Network Working Group
Request for Comments: 2491
Category: Standards Track

G. Armitage
Lucent Technologies
P. Schulter
Bright Tiger Technologies
M. Jork
Digital Equipment GmbH
G. Harter
Compaq

January 1999

IPv6 over Non-Broadcast Multiple Access (NBMA) networks

IPv6 в сетях с множественным доступом без широковещания

Статус документа

В этом документе содержится спецификация 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

Аннотация

Этот документ описывает общую архитектуру IPv6 в сетях NBMA и служит основой для сопровождающих документов, описывающих конкретные технологии NBMA (такие как ATM или Frame Relay). Архитектура IPv6 в сетях NBMA обеспечивает нормальную работу протокола обнаружения соседей IPv6 (Neighbor Discovery или ND) на стороне хоста, а также поддерживает создание «коротких» путей пересылки NBMA при наличии каналов NBMA с динамической сигнализацией. Описана также работа по административно настраиваемым каналам «точка-точка» (Point to Point).

Динамические короткие каналы NBMA (shortcut) обеспечиваются за счёт работы протокола IPv6 ND внутри логических каналов, а также протокола NHRP между маршрутизаторами для обнаружения адресатов NBMA вне канала (off-Link). Поддерживаются как создаваемые потоками, так и явно задаваемые источником короткие соединения.

Оглавление

1. Введение

Сети с множественным доступом без широковещания (Non-Broadcast Multiple Access или NBMA) могут использоваться разными способами. Одним из крайних случаев является применение для простых, административно настраиваемых услуг «точка-точка», достаточных для соединения маршрутизаторов IPv6 (в некоторых случаях и хостов IPv6). Другой крайностью являются применения сетей NBMA с поддержкой динамической организации и удаления виртуальных соединений (Virtual Circuit или VC) или их функциональных эквивалентов для эмуляции услуг, предоставляемых уровню IPv6 традиционными широковещательными средами, такими как Ethernet. Обычно для такой эмуляции нужны сложные протоколы сходимости, особенно для поддержки групповой адресации IPv6.

В этом документе описана базовая архитектура IPv6 в сетях NBMA. Документ служит основой для подробных спецификаций работы с конкретными технологиями NBMA (например, ATM [17] или Frame Relay). Архитектура IPv6 в сети NBMA поддерживает обычные операции протокола IPv6 ND на стороне хоста, а также организацию коротких путей пересылки NBMA (при доступности каналов NBMA с динамической сигнализацией).

Большая часть документа посвящена использованию динамически управляемых соединений «точка-точка» и «один со многими» (point to multipoint) между интерфейсами сети NBMA. Такие соединения обозначаются в документе как SVC. Рассмотрены также административно настраиваемые соединения «точка-точка», обозначаемые как PVC. В зависимости от контекста для обоих типов соединений может применяться обозначение VC.

Некоторые сети NBMA поддерживают форму услуг без организации явных соединений (например, SMDS). В этих случаях «соединения» или VC считаются существующими неявно, если у отправителя есть адрес получателя NBMA, по которому он может передавать пакеты.

1.1. Обнаружение соседей

Основным различием между этой архитектурой и прежними протоколами IP в сетях NBMA является механизм обнаружения соседей IPv6 ND.

Мир IPv4 развивался на основе подхода к распознаванию адресов на основе дополнительного протокола, работающего на уровне логического канала, начиная с Ethernet ARP (RFC 826 [14]). В мире сетей NBMA протокол ARP применялся в IPv4 для сетей SMDS (RFC 1209 [13]) и ATM (RFC 1577 [3]). Позднее рабочая группа ION выпустила протокол обнаружения следующего узла пересылки (Next Hop Resolution Protocol или NHRP [8]), для распознавания адресов внутри подсети и между подсетями, пригодный для многих технологий NBMA.

Разработчики IPv6 решили уйти от модели, привязанной к канальному уровню, выбрав для объединения множества задач протокол, известный как обнаружение соседей (ND) [7], способный работать с разными технологиями канального уровня. Ключевым допущением протокола обнаружения соседей является то, что технология канального уровня, используемая данным интерфейсом IP, поддерживает групповую адресацию (multicast) естественным способом. Это не всегда верно в большинстве сетей NBMA и обычно требует протокола сближения для эмуляции нужных услуг (примером такого протокола сближения является MARS, RFC 2022 [5]). Этот документ дополняет и оптимизирует протокол MARS для использования с IPv6 ND, расширяя применимость RFC 2022 за пределы сетей ATM.

1.2. Короткие пути NBMA

Коротким путём (shortcut) является соединение уровня NBMA (VC), напрямую связывающее две конечные точки IP, логически разделённые одним или несколькими маршрутизаторами на уровне IP. О пакетах IPv6, проходящих через такие VC, говорят, что они «закорачивают» маршрутизаторы на логическом пути IPv6 между конечными точками VC.

Короткие пути (перемычки) NBMA являются механизмом, минимизирующим расход ресурсов в облаке IP через NBMA cloud (например, интервалы пересылки и NBMA VC).

Важно, что перемычки NBMA поддерживаются при любом развёртывании IP через сети NBMA с поддержкой динамических соединений (SVC или функциональный эквивалент). Для IPv6 в сети NBMA обнаружение и поддержка коротких путей обеспечивается комбинацией ND и NHRP.

1.3. Основные компоненты архитектуры IPv6 в сети NBMA

1.3.1. Сети NBMA с поддержкой PVC

При использовании сети NBMA в режиме PVC каждый PVC соединяет два узла и применение ND и других свойств IPv6 ограничено. Интерфейсы IPv6/NBMA имеют лишь одного соседа на каждом канале. Протоколы MARS и NHRP **не** требуются, поскольку групповые и широковещательные операции сведены к индивидуальным (unicast) операциям на уровне NBMA. Динамически обнаруживаемые перемычки не поддерживаются.

Фактические детали инкапсуляции и создания маркеров канала **нужно** описывать в документах для конкретной технологии NBMA. Документам **нужно** следовать приведённым ниже рекомендациям.

Индивидуальные и групповые пакеты IPv6 **нужно** передавать через PVC с инкапсуляцией, описанной в параграфе 4.4.1.

Маркеры интерфейсов для каналов PVC **нужно** создавать в соответствии с разделом 5 и они должны быть разными у двух узлов на канале PVC.

Такое использование каналов PVC не требует и не запрещает применение расширений протокола ND, которые могут быть разработаны для общего случая или специально для PVC (например, Inverse Neighbor Discovery).

Относящиеся к NBMA дополнительные документы **могут** задавать механизмы IPv6 over PPP и PPP over NBMA как **необязательное** дополнение к IPv6 на каналах «точка-точка».

Остальная часть документа относится к соединениям SVC, если явно не указано иное.

1.3.2. Сети NBMA с поддержкой SVC

Ключевые компоненты при использовании сети NBMA в режиме SVC перечислены ниже.

- Модель соседства IPv6, где соседи обнаруживаются с помощью сообщений, предаваемых по групповому адресу Link-local интерфейса IPv6.
- Модель MARS, позволяющая эмулировать групповую адресацию с использованием multipoint-соединений, предоставляемых базовой сетью NBMA.
- Служба NHRP для поиска отождествлений NBMA интерфейсов IP, логически удалённых в топологии IP.
- Моделирование трафика IP «потоками» и необязательное использование наличия потока в качестве основы для попытки организации короткого соединения на канальном уровне.

В результате применения описанных выше свойств обеспечивается ряд возможностей.

«Канал» IPv6 обобщается как «логический канал» (Logical Link или LL) в средах NBMA (по аналогии с обобщением подсети IPv4 в логическую подсеть IP в RFC 1209 и RFC 1577).

Интерфейсы IPv6/NBMA используют RFC 2022 (MARS) для базовой групповой передачи внутри LL. Сам протокол MARS служит для оптимального распространения сообщений об обнаружении внутри LL.

Для адресатах, в настоящее время не считающихся соседями, хост передаёт пакеты одному из принятых по умолчанию маршрутизаторов.

При подобающей настройке выходной маршрутизатор LL отвечает за обнаружение потока пакетов IP через него, что может давать преимущества при использовании коротких соединений.

Продолжая традиционно пересылать пакеты потока, маршрутизатор инициирует запрос NHRP для IP-адреса получателя потока.

Последний маршрутизатор или NHS (Next Hop Server) перед адресатом запроса NHRP уточняет предпочтительный адрес NBMA целевого интерфейса.

Затем передавший исходный запрос маршрутизатор передаёт по IP-адресу отправителя сообщение Redirect, указывающее адресата потока как временного соседа.

Инициированное хостом обнаружение короткого пути, независимо от наличия потока пакетов, поддерживается также с помощью конкретных запросов соседства (Neighbor Solicitation или NS), передаваемых принятому по умолчанию маршрутизатору хоста-источника.

Для этого подхода заявлен ряд важных преимуществ.

Стеки протокола IPv6 на хостах не реализуют отдельных протоколов ND для каждой технологии канального уровня.

Когда адресат потока запрошен как временный сосед, возвращаемый адрес NBMA будет одним из выбранных адресатом при исходной организации потока путём поэтапной (hop-by-hop) обработки. Это поддерживает имеющуюся возможность ND выполнять динамическую балансировку между интерфейсами IPv6.

1.4. Терминология

Битовые последовательности и численные значения, применяемые для идентификации конкретного интерфейса NBMA на уровне NBMA называются адресами NBMA (примерами служат ATM End System Address - AESA при использовании архитектуры в сети ATM или E.164 для сетей SMDS).

Организованное когда-либо соединение, служащее для передачи пакетов IP от одного интерфейса NBMA к другому, называется SVC или PVC в зависимости от способа организации соединения (динамически по сигнальному протоколу или административно). Конкретные механизмы сигнализации для организации и разрыва SVC будут определяться соответствующими спецификациями NBMA. Некоторые сети NBMA могут предоставлять услуги без организации явных соединений (например, SMDS) и в таких случаях «соединение» или SVC считается существующим неявно, если отправитель знает адрес получателя NBMA для отправки тому пакетов, когда это нужно.

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не нужно** (SHALL NOT), **следует** (SHOULD), **не следует** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе должны интерпретироваться в соответствии с RFC2119 [16].

1.5. Структура документа

В разделе 2 этого документа рассматривается обобщение канала IPv6 как LL при использовании в сетях NBMA и вводится понятие временного соседа (Transient Neighbor). Раздел 3 посвящён изменениям в протоколе MARS для эффективного распространения сообщений ND внутри LL, а также правилам и механизмам обнаружения временных соседей. Раздел 4 посвящён базовым правилам инициализации интерфейсов IPv6/NBMA, инкапсуляции пакетов и управляющих сообщений, а также правилам управления SVC. В разделе 5 описаны общие правила создания маркеров интерфейса (Interface Token), а опция Link Layer Address и адреса Link-Local. Раздел 6 завершает нормативную часть документа. В Приложении A приведено ненормативное описание работы IPv6 ND, Приложение B описывает некоторые неоптимальные решения для эмуляции групповой передачи сообщений ND через LL. В Приложении C обсуждается подавление коротких путей и приведён краткий обзор будущего взаимодействия обнаружения потоков и их сопоставления с SVC, обеспечивающими разное качество обслуживания.

