

ROUTING

Part 3.

EIGRP

Enhanced Interior Gateway Routing Protocol

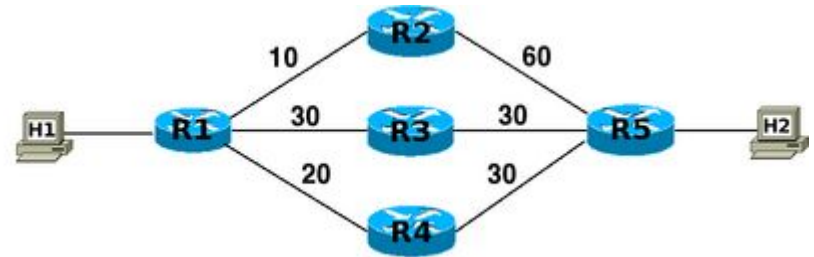
EIGRP использует 5 типов сообщений

- **Hello** — маршрутизаторы используют hello-пакеты для обнаружения соседей. Пакеты отправляются multicast и не требуют подтверждения о получении.
- **Update** — содержится информация об изменении маршрутов. Они отправляются только маршрутизаторам, которых касается обновление. Эти пакеты могут быть отправлены конкретному маршрутизатору (unicast) или группе маршрутизаторов (multicast). Получение update-пакета подтверждается отправкой ACK.
- **Query** — когда маршрутизатор выполняет подсчет маршрута и у него нет feasible successor, он отправляет query-пакет своим соседям для того чтобы определить нет ли feasible successor для этого destination у них. Обычно query-пакеты отправляются multicast, но могут быть и unicast. Получение query-пакета подтверждается отправкой ACK получателем пакета.
- **Reply** — маршрутизатор отправляет reply-пакет в ответ на query-пакет. Reply-пакеты отправляются unicast тому, кто отправил query-пакет. Получение reply-пакета подтверждается отправкой ACK.
- **ACK** — пакет, который подтверждает получение пакетов update, query, reply. ACK-пакеты отправляются unicast и содержат в себе acknowledgment number. Фактически это hello-пакеты, которые не передают данных. Используется негарантированная доставка.

Diffusing Update Algorithm (DUAL)

- Терминология:
- **Advertised distance (AD)** — стоимость расстояния между соседним маршрутизатором, который анонсирует маршрут и сеть назначения.
- **Feasible distance (FD)** — стоимость расстояния от локального маршрутизатора до сети назначения = AD, которое анонсирует соседний маршрутизатор + стоимость расстояния между локальным маршрутизатором и соседним маршрутизатором.
- **Successor** — соседний маршрутизатор с путем без петель и с наименьшей стоимостью пути к сети назначения.
- **Feasible successor** — резервный маршрутизатор с путем без петель (AD feasible successor должно быть меньше чем FD текущего маршрута successor).
- **Feasible condition** — AD feasible successor должно быть меньше чем FD текущего маршрута successor.

Пример выбора лучшего маршрута и feasible successor



| Соседний маршрутизатор | Роль соседнего маршрутизатора | Статус маршрута | FD | AD | Feasible condition |
|------------------------|-------------------------------|-------------------|--------------------|-------------|--|
| R2 | Не выбран | Не выбран | $R1R2 + R2R5 = 70$ | $R2R5 = 60$ | Не выполняется |
| R3 | Feasible successor | Резервный маршрут | $R1R3 + R3R5 = 60$ | $R3R5 = 30$ | Выполняется $AD < FD$ ($R3R5 < R1R4R5$) |
| R4 | Successor | Лучший маршрут | $R1R4 + R4R5 = 50$ | $R4R5 = 30$ | Лучший маршрут |

Метрика EIGRP

Метрика EIGRP основана на таких 5 компонентах (по умолчанию используются только два):

- Bandwidth — наименьшая bandwidth между source и destination (используется по умолчанию);
- Delay — cumulative interface delay всего пути;
- Reliability — наихудший показатель надежности на всем пути, на основании keepalive;
- Loading — наихудший показатель загрузки линка на всем пути, на основании packet rate и настроенной bandwidth на интерфейсе;
- MTU — наименьшее MTU на всем пути. MTU включается в обновления EIGRP, но фактически не используется для подсчета метрики.

