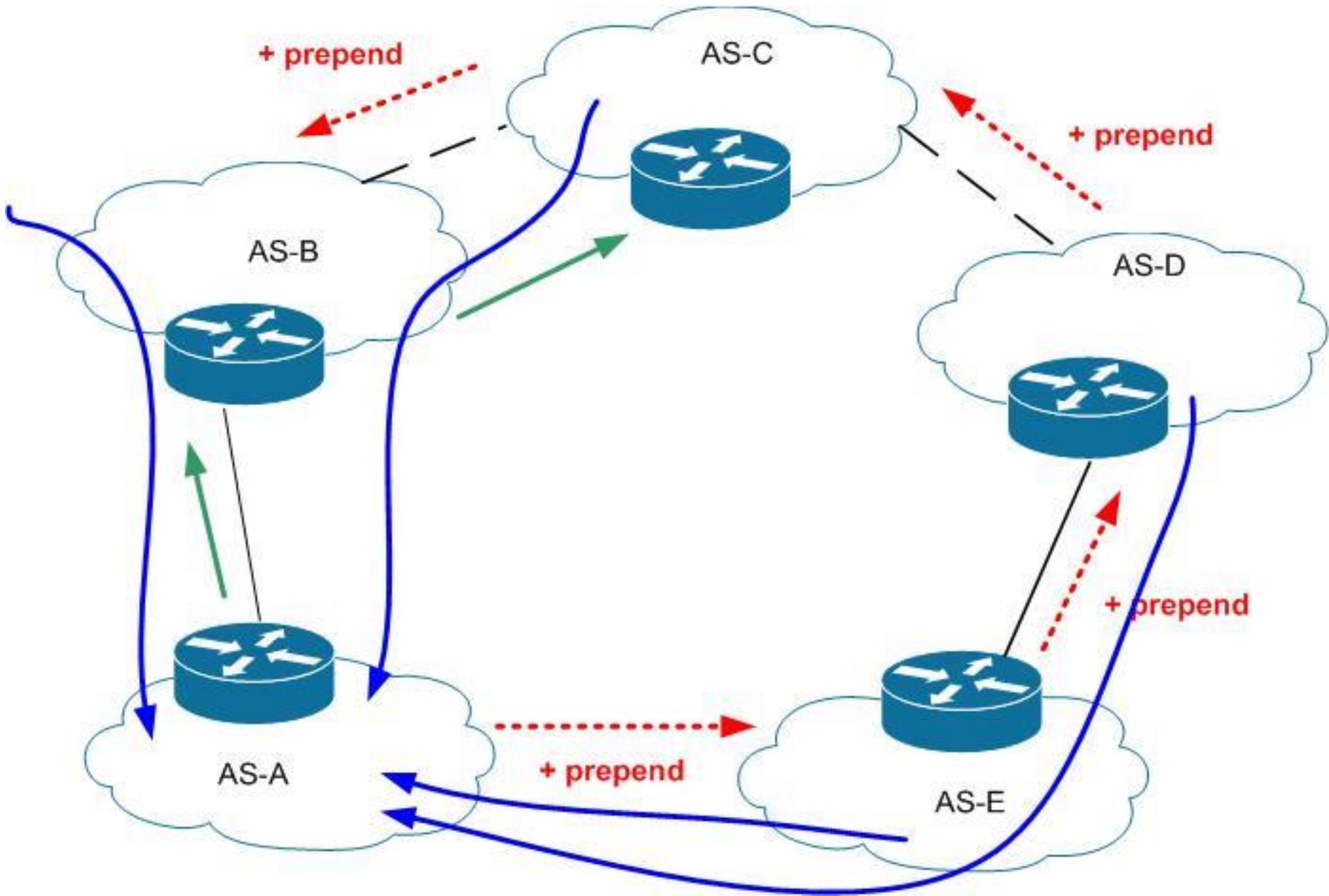
Обрыв канала с основным провайдером



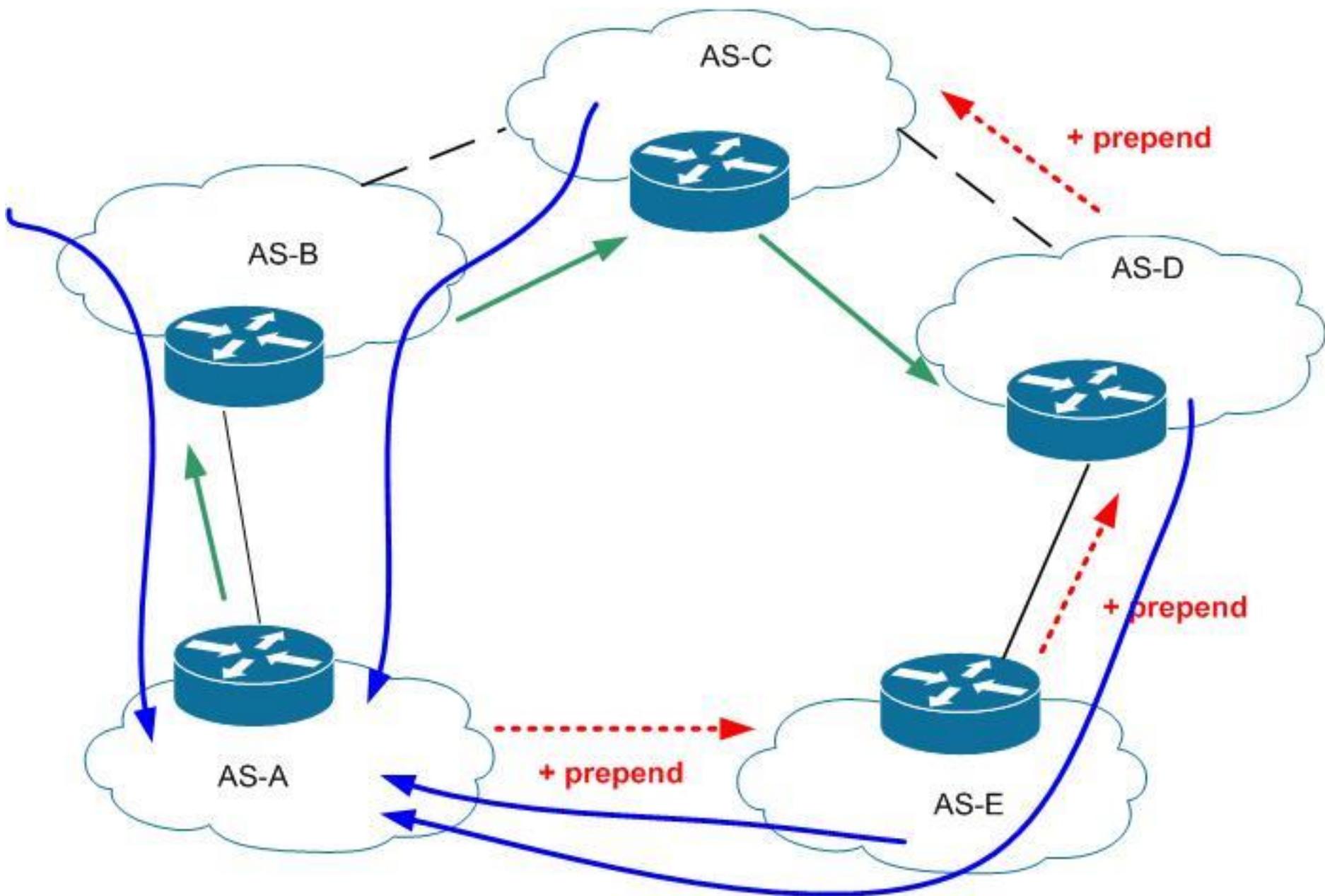
Восстановился канал до AS-B