2. Логические каналы и временные соседи

В IPv6 используется концепция on-link (на канале) и off-link (вне канала). Соседями считаются находящиеся на канале узлы, что обеспечивает возможность обнаружения их адреса на канальном уровне с помощью ND. Ниже приведены определения, заимствованные из текста ND.

on-link

Адрес, назначенный интерфейсу соседа в общем канале. Хост считает адрес относящимся к каналу при выполнении любого из условий:

- адрес охватывается тем же префиксом канала;
- соседний маршрутизатор указывает адрес как цель сообщений Redirect;
- получен анонс соседа (Neighbor Advertisement или NA) для этого адреса;
- с этого адреса принято сообщение ND.

off-link

В отличие от on-link это адрес, не назначенный какому-либо из интерфейсов, подключённых к общему каналу.

Узлы вне канала считаются доступными только через один из подключённых к каналу маршрутизаторов.

Среда NBMA усложняет трактовку термина канал (link) так же, как усложняется трактовка подсети (subnet) в случае IPv4. Для IPv4 это потребовало определения логической подсети (Logical IP Subnet или LIS) - административно заданного набора хостов, использующих общий префикс маршрутизации (маски сети и подсети).

Этот документ рассматривает аналог логического канала (LL) в IPv6. LL состоит из узлов, административно настроенных на одном канале (on link).

Членами LL являются соседи интерфейса IPv6 из начального набора и каждый адрес интерфейсе в LL должен быть уникальным среди этих соседей.

Следует отметить, что члены LL являются соседями IPv6, однако могут быть соседи, которые административно не являются членами одного LL.

События ND могут приводить к прекращению соседства интерфейса IPv6., однако это не меняет набор интерфейсов, образующих LL. В результате возможны три варианта отношений между любой парой интерфейсов IPv6:

- соседи по LL;
- соседи вне LL;
- вне LL, не соседи.

Соседи вне LL представляют короткие соединения (перемычки), где установлена возможность прямого соединения на уровне NBMA с целью, которая не является членом LL источника.

Соседи, найденные с помощью незапрошенных соединений, таких как Redirect, называются временными (Transient Neighbor).

3. Обнаружение внутри и между LL

Этот документ различает обнаружение соседей в логическом канале (intra-LL) и вне LL (inter-LL). Цель состоит в раскрытии топологии в логическом канале и за его пределами, чтобы не менять на стороне хоста стек IPv6 для интерфейсов NBMA.

Отметим, что при поддержке в сети лишь постоянных соединений «точка-точка» (PVC) применяется параграф 1.3.1.

3.1. ND с multicast-эмуляцией внутри LL

Базовая модель ND предполагает, что интерфейс канального уровня будет делать что-либо значимое с пакетом ICMPv6, переданным по групповому адресу IP (IPv6 считает групповую адресацию неотъемлемой частью Internet). В этом документе предполагается поддержка групповой адресации с использованием RFC 2022 (MARS) [5] (обобщён для применения с NBMA в дополнение к ATM). IPv6 LL отображается напрямую в IPv6 MARS Cluster так же, как IPv4 LIS отображается в IPv4 MARS Cluster.

Целью работы внутри LL является в обеспечении уровню IPv6 возможности простой передачи групповых пакетов ICMPv6 в нисходящем направлении драйверу IPv6/NBMA без специальной обработки для NBMA. В этом случае базовый механизм распространения сообщений ND и Router Discovery работает как обычно.

В параграфе 3.1.1 описаны дополнительные функции, которые **нужно** требовать от протокола MARS, используемого в соответствии с этим документом. Эти дополнения рассматриваются в Приложении B.

3.1.1. Обязательное дополнение MARS и поведение клиента MARS

Интерфейсы IPv6/NBMA **нужно** регистрировать как членов MARS Cluster в соответствии с параграфом 4.1 и определённые классы исходящих пакетов IPv6 **нужно** отправлять напрямую в их локальный MARS, как указано в параграфе 4.4.2.

Протоколу MARS затем **нужно** заново передавать эти пакеты в соответствии с приведёнными ниже правилами.

- Когда MARS принимает пакет IPv6, он просматривает базу данных о принадлежности к группам для поиска адресов NBMA членов целевой группы IPv6.
- Затем MARS проверяет наличие у каждого члена группы управляющего VC pt-pt[1], открытого для MARS. При наличии MARS передаёт копию пакета данных каждому члену группы через имеющиеся pt-pt VC.
- Если хотя бы у одного из найденных членов группы нет открытого pt-pt VC для MARS или члены группы не найдены, пакет передаётся в ClusterControlVC. Копии пакета через имеющиеся pt-pt VC не передаются.

3.2. Сообщения Redirect вне LL

Короткие соединения оправданы наличием потоков пакетов IP между отправителем и получателем которых присутствуют маршрутизаторы IP. Короткие соединения организуются между временными соседями.

Временное соседство организуется по сообщениям Redirect (раздел 8 в [7]). IPv6 позволяет маршрутизатору информировать членов LL о наличии лучшего маршрутизатора (first hop) для данного адресата (параграф 8.2 в [7]). Сам анонс обеспечивается сообщением Router Redirect, которое может содержать адрес канального уровня более подходящего маршрутизатора.

Передающий хост слушает сообщения Router Redirect лишь от маршрутизатора, который принят в данный момент по умолчанию для получателя IP, к которому относится сообщение Redirect. При получении Redirect, указывающего лучший первый маршрутизатор для данного адресата и адрес канального уровня (NBMA) для использования в качестве first hop, соответствующая запись Neighbor Cache на хосте-источнике обновляется и для неё устанавливается состояние доступности STALE. Обновление кэша в этом контексте включает организацию нового VC с новым адресом NBMA. При успешном создании старое соединение VC отключается, если оно больше не нужно (поскольку оно вело к маршрутизатору, потребность в нем может сохраняться для других пакетов от хоста).

