

Как выбрать наилучший путь коммутации маршрутизатора для вашей сети

Содержание

[Введение](#)

[Коммутация в контексте процесса](#)

[Коммутация в контексте прерывания](#)

[Быстрая коммутация](#)

[Оптимальная коммутация](#)

[Технология Cisco Express Forwarding](#)

[Какой путь коммутации следует выбрать?](#)

[Дополнительные сведения](#)

Введение

Различные маршрутизаторы Cisco и выпуски Cisco IOS предусматривают большое многообразие путей коммутации. Какой самый лучший для вашей сети, и как все они работают? Это описание технологических решений является попыткой объяснить каждый из следующих путей коммутации так, чтобы можно было принять лучшее решение о том, какой путь коммутации подходит вашей сети.

Вначале проверьте сам процесс пересылки. Пересылка пакетов через маршрутизатор выполняется в три этапа:

1. Определяется доступность получателя пакета.
2. Определяется следующий переход на маршруте к получателю и интерфейс, через который доступен этот следующий переход.
3. Выполняется перезапись заголовка управления доступом к среде (MAC) в пакете так, чтобы пакет успешно достиг следующего узла.

Каждый из этих шагов очень важен для достижения пакетом его места назначения.

Примечание: Всюду по этому документу Путь IP-коммутации используется в качестве примера; практически вся приведенная информация является применимой к эквивалентным путям коммутации для других протоколов, если они существуют.

Коммутация в контексте процесса

Коммутация в контексте процесса – наименьший общий знаменатель в путях коммутации; она доступна во всех версиях IOS, на всех платформах и для всех типов коммутируемого трафика. Коммутация в контексте процесса строится на двух основных понятиях:

- Решение о пересылке и информация, используемая для перезаписи заголовка MAC

пакета, берутся из таблицы маршрутизации (из базы данных маршрутизации, или RIB) и кэша протокола разрешения адресов (ARP), либо из другой таблицы, содержащей данные заголовка MAC, сопоставленные с адресом IP каждого хоста, подключенного напрямую к маршрутизатору.

- Коммутация пакетов осуществляется обычным процессом, выполняемым в операционной системе IOS. Другими словами, решение о пересылке принимается процессом, время для которого выделяется планировщиком IOS и который имеет равные полномочия с другими процессами в маршрутизаторе, такими как протоколы маршрутизации. Процессы, выполняемые на маршрутизаторе в штатном режиме, не прерываются для коммутации пакета в контексте процесса.

Путь коммутации в контексте процесса проиллюстрирован на следующем рисунке.

Рассмотрим эту схему более подробно:

1. Интерфейсный процессор сначала обнаруживает наличие пакета в сетевой среде, а затем передает этот пакет в память ввода/вывода на маршрутизаторе.
2. Интерфейсный процессор порождает прерывание приема. Во время этого прерывания центральный процессор определяет тип пакета (предположим, что пакет является пакетом IP), и копирует его в память процессора в случае необходимости (это решение не зависит от платформы). Наконец, процессор размещает пакет во входной очереди соответствующего процесса, и прерывание снимается.
3. При следующем запуске планировщик записывает пакет во входящей очереди `ip_input` и планирует запуск этого процесса.
4. При выполнении процесса `ip_input` этот процесс сверяется с маршрутной базой данных (RIB) для определения следующего перехода и выходного интерфейса, а затем обращается к кэшу ARP, чтобы определить действительный адрес физического уровня для этого следующего перехода.
5. далее процесс `ip_input` перезаписывает заголовок MAC пакета и помещает пакет в исходящую очередь соответствующего исходящего интерфейса.
6. Пакет копируется из выходной очереди вывода исходящего интерфейса в очередь передачи исходящего интерфейса; между этими двумя очередями применяется выходной механизм качества обслуживания (QoS), если таковой настроен.
7. Выходной интерфейсный процессор обнаруживает пакет в своей очереди передачи и передает его в сетевую среду.

Почти все функциональные возможности, связанные с коммутацией пакетов, например преобразование сетевых адресов (NAT) и маршрутизация на основе политик, первоначально появляются в составе пути коммутации в контексте процесса. [После их апробирования и оптимизации эти функциональные возможности становятся доступны, но не всегда, для коммутации в контексте прерывания.](#)

[Коммутация в контексте прерывания](#)

Коммутация в контексте прерывания – второй из основных методов коммутации, используемых маршрутизаторами Cisco. Основные различия между коммутацией в контексте прерывания и процесса состоят в следующем:

- Процесс, выполняющийся на процессоре в данный момент, прерывается для коммутации пакета. **Коммутация пакетов выполняется по запросу, а не только в**

интервалы времени, выделяемые процессу ip_input.

- Для отыскания необходимых данных, позволяющих выполнить маршрутизацию пакета, процессор использует определенную форму кэша маршрутов.

