

Network Working Group  
Request for Comments: 4605  
Category: Standards Track

B. Fenner  
AT&T Research  
H. He  
Nortel  
B. Haberman  
JHU-APL  
H. Sandick  
Little River Elementary School  
August 2006

## Пересылка групповых пакетов на основе IGMP/MLD Proxy Internet Group Management Protocol (IGMP) / Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")

### Статус документа

В этом документе содержится спецификация стандарта для протокола, предложенного сообществу Internet. Документ служит приглашением к дискуссии в целях развития и совершенствования протокола. Текущее состояние стандартизации протокола вы можете узнать из документа Internet Official Protocol Standards (STD 1). Документ может распространяться без ограничений.

### Авторские права

Copyright (C) The Internet Society (2006).

### Тезисы

В некоторых вариантах топологии нет потребности в использовании протокола групповой маршрутизации. Промежуточному устройству достаточно получить данные о принадлежности к группе и просто пересылать групповые пакеты на основе этой информации. В данном документе описан механизм рассылки на основе лишь информации IGMP<sup>1</sup> или MLD<sup>2</sup>.

## 1. Введение

Этот документ использует групповую маршрутизацию на базе остовного дерева [MCAST] для сред IGMP или MLD. Топология среды ограничена деревом, поскольку не задается протокола для построения дерева в более сложной топологии. Предполагается, что корень дерева соединен с более широкой инфраструктурой групповой адресации.

### 1.1. Мотивация

В простой древовидной топологии не требуется использовать протокол групповой маршрутизации. Промежуточным устройствам достаточно получить информацию о принадлежности к группам и просто использовать ее для пересылки групповых пакетов. Одним из типичных примеров такой топологии может служить краевое агрегирующее устройство типа мультиплексора DSL (DSLAM<sup>3</sup>). В большинстве вариантов применения такое устройство имеет единственное соединение с опорной сетью и множество соединений с абонентами.

Использование пересылки на основе IGMP/MLD для репликации группового трафика на краевых устройствах можно значительно упростить устройство и реализацию такого оборудования. За счет отказа от поддержки сложных протоколов групповой маршрутизации типа PIM<sup>4</sup> или DVMRP<sup>5</sup> снижается не только стоимость устройств, но и операционные расходы. Другим преимуществом является независимость промежуточных устройств (проxy device) от протоколов групповой маршрутизации, используемых в ядре сети. Следовательно, такие устройства-посредники можно легко развернуть в любой сети с групповой передачей.

Отказоустойчивость краевых устройств обычно обеспечивается за счет «горячего» резервирования соединения с ядром сети. При отказе основного соединения такое устройство просто переключается на резервный канал. Пересылка на основе IGMP/MLD может выиграть от такого решения, используя избыточное соединение для резервирования связи с multicast-маршрутизаторами. При переключении краевого устройства на резервный канал восходящее проxy-соединение также будет работать через этот канал.

<sup>1</sup>Internet Group Management Protocol — протокол управления группами.

<sup>2</sup>Multicast Listener Discovery — обнаружение групповых получателей.

<sup>3</sup>Digital Subscriber Line Access Multiplexer.

<sup>4</sup>Protocol Independent Multicast — независимая от протокола групповая пересылка.

<sup>5</sup>Distance Vector Multicast Routing Protocol — протокол групповой маршрутизации на основе «дистантного» вектора.

## 1.2. Заявление о применимости

Пересылка на основе IGMP/MLD работает только в топологии простого дерева. Дерево настраивается вручную путем указания восходящих и нисходящих интерфейсов на каждом промежуточном устройстве. В дополнение к этому применяемую схему адресации IP **следует** настроить так, чтобы устройство-посредник могло быть выбранным в качестве IGMP/MLD Querier для того, чтобы иметь возможность пересылки группового трафика. Кроме промежуточных устройств не используется каких-либо multicast-маршрутизаторов и предполагается, что корень дерева соединен с более широкой инфраструктурой групповой передачи. Работа данного протокола ограничена одним административным доменом.

В более сложной топологии, где желателен более высокий уровень отказоустойчивости и поддержка работы в множестве административных доменов, следует применять протоколы групповой маршрутизации.

## 1.3. Соглашения

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не следует** (SHALL NOT), **следует** (SHOULD), **не нужно** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с RFC 2119 [RFC2119].

Этот документ является результатом работы группы MAGMA<sup>1</sup> в рамках IETF<sup>2</sup>. Комментарии приветствуются и их следует направлять по адресу списка рассылок группы [magma@ietf.org](mailto:magma@ietf.org) и/или непосредственно авторам.

