

Лекция XI. Протоколы маршрутизации: RIP, OSPF, BGP

Курс читает Рогозин Н.О., кафедра ИУ-7

Основная задача алгоритма маршрутизации

- Имея набор маршрутизаторов, которые соединены каналами, найти оптимальный путь от начального маршрутизатора к конечному.

Распределенный подход

- При распределенном подходе все маршрутизаторы сети находятся в равных условиях, они находят маршруты и строят собственные таблицы маршрутизации, работая в тесной кооперации друг с другом, постоянно обмениваясь информацией о конфигурации сети.

Централизованный подход

- При централизованном подходе в сети существует один выделенный маршрутизатор, который собирает всю информацию о топологии и состоянии сети от других маршрутизаторов.
- На основании этих данных выделенный маршрутизатор (который иногда называют сервером маршрутов) строит таблицы маршрутизации для всех остальных маршрутизаторов сети, а затем распространяет их по сети, чтобы каждый маршрутизатор получил собственную таблицу и в дальнейшем самостоятельно принимал решение о продвижении каждого пакета.

Дистанционно-векторные алгоритмы

- Каждый маршрутизатор периодически и широковещательно рассылает по сети вектор, компонентами которого являются расстояния (дистанции), измеренные в той или иной метрике, от данного маршрутизатора до всех известных ему сетей.
- Пакеты протоколов маршрутизации обычно называют объявлениями о расстояниях, так как с их помощью маршрутизатор объявляет остальным маршрутизаторам известные ему сведения о конфигурации сети.
- Получив от некоторого соседа вектор расстояний до известных тому сетей, маршрутизатор наращивает компоненты вектора на величину расстояния от себя до данного соседа и дополняет вектор информацией об известных ему самому других сетях, о которых он узнал непосредственно (если они подключены к его портам) или из аналогичных объявлений других маршрутизаторов.
- Обновленное значение вектора маршрутизатор рассылает своим соседям.

Алгоритмы состояния связей

- Обеспечивают каждый маршрутизатор информацией, достаточной для построения точного графа связей сети.
- Все маршрутизаторы работают на основании одного и того же графа, что делает процесс маршрутизации более устойчивым к изменениям конфигурации.
- Каждый маршрутизатор использует граф сети для нахождения оптимальных по некоторому критерию маршрутов до каждой из сетей, входящих в составную сеть.
- Чтобы понять, в каком состоянии находятся линии связи, подключенные к его портам, маршрутизатор периодически обменивается короткими пакетами HELLO со своими не посредственными соседями.
- В отличие от протоколов DVA, которые регулярно передают вектор расстояний, протоколы LSA ограничиваются короткими сообщениями, а передача более объемных сообщений происходит только в тех случаях, когда с помощью сообщений HELLO был установлен факт изменения состояния какой-либо связи.

Сбалансированный-гибридный алгоритм

- Такое название применяется в компании Cisco для обозначения алгоритма маршрутизации, применяемого в протоколе EIGRP.
- Применение для данных алгоритмов названия "сбалансированные гибридные" основано на том, что они воплощают в себе свойства и дистанционно-векторных алгоритмов маршрутизации, и алгоритмов маршрутизации с учетом состояния каналов.
- Например, хотя по существу в основе протокола EIGRP лежит дистанционно-векторный алгоритм, этот протокол предусматривает передачу дополнительной информации о топологии для формирования "схемы" распределенной сети, как предусмотрено алгоритмом маршрутизации с учетом состояния каналов.

Показатели алгоритмов (метрики)

- Длина маршрута
- Надежность
- Задержка
- Ширина полосы пропускания
- Нагрузка
- Стоимость связи

Длина маршрута

- Длина маршрута является наиболее общим показателем маршрутизации. Некоторые протоколы маршрутизации позволяют администраторам сети назначать произвольные цены на каждый канал сети. В этом случае длиной тракта является сумма расходов, связанных с каждым каналом.
- Другие протоколы маршрутизации определяют "количество пересылок", т.е. показатель, характеризующий число проходов, которые пакет должен совершить на пути от источника до пункта назначения через изделия объединения сетей (такие как роутеры).

Надежность

- Надежность каждого канала сети в контексте алгоритмов маршрутизации обычно описывается в терминах отношения бит/ошибка. Некоторые каналы сети могут отказывать чаще, чем другие. Отказы одних каналов сети могут быть устранены легче или быстрее, чем отказы других каналов. При назначении оценок надежности могут быть приняты в расчет любые факторы надежности.
- Оценки надежности обычно назначаются каналам сети администраторами сети. Как правило, это произвольные цифровые величины.

Задержка

- Под задержкой маршрутизации обычно понимают отрезок времени, необходимый для передвижения пакета от источника до пункта назначения через объединенную сеть.
- Задержка зависит от многих факторов, включая полосу пропускания промежуточных каналов сети, очереди в порт каждого роутера на пути передвижения пакета, перегруженность сети на всех промежуточных каналах сети и физическое расстояние, на которое необходимо переместить пакет.
- Т.к. здесь имеет место конгломерация нескольких важных переменных, задержка является наиболее общим и полезным показателем.

Полоса пропускания

- Полоса пропускания относится к имеющейся мощности трафика какого-либо канала. При прочих равных показателях, канал Ethernet 10 Mbps предпочтителен любой арендованной линии с полосой пропускания 64 Кбайт/сек.
- Хотя полоса пропускания является оценкой максимально достижимой пропускной способности канала, маршруты, проходящие через каналы с большей полосой пропускания, не обязательно будут лучше маршрутов, проходящих через менее быстродействующие каналы.

Автономные системы

Для поддержки масштабируемости глобальная сеть делится на специальные зоны - “Автономные системы” (АС) (или “**домены**”).

Внутри АС (Протоколы внутреннего шлюза):

- Все маршрутизаторы должны использовать один внутридоменный протокол
- В разных АС могут использоваться разные внутридоменные протоколы
- Шлюзовый/граничный маршрутизатор - находится на границе АС для доступа извне

Между АС (Протоколы внешнего шлюза)

- маршрутизация между АС
- выполняется шлюзами

Автономные системы

- **Автономная система** (Autonomous System, AS) — это совокупность сетей под единым административным управлением, обеспечивающим общую для всех входящих в автономную систему маршрутизаторов политику маршрутизации.
- Обычно автономной системой управляет один поставщик услуг Интернета, самостоятельно выбирая протоколы маршрутизации
- Крупные поставщики услуг и корпорации могут представить свою составную сеть как набор нескольких автономных систем. Регистрация автономных систем происходит централизованно, как и регистрация IP-адресов и DNS-имен.
- Номер автономной системы состоит из 16 разрядов и никак не связан с префиксами IP-адресов входящих в нее сетей.

