

СЕТЕВОЙ УРОВЕНЬ

Лекция состоит из разделов:

1. IP-адресация, IPv4, IPv6
2. Маршрутизация
 1. Статическая и динамическая
 2. И ещё классификация
 3. RIP кратко + как настроено
 4. OSPF много + как настроено
 5. BGP кратко + как настроено
3. Попробуем настроить сетку сами

IP

IP-адреса делятся на 5 классов (A, B, C, D, E). A, B и C — это классы коммерческой адресации. D - для многоадресных рассылок, а класс E - для экспериментов.

Класс A: 1.0.0.0 — 126.0.0.0, маска 255.0.0.0

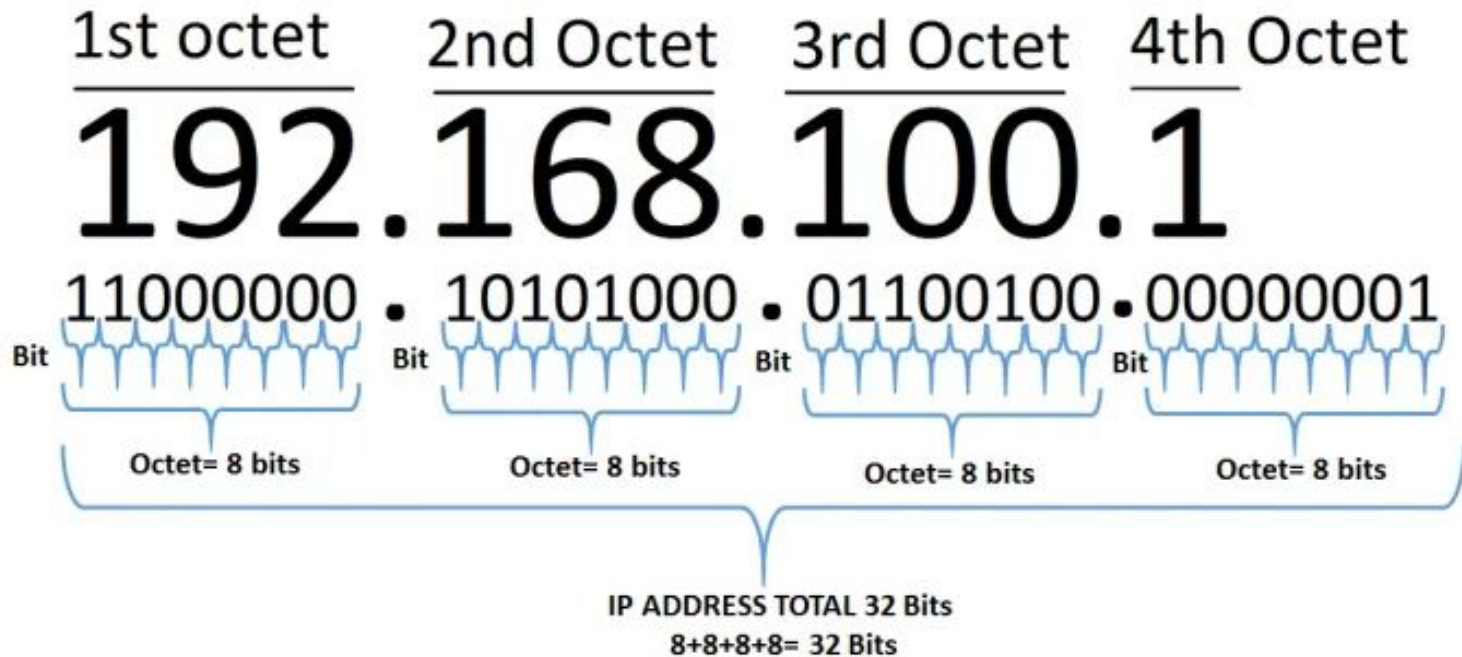
Класс B: 128.0.0.0 — 191.255.0.0, маска 255.255.0.0

Класс C: 192.0.0.0 — 223.255.255.0, маска 255.255.255.0

Класс D: 224.0.0.0 — 239.255.255.255, маска 255.255.255.255


Класс E: 240.0.0.0 — 247.255.255.255, маска 255.255.255.255

IP





IP

Serial No.	Binary Format	Decimal		
1	00000000 . 00000000 . 00000000 . 00000000	0.0.0.0		
2	00000000 . 00000000 . 00000000 . 00000001	0.0.0.1		
3	00000000 . 00000000 . 00000000 . 00000010	0.0.0.2		
4	00000000 . 00000000 . 00000000 . 00000011	0.0.0.3		
5	00000000 . 00000000 . 00000000 . 00000100	0.0.0.4		
6	00000000 . 00000000 . 00000000 . 00000101	0.0.0.5		
-	-			
-	-			
-	-			
256	00000000 . 00000000 . 00000000 . 11111111	0.0.0.255		
257	00000000 . 00000000 . 00000001 . 00000000	0.0.1.0		
258	00000000 . 00000000 . 00000001 . 00000001	0.0.1.1		
259	00000000 . 00000000 . 00000000 . 00000010	0.0.1.2		
260	00000000 . 00000000 . 00000001 . 00000011	0.0.1.3		
261	00000000 . 00000000 . 00000001 . 00000100	0.0.1.4		



IP

262	00000000 . 00000000 . 00000001 . 00000101	0.0.1.5	
-	-		
-	-		
-	-		
512	00000000 . 00000000 . 00000001 . 11111111	0.0.1.255	
512	00000000 . 00000000 . 00000010 . 00000000	0.0.2.0	
512	00000000 . 00000000 . 00000010 . 00000001	0.0.2.1	
512	00000000 . 00000000 . 00000010 . 00000010	0.0.2.2	
512	00000000 . 00000000 . 00000010 . 00000011	0.0.2.3	
512	00000000 . 00000000 . 00000010 . 00000100	0.0.2.4	
512	00000000 . 00000000 . 00000010 . 00000101	0.0.2.5	
	-		
	-		
	-		
	-		
Last IP add	11111111 . 11111111 . 11111111 . 11111111	255.255.255.255	

IP

Маска подсети	Маска в двоичной системе	Префикс	Количество адресов	Обратная маска
255.255.255.255	11111111.11111111.11111111.11111111	/32	1	0.0.0.0
255.255.255.254	11111111.11111111.11111111.11111110	/31	2	0.0.0.1
255.255.255.252	11111111.11111111.11111111.11111100	/30	4	0.0.0.3
255.255.255.248	11111111.11111111.11111111.11111000	/29	8	0.0.0.7
255.255.255.240	11111111.11111111.11111111.11110000	/28	16	0.0.0.15
255.255.255.224	11111111.11111111.11111111.11100000	/27	32	0.0.0.31
255.255.255.192	11111111.11111111.11111111.11000000	/26	64	0.0.0.63
255.255.255.128	11111111.11111111.11111111.10000000	/25	128	0.0.0.127
255.255.255.0	11111111.11111111.11111111.00000000	/24	256	0.0.0.255
255.255.254.0	11111111.11111111.11111110.00000000	/23	512	0.0.1.255
255.255.252.0	11111111.11111111.11111100.00000000	/22	1024	0.0.3.255
255.255.248.0	11111111.11111111.11111000.00000000	/21	2048	0.0.7.255
255.255.240.0	11111111.11111111.11110000.00000000	/20	4096	0.0.15.255
255.255.224.0	11111111.11111111.11100000.00000000	/19	8192	0.0.31.255
255.255.192.0	11111111.11111111.11000000.00000000	/18	16384	0.0.63.255
255.255.128.0	11111111.11111111.10000000.00000000	/17	32768	0.0.127.255
255.255.0.0	11111111.11111111.00000000.00000000	/16	65536	0.0.255.255
255.254.0.0	11111111.11111110.00000000.00000000	/15	131072	0.1.255.255
255.252.0.0	11111111.11111100.00000000.00000000	/14	262144	0.3.255.255
255.248.0.0	11111111.11111000.00000000.00000000	/13	524288	0.7.255.255
255.240.0.0	11111111.11110000.00000000.00000000	/12	1048576	0.15.255.255
255.224.0.0	11111111.11100000.00000000.00000000	/11	2097152	0.31.255.255
255.192.0.0	11111111.11000000.00000000.00000000	/10	4194304	0.63.255.255
255.128.0.0	11111111.10000000.00000000.00000000	/9	8388608	0.127.255.255
255.0.0.0	11111111.00000000.00000000.00000000	/8	16777216	0.255.255.255
254.0.0.0	11111110.00000000.00000000.00000000	/7	33554432	1.255.255.255
252.0.0.0	11111100.00000000.00000000.00000000	/6	67108864	3.255.255.255
248.0.0.0	11111000.00000000.00000000.00000000	/5	134217728	7.255.255.255
240.0.0.0	11110000.00000000.00000000.00000000	/4	268435456	15.255.255.255
224.0.0.0	11100000.00000000.00000000.00000000	/3	536870912	31.255.255.255
192.0.0.0	11000000.00000000.00000000.00000000	/2	1073741824	63.255.255.255
128.0.0.0	10000000.00000000.00000000.00000000	/1	2147483648	127.255.255.255
0.0.0.0	00000000.00000000.00000000.00000000	/0	4294967296	255.255.255.255

Списки зарезервированных сетей, которые не предназначены для глобальной (В Интернет) маршрутизации

Подсеть	Назначение
0.0.0.0/8	Адреса источников пакетов «этой» («своей») сети, предназначены для локального использования на хосте при создании сокетов IP. Адрес 0.0.0.0/32 используется для указания адреса источника самого хоста.
10.0.0.0/8	Для использования в частных сетях.
127.0.0.0/8	Подсеть для коммуникаций внутри хоста (localhost).
169.254.0.0/16	Канальные адреса; подсеть используется для автоматического конфигурирования адресов IP в случае отсутствия сервера DHCP .
172.16.0.0/12	Для использования в частных сетях.
192.168.0.0/16	
100.64.0.0/10	Для использования в сетях сервис-провайдера.
192.0.0.0/24	Регистрация адресов специального назначения.
192.0.2.0/24	Для примеров в документации.
198.51.100.0/24	
203.0.113.0/24	
198.18.0.0/15	Для стендов тестирования производительности.
240.0.0.0/4	Зарезервировано для использования в «будущем».
255.255.255.255	Ограниченный широковещательный адрес.

Зарезервированные сети, которые могут маршрутизироваться глобально

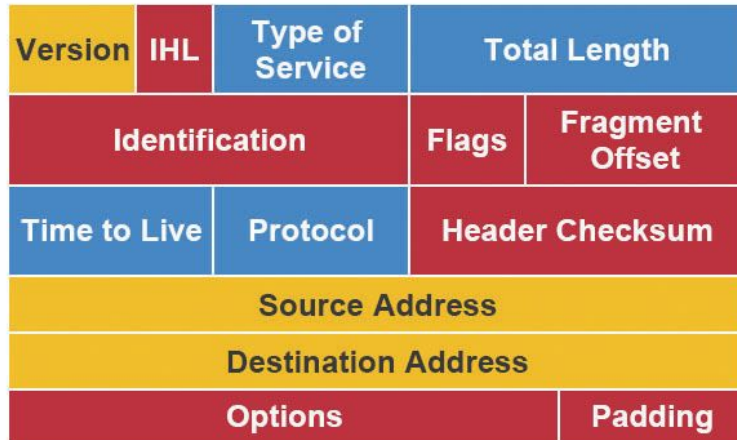
Подсеть	Назначение
192.88.99.0/24	Используются для рассылки ближайшему узлу. Адрес 192.88.99.1/32 применяется в качестве ретранслятора при инкапсуляции IPv6 в IPv4.
224.0.0.0/4	Используются для многоадресной рассылки.