## 2. Определения

### 2.1. Восходящий интерфейс

Интерфейс промежуточного устройства (проху device) в направлении корня дерева. Используется также термин «хост-интерфейс» (Host interface).

### 2.2. Нисходящий интерфейс

Каждый из интерфейсов промежуточного устройства, который не направлен в сторону корня дерева. Используется также термин «интерфейс маршрутизатора» (Router interface).

### 2.3. Групповой режим

В среде IPv4 multicast-группа относится к IGMP версии 1 (IGMPv1) [RFC1112], если она использует отчеты IGMPv1, к IGMP версии 2 (IGMPv2) [RFC2236] при использовании отчетов IGMPv2, но не IGMPv1, и к IGMP версии 3 (IGMPv3) [RFC3376], если она слышит отчеты IGMPv3, но не слышит IGMPv1 и IGMPv2.

В среде IPv6 multicast-группа относится к MLD версии 1 (MLDv1) [RFC2710], если в ней используются отчеты MLDv1. MLDv1 является эквивалентом IGMPv2. Группа относится к MLD версии 2 (MLDv2) [MLDv2], если слышны отчеты MLDv2, но не слышны MLDv1. MLDv2 является эквивалентом IGMPv3.

### 2.4. Подписка

Когда группа работает в режиме IGMPv1 или IGMPv2/MLDv1, подпиской на группу является принадлежность интерфейса к данной группе. Для групп с режимом IGMPv3/MLDv2 подпиской является элемент состояния IGMPv3/MLDv2, т. е., неупорядоченный список (групповой адрес, групповой таймер, режим фильтрации, список источников) на интерфейсе.

### 2.5. База данных о принадлежности

Поддерживаемая на каждом промежуточном устройстве база данных, где хранится информация о принадлежности к группам для каждого нисходящего интерфейса устройства. Эта база представляет собой набор записей вида:

*(групповой адрес, режим фильтрации, список источников)*

Определения полей «режим фильтрации» (filter-mode) и «список источников» (source-list) приведены в спецификациях IGMPv3/MLDv2 [RFC3376, MLDv2]. Операционное поведение базы данных описано в параграфе 4.1.

## 3. Абстрактное определение протокола

Промежуточное устройство с пересылкой на основе IGMP/MLD имеет один восходящий и один или множество нисходящих интерфейсов. Роли интерфейсов явно задаются в конфигурации устройства — протокола, определяющего тип интерфейса, не предусмотрено. Устройство выполняет связанную с маршрутизацией часть протокола IGMP [RFC1112, RFC2236, RFC3376] или MLD [RFC2710, MLDv2] на своих исходящих интерфейсах и связанную с хостом часть IGMP/MLD — на восходящем интерфейсе. Промежуточному устройству **недопустимо** выполнять связанную с маршрутизацией часть IGMP/MLD на его восходящем интерфейсе.

Промежуточное устройство поддерживает базу данных, указывающую все подписки на всех нисходящих интерфейсах. Устройство и поддержка базы данных о принадлежности к группам описаны в разделе 4.

Промежуточное устройство передает отчеты о принадлежности IGMP/MLD на восходящем интерфейсе по запросам, а также незапрошенные отчеты при изменении базы данных.

Когда промежуточное устройство получает пакет, адресованный в multicast-группу (канал в SSM<sup>3</sup>), оно использует список, состоящий из его восходящего интерфейса и всех входящих интерфейсов, связанных с подпиской на относящейся к этому пакету, для которых имеется статус IGMP/MLD Querier. Этот список может создаваться

<sup>1</sup>Multicast & Anycast Group Membership — принадлежность к группам Multicast и Anycast.

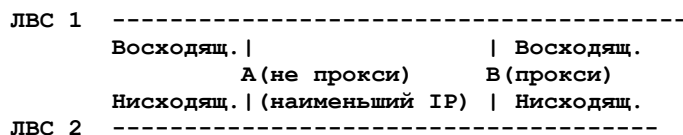
<sup>2</sup>Internet Engineering Task Force.

<sup>3</sup>Source-Specific Multicast — связанная с источником групповая адресация.

динамически или кэшироваться. Устройство исключает из списка принявший пакет интерфейс и пересылает этот пакет во все оставшиеся в списке интерфейсы (возможно, и в восходящий).