По умолчанию для подсчета метрики используются bandwidth и delay. Остальные критерии не рекомендуется использовать, так как это приведет к частым пересчетам маршрутов.

Вычисление метрики маршрута

- EIGRP подсчитывает метрику с использованием коэффициентов. По умолчанию значения коэффициентов такие: $K1 = K3 = 1$, $K2 = K4 = K5 = 0$.
- Общая метрика вычисляется при помощи значений bandwidth (пропускной способности) и delay (задержки). Используется следующая формула для вычисления значения **bandwidth** :

$$\text{bandwidth} = (10000000/\text{bandwidth}(i)) * 256$$

Где bandwidth(i) является **наименьшей** пропускной способностью (bandwidth) из всех исходящих интерфейсов по пути в сеть назначения представленная в килобитах.

- Формула для вычисления значения **delay**:
- $\text{delay} = \text{delay}(i) * 256$

где $\text{delay}(i)$ является **суммой** всех задержек (delays) сконфигурированных на исходящих интерфейсах по пути в сеть назначения в **десятках микросекунд** (microseconds).
Задержка (delay) показываемая командой **ip eigrp topology** или **show interface** указана в **микросекундах**, соответственно это значение нужно поделить на 10 перед использованием в этой формуле.

- При вычислении метрики, когда $K5 = 0$ (значение по умолчанию), используется такая формула:

$$\text{Metric} = (K1 * \text{bandwidth}) + [(K2 * \text{bandwidth}) / (256 - \text{load})] + (K3 * \text{delay})$$

Если значения коэффициентов $K1$, $K2$, $K3$ равны значениям по умолчанию, то формула превращается в такую:

$$\text{Metric} = \text{bandwidth} + \text{delay}$$

- Если $K5$ не равно 0, то дополнительно выполняется такая операция:

$$\text{Metric} = \text{metric} * [K5 / (\text{reliability} + K4)]$$

r0#show int e0

Ethernet0 is up, line protocol is up

Hardware is Lance, address is 0060.5cf3.bb1e (bia 0060.5cf3.bb1e)

Internet address is 10.50.0.1/24

MTU 1500 bytes, BW 10000 Kbit, DLY 1000 usec,
reliability 255/255, txload 1/255, rxload 1/255

Encapsulation ARPA, loopback not set

Keepalive set (10 sec)

ARP type: ARPA, ARP Timeout 04:00:00

Last input 00:00:00, output 00:00:00, output hang never

Last clearing of "show interface" counters never

Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0

Queueing strategy: fifo

Output queue: 0/40 (size/max)

5 minute input rate 3000 bits/sec, 4 packets/sec

5 minute output rate 3000 bits/sec, 4 packets/sec

2212623 packets input, 2653204211 bytes, 0 no buffer

Received 2158768 broadcasts, 0 runts, 0 giants, 0 throttles

0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

0 input packets with dribble condition detected

1054688 packets output, 716849118 bytes, 0 underruns

18 output errors, 6 collisions, 22 interface resets

0 babbles, 0 late collision, 109 deferred

18 lost carrier, 0 no carrier

0 output buffer failures, 0 output buffers swapped out

```
r0#show ip eigrp topology 10.50.0.0/24
```

```
IP-EIGRP (AS 44): Topology entry for 10.50.0.0/24
```

```
State is Passive, Query origin flag is 1, 1 Successor(s),FD  
is 281600
```

```
Routing Descriptor Blocks:
```

```
0.0.0.0 (Ethernet0), from Connected, Send flag is 0x0
```

```
Composite metric is (281600/0), Route is Internal
```

```
Vector metric:
```

```
Minimum bandwidth is 10000 Kbit
```

```
Total delay is 1000 microseconds
```

```
Reliability is 255/255
```

```
Load is 1/255
```

```
Minimum MTU is 1500
```

```
Hop count is 0
```

Базовые настройки

Создание процесса EIGRP:

```
dyn1(config)# router eigrp <AS>
```

Номер автономной системы должен быть одинаковым на всех маршрутизаторах, которые должны обмениваться информацией по протоколу EIGRP.

Включение EIGRP на интерфейсах:

```
dyn1(config)# router eigrp 100
```

```
dyn1(config-router)# network <network> [wildcard mask]
```

BGP

Терминология протокола

- **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.
- **Префикс** — IP-префикс (длина префикса).

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.
- **Внешний BGP (External BGP, eBGP)** — BGP работающий между автономными системами. По умолчанию, eBGP-соседи должны быть непосредственно соединены.

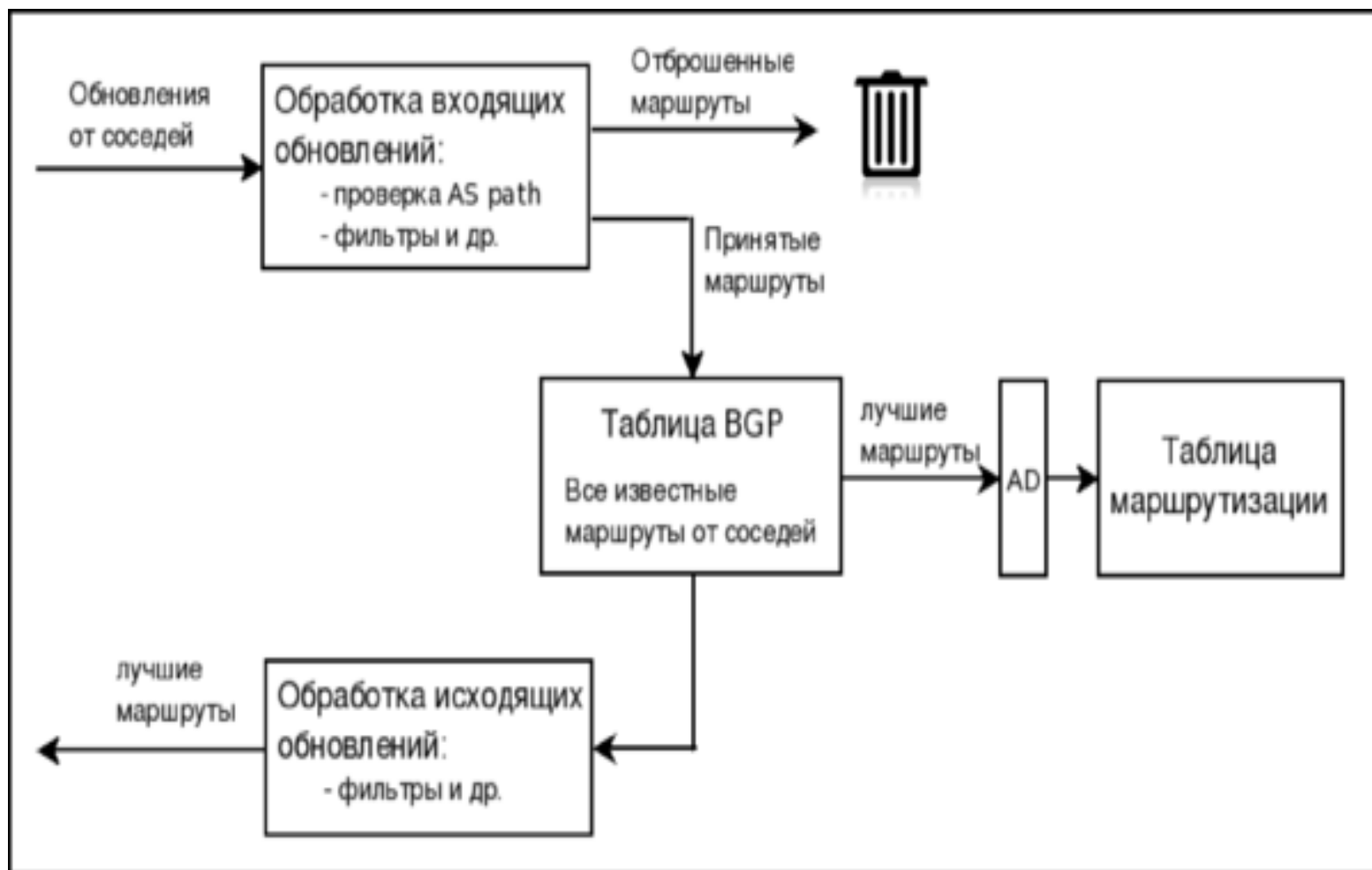
Таймеры протокола

- **Keepalive Interval** — Интервал времени в секундах, между отправкой сообщений keepalive. По умолчанию 60 секунд.
- **Hold Time** — Интервал времени в секундах, по истечении которого сосед будет считаться недоступным. По умолчанию 180 секунд.

Состояния связи с соседями

- Idle
- Connect
- Open sent
- Open confirm
 - active
- Established

| 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 |



BGP Attributes

- **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.

- Well-known mandatory:
 - Autonomous system path
 - Next-hop
 - Origin
- Well-known discretionary:
 - Local preference
- Optional transitive:
 - Communities
 - Multi-exit discriminator (MED)
- Optional non-transitive:
 - Weight

- RTC#
- router bgp 300
- neighbor 1.1.1.1 remote-as 100
- neighbor 1.1.1.1 weight 200