Структура заголовка пакета

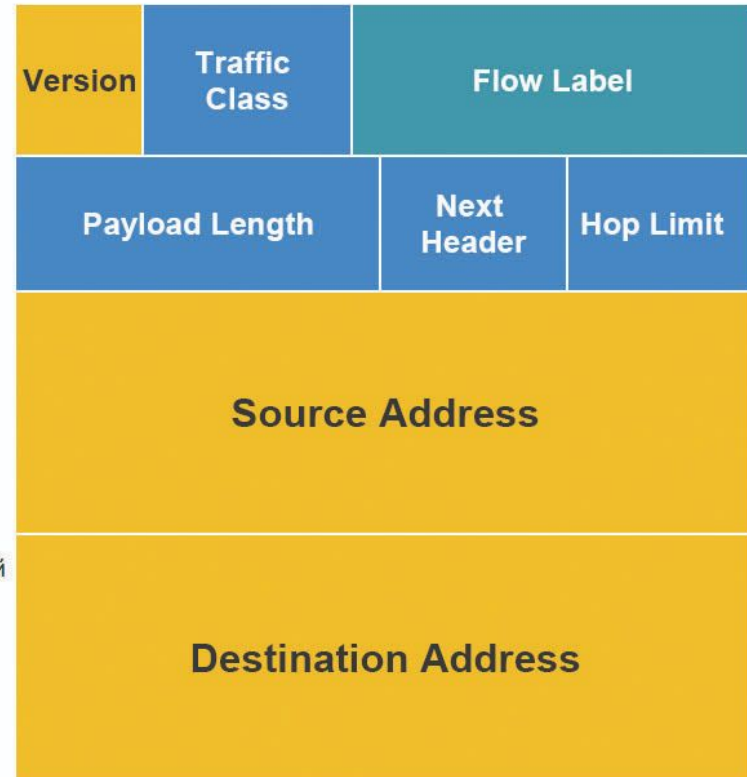
Отступ	Октет	0								1								2								3							
Октет	Бит	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
0	0	Версия				Размер заголовка				Differentiated Services Code Point				Explicit Congestion Notification				Размер пакета (полный)															
4	32	Идентификатор																Флаги		Смещение фрагмента													
8	64	Время жизни								Протокол								Контрольная сумма заголовка															
12	96	IP-адрес источника																															
16	128	IP-адрес назначения																															
20	160	Опции (если размер заголовка > 5)																															
20 или 24+	160 или 192+	Данные																															





Структура заголовка пакета

IPv4 Header



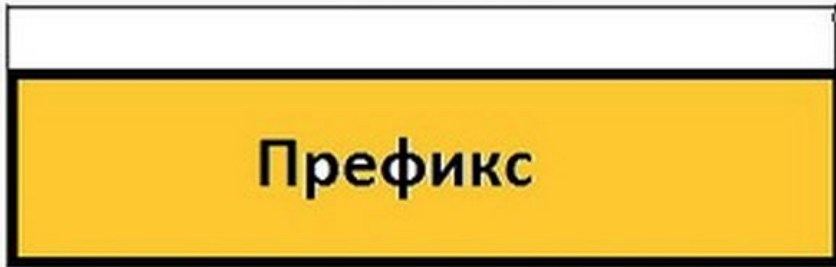
IPv6 Header



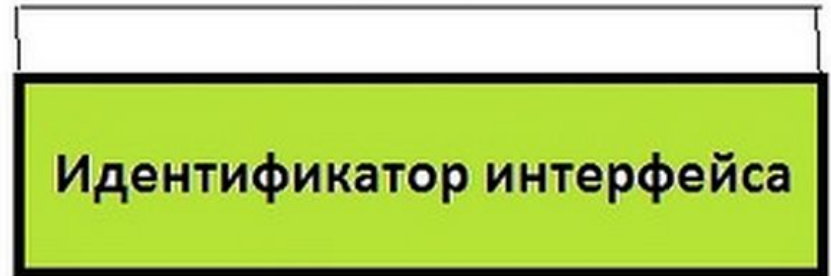
-  Поля которые перешли из IPv4 в IPv6 без изменений
-  Поля убранные из IPv6
-  Поля переименованные в IPv6
-  Новое поле в IPv6

Структура заголовка пакета

64 бита



64 бита



Пример: 2001:0DB8:000B::/64



IPv4	IPv6
Под адреса источника и назначения выделено 32 бита(4 байта) от общей длины заголовка.	Под адреса источника и назначения выделено 128 бит(16 байтов)от общей длины заголовка. Более подробно, см. « IPv6 Адресование.»
Поддержка IPSec не обязательна.	Поддержка IPSec является обязательной опцией. Более подробно,см. « IPv6 заголовок. «
Никакая идентификация потока пакетов, в пределах IPV4 заголовка,для обработки QoS маршрутизаторами не используется.	Идентификация потока пакетов для обработки QoS маршрутизаторами включена в IPV6 заголовок, используя поле ярлыка потока.
Фрагментация осуществляется как маршрутизаторами так и хостами, инициировавшими передачу.	Фрагментация осуществляется хостами, инициировавшими передачу
Заголовок включает контрольную сумму.	Заголовок не включает контрольную сумму.
Заголовок включает опции.	Все опции перемещены в расширенную часть IPV6 заголовка.
Протокол Разрешения Адресов (ARP) использует широковещательные запроса ARP, чтобы установить соответствие сетевого адреса(IPv4), аппаратному.	ARP запросы заменены multicast сообщениями Neighbor Solicitation.Более
Internet Group Management Protocol (IGMP) используется, чтобы управлять группами в	IGMP заменен на Multicast Listener Discovery (MLD) сообщения.
Маршрутизаторы используют ICMP	ICMP
Широковещательный адреса используются посылки сообщения всем узлам в подсети.	В IPv6 И по этому, вместо широковещательной посылки используется адрес для определенной группы. Это помогает снизить нагрузку на среду передачи.
Должен быть сконфигурирован или вручную или через DHCP.	Автоконфигурация Адреса.
Хост и Использует адреса (A) IPv4 которые были выданы Domain Name System (DNS) и привязаны ею к аппаратному адресу хоста.	Хост и Использует адреса которые были выданы Domain Name System (DNS) и привязаны ею к аппаратному адресу хоста.
Использует указатель (PTR) в записях ресурса в IN-ADDR.ARPA DNS доменах карты IPv4 адресов, чтобы обеспечить хосты именами.	Использует указатель (PTR) в записях ресурса в IP6. ARPA DNS доменах карты IPv6 адресов, чтобы обеспечить хосты именами.
Должен поддерживать размер пакета 576 байт (возможно фрагментированный).	Должен поддерживать размер пакета 1280 байт (без фрагментации).

ДЖАМШУТКА

Мальчик сказал маме: “Я хочу кушать”. Мама отправила его к папе.

Мальчик сказал папе: “Я хочу кушать”. Папа отправил его к маме.

Мальчик сказал маме: “Я хочу кушать”. Мама отправила его к папе.

И бегал так мальчик, пока в один момент не упал.

Что случилось с мальчиком?)

Маршрутизация

Маршрутизация - это процесс выбора маршрута следования пакета. Маршрутизация может быть статической или динамической. Маршрутизация работает в сети, состоящей из маршрутизаторов, на третьем уровне модели OSI. Маршрутизатором, или шлюзом, называется узел сети с несколькими IP-интерфейсами (содержащими свой MAC-адрес и IP-адрес), подключенными к разным IP-сетям, осуществляющий на основе решения задачи маршрутизации перенаправления пакетов из одной сети в другую для доставки от отправителя к получателю.

Маршрутизаторы представляют собой либо специализированные вычислительные машины, либо компьютеры с несколькими IP-интерфейсами, работа которых управляется специальным программным обеспечением.

Маршрутизация в IP-сетях

Для маршрутизации маршрутизатор должен владеть следующей информацией:

- Адрес назначения
- Соседний маршрутизатор, от которого он может узнать об удаленных сетях
- Доступные пути ко всем удаленным сетям
- Наилучший путь к каждой удаленной сети
- Методы обслуживания и проверки информации о маршрутизации

Процесс IP-маршрутизации

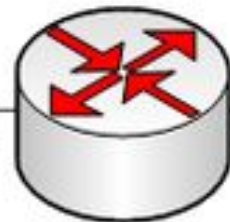


Процесс IP-маршрутизации

Хост А
172.16.10.2/24



Интерфейс E0:
172.16.10.1



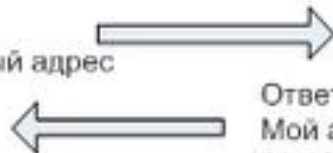
Интерфейс E1:
172.16.20.1

Хост А
172.16.20.2/24



Ping 172.16.20.2

Запрос ARP:
какой аппаратный адрес
172.16.10.1?



Ответ ARP:
Мой аппаратный адрес
00-19-65-76-c7-f6-d0

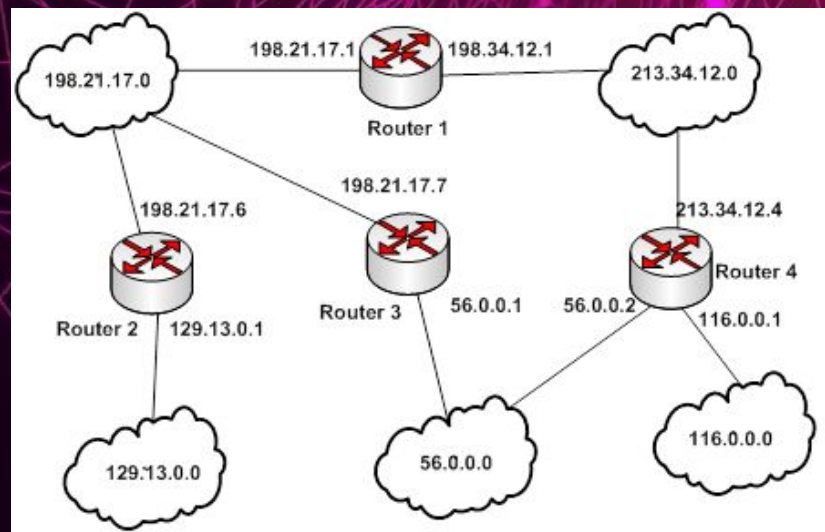
Генерация кадра и пересылка
в нем пакета

Процесс IP-маршрутизации



Процесс IP-маршрутизации

Сетевой адрес	Маска сети	Адрес шлюза	Интерфейс	Метрика
129.13.0.0	255.255.0.0	-	129.13.0.1	подключен
198.21.17.0	255.255.255.0	-	198.21.17.6	подключен
213.34.12.0	255.255.255.0	198.21.17.1	198.21.17.6	1
56.0.0.0	255.0.0.0	198.21.17.7	198.21.17.6	1
116.0.0.0	255.0.0.0			



Статическая маршрутизация

Статическая маршрутизация – вид маршрутизации, при котором маршруты указываются в явном виде при конфигурации маршрутизатора. Вся маршрутизация при этом происходит без участия каких-либо протоколов маршрутизации.

