Network Working Group
Request for Comments: 1853
Category: Informational

W. Simpson
Daydreamer

October 1995

IP in IP Tunneling

Туннелирование IP in IP

Статус документа

Этот документ содержит сведения для сообщества Internet и не задаёт каких-либо стандартов. Документ можно распространять без ограничений.

Замечание IESG

Этот документ является индивидуальной работой автора и отражает текущую информационную практику Internet. В рабочей группе IETF Mobile-IP предпринимаются усилия по разработке подходящего стандарта для решения этой задачи.

Аннотация

В документе обсуждаются методы реализации для применения инкапсуляции пакетов с IP Protocol 4 для туннелирования с IP Security и другими протоколами.

Оглавление

1. Ведение

Инкапсуляция IP в IP Protocol 4 [RFC-1700] уже давно примееняется для мостов между частями Internet с несвязанными возможностями или правилами. Этот документ описывает методы реализации, применяемые много лет в сети Amateur Packet Radio для объединения большой мобильной сети, а также для реализации протоколов IP Security.

Применение инкапсуляции IP в IP отличается от более поздних методов туннелирования (например, с протоколом 98 [RFC-1241], 94 [IDM91a], 53 [swIPe], 47 [RFC-1701]) тем, что она не добавляет своего связывающего заголовка между двумя заголовкаим IP. Вместо этого сохраняется исходный заголовок IP и просто помещается после другого стандартного заголовка IP.

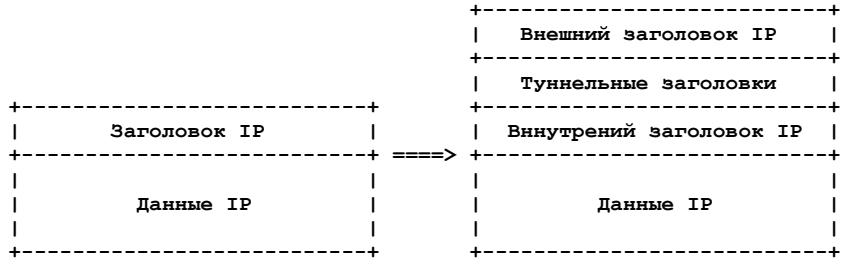
Эта информация относится в основном к инкапсуляции IP версии 4. Другие версии IP будут описаны в отдельных документах.

2. Инкапсуляция

Метод инкапсуляции довольно прост. Внешний заголовок IP помещается перед исходным заголовком IP. Между ними могут находиться другие заголовки для пути, такие как заголовки защиты, связанные с конфигурацией туннеля.

Во внешнем заголовке IP поля Source и Destination указывают конечные точки туннеля, во внутреннем - исходного отправителя и получателя дейтаграммы.

Каждый заголовок связывается с последующим значением поля IP Protocol [RFC-1700].



Формат заголовков IP описан в [RFC-791].

Type Of Service

Копируется из внутреннего заголовка IP. Возможно согласование своего значения TOS между партнерами.

Это согласуется с принципом прозрачности - если пользователь ожидает определённого уровня обслуживания, туннелю следует обеспечивать такой же уровень. Однако некоторые туннели могут создаваться специально для обеспечения другого уровня обслуживания в соответствии с политикой.

Identification

Новый номер, генерируемый для каждого внешнего заголовка IP.

Инкапсулированная дейтаграмма может уже быть фрагментированной, а туннельная инкапсуляция может привести к новому фрагментированию. Такие туннельные фрагменты будет собирать декапсулятор, а не конечный получатель.

Reserved

Игнорируется (должен быть 0).

Этот неофициальный флаг оставлен для экспериментов и не влияет на туннель, хотя и сохраняется во внутреннем заголовке IP.

Don't Fragment

Копируется из внутреннего заголовка IP. Это позволяет отправителю контролировать уровень компромиссов производительности (см. 3.1. Определение MTU в туннеле).

More Fragments

Устанавливается должным образом при фрагментировании.

Флаг не копируется по той же причине, что использование отдельного поля Identification.

Time To Live

Принятое по умолчанию значение в соответствии с последним документом Assigned Numbers [RFC-1700]. Это гарантирует, что неожиданно длинные туннели не прервут поток дейтаграмм между конечными точками.

Внутреннее значение TTL декрементируется перед инкапсуляцией и не меняется при декапсуляции.

Protocol

Следующий заголовок - 4 для внутреннего заголовка IP при отсутствии промежуточных туннельных заголовков.

Source

Адрес IP, связанный с интерфейсом, служащим для передачи дейтаграмм.

Destination

IP-адрес декапсулятора туннеля.

Options

Не копируются из внутреннего заголовка IP, однако **могут** добавляться новые опции, относящиеся к пути.

Опции Timestamp, Loose Source Route, Strict Source Route, Record Route намеренно скрыты внутри туннеля. Туннели часто создаются для преодоления связанной с этими опциями неадекватности.

Любые поддерживаемые разновидности опций безопасности внутреннего заголовка IP **могут** влиять на выбор опций безопасности для туннеля. Не предполагается однозначного отображения таких опций на опции или заголовки безопасности, выбранные для туннеля.

3. Управление туннелем

Возможно возникновение на одном из внутренних маршрутизаторов теннуля связанной с обработкой дейтаграммы ошибки, вызывающей отправку сообщения ICMP об ошибке [RFC-792] инкапсулятору по полю IP Source для туннеля. К сожалению ICMP требует от маршрутизаторов IP возврата лишь 8 байтов (64 бита) дейтаграммы после заголовка IP, что недостаточно для включения инкапсулированного заголовка. Поэтому инкапсулирующий маршрутизатор обычно не может передать сообщение ICMP из туннеля хосту-отправителю. Однако при аккуратной поддержке soft state для своих туннелей инкапсулятор в большинстве случаев может возвращать точные сообщения ICMP. Маршрутизатору **следует** поддерживать для каждого туннеля хотя бы следующий сведения:

- доступность удалённого конца туннеля;
- перегрузка в туннеле;
- MTU для туннеля.