Маршрутизация внутри АС

- **RIP: Routing Information Protocol** [RFC 1723]
 - классический дистанционно-векторный протокол, сообщения каждые 30 с.
 - практически не используется
- **OSPF: Open Shortest Path First** [RFC 2328]
 - классический протокол состояния связей
 - Протокол IS-IS (стандарт ISO, не RFC) аналогичен OSPF
- **EIGRP: Enhanced Interior Gateway Routing Protocol**
 - Дистанционно-векторный
 - Ранее был проприетарным (Cisco)
 - Стал открытым в 2013 [RFC 7868]

Протокол RIP

- Для измерения расстояния до сети стандарты протокола RIP допускают различные метрики: хопы (транзитные переходы), значения пропускной способности, вносимые задержки, надежность сетей (то есть соответствующие признакам **D**, **T** и **R** в поле качества сервиса IP-пакета), а также любые комбинации этих метрик.
- Метрика должна обладать **свойством аддитивности** — метрика составного пути должна быть равна сумме метрик составляющих этого пути.
- В большинстве реализаций RIP используется простейшая метрика — количество хопов/ промежуточных маршрутизаторов

Порядок работы

- **Этап 1 — создание минимальной таблицы.** Маршрутизаторы, работающие по протоколу RIP, могут иметь идентификаторы, однако в RIP-сообщениях эти идентификаторы не передаются. В минимальная таблица маршрутизации, в которой учитываются только непосредственно подсоединенные сети.
- **Этап 2 — рассылка минимальной таблицы соседям.** Данные передаются в UDP-дейтаграммах и включают
- **Этап 3 — получение RIP-сообщений от соседей и обработка полученной информации.** Новые записи попадают в таблицу **только если метрика лучше.**

Схема работы RIP

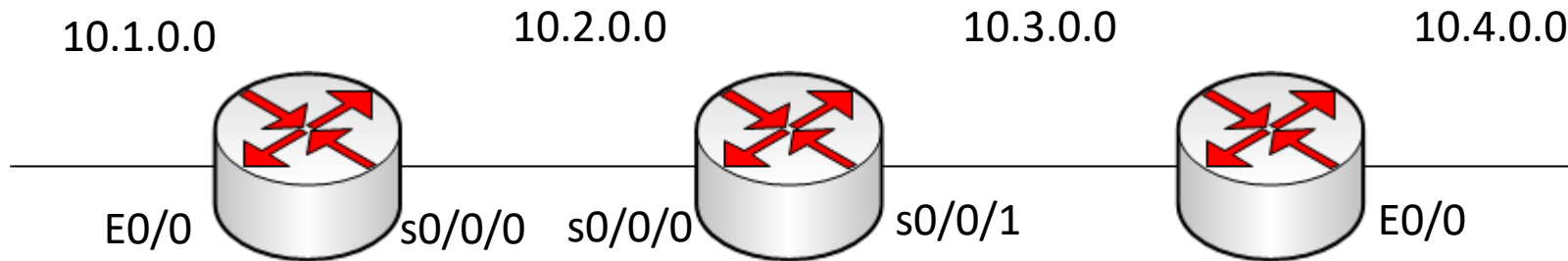


Таблица маршрутизации		
10.1.0.0	E0/0	0
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/0	1
10.4.0.0	S0/0/0	2

Таблица маршрутизации		
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/1	0
10.4.0.0	S0/0/1	1
10.1.0.0	S0/0/0	1

Таблица маршрутизации		
10.3.0.0	S0/0/0	0
10.4.0.0	E0/0	0
10.2.0.0	S0/0/0	1
10.1.0.0	S0/0/0	2

Адаптация RIP к изменениям

- К новым маршрутам маршрутизаторы RIP приспосабливаются просто — они передают новую информацию в очередном сообщении своим соседям, и постепенно эта информация становится известна всем маршрутизаторам сети.
- К изменениям, связанным с потерей какого-либо маршрута, маршрутизаторы RIP адаптируются сложнее, поскольку в формате сообщений протокола RIP нет поля, которое бы указывало на то, что путь к данной сети больше не существует.
- Для уведомления о том, что некоторый маршрут недействителен, используются два механизма:
 - истечение времени жизни маршрута;
 - указание специального (бесконечного) расстояния до сети, ставшей недоступной.

Счет до бесконечности

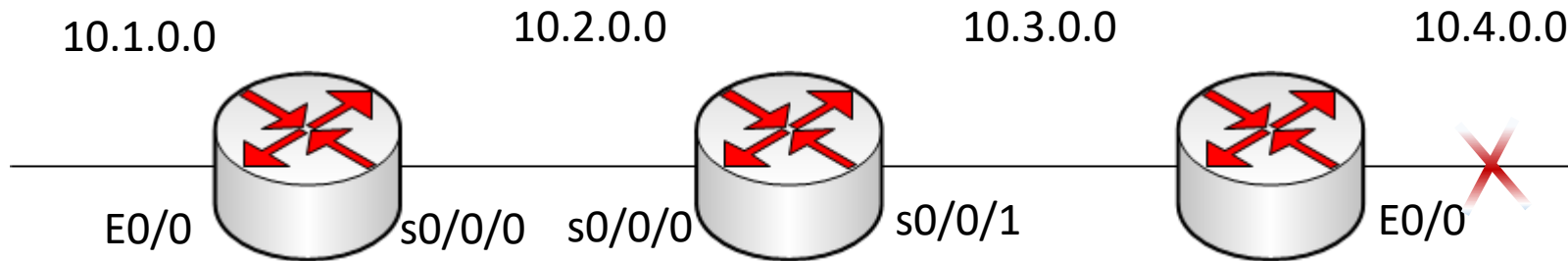


Таблица маршрутизации		
10.1.0.0	E0/0	0
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/0	1
10.4.0.0	S0/0/0	16

Таблица маршрутизации		
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/1	0
10.4.0.0	S0/0/1	16
10.1.0.0	S0/0/0	1

Таблица маршрутизации		
10.3.0.0	S0/0/0	0
10.4.0.0	E0/0	16
10.2.0.0	S0/0/0	1
10.1.0.0	S0/0/0	2

Решения проблем в RIP

- **Метод расщепления горизонта** : маршрутная информация о некоторой сети, хранящаяся в таблице маршрутизации, никогда не передается тому маршрутизатору, от которого она получена
- **Триггерные обновления** : маршрутизатор, получив данные об изменении метрики до какой-либо сети, не ждет истечения периода передачи таблицы маршрутизации, а передает данные об изменившемся маршруте немедленно. Во многих случаях это позволяет предотвратить передачу устаревших сведений об отказавшем маршруте, но он перегружает сеть служебными сообщениями, поэтому триггерные объявления также делаются с некоторой задержкой.
- **Замораживание изменений** — позволяет исключить подобные ситуации. Он связан с введением тайм-аута на принятие новых данных о сети, которая только что стала недоступной. Этот тайм-аут предотвращает принятие устаревших сведений о некотором маршруте от тех маршрутизаторов, которые находятся на некотором расстоянии от отказавшей связи и передают устаревшие сведения о ее работоспособности.

Формат пакета RIP v.1

- Поле **«Команда»** длиной 1 байт определяет, является ли пакет запросом (1) или ответом (2).
- Поле **«Версия»** длиной 1 байт содержит информацию о версии протокола. Для RIP v.1 значение поля должно быть равно 1.
- Поле **«Идентификатор семейства адресов»** длиной 2 байта определяет протокол, используемый для сетевой адресации. Значение 2 соответствует протоколу IPv4.
- Поле **«Сетевой адрес»** длиной 4 байта содержит IP-адрес определенной сети.
- Поле **«Метрика»** длиной 4 байта определяет расстояние (в переходах) до указанной сети. (В запросах поле не используется.)

Отличия RIP v2

- **Поддержка VLSM.** Вместе с обновлениями RIP 2 передаются маски подсети.
- **Многоадресатные обновления.** Обновления передаются с помощью многоадресатной, а не широковещательной рассылки, поэтому сокращаются непроизводительные затраты процессорного времени в хостах, не участвующих в работе протокола RIP.
- **Поддержка аутентификации.** В маршрутизаторах, совместимых с требованиями **RFC 2453**, поддерживается аутентификация на основе открытого текста. (В маршрутизаторах Cisco поддерживается также аутентификация с шифрованием по алгоритму MD5.)