При задании статического маршрута указывается:

- Адрес сети (на которую маршрутизируется трафик), маска сети
- Адрес шлюза (узла), который способствует дальнейшей маршрутизации (или подключен к маршрутизируемой сети напрямую)
- (опционально) метрика (иногда именуется также «ценой») маршрута. При наличии нескольких маршрутов на одну и ту же сеть некоторые маршрутизаторы выбирают маршрут с минимальной метрикой

Статическая маршрутизация

Достоинства

- Лёгкость отладки и конфигурирования в малых сетях.
- Отсутствие дополнительных накладных расходов (из-за отсутствия протоколов маршрутизации)
- Мгновенная готовность (не требуется интервал для конфигурирования/подстройки)
- Низкая нагрузка на процессор маршрутизатора
- Предсказуемость в каждый момент времени

Статическая маршрутизация

Недостатки

- Очень плохое масштабирование (добавление (N+1)-ой сети потребует сделать $2*(N+1)$ записей о маршрутах, причём на большинстве маршрутизаторов таблица маршрутов будет различной, при $N > 3-4$ процесс конфигурирования становится весьма трудоёмким).
- Низкая устойчивость в ситуациях, когда обрыв происходит между устройствами второго уровня и порт маршрутизатора не получает статус down.
- Отсутствие динамического балансирования нагрузки
- Необходимость в ведении отдельной документации к маршрутам, проблема синхронизации документации и реальных маршрутов.

Динамическая маршрутизация

Динамическая маршрутизация – вид маршрутизации, при котором таблица маршрутизации редактируется программно. В случае UNIX-систем демонами маршрутизации; в других системах – служебными программами, которые называются иначе, но фактически играют ту же роль.

Преимущества протоколов динамической маршрутизации:

- автоматическое построение таблиц маршрутизации;
- отказоустойчивость сети;
- возможность балансировки трафика;
- масштабирование сети.

Недостатки:

- более низкая безопасность;
- использование ресурсов ЦП, ОЗУ маршрутизаторов.

Классификация протоколов

Все протоколы динамической маршрутизации можно разделить на две большие группы:

- внешние EGP – Exterior Gateway Protocol;
- внутренние IGP – Interior Gateway Protocol.

Автономная система (AS - Autonomobous System) – это система IP-сетей и маршрутизаторов, управляемых одним или несколькими операторами, имеющими единую политику маршрутизации с Интернетом.

Протоколы внутренней маршрутизации используются внутри автономной системы, а внешние – для соединения автономных систем между собой.

Отличия

В свою очередь, внутренние протоколы маршрутизации подразделяются на:

- Distance-Vector (RIP, EIGRP) - дистанционно-векторные;
- Link State (OSPF, IS-IS) - отслеживающие состояние канала.

Коренные различия между этими двумя видами состоят в следующем:

- типе информации, которой обмениваются роутеры: таблицы маршрутизации у Distance-Vector и таблицы топологии у Link State;
- процессе выбора лучшего маршрута;
- количестве информации о сети, которое “держит в голове” каждый роутер: Distance-Vector знает только своих соседей, Link State имеет представление обо всей сети.

RIP - Routing Information Protocol

RIP - протокол маршрутной информации - протокол дистанционно-векторной маршрутизации, который оперирует транзитными участками в качестве метрики. Алгоритм маршрутизации RIP (алгоритм Беллмана — Форда) был впервые разработан в 1969 году, как основной для сети ARPANET.

Характеристики протокола:

- RIPv1 и RIPv2 используют UDP порт 520.
- RIPng использует UDP порт 521.
- Для передачи сообщений RIPv1 в адресе получателя используется широковещательный адрес 255.255.255.255, а RIPv2 — мультикаст адрес 224.0.0.9.

Работа в сети Уфанет, RIP

В сети Уфанет используется для анонсирования маршрутной информации на серые квартальные сети клиентов ЮЛ VPN с D-Link DGS-36xx на: Cisco-76xx через интерфейс 3024-3031, выделенный под линки между Cisco 76xx и DGS-36xx в стеке кварталов. Cisco-76xx анонсирует данный маршрут по iBGP на route-рефлекторы (smir-core и cdc-core). Ретранслирует запрос на сервере DHCP

```
ckm53#show ip route rip
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
       a - application route
       + - replicated route, % - next hop override

Gateway of last resort is 145.255.7.96 to network 0.0.0.0

10.0.0.0/8 is variably subnetted, 5422 subnets, 20 masks
R   10.6.252.0/23 [120/1] via 10.0.7.102, 00:00:18, Vlan3026
                               [120/1] via 10.0.7.14, 00:00:24, Vlan3027
                               [120/1] via 10.0.7.6, 00:00:04, Vlan3025
                               [120/1] via 10.0.6.66, 00:00:23, Vlan3024
R   10.66.1.0/24 [120/1] via 10.0.6.26, 00:00:04, Vlan3028
R   10.66.2.0/24 [120/1] via 10.0.6.26, 00:00:04, Vlan3028
R   10.66.3.0/24 [120/1] via 10.0.6.66, 00:00:24, Vlan3024
R   10.66.4.0/24 [120/1] via 10.0.7.102, 00:00:19, Vlan3026
R   10.66.5.0/24 [120/1] via 10.0.7.102, 00:00:19, Vlan3026
R   10.66.6.0/24 [120/1] via 10.0.7.102, 00:00:19, Vlan3026
R   10.66.7.0/24 [120/1] via 10.0.7.102, 00:00:19, Vlan3026
R   10.66.9.0/24 [120/1] via 10.0.6.26, 00:00:04, Vlan3028
R   10.66.29.0/24 [120/1] via 10.0.6.66, 00:00:24, Vlan3024
R   10.66.30.0/24 [120/1] via 10.0.7.102, 00:00:19, Vlan3026
R   10.66.31.0/24 [120/1] via 10.0.6.26, 00:00:04, Vlan3028
R   10.66.32.0/24 [120/1] via 10.0.6.26, 00:00:04, Vlan3028
R   10.66.33.0/24 [120/1] via 10.0.7.102, 00:00:19, Vlan3026
R   10.66.34.0/24 [120/1] via 10.0.6.26, 00:00:04, Vlan3028
R   10.66.35.0/24 [120/1] via 10.0.7.22, 00:00:08, Vlan3029
R   10.66.56.0/24 [120/1] via 10.0.7.22, 00:00:08, Vlan3029
```


Работа в сети Уфанет, RIP

```
DGS-3627G:admin#show rip
Command: show rip
```

```
RIP Global State : Enabled
```

```
RIP Interface Settings
```

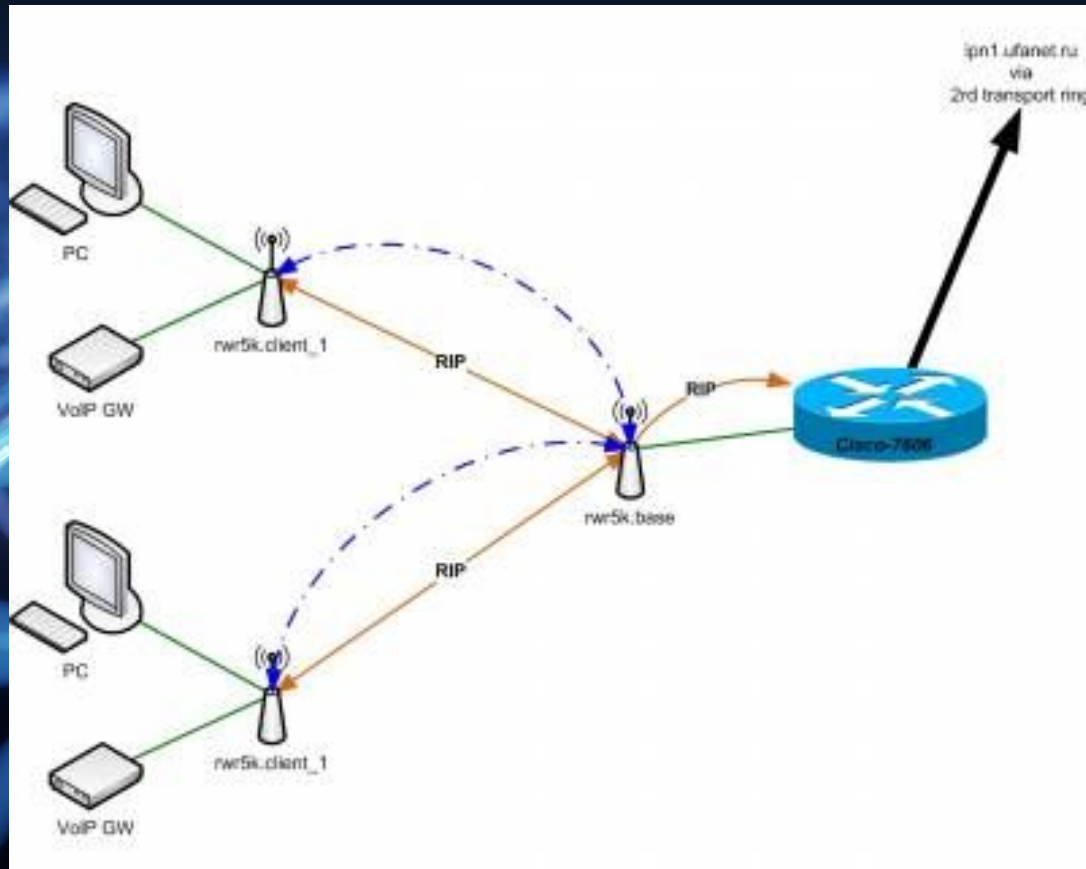
Interface	IP Address	TX Mode	RX Mode	Authen- tication	State
3029	10.0.7.22/30	V2 Only	Disabled	Disabled	Enabled
3715	10.0.9.70/28	Disabled	Disabled	Disabled	Disabled
System	10.1.2.132/27	Disabled	Disabled	Disabled	Disabled
100	10.6.252.236/23	V2 Only	Disabled	Disabled	Enabled
4055	10.66.35.1/24	Disabled	Disabled	Disabled	Enabled
4053	10.66.56.1/24	Disabled	Disabled	Disabled	Enabled
4057	10.66.61.1/24	Disabled	Disabled	Disabled	Enabled
4059	10.66.62.1/24	Disabled	Disabled	Disabled	Enabled
4060	10.66.68.1/24	Disabled	Disabled	Disabled	Enabled
4061	10.66.70.1/24	Disabled	Disabled	Disabled	Enabled
4046	10.66.88.1/24	Disabled	Disabled	Disabled	Enabled
4007	10.66.250.1/24	Disabled	Disabled	Disabled	Enabled
4047	10.67.35.1/24	Disabled	Disabled	Disabled	Enabled
4056	10.67.58.1/24	Disabled	Disabled	Disabled	Enabled
4054	10.67.59.1/24	Disabled	Disabled	Disabled	Enabled

Revolution 5000

Revolution 5000. В схеме работы клиентов IPN via Revolution 5000.