Обеспечивается два механизма запуска обнаружения лучшего первого маршрутизатора:

- идентификация/обнаружение потока по маршрутизаторам;
- запрос короткого соединения от хоста.

В параграфе 3.2.1 описан запуск на основе потоков, в параграфе 3.2.2 - запуск по инициативе хоста, а в параграфе 3.2.3 - применение NHRP для обнаружения отображения для целей IPv6 в удалённых LL.

3.2.1. Вызываемое потоком перенаправление

Изменение путей пересылки на основе динамического обнаружения потоков пакетов IP является основой таких моделей, как Cell Switch Router [11] и IP Switch [12]. Ответственность за обнаружение потоков принимают маршрутизаторы, на которых пакеты пересекают границу маршрутизации IP.

В соответствии с этим документом маршрутизатор **может** инициировать обнаружение лучшего first-hop для проходящего через него потока пакетов IP. Такому маршрутизатору **нужно**:

- **нужно** отслеживать только потоки, исходящие лишь от подключённых напрямую хостов (хосты в LL для интерфейса маршрутизатора);
- **не** использовать пакеты IP, приходящие от других маршрутизаторов, для создания Router Redirect;
- рассматривать лишь пакеты IPv6 с FlowID = 0 для обнаружения потоков в соответствии с этим параграфом;
- использовать NHRP в соответствии с параграфом 3.2.3 для определения лучшего first-hop при обнаружении подходящего потока и анонсирования информации в Router Redirect.

Маршрутизаторам IPv6, поддерживающим **необязательное** обнаружение потоков, описанное выше, **нужно** поддерживать административные механизмы для его отключения. Они **могут** поддерживать механизмы дополнительного ограничения по категориям пакетов IPv6, образующих «поток».

Фактические алгоритмы идентификации последовательности пакетов IPv6 как «потока» выходят за рамки документа. В Приложении C рассмотрена причина использования ненулевого FlowID для подавления обнаружения потоков.

3.2.2. Вызываемое хостом перенаправление

Хост-источник также **может** вызвать перенаправление к временному соседу. Для поддержки этого маршрутизаторам, соответствующим данному документу, **нужно** распознавать определённые сообщения Neighbor Solicitation, передаваемые хостами, как запрос распознавания адреса вне канала.

Для перенаправления хосту-источнику **нужно**:

- создать сообщение Neighbor Solicitation, указывающее получателя вне LL, для которого нужна «перемычка»;
- адресовать сообщение NS маршрутизатору, который является следующим узлом пересылки на пути к цели off-LL (вместо группового адреса запрашиваемой цели);
- использовать стандартное для ND значение Hop Limit = 255, чтобы маршрутизатор не отбросил сообщение NS;

[1]Point-to-point - соединение «точка-точка».

- включить опцию Shortcut Limit, определённую в Приложении D; в опции следует установить значение, совпадающее с hop limit потока данных, для которого передан этот триггер; это позволит маршрутизатору ограничить попытки создания перемычки, не связанные с потоком данных;

- переслать пакет NS маршрутизатору, который является следующим на пути к цели off-LL.

Маршрутизаторам **нужно** считать индивидуальное сообщение NS с опцией Shortcut Limit запросом перенаправления по инициативе хоста. Однако фактическое обнаружение перемычек **необязательно** для маршрутизаторов IPv6. Когда такое обнаружение не поддерживается, маршрутизатору **нужно** создать сообщение Redirect, указывающее этот маршрутизатор как «лучшую перемычку», и вернуть его запросившему хосту. Если обнаружение перемычек поддерживается, отклику маршрутизатора **нужно**:

- создать и передать подходящий запрос NHRP, в соответствии с параграфом 3.2.3; исходное сообщение NS **следует** отбросить;

- после получения маршрутизатором отклика NHRP Reply ему **нужно** создать сообщение Redirect с адресом IPv6 временного соседа и адресом канального уровня NBMA, возвращённым процессом распознавания NHRP;

- полученное сообщение Redirect **нужно** отправить хосту-источнику, которому при получении этого сообщения **нужно** обновить свои кэше соседей и адресатов;

- цель off-LL после этого считается временным соседом и следующий пакет, переданный ему будет приводить к созданию прямого VC (к самой цели вне LL или лучшему выходному маршрутизатору в её направлении, определённому NHRP);

- при получении NHRP NAK или сведений об ошибке для попытки организации перемычки по инициативе хоста, запрашивающему маршрутизатору **нужно** создать сообщение Redirect, указывающее его как «лучшую перемычку» и передать это сообщение запрашивающему хосту.

3.2.3. Применение NHRP между маршрутизаторами

После обнаружения потока или перенаправления от хоста маршрутизаторам **нужно** использовать NHRP в режиме NHS-to-NHS для сопоставления адреса IPv6 с адресом канального уровня лучшего следующего маршрутизатора. Маршрутизаторам IPv6/NBMA, поддерживающим обнаружение перемычек, потребуется выполнить одну или несколько перечисленных ниже функций.

- Создание запросов и откликов NHRP.

- Анализ запросов и откликов NHRP от других NHS (маршрутизаторов).

- Пересылка запросов NHRP в направлении NHS, топологически наиболее близкого к цели IPv6.

- Пересылка откликов NHRP в направлении NHS, топологически наиболее близкого к запрашивающему.

- Синтаксическая трансляция между NS и исходящими запросами NHRP.

- Синтаксическая трансляция между входящими откликами NHRP и Redirect.