OSPF

- Протокол динамической маршрутизации, основанный на технологии отслеживания состояния канала (link-state technology)
- Стандарт описан в **RFC 2328**
- Для нахождения кратчайшего пути используется алгоритм Дейкстры.
- Данные о маршрутах передаются между маршрутизаторами в рамках одной автономной системы

Принцип работы

- Процесс похож на процесс распространения векторов расстояний до сетей в протоколе RIP, но сама информация качественно иная — это информация о топологии сети.
- Сообщения, с помощью которых распространяется топологическая информация, называются объявлениями о состоянии связей (**Link State Advertisement, LSA**) сети
- При транзитной передаче объявлений LSA маршрутизаторы не модифицируют информацию, как это происходит в дистанционно-векторных протоколах, в частности в RIP, а передают ее в неизменном виде.
- В результате все маршрутизаторы сети сохраняют в своей памяти идентичные сведения о текущей конфигурации графа связей сети.

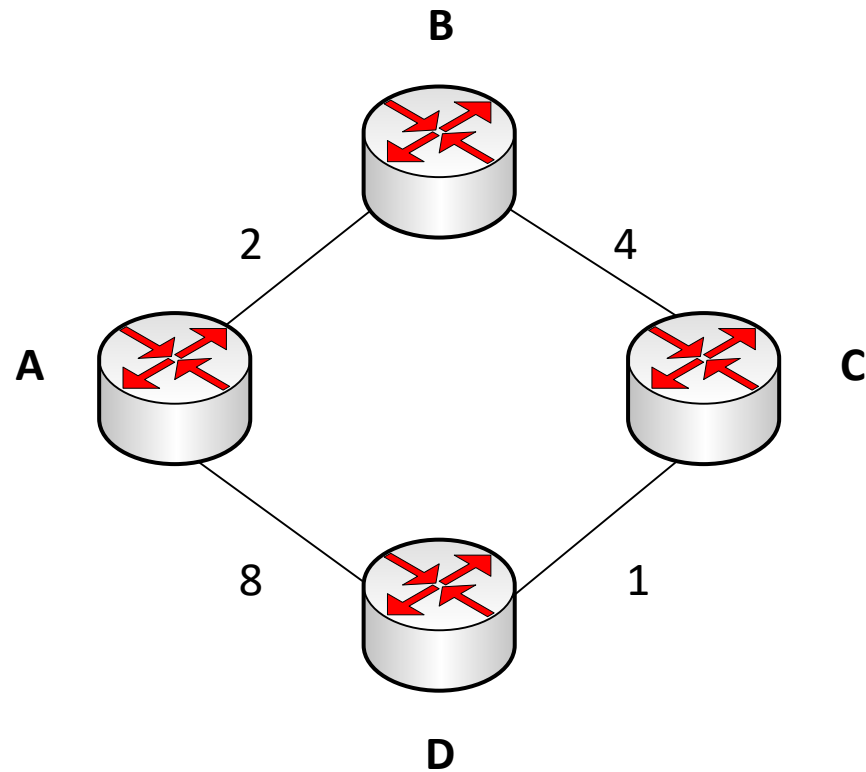
Принцип работы

- Для контроля состояния связей и соседних маршрутизаторов маршрутизаторы OSPF передают друг другу особые сообщения HELLO каждые 10 секунд. Небольшой объем этих сообщений делает возможным частое тестирование состояния соседей и связей с ними.
- Если сообщения HELLO перестают поступать от какого-либо непосредственного соседа, то маршрутизатор делает вывод о том, что состояние связи изменилось с работоспособного на неработоспособное, после чего вносит соответствующие коррективы в свою топологическую базу данных.

Алгоритм Дейкстры

- Пусть $D(v)$ равно сумме весов связей для данного пути.
- Пусть $c(i,j)$ равно весу связи между узлами с номерами i и j .
- Устанавливаем множество узлов $N = \{1\}$.
- Для каждого узла v не из множества n устанавливаем $D(v) = c(1,v)$.
- Для каждого шага находим узел w не из множества N , для которого $D(w)$ минимально, и добавляем узел w в множество N .
- Актуализируем $D(v)$ для всех узлов не из множества N
 $D(v) = \min\{D(v), D(v) + c(w,v)\}$.
- Повторяем действия, пока все узлы не окажутся в множестве N .

Схема работы OSPF (алгоритм Дейкстры)

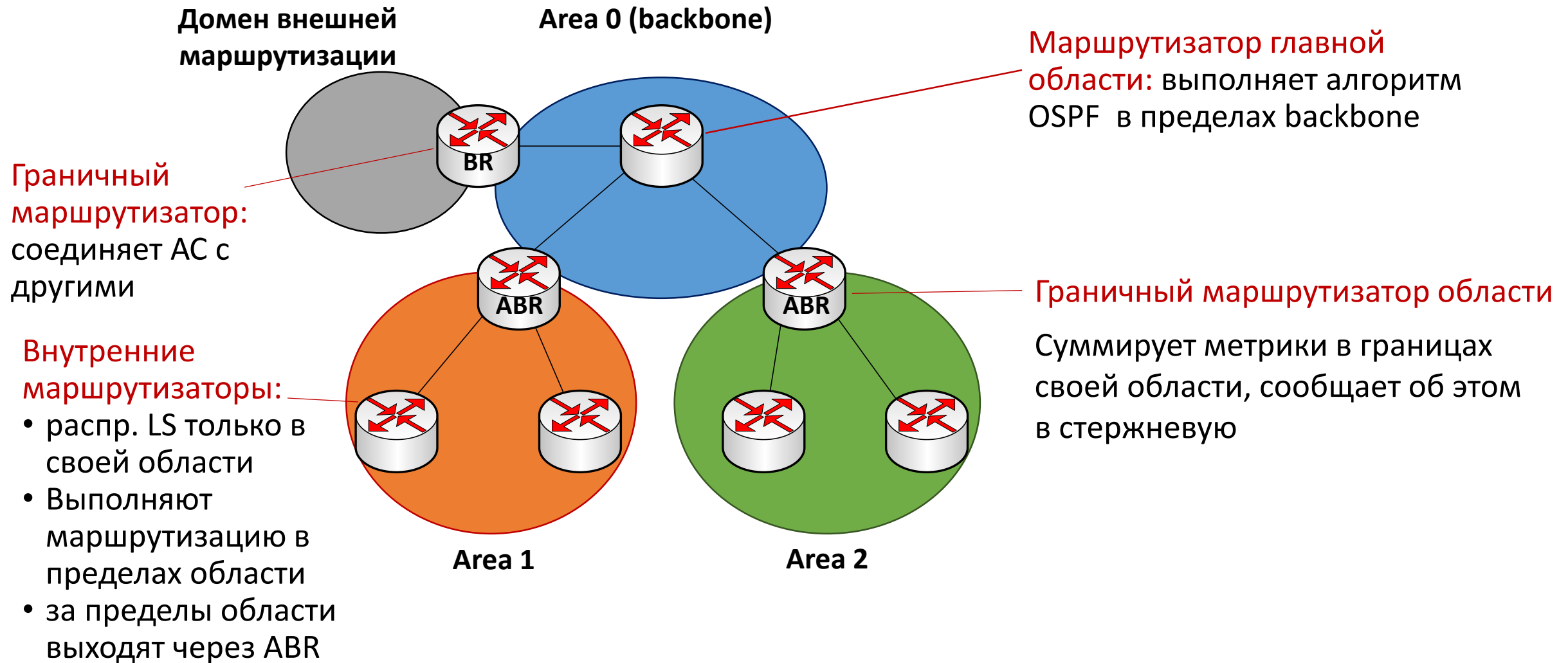


Множество	Метрика связи узла а с узлами			
	Шаг	N	B	C
0	{A}	2	-	8
1	{A,B}	(2)	6	8
2	{A,B,C}	2	(6)	7
3	{A,B,C,D}	2	4	(7)