- !--- The route to 175.10.0.0 from RTA has a 200 weight.

- neighbor 2.2.2.2 remote-as 200
- neighbor 2.2.2.2 weight 100

- !--- The route to 175.10.0.0 from RTB has a 100 weight.

- RTC#
- router bgp 256
- neighbor 1.1.1.1 remote-as 100
- neighbor 128.213.11.2 remote-as 256
- bgp default local-preference 150

- RTD#
- router bgp 256
- neighbor 3.3.3.4 remote-as 300
- neighbor 128.213.11.1 remote-as 256
- bgp default local-preference 200

```
route-map RMAP-1 permit 30
  set local-preference 200
```

```
ip prefix-list PREF-IN seq 100 deny 10.1.1.120/30
ip prefix-list PREF-IN seq 110 deny 10.1.1.116/30
ip prefix-list PREF-IN seq 200 permit 0.0.0.0/0 le 32
!
```

```
ip prefix-list PREF-OUT seq 40 deny 10.1.3.172/30
ip prefix-list PREF-OUT seq 50 deny 10.1.1.116/30
ip prefix-list PREF-OUT seq 100 permit 0.0.0.0/0 le 32
```

```
router bgp 256
neighbor 10.10.3.234 prefix-list PREF-IN in
neighbor 10.10.3.234 prefix-list PREF-OUT out
neighbor 10.10.3.234 route-map RMAP-1 in
```



```
router bgp 256
```

```
bgp always-compare-med
```

```
neighbor 4.4.4.4 remote-as 100
```

```
neighbor 4.4.4.4 route-map setmetricout out
```

```
route-map setmetricout permit 10
```

```
set metric 50
```

```
router bgp 256
```

```
(bgp bestpath med confed)
```

```
neighbor 10.1.14.2 route-map INC_METRIC in
```

```
route-map INC_METRIC permit 10
```

```
set metric +1
```