Анонс распространился до AS-C

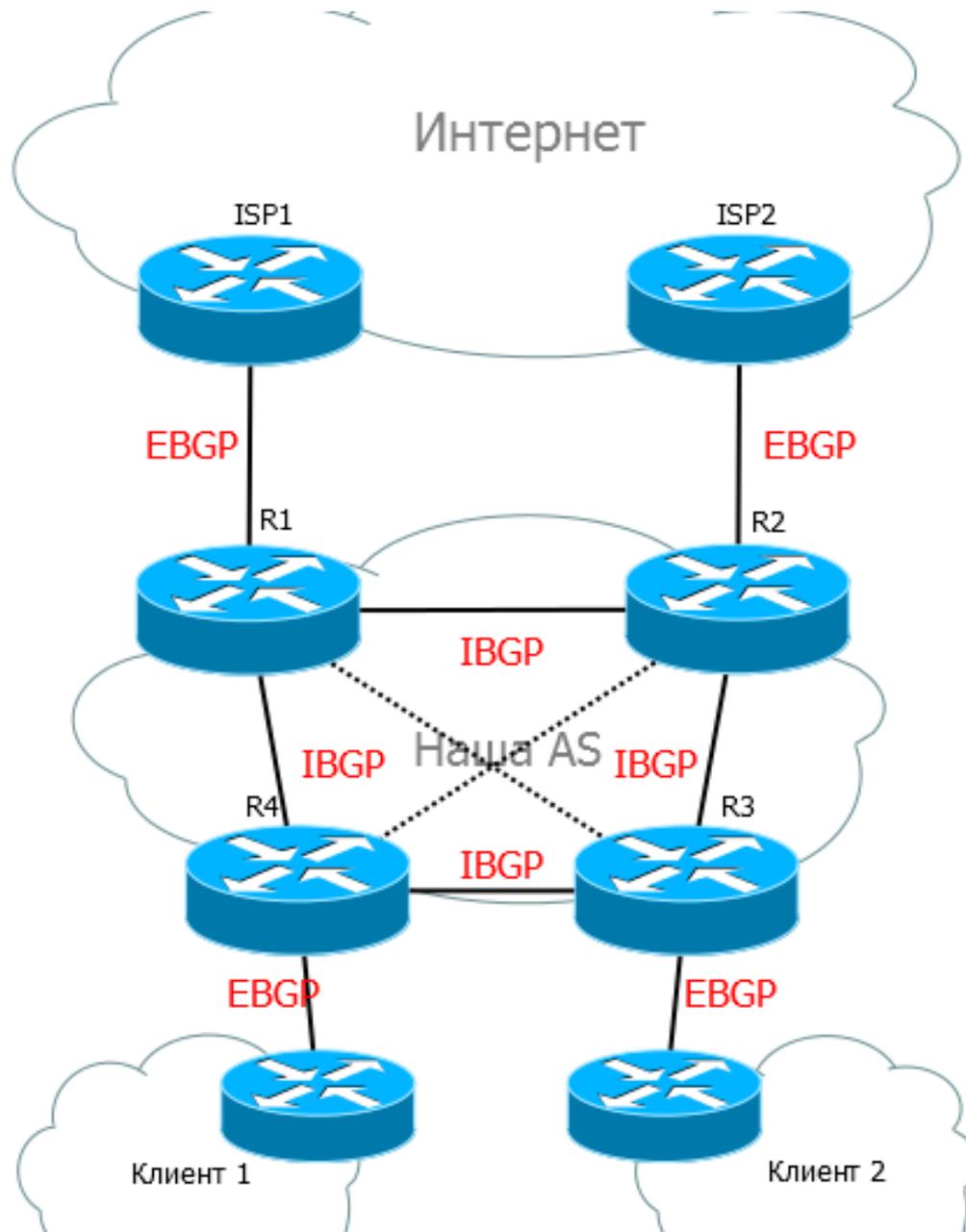


AS-D продолжает отправлять через AS-E

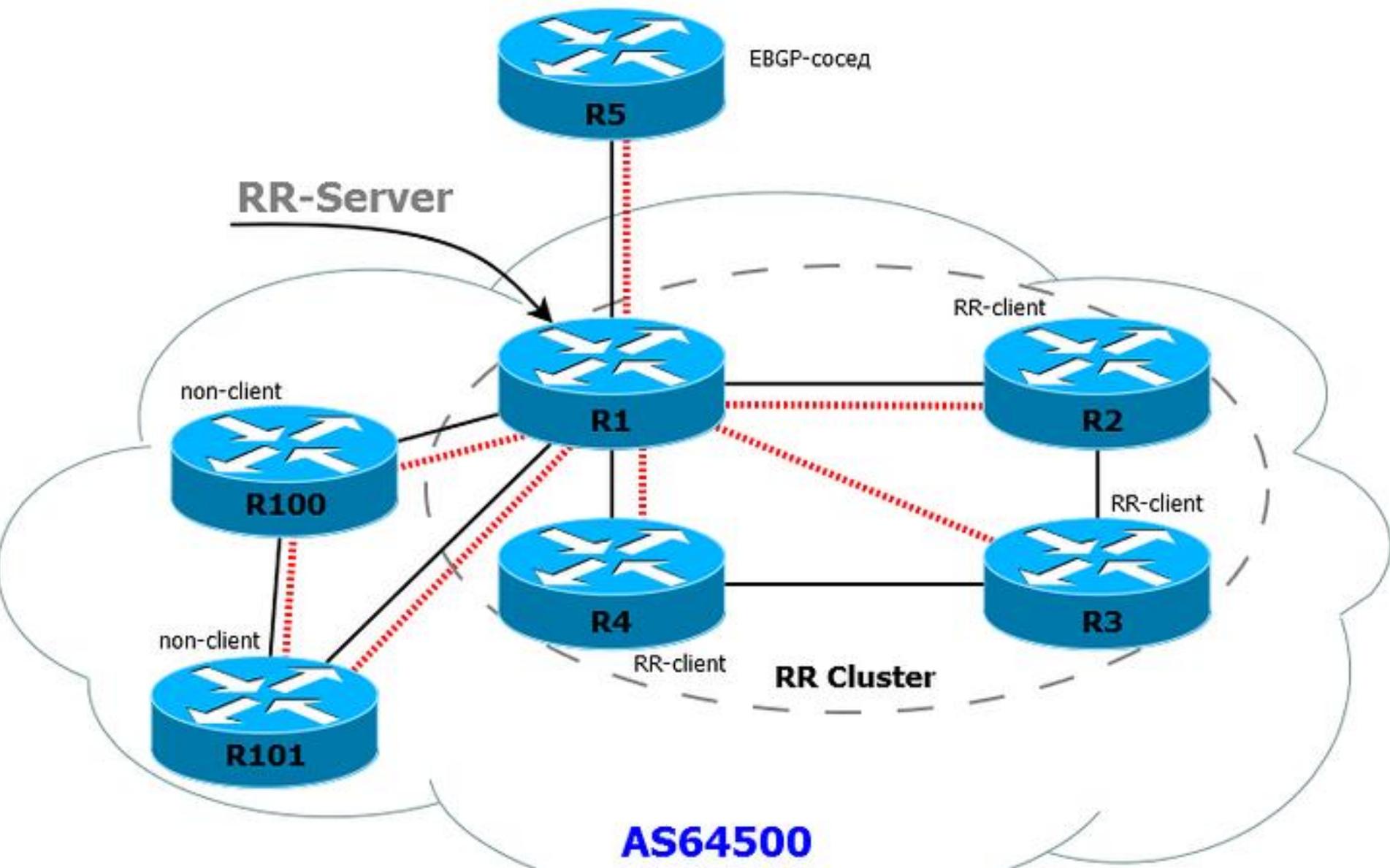


Internal BGP

1. Связность Full mesh
2. Next hop на выход из автономной системы



Route Reflector



Конфедерации

Когда маршруты передаются внутри АС между конфедерациями в их AS-Path добавляется номер конфедерации (сегменты AS_CONFED_SEQ и AS_CONFED_SET) для избежания петель.