Принцип работы

- Маршрутизатор строит граф оптимальных маршрутов, в котором он сам является корнем.
- На основании этого графа маршрутизаторы и производят свои расчеты для каждого маршрута информационного пакета.
- В свою очередь, по дереву оптимальных маршрутов строится маршрутная таблица, которая служит основой оценок и выбора маршрутов.
- В зависимости от того, какую функцию выполняет тот или иной маршрутизатор (т. е. относится к области **backbone**, **смежной области**, **периферии** или **внутренней области**), он выполняет те или иные функции (инициализации, согласования, корректировки и т. п.) с одной или несколькими топологическими базами данных, находящихся под его контролем.

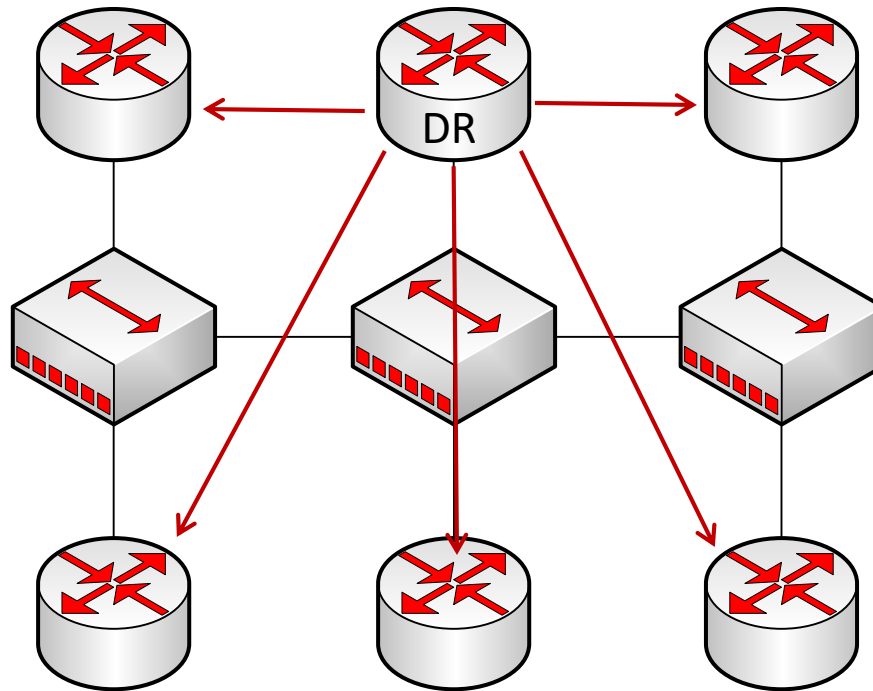
Разделение на области (areas)



Назначенный маршрутизатор (Designated Router)

- Особая роль, позволяющая строго определенному маршрутизатору распределять LSA, LSA от других маршрутизаторов при такой схеме отбрасываются
- Используется в условиях широковещания в сетях множественного доступа
- Присваивается на основе выборов, проводимых маршрутизаторами
- По умолчанию голоса отдаются за маршрутизатор с наибольшим значением приоритета
- В случае равного приоритета у двух и более маршрутизаторов - на основе Router ID (генерируется автоматически, может быть изменен вручную)

Designated Router, пример



Заголовок OSPF

1	1	2	4	4	2	2	8	Var
Номер версии	Тип	Длина пакета	ИД роутера	ИД области	Контр. сумма	Тип аутен-тиф.	Аутен-тификация	Данные

Заголовок OSPF

- Первое поле в заголовке OSPF - это номер версии OSPF (**version number**).
- За номером версии идет поле типа (**type**).
- Поле длины пакета (***packet length***). Это поле обеспечивает длину пакета вместе с заголовком OSPF в байтах.
- Поле идентификатора роутера (***router ID***) идентифицирует источник пакета.
- Поле идентификатора области (***area ID***) идентифицирует область, к которой принадлежит данный пакет
- Поле контрольной суммы IP (***checksum***) проверяет содержимое всего пакета для выявления потенциальных повреждений, имевших место при транзите.

Заголовок OSPF

- Поле типа удостоверения (***authentication type***). Примером типа удостоверения является "простой пароль". Все обмены протокола OSPF проводятся с установлением достоверности.
- Тип удостоверения устанавливается по принципу "отдельный для каждой области".
- Поле удостоверения (***authentication***). Это поле длиной 64 бита и содержит информацию удостоверения.

Типы пакетов LSA

- **Hello** - Отправляется через регулярные интервалы времени для установления и поддержания соседских взаимоотношений. Ключевые параметры пакетов этого типа — маски сети, периоды приветствования и сигнализации обрыва контакта. Эти и другие параметры входят в состав Hello-пакетов.
- **Database Description** - Описывает содержимое базы данных; обмен этими пакетами производится при инициализации смежности. Для обработки таких пакетов используется процедура "переклички" (poll-response), в которой один из маршрутизаторов определяется как master, а другой как slave. Соответственно, master отправляет эти пакеты, а slave должен отвечать за их получение.
- **Link-State Request** - Запрос о состоянии канала. Запрашивает части топологической базы данных соседа. Обмен этими пакетами производится после того, как какой-нибудь роутер обнаруживает, (путем проверки пакетов описания базы данных), что часть его топологической базы данных устарела.

Типы пакетов LSA

- **Link-State Update** - Корректировка состояния канала. Отвечает на пакеты запроса о состоянии канала. Эти пакеты также используются для регулярного распределения LSA. В одном пакете могут быть включены несколько LSA.
- **Link-State Acknowledgement** - Подтверждение состояния канала. Подтверждает пакеты корректировки состояния канала. Пакеты корректировки состояния канала должны быть четко подтверждены, что является гарантией надежности процесса лавинной адресации пакетов корректировки состояния канала через какую-нибудь область.

Типы LSA

- Router links advertisements (RLA)
- Network links advertisements (NLA)
- Summary links advertisements (SLA)
- AS external links advertisements

Router links advertisements (RLA)

- Объявления о каналах роутера. Описывают собранные данные о состоянии каналов роутера, связывающих его с конкретной областью.
- Любой роутер отправляет RLA для каждой области, к которой он принадлежит.
- RLA направляются лавинной адресацией через всю область, но они не отправляются за ее пределы.

Network links advertisements (NLA)

- Объявления о сетевых каналах. Отправляются назначенными роутерами.
- Они описывают все роутеры, которые подключены к сети с множественным доступом, и отправляются лавинной адресацией через область, содержащую данную сеть с множественным доступом.

Summary links advertisements (SLA)

- Суммарные объявления о каналах. Суммирует маршруты к пунктам назначения, находящимся вне какой-либо области, но в пределах данной AS.
- Они генерируются роутерами границы области, и отправляются лавинной адресацией через данную область. В стержневую область посылаются объявления только о внутриобластных роутерах. В других областях рекламируются как внутриобластные, так и межобластные маршруты.

AS external links advertisements

- Объявления о внешних каналах AS.
- Описывают какой-либо маршрут к одному из пунктов назначения, который является внешним для данного AS.
- Объявления о внешних каналах AS вырабатываются граничными роутерами AS.
- Этот тип объявлений является единственным типом объявлений, которые продвигаются во всех направлениях данной AS; все другие объявления продвигаются только в пределах конкретных областей.

Протокол внешнего шлюза BGP [RFC 4271]

- Маршрутизаторы, использующие протокол BGP, обмениваются информацией о доступности сетей, а также атрибутами 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.
- 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.

Принцип работы BGP

- Для установления сеанса с указанными соседями маршрутизаторы BGP используют протокол TCP (порт 179).
- При установлении BGP-сеанса могут применяться разнообразные способы аутентификации маршрутизаторов, повышающие безопасность работы автономных систем.
- Основное сообщение протокола BGP — сообщение **UPDATE** (обновить), с помощью которого маршрутизатор сообщает маршрутизатору соседней автономной системы о достижимости сетей, относящихся к его собственной автономной системе.
- Само название этого сообщения говорит о том, что это триггерное объявление, посылаемое соседу, если в топологии автономной системы происходят изменения: появляются новые сети или новые пути к сетям либо же, напротив, исчезают существовавшие сети или пути. В одном сообщении UPDATE можно объявить об одном новом маршруте или аннулировать несколько маршрутов, переставших существовать.

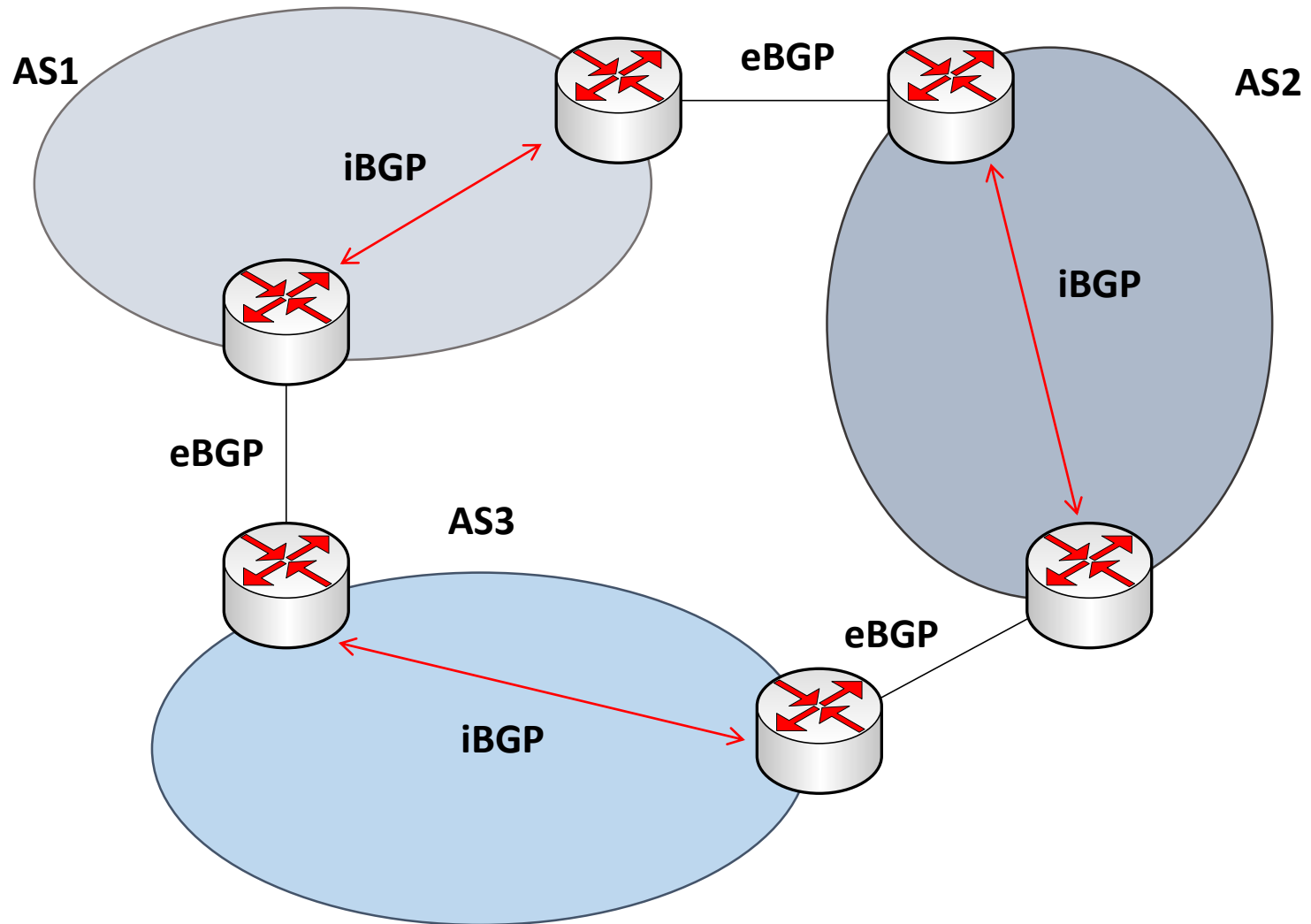
Принцип работы BGP

- Автономный пограничный маршрутизатор, который участвует в маршрутизации с использованием вектора путей, извещает о достижимости сетей в их собственной автономной системе для соседних автономных пограничных маршрутизаторов.
- Концепция окружения здесь та же самая, как в уже рассмотренных протоколах RIP и OSPF. Два пограничных маршрутизатора автономных систем, подключенные к той же самой сети, — соседи.

BGP

- Каждый маршрутизатор, который получает вектор пути, проверяет, что предложенный путь согласован с его политикой (набором правил, назначаемых администратором, который управляет маршрутизатором).
- Если политика маршрутизации соответствует записанной в программе, маршрутизатор обновляет таблицы маршрутизации и модифицирует сообщение, прежде чем послать его к следующему соседу.
- Модификация содержит дополнение номера своего АС для пути и замещающий номер следующего маршрутизатора, входящего со своим собственным идентификатором.

Схема работы BGP



eBGP и iBGP

- Для того, чтобы передавать данные, полученные от протокола OSPF через один из своих внутренних интерфейсов, имеющие другой формат и не содержащие, например, сведений о номере автономной системы, в которой находится сеть, маршрутизаторы с общей АС также устанавливают между собой BGP-сеанс.
- Такая реализация протокола BGP называется внутренней версией BGP (**Interior BGP, iBGP**), в отличие от основной, внешней версии (**Exterior BGP, eBGP**).