Маршрутизатор использует получаемые от туннеля сообщения ICMP для обновления данных soft state этого туннеля. По прибытии следующих дейтаграмм для передачи через туннель маршрутизатор проверяте состояние туннеля. Если дейтаграмма будет нарушать это состояние (например, MTU больше, чем MTU в туннеле и установлен флаг Don't Fragment), маршрутизатор передаст отправителю подходящее сообщение ICMP об ошибке, но перешлёт дейтаграмму в туннель. Пересылка дейтаграммы, несмотря на возврат сообщения об ошибке обеспечивает возможность узнать состояние туннеля.

При использовании этого метода сообщения ICMP о ошибках от инкапсулирующих маршрутизаторов не всегда будут однозначно сопоставляться с ошибками в туннеле, но достаточно точно отразят состояние сети.

3.1. Определение MTU в туннеле

Когда бит Don't Fragment установлен источником и копируется во внешний заголовок IP, корректное значение MTU в туннеле будут определено из сообщений ICMP Datagram Too Big (тип 3, код 4), возвращаемых инкапсулятору. Для поддержки хостов-источников, использующих такую возможность, все реализации **должны** поддерживать Path MTU Discovery [RFC-1191, RFC-1435] внутри своих туннелей.

Преимущество Tunnel MTU Discovery состоит в том, что любая фрагментация связанная с заголовком инкапсуляции, будет выполняться лишь один рах после инкапсуляции. Это предотвращает многократную фрагментация и повышает эффективность обработки на маршрутизаторах пути и декапсуляторе туннеля.

3.2. Перегрузки

В soft state для туннеля будет включаться индикация перегрузки, такая как ICMP Source Quench (тип 4) в дейтаграммах от декапсулятора (партнёр по туннелю). При пересылке другой дейтаграммы в туннель уместно передавать инициатору сообщения Source Quench.

3.3. Отказы маршрутизации

Поскольку TTL «сбрасывается» при инкапсуляции дейтаграммы, петли внутри туннеля особенно опасны, когда они возвращают пакеты инкапсулятору. Если IP Source соответствует любому из интерфейсов инкапсулятора, реализации **недопустимо** инкапсулировать дейтаграмму ещё раз, её нужно просто переслать.

Сообщения ICMP Time Exceeded (тип 11) сообщают о маршрутной петле внутри самого туннеля, ICMP Destination Unreachable (тип 3) говорят декапсулятору о причинах ошибки. Это состояние (soft state) **должно** сообщаться инициатору как Network Unreachable (тип 3, код 0).

3.4. Другие сообщения ICMP

Большинство сообщений ICMP об ошибках не связано с использованием туннеля. В частности, сообщения о проблемах с параметрами могут говорить о некорректной настройке инкапсулятора и их **недопустимо** передавать инициатору.

Вопросы безопасности

Вопросы безопасности кратко рассматриваются в этом документе. Применение туннелей может исключить некоторые старые опции безопасности IP (маркировка), но лучше поддерживать более новые заголовки IP Security.

Литература

[IDM91a] Ioannidis, J., Duchamp, D., Maguire, G., "IP-based protocols for mobile internetworking", Proceedings of SIGCOMM '91, ACM, September 1991.

[RFC-791] Postel, J., "Internet Protocol", STD 5, RFC 791, USC/Information Sciences Institute, September 1981.

[RFC-792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, USC/Information Sciences Institute, September 1981.

[RFC-1191] Mogul, J., and S. Deering, "Path MTU Discovery", RFC 1191, DECWRL, Stanford University, November 1990.

[RFC-1241] Mills, D., and R. Woodburn, "A Scheme for an Internet Encapsulation Protocol: Version 1", UDEL, July 1991.

[RFC-1435] Knowles, S., "IESG Advice from Experience with Path MTU Discovery", RFC 1435, FTP Software, March 1993.

[RFC-1700] Reynolds, J., and J. Postel, "Assigned Numbers", STD 2, RFC 1700, USC/Information Sciences Institute, October 1994.

[RFC-1701] Hanks, S., Li, T., Farinacci, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 1701, October 1994.

[swIPe] Ioannidis, J., and Blaze, M., "The Architecture and Implementation of Network-Layer Security Under Unix", Fourth Usenix Security Symposium Proceedings, October 1993.

Благодарности

Эти детали реализации туннелирования IP в значительной мере получены из независимой работы Phil Karn (1990 г.) и радиолюбителей TCP-Group, использовавших KA9Q NOS.

Особая благодарность John Ioannidis (тогда из Columbia University) за стимулирование и эксперименты, положившие начало финальному этапу разработки IP Mobility и IP Security. Часть текста заимствовна из [IDM91a] и [swIPe].

Связывание заголовков было описано Steve Deering (Xerox PARC) в работе Simple Internet Protocol.

Общая организация и часть текста заимствованы из [RFC-1241] от David Mills (U Delaware) и Robert Woodburn (SAIC).

Часть текста о состоянии (soft state) туннелей заимствована из IP Address Encapsulation (IPAE) от Robert E. Gilligan, Erik Nordmark, Bob Hinden (все из Sun Microsystems).

Адрес автора

Связанные с документом вопросы можно направлять автору.

William Allen Simpson
Daydreamer
Computer Systems Consulting Services
1384 Fontaine
Madison Heights, Michigan 48071
Bill.Simpson@um.cc.umich.edu
bsimpson@MorningStar.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru