Network Working Group
Request for Comments: 4023
Category: Standards Track

T. Worster
Motorola, Inc.
Y. Rekhter
Juniper Networks, Inc.
E. Rosen, Ed.
Cisco Systems, Inc.

March 2005

Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)

Инкапсуляция MPLS в IP или GRE

Статус документа

В этом документе приведена спецификация проекта стандартного протокола Internet. Документ служит приглашением к дискуссии в целях развития и совершенствования протокола. Текущий статус стандартизации протокола можно узнать из текущей версии документа Internet Official Protocol Standards (STD 1). Документ может распространяться без ограничений.

Авторские права

Аннотация

Различные приложения MPLS используют стеки меток со множеством элементов. В некоторых случаях можно заменить верхнюю метку стека инкапсуляцией на базе IP, что позволяет приложению работать в сетях без поддержки MPLS на маршрутизаторах ядра. Этот документ определяет два варианта инкапсуляции на базе IP - MPLS-in-IP и MPLS-in-GRE[1]. Каждый из этих вариантов применим в определённых ситуациях.

Оглавление

1. Предпосылки

Во многих приложениях MPLS пакеты, проходящие через магистраль MPLS, содержат стек со множеством меток. Как указано в параграфе 3.15 [RFC3031], каждая метка представляет путь LSP[2]. Для каждого LSP имеются маршрутизаторы LSR[3], являющиеся входным (LSP Ingress) и выходным (LSP Egress). Если маршрутизаторы A и B являются входным и выходным (соответственно) для LSP, соответствующего верхней метке в пакете, A и B являются смежными LSR на пути LSP, соответствующем второй метке (метке, непосредственно под верхней).

Назначение (или одна из целей) верхней метки состоит в том, чтобы при доставке пакета от A к B маршрутизатор B мог продолжить обработку на основе второй метки. В этом смысле верхняя метка служит заголовком инкапсуляции для остального пакета. Иногда вместо верхней метки могут применяться другие заголовки инкапсуляции без потери функциональности. Например, вместо верхней метки может присутствовать заголовок IP или GRE. Поскольку инкапсулированный пакет остаётся пакетом MPLS, результатом будет инкапсуляция MPLS-in-IP или MPLS-in-GRE.

При такой инкапсуляции два LSR, будучи смежными на LSP, могут быть разделены сетью IP, даже если эта сеть не поддерживает MPLS.

Для применения любого из указанных выше вариантов инкапсулирующий маршрутизатор LSR должен знать:

- IP-адрес декапсулирующего LSR;

[1]Generic Routing Encapsulation - базовая инкапсуляция маршрутных данных.
[2]Label Switched Path - путь с коммутацией по меткам.
[3]Label Switching Router - маршрутизатор с коммутацией по меткам.

- реальность поддержки конкретной инкапсуляции на декапсулирующем LSR.

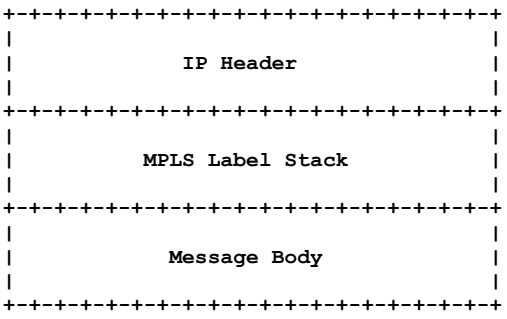
Эта информация может быть передана инкапсулирующему LSR путём ручной настройки или с помощью того или иного протокола обнаружения. В частности, при использовании туннеля для поддержки конкретного приложения, имеющего протокол настройки или обнаружения, этот протокол может предоставить нужную информацию. Способы передачи такой информации выходят за рамки документа.

2. Уровни требований

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не нужно** (SHALL NOT), **следует** (SHOULD), **не следует** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **не рекомендуется** (NOT RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе должны интерпретироваться в соответствии с [RFC2119].

3. Инкапсуляция в IP

Формат сообщений MPLS, инкапсулированных в IP, показан ниже.



IP Header

Это поле содержит заголовок дейтаграммы IPv4 или IPv6 в соответствии с [RFC791] или [RFC2460]. В качестве адресов отправителя и получателя указываются адреса инкапсулирующего и декапсулирующего LSR, соответственно.

MPLS Label Stack

Это поле содержит стек меток MPLS в соответствии с определением [RFC3032].

Message Body

Это поле содержит тело одного сообщения MPLS.

