



Networking
For everyone

IP Routing

Сегодня в выпуске

- Процесс маршрутизации
- Longest Prefix Match (LPM)
- Метрика или Административная дистанция?
- Маршрутизация через интерфейс и через Next-Hop. В чем разница?



Шаг 1. Процесс маршрутизации

- Что делает маршрутизатор при получении IP пакета?
- Ищет LPM
 - `show ip route 1.1.1.1`
 - 1.0.0.0/8, via 2.2.2.2
 - 1.1.0.0/16, via 3.3.3.3
 - 0.0.0.0/0, via 4.4.4.4
- Рекурсивный лупап чтобы найти исходящий интерфейс
 - 2.2.2.2 via 5.5.5.5
 - 5.5.5.5 via 6.6.6.6
 - 6.6.6.6, Directly Connected, Gi0/1



Метрика или Административная дистанция?

- А что, если существуют несколько «одинаковых» LPM?
 - Получены через один протокол маршрутизации
 - Выбираем наименьшую метрику*
 - Правила могут варьироваться
 - OSPF Best Path Selection
 - Получены через разные протоколы маршрутизации
 - Выбираем наименьшую административную дистанцию

Route Source	Default AD
Connected	0
Static	1
EIGRP Summary	5
BGP External	20
EIGRP Internal	90
OSPF	110
IS-IS	115
RIP	120
EIGRP External	170
BGP Internal	200
NHRP	250
Unknown/Not valid	255



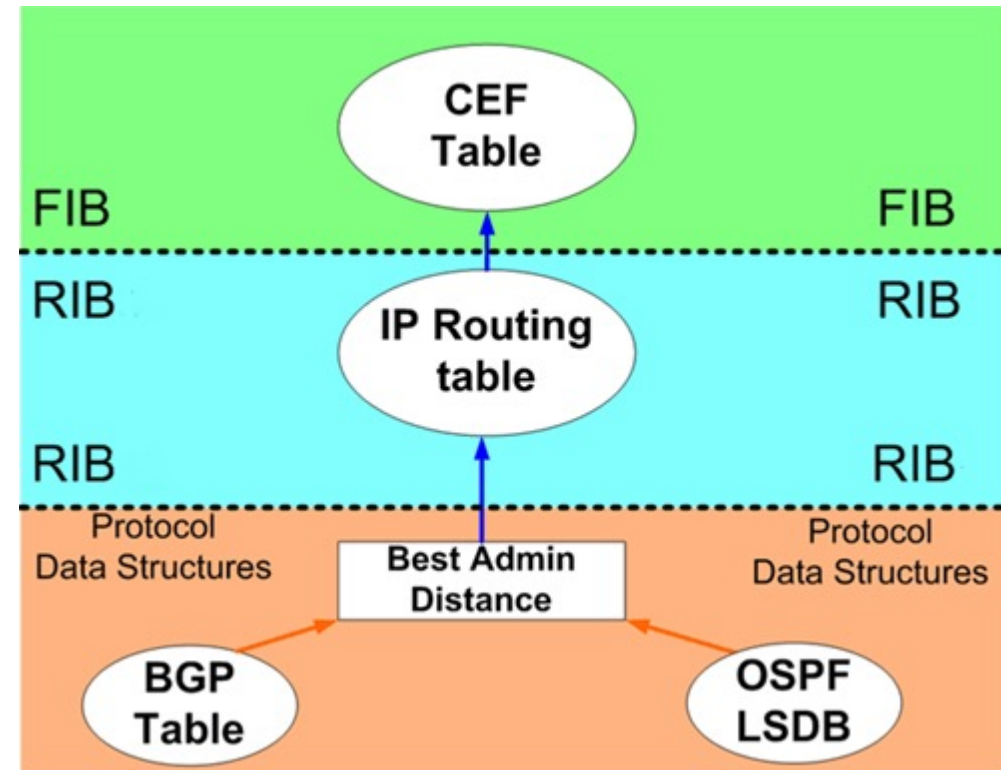
Шаг 2. Коммутация

- После того, как исходящий интерфейс был выбран, необходимо отправить на него пакет
 - В зависимости от устройства могут использоваться разные методы
 - Process switching, Fast switching, CEF и др.
- Где-то на этом этапе происходит балансировка в случае наличия ESMR



RIB vs FIB

- Protocol Data Structure
 - show ip ospf database
 - show ip bgp
 - show ip eigrp topology
- Routing Information Base
 - show ip route
- Forwarding Information Base
 - Ha IOS = CEF
 - show ip cef
 - show adjacency



Cisco Express Forwarding

- Устройства с Cisco IOS используют CEF
- Основная идея – просчитать как можно больше параметров заранее (на основе RIB и ARP) и поместить эти данные в FIB
 - Проще говоря – это таблица для коммутации пакетов
- Хранится либо в TCAM, либо в эмуляции TCAM
- Может использоваться почти всегда, кроме случаев
 - ACL logging
 - Пакеты, предназначенные маршрутизатору
 - Требуется отправка на CPU



Adjacency таблица

- **Cache**: нормальная запись, для которой сформирован L2-header
- **Receive**: виртуальная запись для собственных маршрутов
- **Null**: dev/Null
- **Punt**: CEF not supported для префикса
- **Glean**: не известен L2-header
- **Discard/Drop**
 - записи для отбрасываемых пакетов

```
Gateway#show ip cef
Prefix                Next Hop                Interface
0.0.0.0/0             no route
0.0.0.0/8             drop
0.0.0.0/32            receive
1.1.1.1/32            192.168.1.2            vlan1
2.2.2.2/32            attached               Null0
127.0.0.0/8           drop
192.168.1.0/24        attached               vlan1
192.168.1.0/32        receive                vlan1
192.168.1.1/32        receive                vlan1
192.168.1.2/32        attached               vlan1
192.168.1.255/32      receive                vlan1
224.0.0.0/4           drop
224.0.0.0/24          receive
240.0.0.0/4           drop
255.255.255.255/32    receive
```



CEF и ESMR

- CEF может иметь несколько вариантов next hop для одного префикса
 - Максимальное количество зависит от платформы и версии протокола
 - Трафик будет балансироваться между next hop по некоторому алгоритму
 - `ip cef load-sharing {per-packet|per-flow}`
 - `show ip cef exact-route <src> <dst>`
- Поляризация
 - `ip cef load-sharing algorithm universal <id>`



Шаг 3. Инкапсуляция

- После успешного поиска исходящего интерфейса, к пакету необходимо добавить L2 заголовок
- Берется из CEF (Adj/ARP)



Статические маршруты

- Правильный метод маршрутизации
 - `ip route 10.0.0.0 255.255.0.0 1.1.1.1`
 - Строим L2 Adjacency только для Next-Hop
- Неправильный метод маршрутизации
 - `ip route 10.0.0.0 255.255.0.0 Gi0/1`
 - строим L2 Adjacency для Destination IP



Иногда указать NH недостаточно

- `ip route 0.0.0.0 0.0.0.0 1.1.1.1`
- `ip route 192.168.1.0 255.255.255.0 10.0.0.1`
- `ip route 192.168.1.0 255.255.255.0 10.0.1.1 250`
- Что будет, если упадет Adjacency с 10.0.0.1 ?





Networking
For everyone