Следующий рисунок иллюстрирует коммутацию в контексте прерывания:

Рассмотрим эту схему более подробно:

1. Интерфейсный процессор сначала обнаруживает наличие пакета в сетевой среде, а затем передает этот пакет в память ввода/вывода на маршрутизаторе.
2. Интерфейсный процессор порождает прерывание приема. Во время этого прерывания центральный процессор определяет тип пакета (предположим, что это пакет IP) и затем начинает коммутацию пакета.
3. Процессор выполняет поиск в кэше маршрутов, проверяя доступность получателя пакета и определяя требуемый выходной интерфейс, следующий переход к этому получателю и, наконец, требуемый заголовок MAC для доставки пакета на следующий переход. Руководствуясь этими сведениями, процессор перезаписывает заголовок MAC пакета.
4. Затем пакет копируется либо в очередь передачи, либо в выходную очередь исходящего интерфейса (в зависимости от различных факторов). После этого происходит выход из прерывания по приему пакета, и процесс, который выполнялся на процессоре до возникновения прерывания, продолжает выполняться.
5. Выходной интерфейс обнаруживает пакет в своей очереди передачи и передает его в сетевую среду.

Первый вопрос, который может возникнуть после чтения этого описания: «Что находится в кэше?»? Есть три возможных ответа, в зависимости от конкретного вида коммутации в контексте прерывания:

- [Быстрая коммутация](#)
- [Оптимальная коммутация](#)
- [Технология Cisco Express Forwarding](#)

Быстрая коммутация

При быстрой коммутации используется двоичное дерево для хранения сведений о пересылке и перезаписи заголовка MAC. Двоичное дерево обеспечивает быстрое отыскание и сверку данных. Двоичное дерево показано на следующем рисунке:

В случае быстрой коммутации факт достигаемости подтверждается существованием узла в двоичном дереве для получателя пакета. Заголовок MAC и исходящий интерфейс для каждого получателя сохраняются в составе сведений об узле в дереве. Двоичное дерево может фактически иметь 32 уровня (приведенное выше дерево предельно упрощено для наглядности).

Для поиска в двоичном дереве нужно просто, двигаясь от левой цифры (старшего разряда) двоичного числа, переходить на левую или правую ветвь дерева в зависимости от конкретного числа. Например, если требуется отыскать сведения, связанные с числом 4 в этом дереве, то сначала нужно перейти на правую ветвь, потому что первая двоичная цифра – 1. Далее нужно следовать по дереву вниз, сравнивая следующую цифру в двоичном числе, пока не будет достигнут конец.

Характеристики быстрой коммутации

Быстрая коммутация обладает рядом характеристик, которые обусловлены структурой двоичного дерева и хранением данных о перезаписи заголовка MAC в составе узлов дерева.

- Поскольку корреляция количества между таблицей маршрутизации и содержанием кэша быстрой коммутации (например, данными о перезаписи заголовка MAC) отсутствует, то при формировании записей кэша выполняется вся работа, которая должна быть выполнена в пути коммутации в контексте процесса. Поэтому записи кэша быстрой коммутации формируются при коммутации пакетов в контексте процесса.
- Вследствие отсутствия корреляции между заголовками MAC (используемыми для перезаписи) в кэше ARP и структурой кэша быстрой коммутации, при изменении таблицы ARP часть кэша быстрой коммутации должна стать недействительной (и должна быть создана повторно при коммутации в контексте процесса).
- Кэш быстрой коммутации может создавать записи на одном уровне (длиной в один префикс) для определенного места назначения в таблице маршрутизации.
- Кэш быстрой коммутации не позволяет напрямую от одной записи перейти к другой (предполагается, что заголовок MAC и информация об исходящем интерфейсе находятся в пределах узла), поэтому все рекурсивные маршруты должны быть устранены на этапе формирования записи кэша. Другими словами, рекурсивные маршруты не могут быть разрешены непосредственно в пределах кэша быстрой коммутации.

Устаревание записей быстрой коммутации

Для того, чтобы записи быстрой коммутации не оказались рассинхронизированы с таблицей маршрутизации и кэшем ARP, а неиспользуемые записи в кэше не расходовали память маршрутизатора, каждую минуту из кэша произвольно удаляется 1/20 его часть. Когда объем свободной памяти маршрутизатора опускается ниже критически низкого уровня, каждую минуту удаляется 1/5 записей кэша.