В поле IPv4 Protocol Number или IPv6 Next Header помещается значение 137, указывающее индивидуальный (unicast) пакет MPLS. Применение инкапсуляции MPLS в IP для групповых пакетов MPLS эта спецификация не поддерживает.

После заголовка IP помещается пакет MPLS, как указано в [RFC3032]. Эта инкапсуляция вызывает передачу пакетов MPLS через «туннели IP». Когда конечная точка туннеля принимает пакет, она декапсулирует пакет MPLS, удаляя заголовок IP. Затем пакет обрабатывается как пакет MPLS, в котором «входной меткой» [RFC3031] является верхняя метка декапсулированного пакета.

4. Инкапсуляция в GRE

Инкапсуляция MPLS в GRE помещает пакеты MPLS в пакеты GRE [RFC2784]. Пакет состоит из заголовка IP (IPv4 или IPv6), за которым следует заголовок GRE, а затем стек меток MPLS в соответствии с [RFC3032]. В поле типа протокола заголовка GRE **должно** указываться значение Ethertype для MPLS Unicast (0x8847) или Multicast (0x8848).

Эта инкапсуляция ведёт к передаче пакетов MPLS через «туннели GRE». Когда конечная точка туннеля принимает пакет, она декапсулирует пакет MPLS, удаляя заголовки IP и GRE. После этого принятый пакет MPLS обрабатывается с использованием в качестве «входной» [RFC3031] верхней метки декапсулированного пакета.

[RFC2784] задаёт необязательную контрольную сумму GRE, а [RFC2890] - необязательные поля ключа GRE и порядкового номера. Эти поля не очень полезны для инкапсуляции MPLS в GRE. Поля порядкового номера и контрольной суммы не нужны по причине отсутствия соответствующих полей в туннелируемых пакетах MPLS. Поле GRE key не требуется для демультиплексирования, поскольку для этого служит верхняя метка MPLS инкапсулированного пакета. Иногда поле ключа GRE рассматривают как средство защиты, позволяющее передавать в открытом виде 32-битовый пароль, однако такая защиты очень слаба. Для (a) упрощения высокоскоростных реализаций и (b) обеспечения взаимодействия мы требуем от всех реализаций способность корректно работать без этих необязательных полей.

Точнее говоря, реализации декапсуляторов MPLS-in-GRE **должны** быть способны корректно обрабатывать пакеты без необязательных полей. Они **могут** также корректно обрабатывать пакеты с этими необязательными полями.

Реализации инкапсуляторов MPLS-in-GRE **должны** быть способные генерировать пакеты без этих дополнительных полей. Они **могут** поддерживать генерацию пакетов с такими полями, но по умолчанию пакеты **должны** создаваться без этих необязательных полей. Инкапсуляторам недопустимо включать любое из этих полей, пока он не уверен в корректности его обработки декапсулятором. Методы согласования этого выходят за рамки спецификации.

5. Общие процедуры

Некоторые процедуры являются общими для инкапсуляции MPLS-in-IP и MPLS-in-GRE. Далее инкапсулятор, чей адрес IP появляется в поле отправителя в заголовке IP, называется «началом туннеля» (tunnel head). Декапсулятор, чей адрес указан в поле получателя в декапсулируемом заголовке IP, называется «концом туннеля» (tunnel tail).

При использовании IPv6 (для MPLS-in-IPv6 или MPLS-in-GRE-in-IPv6) процедуры [RFC2473] остаются применимыми.

5.1. Предотвращение фрагментации и сборки

Если пакет MPLS-in-IP или MPLS-in-GRE фрагментируется («обычная» фрагментация IP), конечная точка туннеля должна собрать его до того, как станет возможной декапсуляция MPLS. Когда туннель заканчивается на маршрутизаторе, сборка явно не желательна, поскольку у маршрутизатора может не быть ресурсов для сборки с требуемой производительностью.

Возможность фрагментации туннелируемых пакетов **должна** быть настраиваемой в начале туннеля. По умолчанию фрагментирование пакетов **должно** быть запрещено. Принятое по умолчанию значение можно менять лишь при наличии уверенности в адекватной сборке фрагментов конечной точкой туннеля.

Процедуры, описанные в оставшейся части параграфа, применимы лишь к пакетам, которые не будут фрагментированы.

Обычно, если пакет не был фрагментирован, начальной точке туннеля **недопустимо** фрагментировать пакет перед его инкапсуляцией.

При использовании IPv4 для туннеля **должен** быть установлен бит DF. Это будет препятствовать фрагментированию пакетов на промежуточных узлах туннеля (при использовании IPv6 промежуточные узлы никогда не будут фрагментировать пакеты).

В начальной точке **следует** выполнить процедуру Path MTU Discovery ([RFC1191] для IPv4 или [RFC1981] для IPv6).

В начале туннеля **должно** поддерживаться значение Tunnel MTU для каждого имеющегося туннеля. Это меньшее из двух значений - (a) заданное административно и (b) Path MTU за вычетом издержек инкапсуляции.

Если начальная точка туннеля получает для инкапсуляции пакет MPLS, размер которого превосходит Tunnel MTU, такой пакет **должен** отбрасываться. Однако отбрасывание таких пакетов без уведомления отправителей может создать существенные проблемы в работе, поскольку создатель пакета заметит, что пакет не прошёл, но может не понять, что причиной этого послужил избыточный размер пакета. В результате он может продолжить передачу пакетов, которые будут отбрасываться. Механизм определения Path MTU может помочь (если туннель будет возвращать ошибки ICMP), но зачастую в начальной точке туннеля нет достаточной информации для определения исходного отправителя. Чтобы свести проблемы к минимуму, предлагается делать MTU достаточно большими, предотвращая фрагментацию на практике.

В некоторых случаях начальная точка туннеля будет получать для инкапсуляции пакеты IP, которые были сначала инкапсулированы в MPLS, а затем в MPLS-in-IP или MPLS-in-GRE. Если отправитель пакета IP доступен из начальной точки туннеля и в результате пакета в MPLS его размер превышает Tunnel MTU, значением, которое этой точке **следует** применять для фрагментации и определения PMTU за пределами туннеля, будет значением Tunnel MTU за вычетом размера инкапсуляции MPLS (т. е. Tunnel MTU минус размер инкапсуляции MPLS будет определять значение MTU, передаваемое в сообщениях). Пакет будет отбрасываться, но начальной точке туннеля следует указать IP-адрес его источника в подходящем сообщении ICMP, как указано в [RFC1191] или [RFC1981].

5.2. TTL или Hop Limit

Начальная точка туннеля **может** помещать значение TTL из стека меток MPLS в поле TTL заголовка инкапсуляции IPv4 или поле Hop Limit инкапсулирующего заголовка IPv6. В конце туннеля **можно** помещать значение TTL из инкапсулирующего заголовка IPv4 или Hop Limit из заголовка IPv6 в поле TTL заголовка MPLS, если это не будет увеличивать значение TTL в заголовке MPLS.

Возможность копирования значений и способы этого зависят от конфигурации конечных точек туннеля.

5.3. Дифференцированные услуги

Описанные в документе процедуры позволяют организовать LSP через туннель IP или GRE. В [RFC2983] подробно рассмотрены многочисленные вопросы и процедуры, связанные с поддержкой архитектуры дифференцированных услуг при наличии туннелей IP-in-IP. Эти соображения и процедуры применимы к туннелям MPLS-in-IP и MPLS-in-GRE.

Соответственно, при сборке пакета MPLS в MPLS-in-IP или MPLS-in-GRE начальной точкой туннеля установка поля DS в инкапсулирующем заголовке IPv4 или IPv6 **может** определяться (по крайней мере частично) «агрегатом поведения» пакета MPLS. Процедуры определения Behavior Aggregate для пакетов MPLS заданы в [RFC3270].

Аналогично, в конечной точке туннеля поле DS в заголовке инкапсуляции IPv4 или IPv6 **может** определять Behavior Aggregate для инкапсулированного пакета MPLS. В [RFC3270] указаны связи между агрегатом поведения и последующим размещением пакета.

6. Применимость

Инкапсуляция MPLS-in-IP более эффективна и при прочих равных условиях обычно считается предпочтительной. Однако в некоторых ситуациях может применяться инкапсуляция MPLS-in-GRE, как указано ниже.

- Два маршрутизатора являются «смежными» через туннель GRE, организованный по не связанным с этим документом причинам, и эти маршрутизаторы передают пакеты MPLS через такой туннель. Для всех отправляемых в туннель пакетов применяется инкапсуляция GRE и вариант MPLS-in-GRE будет более предпочтительным, поскольку инкапсуляция MPLS-in-IP все равно будет инкапсулироваться в GRE.
- Особенности реализации могут требовать применения MPLS-in-GRE. Например, то или иное устройство может оказаться способным обрабатывать лишь инкапсуляцию GRE на своём «быстром пути» (fastpath).

7. Взаимодействие с IANA

Агентство IANA выделило значение 137 в качестве IP Protocol Number для инкапсуляции MPLS-in-IP, как указано в разделе 3. Других действий от IANA не требуется. Для инкапсуляции MPLS-in-GRE не требуется действий IANA.