Адресат потока, послужившего триггером (или цель инициированного хостом триггера) служит целью для запроса NHRP Request. Маршрутизатор пересылает NHRP Request ближайшему следующему NHS и этот процесс продолжается (как для обычного NHRP), пока запрос не достигнет NHS, считающего, что цель IP находится в области действия локального канала одного из его интерфейсов (весь процесс может ограничиваться одним маршрутизатором).

Поскольку запросы распознавания NHRP всегда следуют по маршрутизируемому пути к целевому протокольному адресу, область действия запрошенного короткого соединения будет автоматически привязываться к области действия целевого адреса IPv6 (например, запросы распознавания для локальных адресов сайта не выйдут за границы сайта).

Последнему маршрутизатору на пути **нужно** разрешить NHRP Request по сведениям о сопоставлении из своего кэша соседей для интерфейса, через который доступна указанная цель. Если подходящей записи в кэше соседей нет или адресат в данный момент считается недоступным, маршрутизатору последнего интервала **нужно** выполнить процедуру ND на локальном интерфейсе и создать на основе её результата отклик NHRP Reply. Отметим, что в случае отправки NHRP Request по факту обнаружения потока уже должен быть поток пакетов hop-by-hop, проходящий через последний маршрутизатор в направлении цели. В этом типичном случае нужные сведения имеются в кэше соседей.

NHRP Reply распространяется к источнику NHRP Request с использованием пути hop-by-hop как в обычном NHRP.

Если процесс обнаружения был вызван обнаружением потока на создавшем запрос маршрутизаторе, возврат NHRP Reply вызовет указанные ниже события.

- Создаётся сообщение Redirect с использованием отображения IPv6/NBMA из NHRP Reply.

- Redirect передаётся по индивидуальному IP-адресу источника потока (с применением VC, по которому поток пришел в маршрутизатор, если это двухсторонний pt-pt VC).

- Любое сообщение Redirect, переданное маршрутизатором, **должно** соответствовать всем правилам, заданным в [7], чтобы пакет был должным образом проверен принимающим хостом. Е частности, если целью перемычки является хост-получатель, поле ICMP Target Address **должно** совпадать с ICMP Destination Address в исходном запросе. Если целью перемычки является выходной маршрутизатор, в поле ICMP Target Address **должен** быть адрес Link-Local этого маршрутизатора, уникальный в облаке NBMA, к которому присоединён интерфейс NBMA этого маршрутизатора.

- Следует также отметить, что выходной маршрутизатор может потом перенаправить хост-источник. Для этого Link-Local ICMP Source Address в сообщении Redirect **должен** совпадать с Link-Local ICMP Target Address в исходном сообщении Redirect.

Отметим, что создающий NHRP Reply маршрутизатор использует адрес NBMA, возвращённый целевым хостом, когда тот в первый раз воспринял поток трафика IP. Это сохраняет полезную функцию ND - распределение нагрузки целевого интерфейса.

При получении отклика NHRP NAK или сообщения об ошибке при вызванной потоком попытке создать короткое соединения, эта информация не передаётся источнику потока.

3.2.3.1. Правила трансляции пакетов NHRP/ND

Ниже приведены правила, предназначенные для дополнения спецификации формата пакетов в разделе 5 NHRP [8], охватывающие поля, используемые архитектурой IPv6/NBMA.

Сообщения NHRP создаются и отправляются в соответствии с правилами [8]. Значения связанных с технологией NBMA полей, таких как arafn, arpro.type, ar$pro.snap, и формат адреса канального уровня определены в соответствующих документах NBMA. Адреса отправителя и получателя, протокольный адрес клиента в базовом заголовке или CIE в сообщении NHRP всегда являются адресами IPv6 размером 16 байтов.

При создании инициированного хостом запроса распознавания NHRP в ответ на NS выполняются указанные ниже действия.

- Поле ar$hopcnt **должно** быть меньше значения Shortcut Limit в опции Shortcut Limit, включённой в вызвавшее сообщение NS. Это обеспечивает хостам контроль доступа к их запросам на создание перемычек. Отметим, что значение Shortcut Limit в опции задаётся относительно запрашивающего хоста, поэтому ar$hopcnt будет меньше данного Shortcut Limit.
- В поле Flags в базовом заголовке запроса распознавания NHRP **следует** устанавливать биты Q и S.
- **Следует** установить бит U. Адрес NBMA и протокольный адрес отправителя относятся к создающему запрос маршрутизатору.
- Адрес цели из сообщения NS используется как протокольный адрес получателя NHRP. Задавать CIE **не нужно**.

При создании запроса распознавания NHRP по факту обнаружения потока выбор значений зависит от конфигурации.

Отклик распознавания NHRP создаётся в соответствии с правилами [8].

- Для каждого возвращённого CIE время удержания составляет 10 минут.
- MTU может иметь значение 0 или значение, заданное соответствующим документом NBMA.

Успешный отклик распознавания NHRP для инициированной хостом попытки создать перемычку транслируется в сообщение IPv6 Redirect, как показано ниже.

Поля IP

Source Address
Адрес Link-Local, назначенный интерфейсу маршрутизатора, с которого передано сообщение.
Destination Address
IPv6 Source Address вызвавшего запрос сообщения NS.
Hop Limit
255
Поля ICMP

Target Address
NHRP Client Protocol Address
Destination Address
Цель триггерного NS (эквивалент NHRP Destination Protocol Address)
Target link-layer address
NHRP Client NBMA Address
Определённые в настоящее время расширения NHRP [8] не влияют на трансляцию NHRP/ND и **могут** использоваться в сообщениях NHRP для IPv6.