Устройство `rwr5k.client` по RIP анонсирует информацию о маршруте на сеть клиента устройству `rwr5k.base`,

который в свою очередь перераспределяет маршрут (так же по RIP) на `Cisco-7606` в `vrf Dipn1`, при этом `rwr5k.base` анонсирует для всех `rwr5k.client` маршрут по умолчанию через себя.



OSPF - Open Shortest Path First

OSPF (англ. Open Shortest Path First) — протокол динамической маршрутизации, основанный на технологии отслеживания состояния канала (link-state technology) и использующий для нахождения кратчайшего пути алгоритм Дейкстры.

В Link-state протоколах каждый маршрутизатор должен не просто знать самые лучшие маршруты во все удалённые сети, но и иметь в памяти полную карту сети со всеми существующими связями между другими маршрутизаторами в том числе. Это достигается за счет построения специальной базы LSDB.

OSPF инкапсулируется в IP. Номер протокола 89.

Для передачи пакетов использует мультикаст адреса:

- 224.0.0.5 все маршрутизаторы OSPF;
- 224.0.0.6 все DR.

OSPF - Open Shortest Path First

Базовые термины:

Канал/интерфейс (link/interface) — соединение маршрутизатора и одной из подключенных к нему сетей. При обсуждении OSPF термины интерфейс и канал (link) часто употребляются как синонимы

Метрика (metric) — условный показатель расстояния до сети назначения

Стоимость (cost) — условный показатель "стоимости" пересылки данных по каналу. В OSPF зависит от пропускной способности интерфейса (bandwidth)

Автономная система (autonomous system) — группа маршрутизаторов, обменивающаяся маршрутизирующей информацией с помощью одного протокола маршрутизации (определение соответствует тому, как этот термин используется в протоколах IGP)

Базовые термины OSPF

Идентификатор маршрутизатора (router ID, RID) — уникальное 32-битовое число, которое уникально идентифицирует маршрутизатор в пределах одной автономной системы

Зона (area) — совокупность сетей и маршрутизаторов, имеющих один и тот же идентификатор зоны

Объявление о состоянии канала (link-state advertisement, LSA) — единица данных, которая описывает локальное состояние маршрутизатора или сети.

Например, для маршрутизатора LSA включает описание состояния каналов и отношений соседства. Множество всех LSA, описывающих маршрутизаторы и сети, образуют базу данных состояния каналов (LSDB).

База данных состояния каналов (link state database, LSDB) — список всех записей о состоянии каналов (LSA). Встречается также термин топологическая база данных (topological database), употребляется как синоним базы данных состояния каналов

Соседи OSPF

Соседи (neighbours) — два маршрутизатора, интерфейсы которых находятся в одном широковещательном сегменте (и на которых включен OSPF на этих интерфейсах)

Отношения соседства (adjacency) — взаимосвязь между соседними маршрутизаторами, установленная с целью синхронизации информации

Hello-протокол (hello protocol) — протокол, использующийся для установки и поддержания соседских отношений

База данных соседей (neighbours database) — список всех соседей (также используется термин neighbour table)

Пакеты OSPF

Hello — пакеты, которые используются для обнаружения соседей, установки отношений соседства и мониторинга их доступности (keepalive)

DBD (Database Description) — пакеты, которые описывают содержание LSDB

LSR (Link State Request) — пакеты, с помощью которых запрашивается полная информация об LSA, которых не хватает в LSDB локального маршрутизатора

LSU (Link State Update) — пакеты, которые передают полную информацию, которая содержится в LSA

LSAck — пакеты, с помощью которых подтверждается получение других пакетов

Терминология протокола OSPF

При запуске процесса OSPF на любом маршрутизаторе, обязательно должен быть выбран Router ID. Router ID – это уникальное имя маршрутизатора, по которому он известен в AS. В зависимости от реализации, Router ID может выбираться по-разному:

- минимальный IP-адрес или максимальный IP-адрес, который назначен на интерфейсах маршрутизатора
- также обычно есть способ задания Router ID вручную
- главное, чтобы Router ID был уникален в AS

После изменения Router ID, процесс OSPF должен быть перезагружен, а все LSA, которые сгенерировал этот маршрутизатор, должны быть удалены из AS, до перезагрузки.

OSPF, отношения

Различают понятия сосед и отношения соседства:

Сосед (neighbor) – два маршрутизатора, которые находятся в одном широковещательном сегменте и у которых совпали нужные поля в hello-пакетах

Отношения соседства (adjacency или full adjacency) – два соседа, которые завершили процесс синхронизации LSDB между собой.

OSPF, соседство

необходимо чтобы в hello-пакетах совпали значения таких полей:

Hello Interval – частота отправки сообщений Hello

Router Dead Interval – период времени, по прохождению которого, сосед считается недоступным, если не было Hello

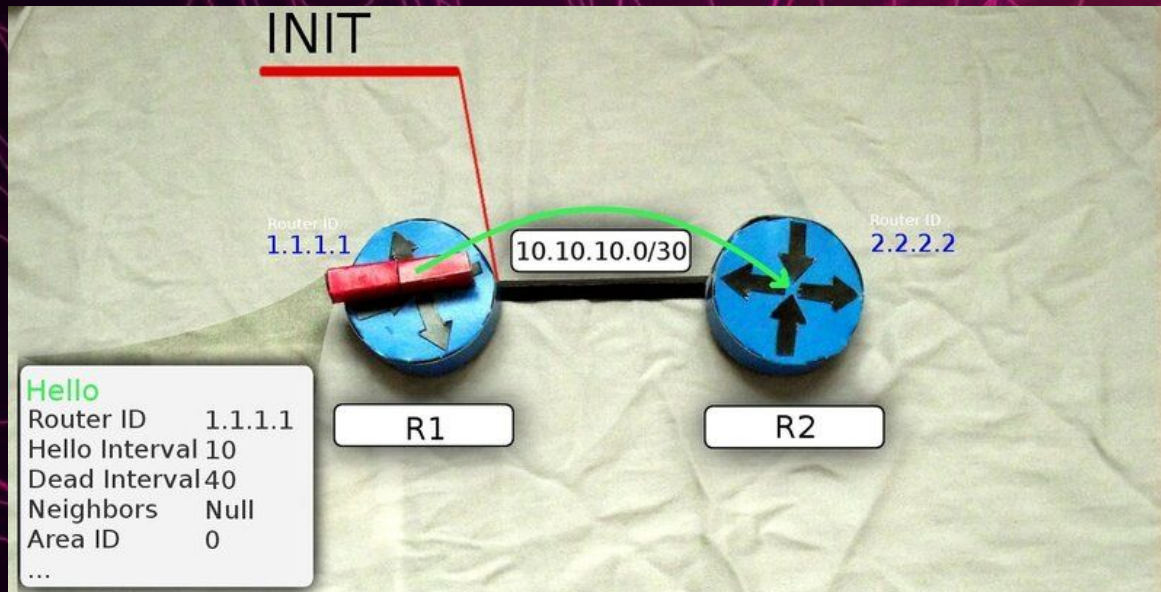
Area ID – так как в OSPF граница зоны проходит через маршрутизатор, то маршрутизаторы в одном широковещательном сегменте, должны быть в одной зоне (подробнее про зоны ниже)

Authentication – пароль использующийся для аутентификации и тип аутентификации. Маршрутизаторы не обязательно должны использовать аутентификацию, но если она используется, то пароли и тип должны совпадать

Stub area flag – не обязательный флаг, который устанавливается на всех маршрутизаторах, которые принадлежат тупиковой зоне (stub area)

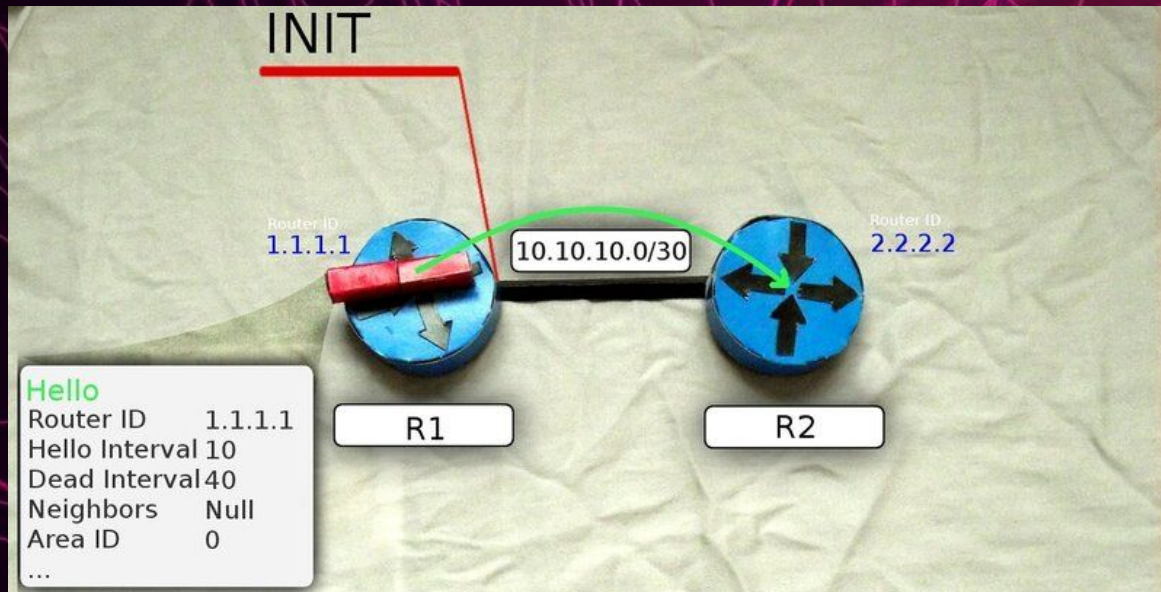
у маршрутизаторов должны совпадать сеть и маска сети

OSPF, пример



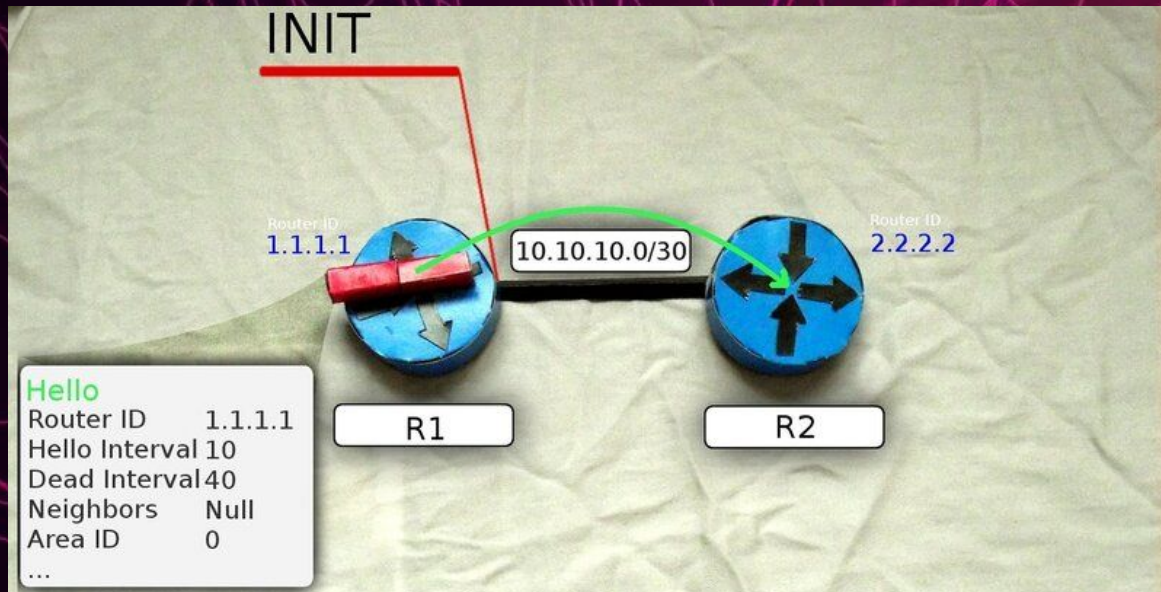
2) Поднимается ветер: маршрутизатор рассылает Hello-пакеты на мультикастный адрес 224.0.0.5 со всех интерфейсов, где запущен OSPF. TTL таких сообщений равен одному, поэтому их получают только маршрутизаторы, находящиеся в том же сегменте сети. R1 переходит в состояние INIT.