Отметим, что необходимость иметь статус запрашивающего (querier) на прокси-устройстве для пересылки пакетов вносит ограничения в схему используемой адресации IP — в частности, пересылающие на базе IGMP/MLD устройства должны иметь адрес, который меньше IP-адреса любого потенциального IGMP/MLD Querier для выбора этого устройства в качестве IGMP/MLD Querier. Правило выбора IGMP/MLD Querier определяет в качестве запрашивающего устройство с наименьшим адресом IP (это правило определено в спецификациях IGMP/MLD, как элемент поведения IGMP/MLD). По этой причине, если выбор IGMP/MLD Querier отдается устройству, не являющемуся прокси, пересылка пакетов не будет выполняться.

На рисунке приведен пример среды с пересылкой исключительно на основе IGMP/MLD.



Устройство А имеет наименьший адрес IP в сети ЛВС 2, но не является прокси-устройством. В соответствии с правилом выбора IGMP/MLD Querier устройство А будет выбрано для ЛВС 2, поскольку оно имеет наименьший адрес IP. Устройство В не будет пересылать трафик в ЛВС 2, поскольку оно не является запрашивающим для этой сети.

Выбор единственного пересылающего прокси обусловлен необходимостью предотвращения петель и избыточного трафика в каналах, которые считаются нисходящими множеством устройств, пересылающих пакеты на основе IGMP/MLD. Это правило piggy-backs для выбора IGMP/MLD Querier. Использование процесса выбора IGMP/MLD Querier для выбора пересылающего прокси обеспечивает на локальном канале функциональность, похожую на механизм PIM Assert. На канале с единственным пересылающим на базе IGMP/MLD устройством это правило **может** быть отменено (т. е., устройство **может** быть настроено на пересылку пакетов через интерфейс, для которого оно не имеет статуса querier). Однако используемая по умолчанию конфигурация **должна** включать указанное правило для поддержки избыточности, как показано на рисунке ниже.



ЛВС 2 может иметь два промежуточных устройства - А и В. В такой конфигурации одно из прокси-устройств должно быть выбрано для пересылки пакетов. Этот документ требует, чтобы пересылка выполнялась устройством, которое имеет статус IGMP/MLD querier. Следовательно, промежуточное устройство А будет пересылать пакеты в ЛВС 2 только в том случае, когда оно выбрано запрашивающим (querier). На приведенном выше рисунке, если А является единственным прокси-устройством, оно может пересылать пакеты даже при выборе в качестве querier устройства В.

Отметим, что это не обеспечивает предотвращения «восходящей петли» (upstream loop). Пример можно видеть на следующем рисунке.



В будет безусловно пересылать пакеты из ЛВС 2 в ЛВС 1 и А будет безусловно пересылать пакеты из ЛВС 1 в ЛВС 2<sup>1</sup>, что приведет к возникновению петли в восходящем направлении. Для устранения таких петель требуется протокол групповой маршрутизации, поддерживающий алгоритм построения дерева пересылки.

### 3.1. Топологические ограничения

Эта спецификация описывает протокол, который работает только в топологии простого дерева. Дерево должно настраиваться вручную путем указания восходящих и нисходящих интерфейсов на каждом промежуточном устройстве. Корень этого дерева соединяется с более широкой инфраструктурой групповой пересылки.

### 3.2. Supporting Senders

Для обеспечения возможности передачи отправителям, находящимся внутри прокси-дерева, весь трафик пересылается в направлении корня. Multicast-маршрутизатор (маршрутизаторы), соединенный с более широкой инфраструктурой групповой передачи, следует настроить так, чтобы все системы внутри прокси-дерева трактовали, как подключенные напрямую — например, для PIM-SM<sup>2</sup> [PIM-SM]. Этим маршрутизаторам следует инкапсулировать трафик от новых источников внутри прокси-дерева в пакеты Register, как будто эти источники подключены непосредственно к маршрутизатору.

Эти настройки задаются вручную — групповая пересылка на основе IGMP/MLD не поддерживает для восходящих маршрутизаторов прокси-дерева способа детектирования подключенных к дереву сетей. Если прокси-топология совпадает с топологией маршрутизации, эту информацию **можно** получить от протокола маршрутизации (например, маршрутизатор можно настроить так, чтобы он трактовал групповые пакеты из всех префиксов, полученных от протокола маршрутизации X через интерфейс Y, как пакеты от подключенных непосредственно систем).