3.2.3.2. Правила очистки NHRP

Очистка NHRP выполняется при обнаружении изменений, аннулирующих ранее выданный отклик NHRP Reply (это может быть изменение топологии, отключение целевого хоста, смена отождествления). Все короткие соединения IPv6, созданные ранее на основе недавно сброшенной информации, **следует** разорвать.

Маршрутизаторам **нужно** отслеживать записи кэша NHRP, для которых были выпущены сообщения Neighbor Advertisement или Router Redirect. При получении NHRP Purge с аннулированием сведений, выданных ранее локальному хосту, маршрутизатору нужно выдать сообщение Router Redirect, указывающее его в качестве лучшего следующего интервала (next-hop) для соответствующей цели IPv6.

Маршрутизаторам **нужно** отслеживать записи кэша соседей, использованные ранее для генерации NHRP Reply. По завершению срока действия любой записи в кэше соседей **нужно** передавать NHRP Purge в направлении маршрутизатора, исходно запросившего NHRP Reply.

3.3. Обнаружение недоступности соседа

Сообщения NS для обнаружения недоступности соседа (Neighbor Unreachability Detection или NUD) передаются индивидуально (unicast) соответствующему соседу с использованием VC к данному соседу. Это говорит о том, что с точки зрения NUD временный сосед не отличим от соседа на локальном канале (On-LL).

3.4. Обнаружение дубликатов адресов

Обнаружение дубликатов адресов (Duplicate Address Detection или DAD) требуется лишь в области действия Link-Local, которая в данном случае совпадает с областью действия локального LL. Временные соседи не попадают в эту область. Не требуется какого-либо специального взаимодействия между механизмами организации перемычек и обнаружения дубликатов адресов на локальном канале.

4. Концепции работы узла

В этом разделе описаны операции узла при выполнении основных функций (таких как передача и приём данных) на логическом канале (LL). Применение этих операций к различным протоколам IPv6, таким как ND, описано в Приложении A. Большая часть этого раздела относится только к сетям NBMA, применяемым для SVC «точка-точка» и «один со многими». В разделе 7 рассматривается сеть NBMA, используемая лишь для соединения PVC «точка-точка».

4.1. Подключение к LL

До того как узел сможет передавать или принимать дейтаграммы IPv6 его базовый интерфейс IPv6/NBMA должен подключиться к логическому каналу (LL). Драйверу IPv6/NBMA **нужно** организовать pt-pt VC с MARS, связанным с его LL, и зарегистрироваться как Cluster Member [5]. Интерфейс IPv6/NBMA становится членом LL, получает идентификатор члена кластера (Cluster Member ID или CMI) и может начинать операции IPv6 и IPv6 ND.

Если узел является хостом или запускающимся маршрутизатором, ему **нужно** передать групповое сообщение MARS_JOIN в группы:

- Solicited-node с областью действия Link-Local;
- All-nodes с областью действия Link-Local;
- другие настроенные группы с областью действия не уже Link-Local.

Если узел является маршрутизатором, ему **нужно** дополнительно передать:

- одно групповое сообщение MARS_JOIN по адресу All-routers с областью действия Link-Local;
- блок MARS_JOIN для диапазонов групповых адресов IPv6 (с областью действия шире Link-Local), которым нужно получать все (promiscuous reception).

Механизм инкапсуляции и значения ключевых полей управляющих сообщений MARS **нужно** задать в документах для соответствующей технологии NBMA.

4.2. Подключение к Multicast-группе

В этом параграфе описано поведение узла при получении запроса JoinLocalGroup от уровня IPv6. Детали этого поведения зависят от реализации.

При получении JoinLocalGroup для адреса node-local драйверу IPv6/NBMA **нужно** возвращать индикацию успеха без дополнительных действий (пакеты, переданные по адресу node-local не приходят в драйвер IPv6/NBMA).

При получении JoinLocalGroup для адреса с выходящей за пределы узла областью действия драйверу IPv6/NBMA **нужно** передать один групповой запрос MARS_JOIN для регистрации этого адреса в MARS.

4.3. Выход из Multicast-группы

В этом параграфе описано поведение узла при получении запроса LeaveLocalGroup от уровня IPv6. Детали этого поведения зависят от реализации.

При получении LeaveLocalGroup для адреса node-local драйверу IPv6/NBMA **нужно** возвращать индикацию успеха без дополнительных действий (пакеты, переданные по адресу node-local не приходят в драйвер IPv6/NBMA).

При получении LeaveLocalGroup для адреса с выходящей за пределы узла областью действия драйверу IPv6/NBMA **нужно** передать один групповой запрос MARS_ LEAVE для отмены регистрации этого адреса в MARS.

4.4. Передача данных

Для исходящих индивидуальных и групповых пакетов применяются раздельные правила обработки и инкапсуляции.

4.4.1. Передача индивидуальных данных

Применяется следующий интервал уровня IP (next hop) для каждого индивидуального исходящего пакета IPv6, задающий pt-pt VC для пересылки пакета.

Для сетей NBMA с инкапсуляцией LLC/SNAP (например. ATM или SMDS) пакеты IPv6 **нужно** инкапсулировать с показанным ниже заголовком LLC/SNAP и передавать через VC.

```
[0xAA-AA-03][0x00-00-00][0x86-DD][IPv6 packet]
   (LLC)        (OUI)      (PID)
```

Для сетей NBMA, не применяющих инкапсуляцию LLC/SNAP, **нужно** задавать соответствующие правила в спецификации NBMA.

При отсутствии pt-pt VC для адреса следующего маршрутизатора узлу **нужно** сделать вызов для организации VC со следующим маршрутизатором. Каждый раз при получении драйвером IPv6/NBMA индивидуального пакета для передачи уровень IPv6 уже знает адрес канального уровня (NBMA) следующего маршрутизатора. Таким образом, информация для соединения NBMA со следующим маршрутизатором будет доступна.