Длина префикса при быстрой коммутации

Какой длины префикс используется средством скоростной коммутации при построении записей, если для каждого пункта назначения допустимо только одно значение длины? В терминах быстрой коммутации получатель – это одиночный доступный получатель из таблицы маршрутизации или основная сеть. Для определения длин префиксов при построении отдельных записей кэша действуют следующие правила:

- При создании записи кэширования для скоростной политики кэшируются сети вплоть до /32.
- В случае создания записи для мультипротокольного доступа по виртуальному каналу ATM (MPOA VC) всегда кэшируется /32.
- Если сеть не разделена на подсети (запись основной сети): В случае прямого подключения используется сеть /32; В противном случае используется маска основной сети.
- В случае суперсети используется маска суперсети.
- Если сеть разделена на подсети: В случае прямого подключения используется сеть

/32; Если к данной подсети ведет несколько маршрутов, используется сеть /32; В этой основной сети во всех остальных случаях используется наибольшая длина префикса.

Распределение нагрузки

Быстрая коммутация целиком привязана к получателям; распределение нагрузки выполняется в масштабе отдельных получателей. При наличии нескольких путей равной стоимости для конкретной сети получателя кэш быстрой коммутации содержит одну запись для каждого хоста, доступного в этой сети, но весь трафик, адресованный конкретному хосту, проходит по одному и тому же соединению.

Оптимальная коммутация

В случае оптимальной коммутации сведения о пересылке и перезаписи заголовка MAC хранятся в разветвленном дереве (mtree) с 256 ветвями. Использование дерева типа mtree сокращает количество шагов, которые необходимо предпринять для отыскания префикса, как показано на следующем рисунке.

Каждый октет используется для определения конкретной ветви из 256 ветвей дерева, таким образом для отыскания любого получателя требуется не более 4 операций поиска. При меньшей длине префикса может быть достаточно одной-трех операций поиска. Сведения о перезаписи заголовка MAC и выходного интерфейса хранятся в составе узла дерева, поэтому аннулирование и устаревание записей кэша происходят таким же образом, как и при быстрой коммутации.

Оптимальная коммутация определяет длину префикса для каждой записи кэша таким же путем, как и быстрая коммутация.

Технология Cisco Express Forwarding

Технология быстрой коммутации Cisco Express Forwarding также использует структуру с 256 элементами для хранения информации о пересылке и заголовках MAC, но не использует древовидные структуры. Технология Cisco Express Forwarding использует TRIE-структуру. Это означает, что сами искомые сведения отсутствуют в структуре данных, вместо этого данные запоминаются в отдельной структуре, а TRIE-структура просто указывает на нее. Другими словами, вместо хранения информации об исходящем интерфейсе и перезаписи заголовка MAC непосредственно в самом дереве технология Cisco Express Forwarding хранит эту информацию в отдельной структуре данных, называемой таблицей смежности.

Это разделение сведений о доступности (в таблице Cisco Express Forwarding) и сведений о перенаправлении (в таблице соседей) предоставляет следующие преимущества:

- Можно построить таблицу смежности отдельно от таблицы экспресс пересылки Cisco, что позволяет создавать обе таблицы без коммутации в контексте процесса.
- Перезапись MAC-заголовка, используемая для передачи пакета, не сохранена в записях в кэше, таким образом, изменения в строке перезаписи заголовка MAC не требуют аннулирования записей в кэше.
- Можно указать непосредственно к информации перенаправления, а не к рекурсивно вызванному следующему переходу, для решения рекурсивных маршрутов.

Фактически принцип устаревания кэша полностью устраняется, и кэш заранее формируется

на основе сведений, содержащихся в кэше ARP и таблице маршрутизации. Для формирования записей кэша не требуется коммутировать какие-либо пакеты в контексте процесса.

Другие записи в таблице смежности

Таблица смежности может содержать иные записи, отличные от строк перезаписи заголовка MAC и сведений об исходящем интерфейсе. В числе различных типов записей, которые могут размещаться в таблице смежности, можно отметить следующие:

- **cache** – строка перезаписи заголовка MAC в кэше и исходящий интерфейс, используемые для достижения определенного смежного узла или маршрутизатора.
- **receive** – пакеты, предназначенные для этого IP-адреса, должны приниматься маршрутизатором. К подобным адресам относятся широковещательные адреса и адреса, настроенные непосредственно на маршрутизаторе.
- **drop** – пакеты, направляемые на этот IP-адрес, подлежат удалению. Этот параметр может использоваться для трафика, отклоняемого списком контроля доступа, или направляемого на интерфейс NULL.
- **punt** – технология Cisco Express Forwarding не может обработать этот пакет; обработка производится следующим по приоритету методом коммутации (обычно – быстрая коммутация).
- **glean** – следующий переход подключен непосредственно, но в данный момент недоступны строки перезаписи заголовка MAC.

Смежности glean

Запись смежности glean указывает, что конкретный узел следующего перехода должен быть подключен непосредственно, но информация о перезаписи заголовка MAC недоступна. Каким образом формируются и используются эти сведения? Маршрутизатор под управлением Cisco Express Forwarding, подключенный к широковещательной сети, как показано на рисунке, по умолчанию создает несколько записей в таблице смежности.