OSPF, пример



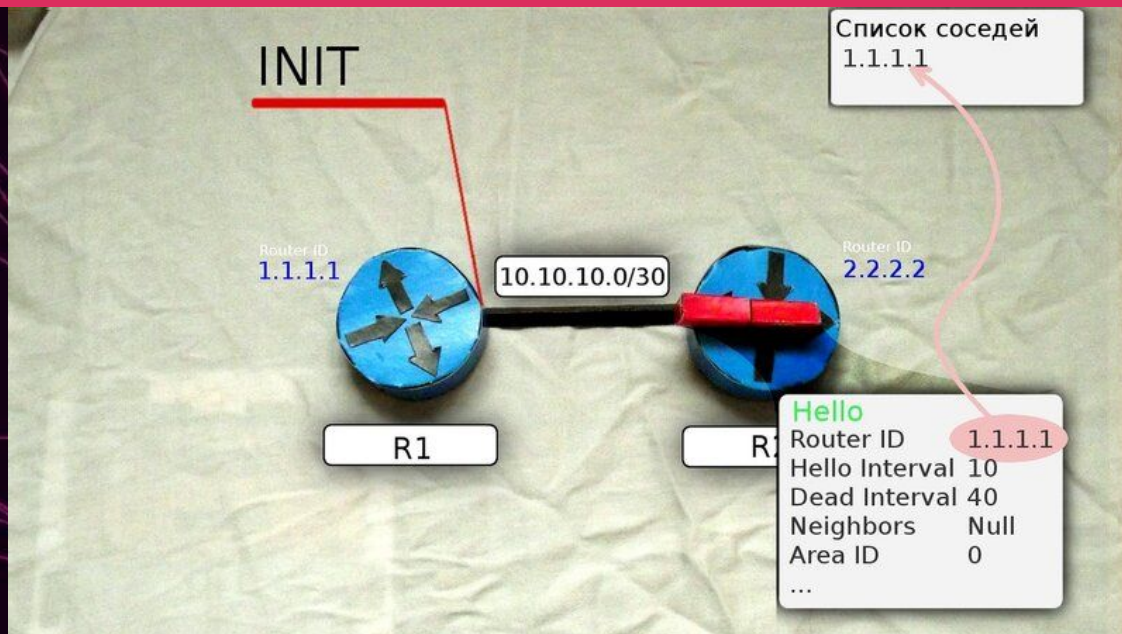
2) Поднимается ветер: маршрутизатор рассылает Hello-пакеты на мультикастный адрес 224.0.0.5 со всех интерфейсов, где запущен OSPF. TTL таких сообщений равен одному, поэтому их получают только маршрутизаторы, находящиеся в том же сегменте сети. R1 переходит в состояние INIT.

OSPF, пример



2) Поднимается ветер: маршрутизатор рассылает Hello-пакеты на мультикастный адрес 224.0.0.5 со всех интерфейсов, где запущен OSPF. TTL таких сообщений равен одному, поэтому их получают только маршрутизаторы, находящиеся в том же сегменте сети. R1 переходит в состояние INIT.

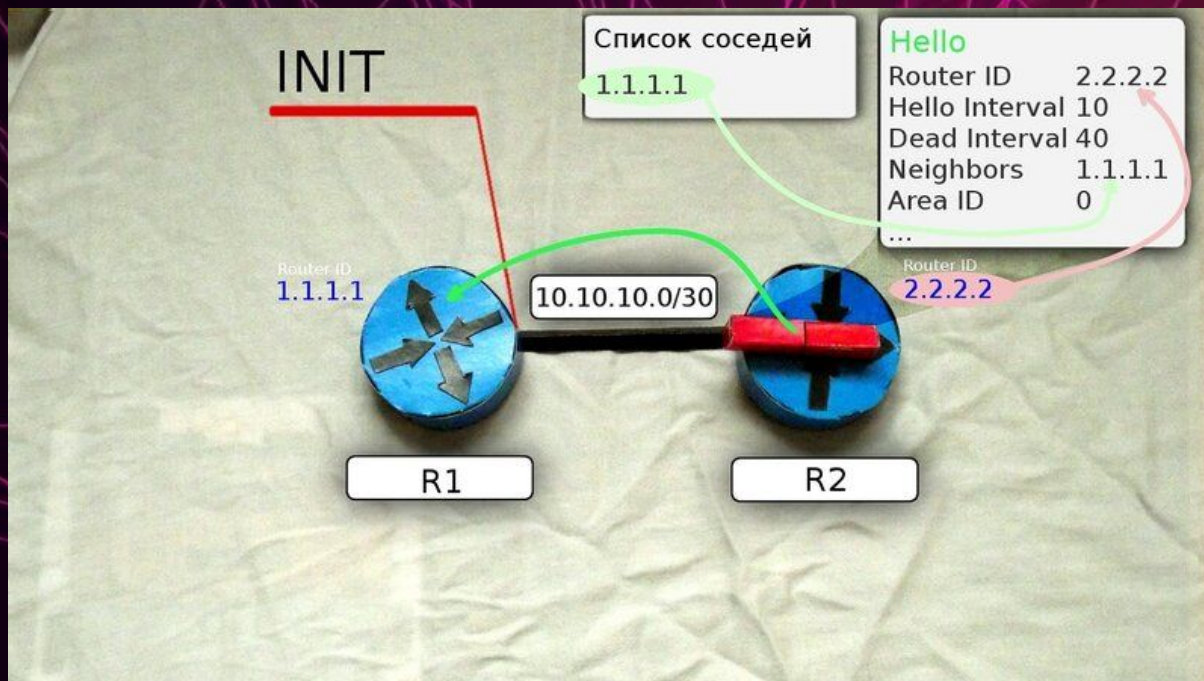
OSPF, пример



Нас интересуют пока первые четыре или точнее вообще только Router ID и Neighbors. Сообщение Hello от маршрутизатора R1 несёт в себе его Router ID и не содержит Neighbors, потому что у него их пока нет.

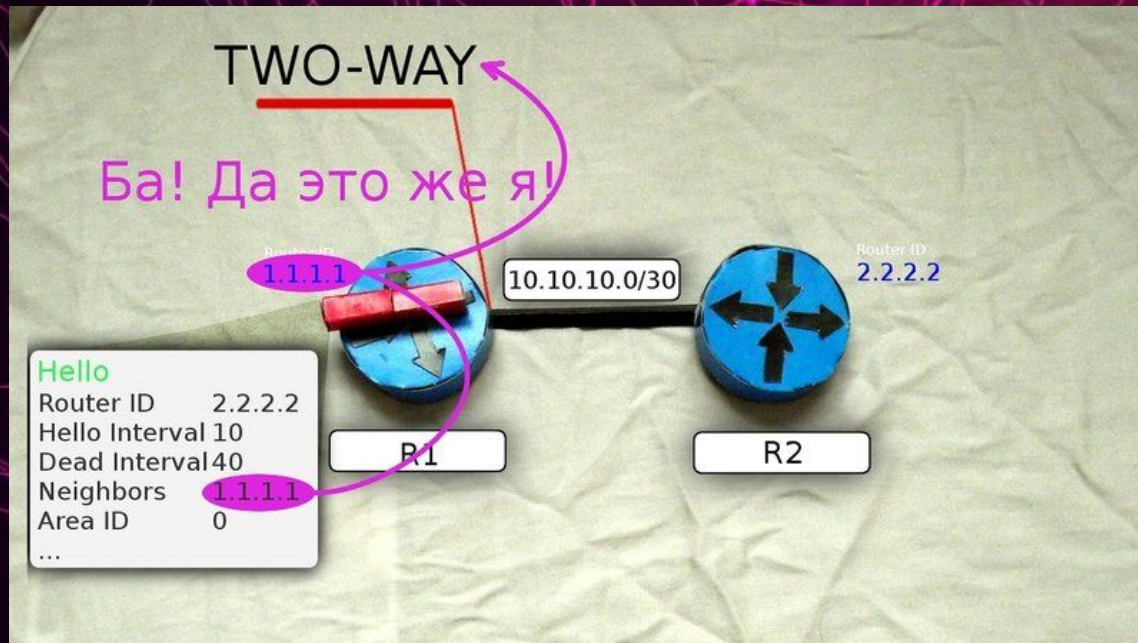
После получения этого мультикастного сообщения маршрутизатор R2 добавляет R1 в свою таблицу соседей (если совпали все необходимые параметры)

OSPF, пример



И отправляет на R1 уже юникастом новое сообщение Hello, где содержится Router ID этого маршрутизатора, а в списке Neighbors перечислены все его соседи. В числе прочих соседей в этом списке есть Router ID R1, то есть R2 уже считает его соседом.

OSPF, пример



Теперь R1 и R2 друг у друга во взаимных соседях — это означает, что между ними установлены отношения смежности и маршрутизатор R1 переходит в состояние TWO WAY.