Передающему узлу **следует** поместить в очередь пакет, вызвавший соединение и передать его после организации соединения. Если организация соединения со следующим маршрутизатором завершится отказом, передающему узлу

нужно отбросить пакет, вызвавший организацию соединения. Сохраняющийся отказ при создании VC со следующим маршрутизатором будет обнаруживаться и обрабатываться сетевым уровнем IPv6 через NUD.

В настоящее время правила сопоставления исходящих пакетов с VC заданы лишь на основании адреса получателя.

4.4.2. Передача групповых данных

Применяется следующий интервал уровня IP (next hop) для каждого исходящего группового пакета IPv6, указывающий VC типа pt-pt или pt-mpt[1] для пересылки пакета.

Для сетей NBMA с инкапсуляцией LLC/SNAP (например. ATM или SMDS) пакеты IPv6 **нужно** инкапсулировать с показанным ниже заголовком.

```
[0xAA-AA-03][0x00-00-5E][0x00-01][pkt$cmi][0x86DD][IPv6 packet]
    (LLC)       (OUI)      (PID)    (инкапсуляция MARS)
```

Значение Cluster Member ID драйвера IPv6/NBMA **нужно** скопировать в 2-октетное поле pkt$cmi до передачи.

Для сетей NBMA, не применяющих инкапсуляцию LLC/SNAP, **нужно** задавать соответствующие правила в спецификации NBMA, указав механизм передачи Cluster Member ID драйвера IPv6/NBMA.

Если получатель пакета указан одним из приведённых ниже групповых адресов, пакет **нужно** передать напрямую через pt-pt VC драйвера IPv6/NBMA в MARS.

- Solicited-node с областью действия Link-Local;
- All-nodes с областью действия Link-Local;
- All-routers с областью действия Link-Local;
- групповой адрес сервера или ретранслятора DHCP-v6.

Протоколу MARS **нужно** распространить пакет IPv6 в соответствии с параграфом 3.1.1 (если VC к MARS отключён по тайм-ауту, соединение **должно** быть восстановлено до передачи пакета протоколу MARS).

Если получатель пакета указан иным адресом, используются обычные механизмы клиента MARS в драйвере IPv6/NBMA для выбора и/или организации pt-mpt VC с целью передачи пакета.

В настоящее время правила сопоставления исходящих пакетов с VC заданы лишь на основании адреса получателя.

4.5. Приём данных

Пакеты, полученные с инкапсуляцией, описанной в параграфе 4.4.1, **нужно** декапсулировать и передать уровню IPv6, который определит дальнейшую обработку пакетов. Для пакетов, полученных с инкапсуляцией, описанной в параграфе 4.4.2, **нужно** сравнивать поле pkt$cmi с идентификатором CMI локального драйвера IPv6/NBMA и при совпадении значений пакет **нужно** просто отбрасывать. В ином случае пакет **нужно** декапсулировать и передать уровню IPv6, который определит дальнейшую обработку пакетов.

Для сетей NBMA, не использующих инкапсуляцию LLC/SNAP, **нужно** задать правила в соответствующей спецификации NBMA.

Драйверу IPv6/NBMA **не нужно** пытаться фильтровать прибывающие групповые пакеты с инкапсуляцией, заданной для индивидуальных пакетов IPv6, или индивидуальные пакеты IPv6 с заданной для групповых пакетов инкапсуляцией.

4.6. Организация и освобождение VC для индивидуальных данных

Индивидуальные VC поддерживаются отдельно от групповых. Организация и поддержка групповых VC обслуживается клиентом MARS в каждом драйвере IPv6/NBMA [5].Здесь описана лишь организация и поддержка pt-pt VC для индивидуального трафика IPv6 и рассматриваются лишь VC класса best effort. Создание VC для других классов выходит за рамки этого документа.

Перед отправкой пакет новому адресату в том же LL узел сначала выполняет процедуру ND для цели внутри LL. Это нужно для определения по адресу получателя IPv6 адреса канального уровня, который отправитель может использовать для передачи индивидуального пакета.

В Приложении A.1.1 приведено ненормативное описание обмена NS/NA и возможная организация нового SVC.

Сообщение Redirect (перенаправление на узел того же LL или в перемычку к узлу вне LL) приводит к созданию передающим (перенаправленным) узлом нового pt-pt VC к другому принимающему узлу. В сообщение Redirect **нужно** включать адрес канального уровня (NBMA) нового приёмного интерфейса IPv6/NBMA. Перенаправленный узел не заботится о местоположении нового принимающего узла в сети NBMA и будет создавать соединение pt-pt VC с новым узлом, если его ещё нет. Затем перенаправленный узел использует новое соединение VC для передачи данных вместо прежнего VC.

Перенаправление является односторонним. Даже после того, как отправитель отреагировал на перенаправление, получатель будет отправлять пакеты IPv6 перенаправленному узлу по прежнему пути. Это обусловлено тем, что у него нет возможности определить адрес IPv6 на другой стороне нового VC без ND. Таким образом, перенаправление не будет приводит к использованию нового VC обеими сторонами. Перенаправления IPv6 не предназначены для обеспечения симметрии. Если узел без перенаправления в конечном итоге тоже будет перенаправлен, он **может** обнаружить имеющееся соединение VC с целевым узлом и использовать его без создания нового VC.

Желательно освобождать VC, ставшие ненужными. Драйверу IPv6/NBMA **нужно** освобождать любые VC после бездействия в течение 20 минут. Это время может быть сокращено в соответствующей спецификации NBMA. Если запись в кэше адресатов или соседей очищается, связанные с ней VC также следует освободить. Если для записи в кэше соседей установлен статус STALE, связанные с ней VC **следует** освободить.