8. Вопросы безопасности

Основной проблемой безопасности при использовании туннелей IP или GRE является возможность получения конечной точкой туннеля пакетов, которые представляются исходящими из туннеля, но на деле отправлены не другой конечной точкой этого туннеля (указанные варианты инкапсуляции сами по себе не обеспечивают декапсулятору возможность проверить подлинность инкапсулятора). Другая проблема связана с возможностью изменения пакета на пути от входа в туннель до выхода из него. Третья проблема заключается в возможности просмотра содержимого пакета при прохождении через туннель (оба варианта инкапсуляции не обеспечивают конфиденциальности). Практическая значимость этих проблем зависит от требований безопасности приложений, трафик которых передаётся через туннель. Например, отсутствие конфиденциальности не будет иметь большого значения для туннелей, через которые передаётся открытая информация.

Из-за различных требований безопасности, вариантов развёртывания и вопросов производительности приложений, использующих описанный механизм инкапсуляции, в этой спецификации поддержка IPsec указана **необязательной**. Основные требования к реализациям, использующим IPsec, приведены в параграфе 8.1. Если IPsec не используется, могут потребоваться дополнительные механизмы защиты. Этот вопрос рассматривается в параграфе 8.2.

8.1. Защита туннелей с помощью IPsec

Всех упомянутых проблем безопасности можно избежать, если туннель MPLS-in-IP или MPLS-in-GRE будет защищён с помощью IPsec. Приведённые ниже требования к реализации применимы для случаев использования IPsec.

При использовании IPsec начало и конец туннеля следует считать конечными точками защищённой связи SA[1]. Для этого один адрес IP начальной точки туннеля будет служить IP-адресом отправителя, а один адрес IP конечной точки туннеля - IP-адресом получателя. Способы, используемые для получения информации об используемом другой стороной туннеля адресе выходят за рамки этого документа. Если для организации туннелей используется протокол управления (например, для информирования оконечной точки туннеля о IP-адресе другой стороны), этот протокол **должен** иметь механизм проверки подлинности и этот механизм **должен** применяться при организации туннеля. Если туннель организован автоматически, например, с помощью распространяемой по протоколу BGP информации, использования механизма аутентификации BGP на основе MD5 будет достаточно.

Пакеты с инкапсуляцией MPLS-in-IP или MPLS-in-GRE следует считать исходящими из начальной точки туннеля и адресованными конечной точке. **Следует** применять транспортный режим IPsec.

Заголовок IP пакета MPLS-in-IP становится внешним заголовком IP результирующего пакета, когда начальная точка туннеля использует транспортный режим IPsec для защиты пакетов MPLS-in-IP. За ним следует заголовок IPsec, а затем - стек меток MPLS. В заголовке IPsec устанавливается тип данных MPLS путём указания номера протокола IP, заданного в разделе 3. Если транспортный режим IPsec применяется для пакетов MPLS-in-GRE, заголовок GRE следует после заголовка IPsec.

В конечной точке туннеля выходная обработка IPsec восстанавливает инкапсулированный пакет MPLS-in-IP/GRE. Затем конечная точка вырезает заголовок IP/GRE для восстановления пакета MPLS, который пересылается в соответствии со стеком меток.

Отметим, что начало и конец туннеля являются смежными на LSP и это означает, что верхняя метка в любом пакете, переданном через туннель, должна быть получена начальной точкой туннеля от его конечной точки. Конечная точка **должна** достоверно знать, какие метки она распространила начальным точкам защищённых с помощью IPsec туннелей. Метки из этого набора конечной точке **недопустимо** распространять через другие смежности LSP. Если полученный без инкапсуляции IPsec пакет MPLS имеет метку из такого набора, пакет **должен** быть отброшен.

Защищённые с помощью IPsec туннели MPLS-in-IP и MPLS-in-GRE **должны** обеспечивать аутентификацию и целостность (отметим, что проверка подлинности и защита целостности применяются ко всему пакету MPLS, включая стек меток). Поэтому реализация **должна** поддерживать ESP с null (пустым) шифрованием. **Может** поддерживаться ESP с реальным шифрованием, если требуется защита конфиденциальности. При использовании ESP конечная точка туннеля **должна** проверять принадлежность IP-адреса всех пакетов, принятых через SA, к числу ожидаемых.

Распространение ключей может выполняться вручную или автоматически, с помощью IKE [RFC2409]. **Должна поддерживаться установка ключей вручную**. При реализации автоматического распространения ключей **должен** поддерживаться в качестве основного механизм IKE с заранее известными ключами. Конкретные приложения могут усиливать требования и запрашивать автоматическое распространение ключей.