<sup>1</sup>В исходном документе указаны обратные направления пересылки, что является ошибкой (см. [https://www.rfc-editor.org/errata\\_search.php?eid=3285](https://www.rfc-editor.org/errata_search.php?eid=3285)). Прим. перев.

<sup>2</sup>Protocol Independent Multicast - Sparse Mode.

## 4. Поведение промежуточного устройства

В этом разделе более подробно описана работа устройства групповой пересылки на основе IGMP/MLD.

### 4.1. База данных о принадлежности

Промежуточное устройство выполняет связанную с маршрутизацией часть протокола IGMP/MLD на каждом из своих нисходящих интерфейсов. Для каждого интерфейса используемая версия IGMP/MLD задается явно и по умолчанию используется старшая из поддерживаемых системой версий.

Результатом работы этого протокола является набор подписок, который поддерживается отдельно для каждого нисходящего интерфейса. В дополнение к этому подписки на всех интерфейсах сливаются в общую базу данных о принадлежности к группам.

База данных о принадлежности к группам представляет собой множество записей вида:

*(групповой адрес, режим фильтрации, список источников)*

Каждая запись является результатом слияния всех подписок для данного группового адреса на нисходящих интерфейсах. Если некоторые подписки относятся к версии IGMPv1 или IGMPv2/MLDv1, они преобразуются к версии IGMPv3/MLDv2. Подписки IGMPv3/MLDv2 и преобразованные подписки сначала обрабатываются на предмет удаления таймеров в подписках, а затем, если используется режим фильтрации EXCLUDE, удаляются все источники, для которых значение таймера источника > 0. После предварительной обработки подписки сливаются с использованием правил слияния для множества принадлежностей к группам на одном интерфейсе (правила указаны в параграфе 3.2 спецификации IGMPv3 [RFC3376] и параграфе 4.2 спецификации MLDv2 [MLDv2]) для создания записи о принадлежности к группам. Например, имеется два нисходящих интерфейса I1 и I2 с подписками на групповой адрес G. I1 имеет подписку IGMPv2/MLDv1, которая имеет вид (G). I2 имеет подписку IGMPv3/MLDv2 с информацией о принадлежности к группе (G, INCLUDE, (S1, S2)). Подписка I1 преобразуется к версии IGMPv3/MLDv2 и будет иметь вид (G, EXCLUDE, NULL). Затем эти подписки обрабатываются и сливаются, давая в результате запись о принадлежности к группе (G, EXCLUDE, NULL).

Промежуточное устройство выполняет относящуюся к хосту часть протокола IGMP/MLD на своем восходящем интерфейсе. Если устройство является запрашивающим (querier) IGMPv1 или IGMPv2/MLDv1 в восходящей сети, оно будет выполнять на восходящем интерфейсе функции IGMPv1 или IGMPv2/MLDv1. В противном случае будет использоваться IGMPv3/MLDv2.

Если прокси-устройство выполняет на восходящем интерфейсе функции IGMPv3/MLDv2, то при изменении базы данных о принадлежности к группам через восходящий интерфейс будет передаваться отчет о таком изменении, как будто это устройство является хостом, выполнившим соответствующие действия. Если промежуточное устройство выполняет на восходящем интерфейсе функции IGMPv1 или IGMPv2/MLDv1, то при создании или удалении записи о принадлежности к группе отчет о таком изменении будет передаваться через восходящий интерфейс. Все прочие изменения в этом случае игнорируются. При отправке промежуточным устройством отчетов, использующих IGMPv1 или IGMPv2/MLDv1 в записи о принадлежности к группе заполняется только поле группового адреса.

### 4.2. Пересылка пакетов

Промежуточное устройство пересылает полученные на восходящем интерфейсе пакеты в свои нисходящие интерфейсы на основе подписок для этого интерфейса с учетом статуса IGMP/MLD Querier на каждом интерфейсе прокси-устройства. Промежуточное устройство пересылает все пакеты, полученные на любом нисходящем интерфейсе в восходящий интерфейс и другие нисходящие интерфейсы (в соответствии с подписками на каждом интерфейсе и статусом IGMP/MLD Querier для него). Промежуточное устройство **может** использовать кэширование, чтобы не приходилось принимать решение о пересылке для каждого отдельного пакета, но этот кэш **должен** обновляться с использованием указанного выше правила при каждом изменении информации, использованной для создания кэша.