По умолчанию создаются следующие четыре записи в таблице смежности:

```
10.1.1.0/24, version 17, attached, connected
0 packets, 0 bytes
  via Ethernet2/0, 0 dependencies
    valid glean adjacency
10.1.1.0/32, version 4, receive
10.1.1.1/32, version 3, receive
10.1.1.255/32, version 5, receive
```

Обратите внимание, что записей четыре: три записи типа receive и одна запись типа glean. Каждая запись типа receive представляет широковещательный адрес или адрес, настроенный на маршрутизаторе, а запись типа receive представляет остальную часть пространства адресов в подключенной сети. В случае приема пакета для хоста 10.1.1.50 маршрутизатор пытается выполнить его коммутацию и обнаруживает, что происходит разрешение пакета на данную смежность glean. Затем Cisco Express Forwarding сигнализирует о необходимости записи кэша ARP для 10.1.1.50, процесс ARP посылает пакет ARP, и соответствующая запись таблицы смежности создается на основе новых сведений из кэша ARP. После завершения этого этапа в таблице смежности будет присутствовать запись для адреса 10.1.1.50.

```
10.1.1.0/24, version 17, attached, connected
0 packets, 0 bytes
  via Ethernet2/0, 0 dependencies
    valid glean adjacency
10.1.1.0/32, version 4, receive
10.1.1.1/32, version 3, receive
10.1.1.50/32, version 12, cached adjacency 208.0.3.2
0 packets, 0 bytes
  via 208.0.3.2, Ethernet2/0, 1 dependency
    next hop 208.0.3.2, Ethernet2/0
    valid cached adjacency
10.1.1.255/32, version 5, receive
```

Следующий полученный маршрутизатором пакет, предназначенный для 10.1.1.50, коммутируется через эту новую смежность.

Распределение нагрузки

Технология Cisco Express Forwarding также использует в своих интересах разделение между таблицей Cisco Express Forwarding и таблицей смежности, что позволяет реализовать более эффективную форму распределения нагрузки по сравнению с другими режимами коммутации в контексте прерывания. Таблица распределения нагрузки вставляется между таблицей Cisco Express Forwarding и таблицей смежности, как показано на рисунке:

Таблица Cisco Express Forwarding указывает на эту таблицу распределения нагрузки, в которой, в свою очередь, содержатся указатели на записи таблицы смежности для доступных параллельных маршрутов. IP-адреса источника и получателя пропускаются через алгоритм хэширования для определения конкретной записи таблицы распределения нагрузки, используемой для каждого пакета. Может быть настроено разделение нагрузки на уровне отдельных пакетов. В этом случае каждый пакет использует отдельную запись из таблицы распределения нагрузки.

Каждая таблица распределения нагрузки содержит 16 записей, между которыми делятся доступные пути в зависимости от счетчика доли трафика в таблице маршрутизации. Если счетчики долей трафика в таблице маршрутизации все равны 1 (как в случае нескольких путей с равной стоимостью), то каждый следующий возможный переход получает равное количество указателей из таблицы распределения нагрузки. Если число доступных путей не будет делиться на 16 без остатка (с учетом того, что таблица распределения нагрузки содержит 16 записей), то некоторые пути будут иметь больше записей, чем другие.

Начиная с выпуска программного обеспечения Cisco IOS 12.0, количество записей в таблице распределения нагрузки уменьшается таким образом, чтобы каждый путь имел пропорциональное количество записей в таблице распределения нагрузки. Например, если в таблице маршрутизации имеются три пути с равной стоимостью, то используются только 15 записей из таблицы распределения нагрузки.

Какой путь коммутации следует выбрать?

По возможности для маршрутизаторов следует использовать коммутацию в контексте прерывания, поскольку этот режим как минимум на порядок быстрее, чем коммутация в контексте процесса. Коммутация Cisco Express Forwarding определенно быстрее и лучше любого другого режима коммутации. Мы рекомендуем использовать Cisco Express Forwarding, если эту технологию поддерживают используемый протокол и версия IOS. Это в

особенности актуально при наличии нескольких параллельных соединений, по которым трафик должен передаваться с распределением нагрузки. [На странице навигатора по функциям Cisco \(доступной только зарегистрированным заказчикам\) можно определить требуемую версию IOS для поддержки CEF.](#)

Дополнительные сведения

- [Проверка коммутации Cisco Express Forwarding](#)
- [Устранение неполадок при выравнивании загрузки по параллельным соединениям с использованием Cisco Express Forwarding](#)
- [Распределение нагрузки с CEF](#)
- [Страница поддержки технологии Cisco Express Forwarding](#)
- [Страница поддержки IP-маршрутизации](#)
- [Руководство по настройке служб коммутации Cisco IOS, выпуск 12.1](#)
- [Техническая поддержка - Cisco Systems](#)