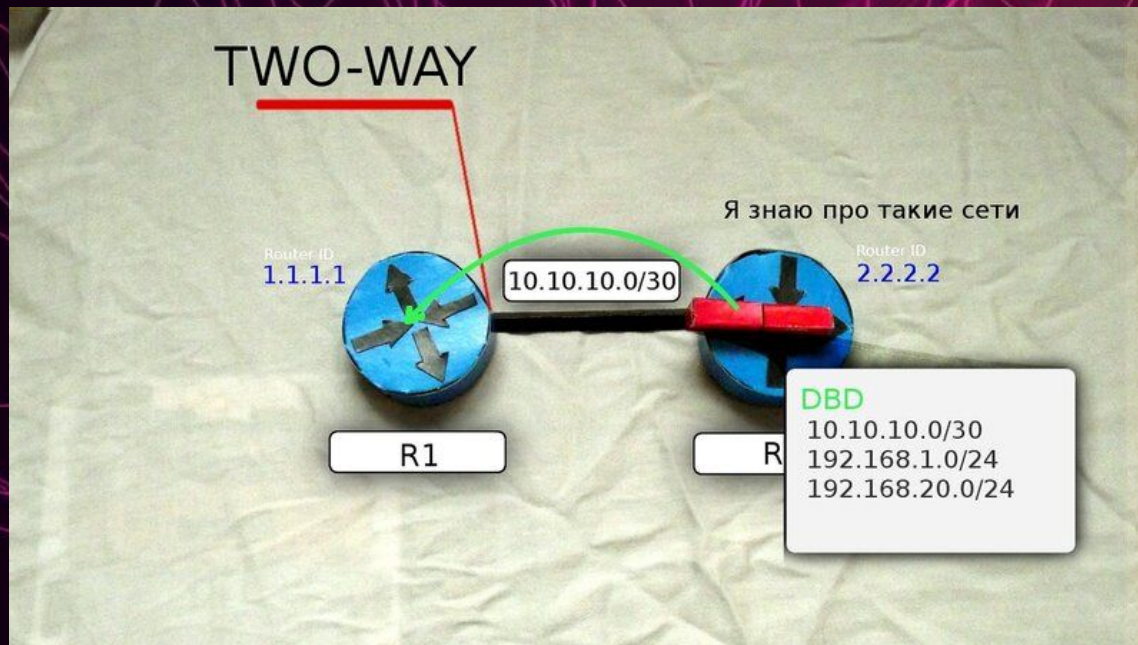
OSPF, пример

4) Затишье перед бурей. Далее все переходят в состояние EXSTART. Здесь все соседи решают между собой, кто босс. Им становится маршрутизатор с наибольшим Router ID – R2.

5) Когда выбран Батяка, соседи переходят в состояние Exchange и обмениваются DBD-сообщениями (или DD) – Data Base Description, которые содержат описание LSDB (Link State Data Base), мол, я знаю про вот такие подсети.

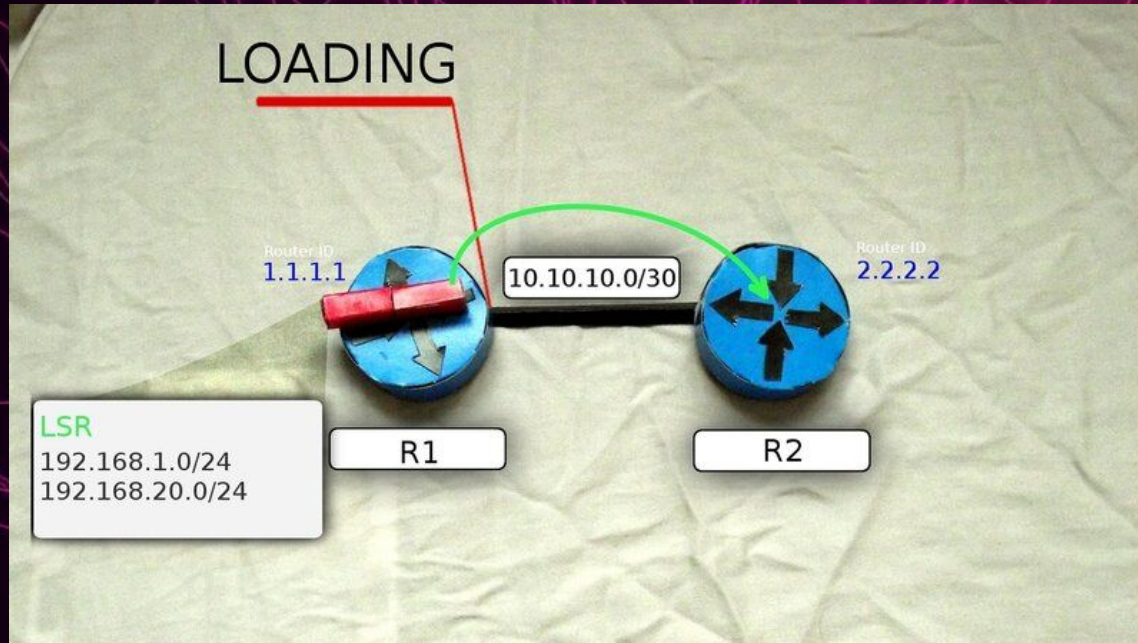
Тут надо пояснить, что такое LSDB. Если перевести на русский дословно: база данных о состоянии линков. В изначальном состоянии маршрутизатор знает только о тех линках (интерфейсах), на которых запущен процесс OSPF. По ходу пьесы, каждый маршрутизатор собирает всю информацию о сети и составляет топологию. Именно она и будет являться LSDB, которая должна быть одинакова на всех членах зоны. Первым отправляет свою DBD маршрутизатор, выбранный главным на данном интерфейсе – 2.2.2.2. Следом за ним то же делает и 1.1.1.1.

OSPF, пример



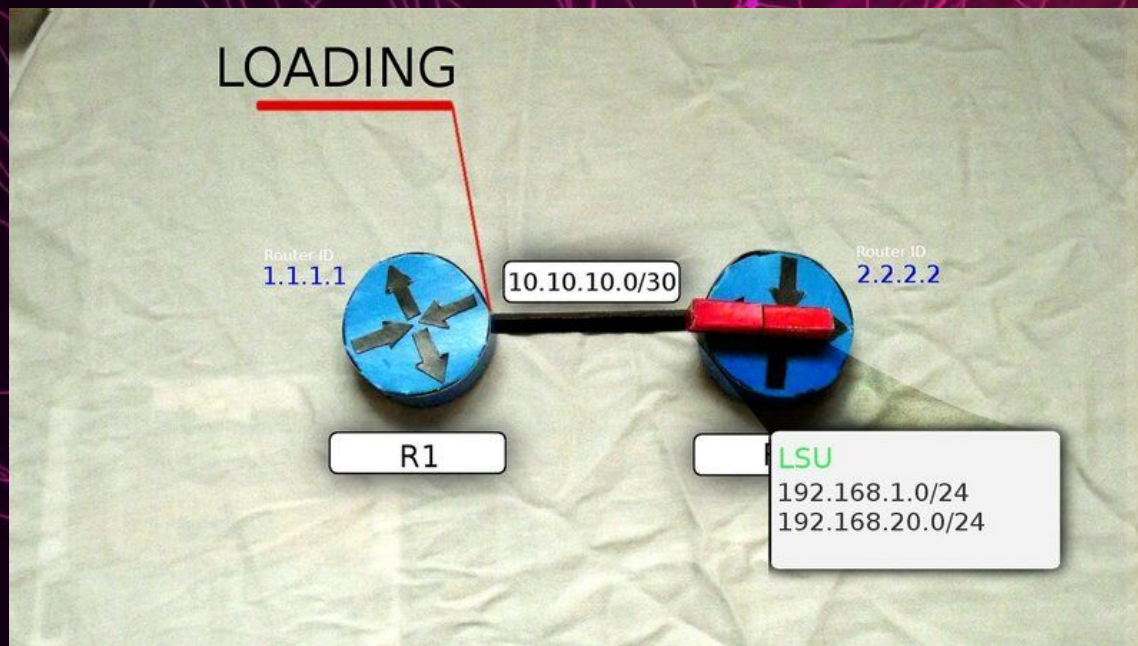
Первым отсылает свою DBD маршрутизатор, выбранный главным на данном интерфейсе — 2.2.2.2. Следом за ним то же делает и 1.1.1.1.

OSPF, пример



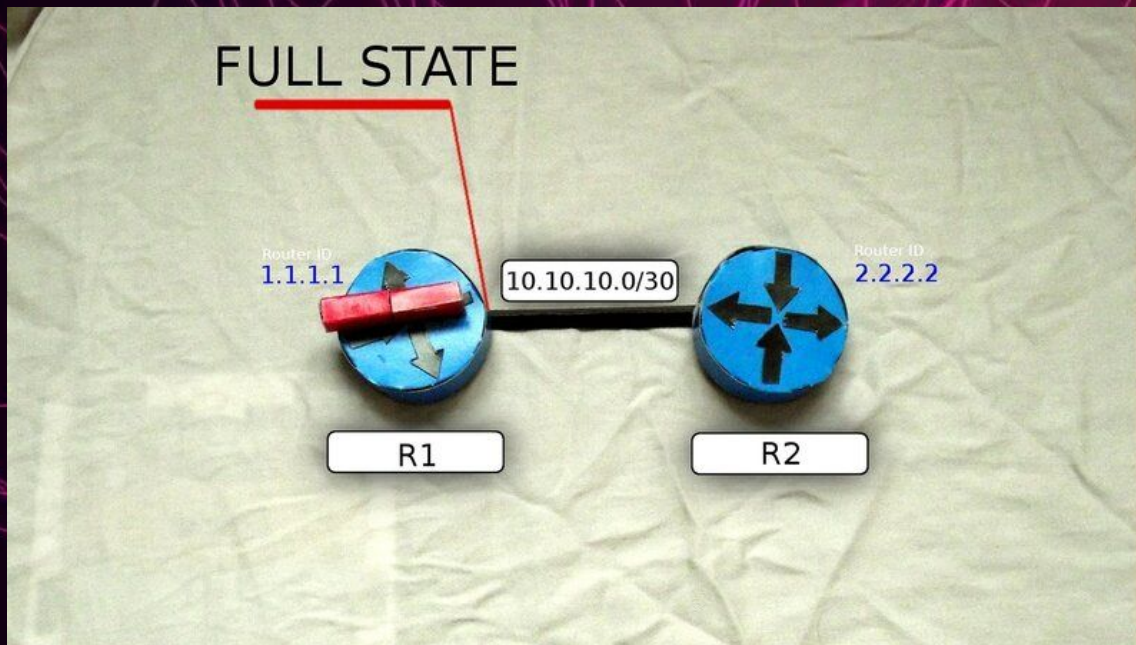
6) Получив сообщение, маршрутизаторы R1 и R2 отправляют подтверждение о приёме DBD (LSAck), а затем сравнивают новую информацию с той, что содержится у них в LSDB и, если есть отличия, посылают LSR (Link State Request) друг другу, тем самым переходя в новое состояние **LOADING**. В LSR они говорят – “Я про вот эту сеть ничего не знаю. Расскажи мне подробнее”.

OSPF, пример



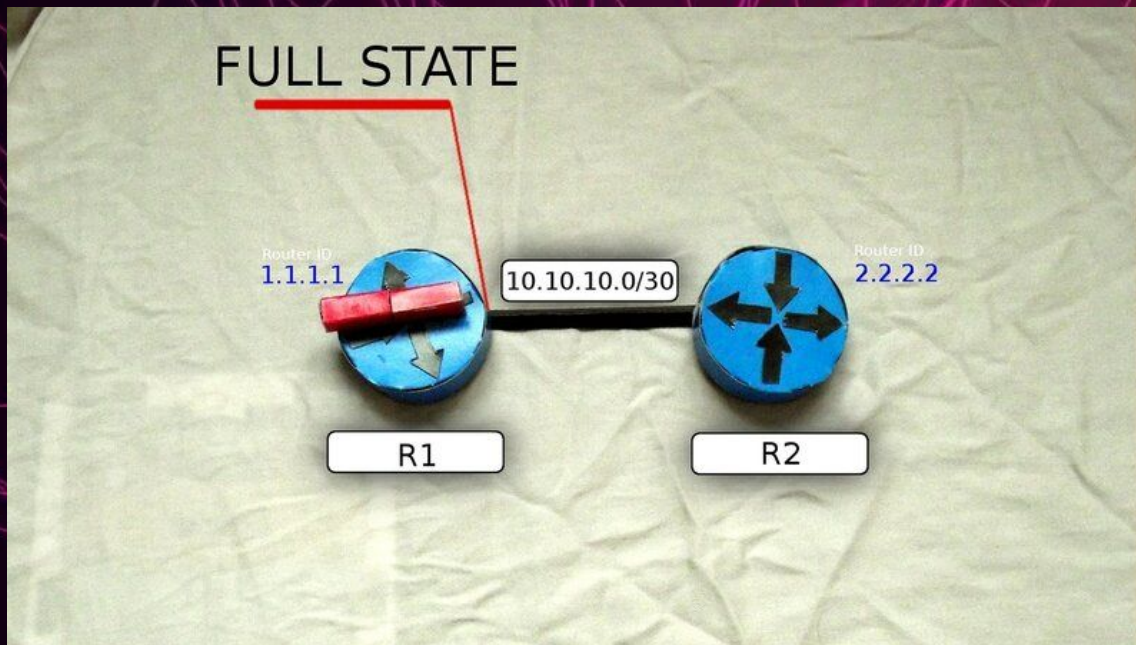
7) R2, получив LSR от R1, высылает LSU (Link State Update), которые содержат в себе LSA (Link State Advertisement) с детальной информацией о нужных подсетях.

OSPF, пример



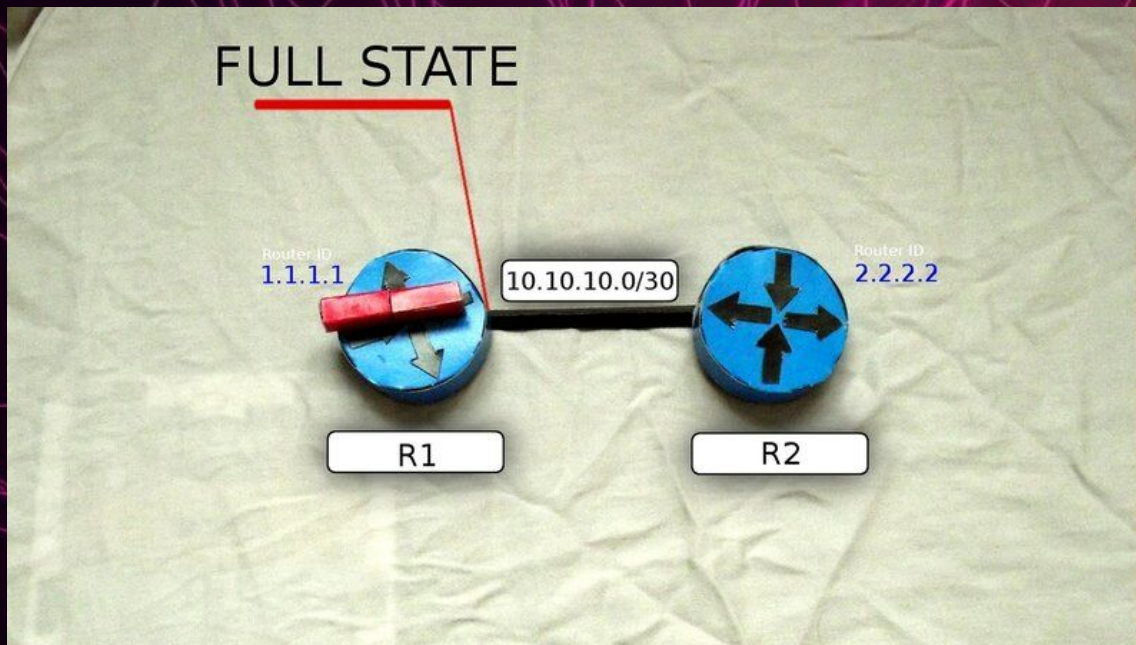
И вот, как только R1 получит последнюю порцию данных о всех подсетях и сформирует свою LSDB, он переходит в своё конечное состояние FULL STATE.

OSPF, пример



И вот, как только R1 получит последнюю порцию данных о всех подсетях и сформирует свою LSDB, он переходит в своё конечное состояние FULL STATE.

OSPF, пример



И вот, как только R1 получит последнюю порцию данных о всех подсетях и сформирует свою LSDB, он переходит в своё конечное состояние FULL STATE.

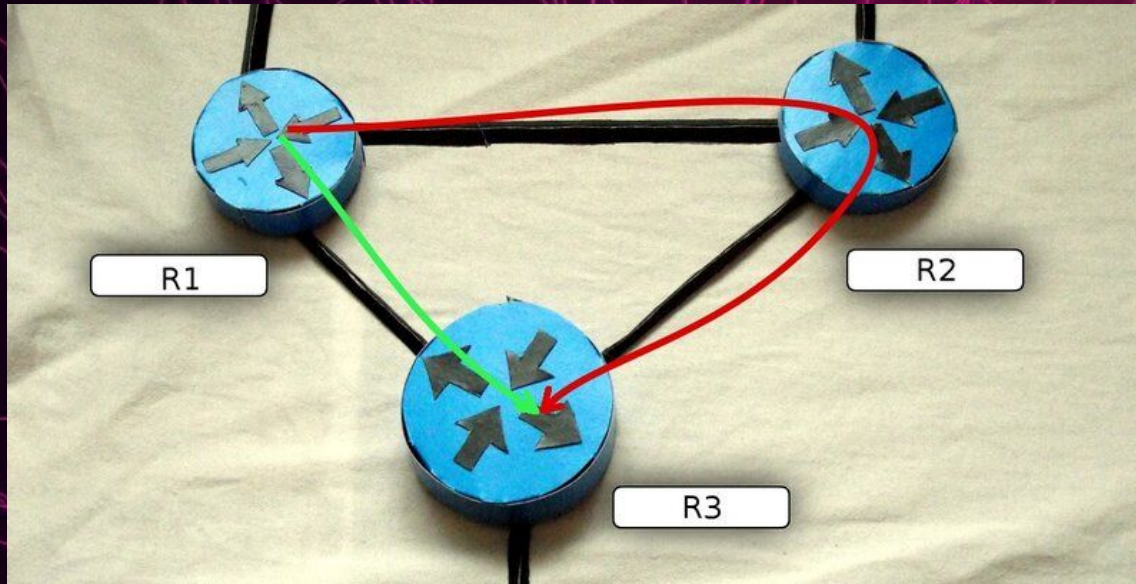
OSPF, пример

К тому моменту, как все маршрутизаторы зоны придут к состоянию Full State на всех на них должна быть полностью одинаковая LSDB — они же одну и ту же сеть изучали. То есть фактически это означает, что маршрутизатор знает всю вашу сеть, что, как и куда подключено.

8) Итак, сейчас у нас все маршрутизаторы знают всё о сети, но это знание не помогает в маршрутизации.

Следующим шагом OSPF, используя алгоритм Дейкстры (или его ещё называют SPF — Shortest Path First), вычисляет кратчайший маршрут до каждого маршрутизатора в зоне — он ведь знает всю топологию. В этом ему помогают метрики. Чем она ниже, тем маршрут лучше. Метрика — это стоимость движения по маршруту.

OSPF, пример



Например, в такой сети из R1 в R3 можно добраться напрямую или через R2. Естественно первый вариант будет стоить меньше. Но это при условии, что у вас везде одинаковый тип интерфейсов. А если, например, между R1 и R3 у вас модемное соединение в 56к или крайне нестабильный GPRS линк? Тогда у них будет очень высокая стоимость и OSPF предпочтёт более длинный, но быстрый путь. Найденный путь потом добавляется в таблицу маршрутизации.

OSPF, пример

Теперь каждые 10 секунд каждый маршрутизатор будет отправлять Hello-пакеты, а каждые 30 минут рассылаются LSA — это типа данные уже считаются устаревшими, надо бы обновить, даже если изменений не было.

Но топология может поменяться. Разумеется, было бы несколько странно ждать 40 секунд (Dead Interval) и только потом начинать перестраивать таблицу. Это было бы простительно ещё RIP'у, но не протоколу, который используется в огромном количестве современных сетей. Итак, как только падает какой-либо из линков (или несколько), маршрутизатор изменяет свою LSDB и генерирует LSU, присваивая ей номер больше, чем он был прежде (у каждой LSDB есть номер, который берётся из последнего полученного LSA).

OSPF, пример

Это LSU сообщение рассылается на мультикастовый адрес 224.0.0.5. Маршрутизаторы получившие его, проверяют номер LSA, содержащихся в LSU.

- 1) Если номер больше, чем номер текущей LSA маршрутизатора — LSDB меняется. (Версия LSDB старая, информация новая),
- 2) Если номер такой же, ничего не происходит. Этот маршрутизатор уже получил данный LSA по какому-то другому пути,
- 3) Если номер полученного LSA меньше локальной LSDB, это означает, что у маршрутизатора уже более актуальная информация, и он посылает новый LSA (на основе своей LSDB) отправителю прежнего.

После произведённых (или непроизведённых) действий соседу, от которого пришёл LSU пересылаются LSAck (мол, «посылку получили — всё в порядке»), а другим соседям отправляется изначальный LSU без изменений. На данном маршрутизаторе снова запускается алгоритм SPF и, при необходимости, обновляется таблица маршрутизации.

В общем, всё это происходит в целях поддержания актуальности информации на всех устройствах — LSDB должна быть одинаковой у всех.