### 4.3. Поддержка SSM

Для поддержки специфической для источника групповой рассылки (SSM<sup>1</sup>) промежуточное устройство должно соответствовать спецификации применения IGMPv3 для SSM [SSM]. Отметим, что прокси-устройство должно соответствовать требованиям IGMP Host и IGMP Router для SSM, поскольку оно выполняет IGMP Host Portion на восходящем интерфейсе и IGMP Router Portion на каждом нисходящем интерфейсе.

Интерфейс может быть настроен на работу в режиме IGMPv1 или IGMPv2. В таких случаях семантика SSM не будет поддерживаться интерфейсом. Однако промежуточным устройствам, соответствующим данной спецификации, следует игнорировать подписки IGMPv1 и IGMPv2, передаваемые по адресам SSM. Более важным является то, что пакеты со специфическими для источника адресами **не следует** пересылать на интерфейсы с подписками IGMPv2 или IGMPv1 для таких адресов.

## 5. Вопросы безопасности

Поскольку пересылку пакетов выполняют только устройства со статусом Querier, процесс выбора IGMP/MLD Querier может приводить к возникновению «черных дыр», если в качестве Querier будет выбрано не пересылающее пакеты устройство. Атакующий из ЛВС на нисходящей стороне может спровоцировать выбор себя в качестве Querier, что приведет к блокированию пересылки пакетов.

Однако этой проблемы можно избежать операционными средствами. Среди операторов достаточно распространена практика выбора для маршрутизаторов адресов из младшей части подсети. Таким образом, для предотвращения описанной выше проблемы операторам **следует** выделять для прокси-устройств младшие адреса из их подсети для того, чтобы эти устройства были выбраны в качестве запрашивающих (querier).

<sup>1</sup>Source-Specific Multicast.

Пересылка на основе IGMP/MLD не обеспечивает механизмов обнаружения «восходящих петель» (upstream loop), описанных в разделе 3. Следовательно, для предотвращения проблем, связанных с такими петлями, **должны** использоваться административные средства, гарантирующие отсутствие таких петель в среде IGMP/MLD Proxy.

Информация IGMP/MLD, представляемая прокси-устройствами восходящим маршрутизаторам, образуется путем агрегирования данных о принадлежности к группам на нисходящих интерфейсах. Никакой из методов контроля доступа к протоколу управления принадлежностью к группам не может быть применен к одному подписчику. Т. е., контроль доступа применяется одинаково для всех заинтересованных хостов, доступных через данное прокси-устройство.

## 6. Благодарности

Авторы благодарят Erik Nordmark, Dave Thaler, Pekka Savola, Karen Kimball и других людей за рецензирование спецификации и комментарии к ней.

## 7. Нормативные документы

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, [RFC 2119](#), March 1997.
- [RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", [RFC 3376](#), October 2002.
- [RFC2236] Fenner, W., "Internet Group Management Protocol, Version 2", [RFC 2236](#), November 1997.
- [RFC1112] Deering, S., "Host extensions for IP multicasting", STD 5, [RFC 1112](#), August 1989.
- [RFC2710] Deering, S., Fenner, W., and B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, October 1999.
- [MLDv2] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.
- [SSM] Holbrook, H., Cain, B., and B. Haberman, "Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast", RFC 4604, August 2006.

## 8. Дополнительная литература

- [MCAST] Deering, S., "Multicast Routing in a Datagram Internetwork", Ph.D. Thesis, Stanford University, December 1991.
- [PIM-SM] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.

### Адреса авторов

#### Bill Fenner

AT&T Labs - Research

1 River Oaks Place

San Jose, CA 95134

Phone: +1 408 493-8505

EMail: [fenner@research.att.com](mailto:fenner@research.att.com)

#### Haixiang He

Nortel

600 Technology Park Drive

Billerica, MA 01821

EMail: [haixiang@nortel.com](mailto:haixiang@nortel.com)

#### Brian Haberman

Johns Hopkins University Applied Physics Lab

11100 Johns Hopkins Road

Laurel, MD 20723-6099

EMail: [brian@innovationslab.net](mailto:brian@innovationslab.net)

#### Hal Sandick

Little River Elementary School

2315 Snow Hill Road

Durham, NC 27712

EMail: [sandick@nc.rr.com](mailto:sandick@nc.rr.com)



Николай Малых

[nmalykh@protocols.ru](mailto:nmalykh@protocols.ru)

#### **Полное заявление авторских прав**

Copyright (C) The Internet Society (2006).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

#### **Интеллектуальная собственность**

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

#### **Подтверждение**

Финансирование функций RFC Editor обеспечено IETF Administrative Support Activity (IASA).