Таблица BGP

Под маршрутом в BGP понимается последовательность автономных систем, которую нужно пройти на пути к указанной в адресе сети. Формальная запись о маршруте (BGP Route) к сети (Network/Mask_length) выглядит так:

BGP Route = AS Path; NextHop; Network/Mask length

Сеть	Следующий маршрутизатор	Путь
N01	R01	AS14,AS23, AS67
N02	R05	AS22,AS67, AS05, AS89
N03	R06	AS67,AS89, AS09, AS34
N04	R12	AS62,AS02, AS09

Пример таблицы

Формат пакета BGP



Формат пакета BGP

- **Marker** (16 байт). Поле маркера. Оно содержит величину, которую получатель сообщения может прогнозировать (например, при использовании механизма аутентификации). Это поле также может быть использовано для обнаружения потерянной синхронизации между парой BGP-роутеров.
- **Length** (16 бит). Поле длины пакета. Это поле содержит полную длину сообщения в байтах, включая заголовок и данные.
- **Type** (8 бит). Поле тип сообщения (Open, Notification, Update, Keepalive, Errors).
- За основным заголовком BGP следует структура данных, определяемая типом сообщения.

Сообщения BGP

- **Открывающие сообщения (open message).** После того, как организовано соединение протокола транспортного уровня, первым сообщением, отправляемым каждой стороной, является открывающее сообщение. Пакет открывающего сообщения содержит, в дополнение к основному заголовку BGP несколько дополнительных параметров.
- **Параметр версия (version)** — номер версии BGP дает возможность получателю проверить, совпадает ли его версия с версией отправителя.
- **Параметр автономная система (*autonomous system*)** содержит номер AS отправителя.
- **Параметр время удерживания (*hold time*)** указывает максимальное число секунд, которые могут пройти без получения какого-либо сообщения от передающего устройства, прежде чем считать его отказавшим.
- **Параметр код аутентификации (*authentication code*)** указывает на используемый код удостоверения, если, конечно, таковой имеется/ Параметр данных аутентификации (*authentication data*) содержит возможные данные аутентификации.

Необходимость маршрутизации внутри и между АС

Политика маршрутизации:

- Между АС: администратору нужен контроль над управлением трафиком, кто получает доступ
- Внутри АС: один администратор (условно), политика имеет меньшее значение

Масштабирование

- иерархическая маршрутизация: ограничивает диапазон полной информации о топологии сетей
- BGP: маршрутизация в CIDR сети(объединенные маршруты)

Производительность:

- Внутри АС: производительность в приоритете
- Между АС: политика маршрутизации превышает производительности

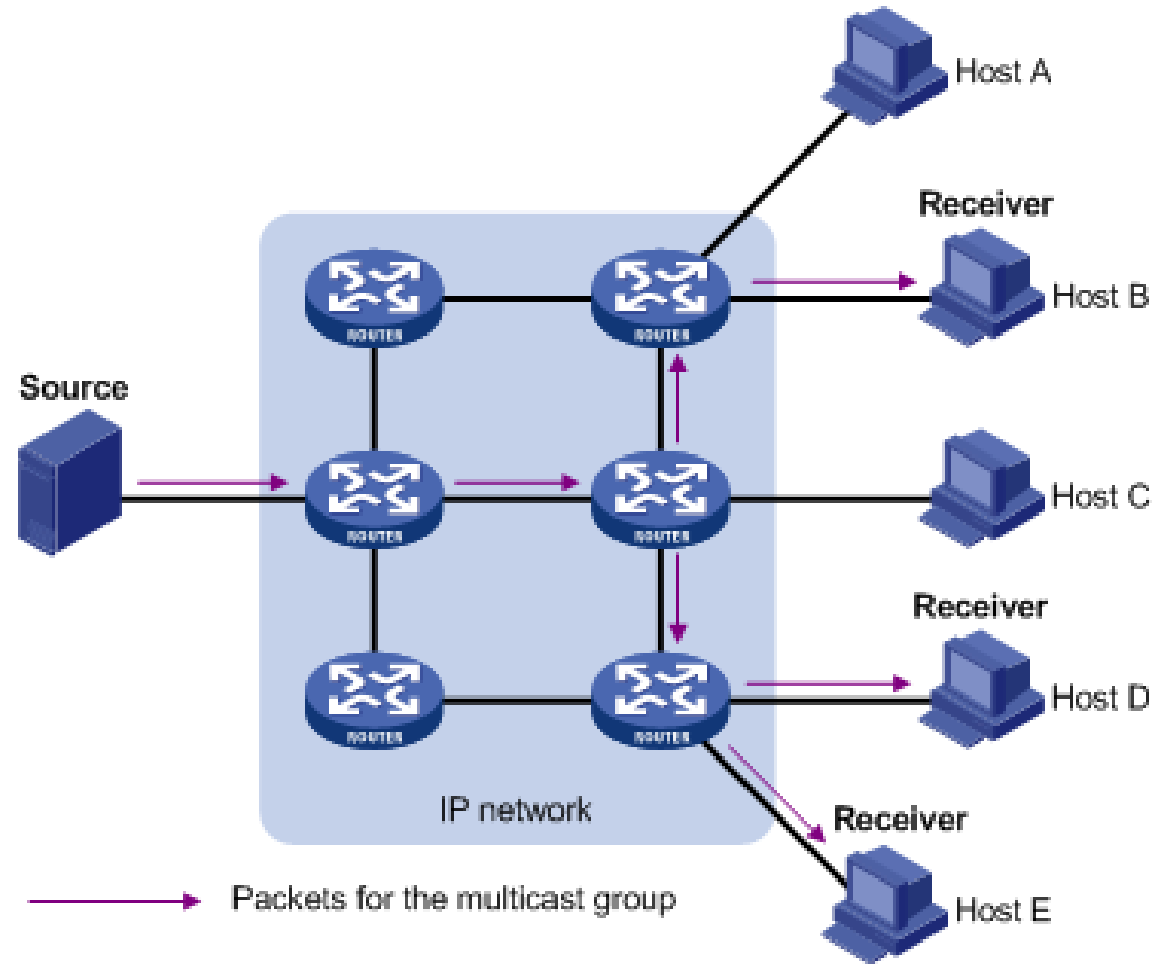
Групповая адресация (Multicast)

- В случае одноадресной рассылки IP-адрес получателя содержится в каждой одноадресной IP-дейтаграмме и означает единственного получателя.
- в архитектуре Интернета групповая дейтаграмма **адресуется косвенно**. То есть для группы получателей используется один идентификатор, а копия дейтаграммы, адресованной группе получателей при помощи этого единственного идентификатора, доставляется **всем членам этой группы**.
- В Интернете единый идентификатор, соответствующий группе получателей, представляет собой групповой адрес класса D

Групповая маршрутизация

- Протоколы групповой маршрутизации распределяют групповой трафик между узлами сети точно так же, как протоколы типа RIP и OSPF управляют маршрутизацией потоков данных в режиме одноадресной передачи.
- Доминирующими протоколами являются Distance Vector Multicast Routing Protocol (DVMRP), Multicast Open Shortest Path First (MOSPF) и Protocol-Independent Multicast (PIM).

Групповая маршрутизация

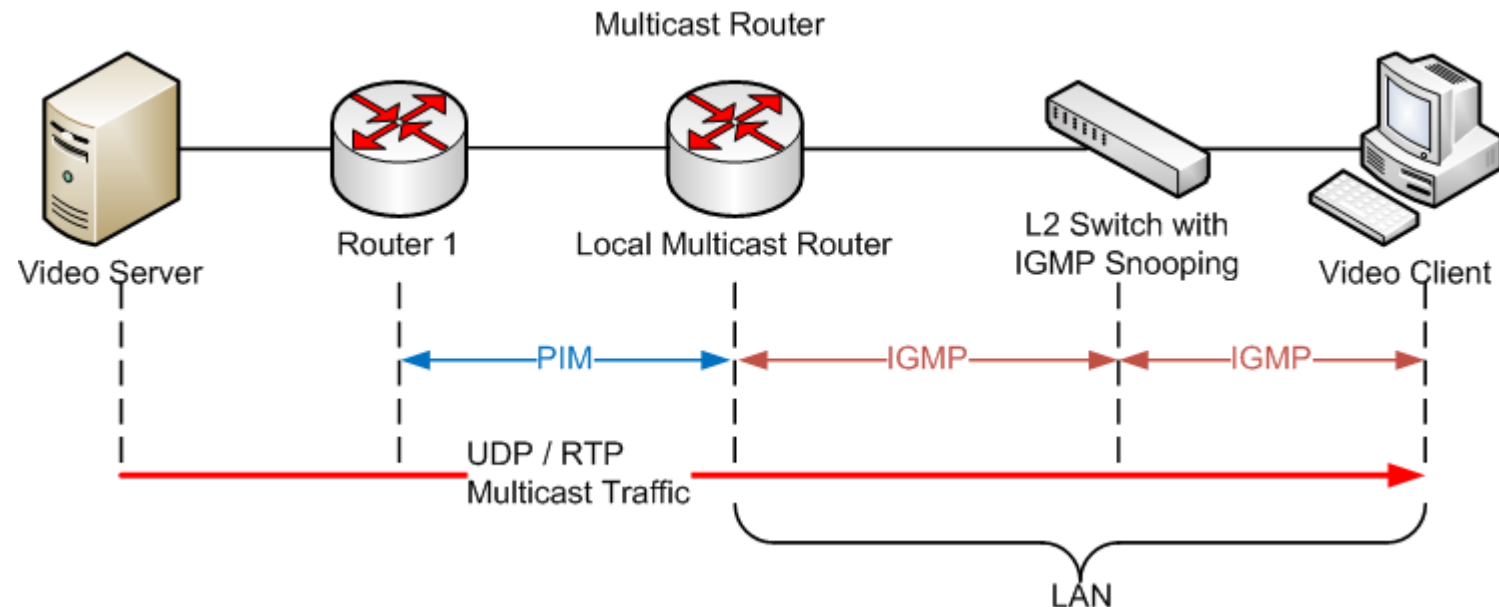


Групповая маршрутизация

- Групповые маршрутизаторы используют эти протоколы для динамической генерации "деревьев" маршрутов от отправителя (передатчика) ко всем приемникам.
- По меньшей мере один из этих протоколов должен поддерживаться используемыми маршрутизаторами.
- Большинство фирм-поставщиков поддерживают один или два протокола, причем каждый из них имеет свои плюсы и минусы.

Протокол IGMP [RFC 3376]

- Протокол управления групповой (multicast) передачей данных в сетях, основанных на протоколе IP.
- IGMP используется маршрутизаторами и IP-узлами для организации сетевых устройств в группы.



Протокол IGMP

- Чтобы стать получателем групповых данных, узел должен «выразить» свою заинтересованность маршрутизатору, к которому непосредственно подсоединена его сеть.
- Для этого хост должен установить взаимодействие с маршрутизатором по протоколу IGMP. Версия IGMP для хоста прямо зависит от типа операционной системы хоста
- Для функции опроса членов группы один из маршрутизаторов локальной сети выбирается доминирующим. Доминирующий маршрутизатор средствами протокола IGMP периодически опрашивает все системы (групповой адрес 224.0.0.1) в непосредственно присоединенных к нему подсетях, проверяя, активны ли члены всех известных ему групп.

Формат IGMP v1 сообщения

0	3	4	7	8	15	16	31
4 бита IGMP версия(1)		4 бита IGMP тип(1-2)		(не используется)		16-бит контрольная сумма	
32 бита адрес группы (ip адрес класса D)							

Формат полей IGMP сообщения.

Формат IGMP v1 сообщения

- Поле **версии** IGMP (IGMP version) установлено в 1.
- Поле **тип** IGMP (IGMP type) устанавливается в 1 для запроса, посылаемого групповым маршрутизатором, и в 2 для ответа, отправляемого хостом.
- **Групповой адрес** (group address) это IP адрес класса D. В запросе поле группового адреса устанавливается в 0. В отчете оно содержит групповой адрес.

Формат IGMP v2 сообщения

- Все типы IGMP-сообщений имеют длину 8 байт и состоят из четырех полей.
- В зависимости от версии протокола IGMP назначение полей может несколько меняться.

1-4 байты	Тип сообщения	Максимальное время ответа	Контрольная сумма
5-8 байты	Адрес группового вещания		

Формат IGMP v2 сообщения

- **Тип** — поле определяет тип сообщения, размер 8 бит. Типы сообщений в IGMPv2:
 - **Membership Query** (значение поля = 0x11) — используется маршрутизаторами для того чтобы определить наличие хостов, которые хотят получать групповую рассылку:
 - **General Membership Query** устанавливает значение поля Group Address в 0.0.0.0.
 - **Group-Specific Query** устанавливает значение поля Group Address равным адресу группы в которой отправляется запрос. Это сообщение отправляет маршрутизатор после того как он получает от хоста сообщение Leave Group. Оно используется для того чтобы определить остались ли хосты в этой группе.
 - **Version 1 Membership Report** (значение поля = 0x12) — используется хостами IGMPv2 для обратной совместимости с IGMPv1.
 - **Version 2 Membership Report** (значение поля = 0x16) — отправляется членом группы для того чтобы сообщить маршрутизатору, что в группе есть как минимум один хост.
 - **Leave Group** (значение поля = 0x17) — отправляется членом группы для того чтобы сообщить маршрутизатору, что хост покидает группу.

Формат IGMP v2 сообщения

- **Максимальное время ответа** — это поле используется только в сообщениях Query, 8 бит. Единицы измерения 1/10 секунды, по умолчанию значение поля 100 (10 секунд). Диапазон значений от 1 до 255 (0.1 до 25.5 секунд).
- **Контрольная сумма** — контрольная сумма вычисленная отправителем, 16 бит. Вычисляется контрольная сумма всего IP-пакета без учета заголовка.
- **Групповой адрес** — устанавливается равным 0.0.0.0 в сообщениях General Query и равным адресу групповой рассылки в сообщениях Group-Specific. В сообщениях Membership Report это группа трафик которой хост хочет получать, а в сообщениях Leave Group это адрес группы которую хост хочет покинуть.

Типы сообщений IGMP

- **Запрос о членстве (membership query).**
 - С помощью этого сообщения маршрутизатор пытается узнать, в каких группах состоят хосты в локальной сети, присоединенной к какому-либо его интерфейсу. (Общий запрос обо всех группах, или запрос о некоторой конкретной группе с помощью ее адреса)
- **Отчет о членстве (membership report).**
 - Этим сообщением хосты отвечают маршрутизатору, который послал в сеть запрос о членстве. В сообщении содержится информация об адресе группы, в которой они состоят. (Важен только факт наличия хостов в группе)
- **Покинуть группу (leave group).**
 - Это сообщение хост может использовать, чтобы сигнализировать «своему» маршрутизатору о желании покинуть некоторую группу, в которой он до этого состоял. После этого маршрутизатор попытается узнать, есть ли еще хосты в группе, если нет, данные для группы больше не передаются

Принципы группового вещания

- Маршрутизация на основе доменов
- Учет плотности получателей группового трафика
- Два подхода к построению маршрутного дерева
- Концепция продвижения по реверсивному пути

Маршрутизация на основе доменов

- Значительный объем хранимой и передаваемой по сети служебной информации, используемой для поддержания группового вещания, стал фактором, ограничивающим масштабируемость данной технологии.
- Для улучшения масштабируемости разработчики технологии группового вещания предложили традиционный для Интернета иерархический подход, основанный на доменах.
- Подобно автономным системам (доменам маршрутизации) и DNS-доменам вводятся домены группового вещания.