Ручное распространение ключей значительно проще, но слабо расширяемо по сравнению с автоматическим. Поэтому выбор подходящего для конкретного приложения метода распространения ключей должен быть внимательно рассмотрен администратором (или парой администраторов), ответственным за конечные точки туннеля. Если требуется защита от повторного использования пакетов (replay) для определённого туннеля, следует настроить автоматическое распространение ключей.

При использовании инкапсуляции MPLS-in-IP селекторами, связанными с SA, будут адреса отправителя и получателя, упомянутые выше, а также номер протокола IP, заданный в разделе 3. Если нужно защитить множество туннелей MPLS-in-IP между данной парой узлов по отдельности, для каждого туннеля должна применяться уникальная пара адресов.

При использовании инкапсуляции MPLS-in-GRE селекторами, связанными с SA, будут адреса отправителя и получателя, упомянутые выше, а также номер протокола IP, заданный для GRE (47). Если нужно защитить множество туннелей MPLS-in-GRE между данной парой узлов по отдельности, для каждого туннеля должна применяться уникальная пара адресов.

8.2. Отсутствие IPsec

Если туннели не защищены с помощью IPsec, следует применять какой-либо иной метод, обеспечивающий декапсуляцию и пересылку конечной точкой туннеля лишь тех пакетов, которые были инкапсулированы в начале

[1]Security Association.

туннеля. Если туннель полностью находится в одном административном домене, можно использовать фильтрацию адресов на границе, чтобы пакеты с адресом отправителя или адресом получателя, относящимся к конечным точкам туннеля, не могли приходить в домен извне.

Однако при размещении конечных точек туннеля в разных административных доменах ситуация усложняется и фильтрация по адресам получателей может стать невозможной при прохождении пакетов через Internet.

Иногда на границе административного домена выполняется фильтрация лишь по адресам отправителей (без фильтрации по адресам получателей). В таких случаях фильтрация совсем не обеспечивает защиты, если декапсулятор не проверяет IP-адрес отправителя в пакете MPLS-in-IP или MPLS-in-GRE. Этот документ не требует от декапсулятора проверки IP-адреса отправителя в туннелируемом пакете, но следует понимать, что отказываться от такой проверки можно лишь при наличии эффективной фильтрации на границах по адресам получателя (или получателя и отправителя).

9. Благодарности

Эта спецификация объединяет раннюю работу по инкапсуляции MPLS в IP, выполненную Tom Worster, Paul Doolan, Yasuhiro Katsube, Tom K. Johnson, Andrew G. Malis и Rick Wilder, с ранней работой по инкапсуляции MPLS в GRE, выполненной Yakov Rekhter, Daniel Tappan и Eric Rosen. Авторы благодарят своих предшественников за их вклад.

Множество людей, включая Rahul Aggarwal, Scott Bradner, Alex Conta, Mark Duffy, Francois Le Feucheur, Allison Mankin, Thomas Narten, Pekka Savola, Alex Zinin, представило полезные замечания и поправки к работе.

10. Нормативные документы

[RFC791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, September 1981.

[RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.

[RFC1981] McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC2463] Conta, A. and S. Deering, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 2463, December 1998.

[RFC2473] Conta, A. and S. Deering, "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

11. Дополнительная литература

[RFC2401] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

[RFC2402] Kent, S. and R. Atkinson, "IP Authentication Header", RFC 2402, November 1998.

[RFC2406] Kent, S. and R. Atkinson, "IP Encapsulating Security Payload (ESP)", RFC 2406, November 1998.

[RFC2409] Harkins, D. and D. Carrel, "The Internet Key Exchange (IKE)", RFC 2409, November 1998.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Service", RFC 2475, December 1998.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE", RFC 2890, September 2000.

[RFC2983] Black, D., "Differentiated Services and Tunnels", RFC 2983, October 2000.

[RFC3260] Grossman, D., "New Terminology and Clarifications for Diffserv", RFC 3260, April 2002.

[RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, May 2002.

Адреса авторов

Tom Worster

Motorola, Inc.

120 Turnpike Road

Southborough, MA 01772

EMail: tom.worster@motorola.com

Yakov Rekhter

Juniper Networks, Inc.

1194 N. Mathilda Ave.

Sunnyvale, CA 94089

EMail: yakov@juniper.net

Eric Rosen

Cisco Systems, Inc.

1414 Massachusetts Avenue

Boxborough, MA 01719

EMail: erosen@cisco.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.