Работа в сети Уфанет OSPF

```
show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
92.50.191.15	1	FULL/BDR	00:00:39	92.50.190.174	TenGigabitEthernet1/4
92.50.191.32	1	FULL/DR	00:00:39	92.50.190.137	Port-channel1
92.50.191.0	1	FULL/BDR	00:00:39	92.50.190.69	TenGigabitEthernet1/1
192.168.152.186	1	FULL/BDR	00:00:33	192.168.152.186	Vlan19
N/A	0	ATTEMPT/DROTHER	-	10.136.242.74	Vlan718
N/A	0	ATTEMPT/DROTHER	-	10.136.241.194	Vlan511

show ip ospf

```
Routing Process "ospf 1" with ID 92.50.191.17
Start time: 00:02:53.612, Time elapsed: 2y5w
Supports only single TOS(TOS0) routes
Supports opaque LSA
Supports Link-local Signaling (LLS)
Supports area transit capability
Supports NSSA (compatible with RFC 3101)
Event-log enabled, Maximum number of events: 1000, Mode: cyclic
Router is not originating router-LSAs with maximum metric
Initial SPF schedule delay 5000 msec
Minimum hold time between two consecutive SPFs 10000 msec
Maximum wait time between two consecutive SPFs 10000 msec
Incremental-SPF disabled
Minimum LSA interval 5 secs
Minimum LSA arrival 1000 msec
LSA group pacing timer 240 secs
Interface flood pacing timer 33 msec
Retransmission pacing timer 66 msec
Number of external LSA 0. Checksum Sum 0x000000
```

show ip ospf

```
Number of opaque AS LSA 0. Checksum Sum 0x000000
Number of DCbitless external and opaque AS LSA 0
Number of DoNotAge external and opaque AS LSA 0
Number of areas in this router is 1. 1 normal 0 stub 0 nssa
Number of areas transit capable is 0
External flood list length 0
IETF NSF helper support enabled
Cisco NSF helper support enabled
Reference bandwidth unit is 200000 mbps
  Area BACKBONE(0)
    Number of interfaces in this area is 4 (1 loopback)
    Area has RRR enabled
    Area has no authentication
    SPF algorithm last executed 3d04h ago
    SPF algorithm executed 80774 times
    Area ranges are
    Number of LSA 179. Checksum Sum 0x5B8705
    Number of opaque link LSA 0. Checksum Sum 0x000000
    Number of DCbitless LSA 54
    Number of indication LSA 0
    Number of DoNotAge LSA 0
    Flood list length 0
```


Информация о настройках OSPF на интерфейсах:

```
ckm53#show ip ospf interface brief
```

Interface	PID	Area	IP Address/Mask	Cost	State	Nbrs	F/C
Lo100	1	0	92.50.191.17/32	1	LOOP	0/0	
Te1/4	1	0	92.50.190.173/30	20	DR	1/1	
Po1	1	0	92.50.190.138/30	100	BDR	1/1	
Te1/1	1	0	92.50.190.70/30	20	DR	1/1	
Vl19	4	0	192.168.152.185/29	1	DR	1/1	
Vl293	3	112	10.136.242.89/30	1	DR	0/0	
Vl718	3	112	10.136.242.73/30	1	DR	0/0	
Vl511	3	112	10.136.241.193/30	1	DR	0/0	
Vl457	3	112	10.136.241.61/30	1	DR	0/0	
Vl393	3	112	10.136.241.177/30	1	DR	0/0	
Vl372	3	112	10.136.241.229/30	1	DR	0/0	
Vl365	3	112	10.136.241.21/30	1	DR	0/0	

База данных состояния каналов: show ip ospf database

Router Link States (Area 0)

Link ID	ADV Router	Age	Seq#	Checksum	Link count
10.14.254.67	10.14.254.67	1805	0x80001A7D	0x00040D	7
79.140.16.240	79.140.16.240	63	0x800015FB	0x007885	15
79.140.16.241	79.140.16.241	786	0x800015BE	0x0099A5	5
79.140.16.242	79.140.16.242	1148	0x800015D0	0x00641E	3
79.140.16.243	79.140.16.243	120	0x80000E40	0x00A0A9	3
79.140.16.244	79.140.16.244	559	0x8000076B	0x009316	3
79.140.16.245	79.140.16.245	1652	0x80000363	0x00C4F5	3
79.140.16.246	79.140.16.246	862	0x80000177	0x0029DD	3
79.140.16.248	79.140.16.248	1165	0x80019A59	0x005970	4
79.140.16.249	79.140.16.249	1527	0x8002391F	0x0071E6	5
79.140.16.252	79.140.16.252	1078	0x8002425C	0x00ADA1	7
79.140.16.253	79.140.16.253	510	0x80018076	0x00E24E	5
84.39.255.253	84.39.255.253	140	0x800000E9	0x006BFE	3
84.39.255.254	84.39.255.254	130	0x8000010D	0x008917	5
84.39.255.255	84.39.255.255	246	0x8000011E	0x00E92C	4
92.50.191.0	92.50.191.0	787	0x80001EF6	0x00B14C	17
92.50.191.2	92.50.191.2	40	0x8001E756	0x003470	10
92.50.191.3	92.50.191.3	168	0x80001F68	0x0042FE	17
92.50.191.4	92.50.191.4	402	0x80001E1D	0x00EA3A	3
92.50.191.5	92.50.191.5	610	0x80000772	0x003500	5

BGP, кратко

BGP (Border Gateway Protocol) — это основной протокол динамической маршрутизации, который используется в Интернете.

Маршрутизаторы, использующие протокол BGP, обмениваются информацией о доступности сетей. Вместе с информацией о сетях передаются различные атрибуты этих сетей, с помощью которых BGP выбирает лучший маршрут и настраиваются политики маршрутизации.

Один из основных атрибутов, который передается с информацией о маршруте — это список автономных систем, через которые прошла эта информация. Эта информация позволяет BGP определять где находится сеть относительно автономных систем, исключать петли маршрутизации, а также может быть использована при настройке политик.

BGP, кратко

Маршрутизация осуществляется пошагово от одной автономной системы к другой. Все политики BGP настраиваются, в основном, по отношению к внешним/соседним автономным системам. То есть, описываются правила взаимодействия с ними.

Так как BGP оперирует большими объемами данных (текущий размер таблицы для IPv4 более 450 тысяч маршрутов), то принципы его настройки и работы отличаются от внутренних протоколов динамической маршрутизации (IGP).

Терминология протокола

- Внутренний протокол маршрутизации (interior gateway protocol) - протокол, который используется для передачи информации о маршрутах внутри автономной системы.
- Внешний протокол маршрутизации (exterior gateway protocol) - протокол, который используется для передачи информации о маршрутах между автономными системами.
- 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.

Терминология протокола

- 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.
- Информация сетевого уровня о доступности сети (Network Layer Reachability Information, NLRI) – IP-префикс и длина префикса.
-

Описание протокола

BGP выбирает лучшие маршруты не на основании технических характеристик пути (пропускной способности, задержки и т.п.), а на основании политик. В локальных сетях наибольшее значение имеет скорость сходимости сети, время реагирования на изменения. И маршрутизаторы, которые используют внутренние протоколы динамической маршрутизации, при выборе маршрута, как правило, сравнивают какие-то технические характеристики пути, например, пропускную способность линков.

При выборе между каналами двух провайдеров, зачастую имеет значение не то, у какого канала лучше технические характеристики, а какие-то внутренние правила компании. Например, использование какого канала обходится компании дешевле. Поэтому в BGP выбор лучшего маршрута осуществляется на основании политик, которые настраиваются с использованием фильтров, анонсирования маршрутов, и изменения атрибутов.

Основные характеристики протокола

BGP это path-vector протокол с такими общими характеристиками:

- Использует TCP для передачи данных, это обеспечивает надежную доставку обновлений протокола (порт 179)
- Отправляет обновления только после изменений в сети (нет периодических обновлений)
- Периодически отправляет keepalive-сообщения для проверки TCP-соединения
- Метрика протокола называется path vector или атрибуты (attributes)

Описание работы протокола

Таблица соседей (neighbor table) – список всех соседей BGP

- Таблица BGP (BGP table, forwarding database, topology database):
- Список сетей, полученных от каждого соседа
- Может содержать несколько путей к destination сетям
- Атрибуты BGP для каждого пути
- Таблица маршрутизации – список лучших путей к сетям

BGP, походу не кратко



Отношения соседства

Для того чтобы установить отношения соседства, в BGP надо настроить вручную каждого соседа.

Когда указывается сосед локального маршрутизатора, обязательно указывается автономная система соседа. По этой информации BGP определяет тип соседа:

- **Внутренний BGP сосед (iBGP-сосед)** – сосед, который находится в той же автономной системе, что и локальный маршрутизатор. iBGP-соседи не обязательно должны быть непосредственно соединены.
- **Внешний BGP сосед (eBGP-сосед)** – сосед, который находится в автономной системе отличной от локального маршрутизатора. По умолчанию, eBGP-соседи должны быть непосредственно соединены.

Looking Glass Ufanet

Router: bgw0.ufa.ufanet.ru

Command: show route protocol bgp 77.88.55.50 terse table inet.0

inet.0: 909115 destinations, 2640040 routes (902264 active, 2 holddown, 607624 hidden)
+ = Active Route, - = Last Active, * = Both

A	V	Destination	P	Prf	Metric 1	Metric 2	Next hop	AS path
*	?	77.88.55.0/24	B	170	160			31133 13238 I
		unverified					>85.26.205.122	
	?		B	170	100	25		13238 I
		unverified					>178.18.225.147	
	?		B	170	100	25		13238 I
		unverified					>178.18.225.147	
	?		B	170	100	25		13238 I
		unverified					>178.18.233.1	
	?		B	170	100	25		13238 I
		unverified					>178.18.233.1	
	?		B	170	100			47775 13238 I
		unverified					145.255.24.230	
							>145.255.24.238	
							145.255.24.246	
							145.255.24.110	
							145.255.24.126	
							145.255.24.214	
							145.255.24.142	
							145.255.24.206	
							145.255.24.222	

Работа в сети Уфанет

```
ckm53#show ip bgp 81.30.198.190
BGP routing table entry for 81.30.198.188/30, version 87105
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  65510 24955, (received & used)
    10.2.3.10 (metric 21) from 92.50.191.3 (92.50.191.3)
      Origin incomplete, localpref 100, valid, internal
      rx pathid: 0, tx pathid: 0
  Refresh Epoch 1
  65510 24955, (received & used)
    10.2.3.10 (metric 21) from 92.50.191.0 (92.50.191.0)
      Origin incomplete, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```


Работа в сети Уфанет

```
ckm53#show ip bgp
BGP table version is 212758, local router ID is 92.50.191.17
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network          Next Hop          Metric LocPrf Weight Path
* i  0.0.0.0         145.255.7.96          500     0 i
*>i  0.0.0.0         145.255.7.96          500     0 i
      0.0.0.0         0.0.0.0                0     0 i
* i  10.0.0.0        92.50.191.3           0     100   0 i
*>i  10.0.0.0        92.50.191.0           0     100   0 i
* i  10.0.0.64/26    10.1.164.10           0     100   0 i
*>i  10.0.0.64/26    10.1.164.10           0     100   0 i
*>i  10.0.0.128/28   10.2.3.10            100    0 65510 24955 ?
* i  10.0.0.128/28   10.2.3.10            100    0 65510 24955 ?
*>i  10.0.0.224/28   10.2.3.10            100    0 65510 24955 ?
* i  10.0.0.224/28   10.2.3.10            100    0 65510 24955 ?
* i  10.0.1.16/30    92.50.191.18          0     100   0 ?
*>i  10.0.1.16/30    92.50.191.18          0     100   0 ?
* i  10.0.3.0/26     10.0.6.130            0     100   0 65500 24955 ?
*>i  10.0.3.0/26     10.0.6.130            0     100   0 65500 24955 ?
* i  10.0.3.64/26    10.0.6.130            0     100   0 65500 24955 ?
*>i  10.0.3.64/26    10.0.6.130            0     100   0 65500 24955 ?
* i  10.0.3.128/26   10.0.6.130            0     100   0 65500 24955 ?
*>i  10.0.3.128/26   10.0.6.130            0     100   0 65500 24955 ?
* i  10.0.3.192/26   10.0.6.130            0     100   0 65500 24955 ?
*>i  10.0.3.192/26   10.0.6.130            0     100   0 65500 24955 ?
* i  10.0.4.0/26     10.0.6.130            0     100   0 65500 24955 ?
```