Community Attribute

```
ip prefix-list ALL-OUT seq 100 permit 1.1.0.0/20
```

```
ip prefix-list ALL-OUT seq 110 permit 2.2.0.0/20
```

```
ip prefix-list ALLOUT seq 120 permit 3.3.0.0/21
```

```
ip prefix-list ALL-OUT seq 130 permit 4.4.0.0/21
```

```
ip prefix-list ALL-OUT seq 140 permit 5.5.0.0/20
```

```
route-map MEGAFON-AS31133-OUT permit 30
```

```
match ip address prefix-list ALL-OUT
```

```
set community 31133:10009 31133:15009 31133:20009  
31133:25009
```

- remarks: | Community to INTERNATIONAL UPSTREAMS
 remarks: +-----+
 remarks: | 31133:1000x - All international upstreams
 remarks: | 31133:1101x - TeliaSonera (AS1299)
 remarks: | 31133:1102x - TelecomItalia (AS6762)
 remarks: | 31133:1104x - Level3 (AS3356)
 remarks: | 31133:1105x - Global Crossing (AS3549)
 remarks: |
 remarks: | x = 0 - advertise to this only (don't advertise to any other Upstream)
 remarks: | x = 1,2,3 or 5 - prepend 31133 (1,2,3 or 5) times
 remarks: | x = 8 - prepend AS31133 3 times to all other International Upstreams
 remarks: | x = 9 - don't advertise
 remarks: +-----+
 remarks: |
 remarks: +-----+
 remarks: | Community to RUSSIAN UPSTREAMS
 remarks: +-----+
 remarks: | 31133:1500x - All russian upstreams
 remarks: | 31133:1604x - GoldenTelecom (AS3216)
 remarks: |
 remarks: | x = 0 - advertise to this only (don't advertise to any other Upstream)
 remarks: | x = 1,2,3 or 5 - prepend 31133 (1,2,3 or 5) times
 remarks: | x = 8 - prepend AS31133 3 times to all other Russian Upstreams)
 remarks: | x = 9 - don't advertise
 remarks: +-----

- remarks: | Community to INTERNATIONAL PEERS
remarks: +-----+
remarks: | 31133:2000x - All International peers
remarks: | 31133:2101x - DE-CIX (AS6695)
remarks: | 31133:2102x - Private Peers at the DE-CIX public vlan
remarks: | 31133:2103x - UPL-Telecom
remarks: | 31133:2104x - NetNOD
remarks: | 31133:2105x - Private Peers at the NetNOD public vlan
remarks: | 31133:2201x - AMS-IX (AS1002)
remarks: | 31133:2202x - Private Peers at the AMS-IX public vlan
remarks: |
remarks: | x = 0 - advertise to this only (don't advertise to any other Peer)
remarks: | x = 1,2,3 or 5 - prepend 31133 (1,2,3 or 5) times
remarks: | x = 9 - don't advertise
remarks: +-----+

remarks: | Community to RUSSIAN PEERS
remarks: | 31133:2500x - All russian peers
remarks: | 31133:2601x - MSK-IX Route Servers (AS8631)
remarks: | 31133:2602x - Private Peers at the MSK-IX public vlan
remarks: | 31133:2605x - NCnet (AS42610)
remarks: | 31133:2606x - Global Net (AS31500)
remarks: | 31133:2608x - Orange Business Services (AS2854)
remarks: | 31133:2610x - RuTube (AS48061)
remarks: | 31133:2611x - Mnogobayt (AS42632)
remarks: | 31133:2612x - Mail.ru (AS47764)
remarks: | 31133:2614x - Google (AS15169)
remarks: | 31133:2619x - EuroTransTelecom (AS35320)
remarks: | 31133:1602x - TTK (AS20485)
remarks: | 31133:2702x - Yandex (AS13238)
remarks: |
remarks: | x = 0 - advertise to this only (don't advertise to any other
Peer)
remarks: | x = 1,2,3 or 5 - prepend 31133 (1,2,3 or 5) times
remarks: | x = 9 - don't advertise

ADVANCED ROUTING

- REDISTRIBUTION
- PBR