Учет плотности

- **Протоколы плотного режима** (Dense Mode, DM) разработаны в предположении, что в сетевом домене существует большое число принимающих узлов. Протокол затапливает сеть пакетами группового вещания по всем направлениям, останавливая продвижение пакетов, лишь когда находящийся на пути распространения трафика маршрутизатор явно сообщит, что ниже по потоку членов данной группы нет
- **Протоколы разряженного режима** (Sparse Mode, SM) рассчитаны на работу в сети, в которой количество маршрутизаторов с подключенными к ним членами групп невелико по сравнению с общим числом маршрутизаторов. Выгоднее не усекать некоторые пути распространения широковещательной рассылки, а использовать явные сообщения о необходимости присоединения подсетей к дереву рассылки

Точка рандеву (Rendezvous Point)

- Центральный элемент в системе разреженных протоколов
- Точка встречи должна существовать для каждой имеющейся в сети группы и быть единственной для группы.
- Все узлы, заинтересованные в получении информации, предназначенной той или иной группе, должны регистрироваться в соответствующей точке встречи.
- Функции точки (или нескольких точек) встречи выполняет специально назначенный для этого маршрутизатор. В сети может быть несколько маршрутизаторов, играющих роли точек встречи.

Подходы к построению маршрутного дерева

- Для всех источников данной группы строится единственный граф связей, называемый **разделяемым деревом**. Этот граф связывает всех членов данной группы (точнее, все маршрутизаторы, к которым подключены локальные сети, имеющие в своем составе членов данной группы).
- Shared tree принято обозначать **(*, G)**:
 - * — обозначает то, что дерево строится не к конкретному источнику, а к общей точке (RP),
 - G — адрес multicast группы.

Подходы к построению маршрутного дерева

- Для каждой группы строятся несколько графов по числу источников, вещающих на каждую из этих групп. Каждый такой граф, называемый деревом с вершиной в источнике, служит для доставки трафика всем членам группы, но только от одного источника.
- Source tree принято обозначать **(S, G)**:
 - S — это IP-адрес источника трафика,
 - G — адрес multicast группы.

Продвижение по реверсивному пути

- Проверка, которую маршрутизатор выполняет, чтобы понять, является ли входной интерфейс, получивший групповой пакет, интерфейсом, через который пролегает кратчайший путь к источнику, с помощью обычной таблицы маршрутизации, которая, как известно, содержит указания о рациональных путях ко всем сетям составной интерсети.
- Название объясняется тем, что эта процедура связана не столько с путями, ведущими вперед от текущего места нахождения пакета к пункту назначения, сколько с обратным (реверсивным) путем, который **уже пройден пакетом от того места, где он находится сейчас, до источника.**

Протокол DVMRP [RFC 1075]

- У протокола DVMRP самая длинная история существования по сравнению с остальными.
- Когда маршрутизатор получает групповой пакет, протокол DVMRP строит дистрибутивное дерево, направляя поток пакетов ко всем интерфейсам сети, кроме уже получившего соответствующий входящий пакет.
- При этом пакеты доходят до всех связанных подсетей. Далее, не нуждающиеся в получении конкретной групповой услуги маршрутизаторы (так как ни один из хостов подсети не направил запрос на участие в соответствующей группе через IGMP) могут послать ответное "отсекающее" сообщение (prune) для исключения ненужного трафика.
- Определение маршрутизаторов, размещенных на пути от приемника к передатчику, осуществляется аналогично тому, как это делается в протоколе RIP, и, как следствие, протокол DVMRP наследует некоторые недостатки протокола RIP.

Дейтаграмма DVMRP

+	0 — 3	4 — 7	8 — 15	16 — 23	24 — 31
0	Версия	Тип	Код	Контрольная сумма	
3 2	Тело DVMRP сообщения				

Дейтаграмма DVMRP

- **Версия** — текущее значение 1;
- **Тип** — должно быть 3;
- **Код** — определяет назначение пакета DVMRP:
 - 1 = Response — сообщение обеспечивает маршрутизацию для нескольких групп;
 - 2 = Request — сообщение запрашивает маршрутизацию для нескольких групп;
 - 3 = Non-membership report — сообщение обеспечивает доклад о неучастии в группах;
 - 4 = Non-membership cancellation — сообщение отменяет предыдущий доклад о неучастии в группах;
- **Контрольная сумма** — рассчитывается при передаче пакета и проверяется при получении; при расчете контрольной суммы, поле контрольной суммы принимается за ноль;
- **Тело сообщения** — набор команд, выровненных по границе 16 бит:
 - **код команды** (8 бит);
 - **данные** (минимум 8 бит).

MOSPF [RFC 1584]

- Multicast extensions to OSPF — расширения протокола OSPF для группового вещания, опирается на обычные механизмы OSPF для поддержки группового вещания.
- Маршрутизаторы добавляют к информации о состоянии связей, распространяемой по протоколу OSPF, данные о членстве в группах узлов в непосредственно присоединенных сетях.
- Эти данные рассылаются по сети в дополнительном сообщении о членстве в группе (**group membership**). В результате помимо топологии связей, маршрутизаторам становится известно о наличии членов каждой из групп в каждой подсети области.
- На основании этой информации маршрутизатор находит дерево кратчайших путей для каждой группы. Это позволяет распространять групповые пакеты не широковещательно, а по кратчайшим путям от источника до подсетей, в которых есть активные члены группы.

MOSPF [RFC 1584]

- Для получения данных о том, в какие группы входят конечные узлы в связанных с ним подсетях, MOSPF-маршрутизатор использует IGMP.
- При каждом подключении узла к группе или исключении узла из группы маршрутизатор рассылает по сети новое сообщение о членстве в группе
- Кроме того, известные положительные свойства протокола OSPF — устойчивое поведение при изменениях топологии сети, меньшие объемы служебного трафика по сравнению с протоколом RIP, а также возможность деления сети на области — полностью наследуются протоколом MOSPF

PIM (Protocol Independent Multicast)

- Группа протоколов, которые занимаются маршрутизацией мультикаст. И, хотя некоторые основы работы протоколов из этой группы одинаковы, каждый конкретный протокол работает по-разному.
- "Protocol independent" в названии протокола означает, что PIM может работать с unicast таблицей маршрутизации, независимо от того, как именно она заполнена.
- То есть, он может использовать, например, маршруты OSPF, EIGRP, статические маршруты и др.
- Кроме того, сам PIM не передает информацию о маршрутах, а строит дерево для передачи мультикаст трафика на основе таблицы маршрутизации unicast.

Стандарты

- **PIM-SM** [RFC 7761] (2016) / [RFC 4601] (2006)
- **PIM-DM** [RFC 3973] (2005)

Использованные источники

- В. Олифер, Н. Олифер “Компьютерные сети. Принципы, технологии, протоколы”
- Д. Куроуз, К. Росс “Компьютерные сети. Нисходящий подход.”
- <https://ru.wikipedia.org/wiki/IGMP>
- https://ru.wikipedia.org/wiki/Distance_Vector_Multicast_Routing_Protocol