[1]Point-to-multipoint - соединение «один со многими».

4.7. Поддержка сигнализации NBMA SVC и MTU

Механизмы сигнализации для организации и разрыва SVC типов pt-pt и pt-mpt для разных сетей NBMA **нужно** задавать в соответствующих документах.

Поскольку данный драйвер IPv6/NBMA не знает, относится ли удалённая сторона VC в тому же LL, **драйверам нужно** реализовать зависящие от NBMA механизмы согласования MTU на уровне VC, которые **нужно** задавать в соответствующих документах.

Однако драйверы IPv6/NBMA могут предполагать, что они всегда взаимодействуют с другим драйвером, подключённым к такой же сети NBMA (например, драйверу IPv6/NBMA не требуется рассматривать возможность организации VC-перемычки напрямую к драйверу IPv6/FR).

5. Маркеры интерфейсов, опции Link Layer Address, адреса Link-Local

5.1. Маркеры интерфейсов

Каждый интерфейс IPv6 должен иметь маркер, по которому можно автоматически настроить адрес IPv6. Этот маркер должен быть уникальным в рамках LL для предотвращения дубликатов адресов при автоматической настройке без учёта состояния. Если два узла на одном LL создают одинаковые идентификаторы интерфейсов в канал, один из интерфейсов **должен** выбрать другой маркер. Все реализации **должны** поддерживать ручную настройку маркеров интерфейсов, позволяющую оператору задавать маркеры на уровне LL. Операторы могут выбрать ручную настройку не только с целью предотвращения дубликатов.

Все маркеры интерфейсов **должны** иметь размер 64 бита и описанный ниже формат на основе идентификаторов EUI-64 [10], как описано в Приложении A к [19].

5.1.1. Один LL на интерфейсе NBMA

Физические интерфейсы NBMA обычно будут иметь некий локальный идентификатор, который можно использовать для создания маркера интерфейса IPv6/NBMA. Механизмы генерации таких маркеров **нужно** задавать в соответствующих документах NBMA.

5.1.2. Несколько LL на интерфейсе NBMA

Физический интерфейс NBMA **может** служить для организации нескольких логических интерфейсов NBMA. Поскольку каждый логический интерфейс NBMA **может** поддерживать независимый интерфейс IPv6, возможны 2 варианта.

- Один хост с раздельными интерфейсами IPv6/NBMA в несколько независимых каналов LL.
- Набор из 2 или более «виртуальных хостов» (vhost), использующих общий драйвер NBMA. Каждый vhost может организовать свои интерфейсы IPv6/NBMA, связанные с одним или разными LL. Однако для vhost применимо требование к обычным хостам, не позволяющее двум интерфейсам в один канал LL иметь одинаковые маркеры.

В первом варианте каждый интерфейс IPv6/NBMA связан со своим LL и внешняя идентификация может различаться по префиксу маршрутизации LL, т. е. хост может применять один уникальный маркер интерфейса для всех своих интерфейсов IPv6/NBMA (на внутреннем уровне хост будет помечать пакеты теми или иными локальными тегами для идентификации принявшего пакет интерфейса IPv6/NBMA, однако это относится к IPv6 и не требует специального рассмотрения здесь).

Второй вариант сложнее и встречается более редко.

При поддержке нескольких логических интерфейсов NBMA на одном физическом интерфейсе NBMA независимые и уникальные идентификаторы **нужно** генерировать для каждого виртуального интерфейса NBMA, чтобы можно было создать уникальные маркеры интерфейсов IPv6/NBMA. Механизм генерации маркеров **нужно** задавать в соответствующих документах NBMA.

5.2. Опции адреса канального уровня

В ND определены 2 поля для передачи адресов отправителя и получателя на канальном уровне. Между интерфейсами IPv6/NBMA эти опции передаются в формате, заимствованном из спецификаций MARS [5] и NHRP [8].

```
[Type][Length][NTL][STL][..NBMA Number..][..NBMA Subaddress..]
| Фиксирован ||            Адрес канального уровня
|
```

[Type]
Однооктетное поле - 1 для адреса отправителя, 2 для адреса получателя.

[Length]
Однооктетное поле размера. Общий размер опции кратен 8 октетам с добавлением нулей в конце для выравнивания по 8-октетной границе.

[NTL]
Один октет с полями Number Type и Length.

[STL]
Один октет с полями SubAddress Type и Length.

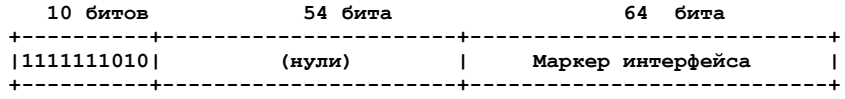
[NBMA Number]
Поле переменного размера, содержащее основной адрес NBMA.

[NBMA Subaddress]
Необязательное поле переменного размера, с субадресами NBMA.

При отсутствии [NBMA Subaddress] опция завершается после поля [NBMA Number] и нулей, добавляемых для выравнивания по 8-байтовой границе. Содержимое и интерпретация полей [NTL], [STL], [NBMA Number], [NBMA Subaddress] зависит от типа сети NBMA и их **нужно** описывать в соответствующих документах.

5.3. Адреса Link-Local

Адрес IPv6 Link-Local формируется добавлением маркера интерфейса, описанного выше, к префиксу FE80::/64.



6. Заключение и нерешенные вопросы

В этом документе описана общая архитектура IPv6 в сетях NBMA и он служит основой для дополнительных документов, описывающих конкретные технологии NBMA (такие как ATM или Frame Relay). Архитектура IPv6 в сети NBMA обеспечивает возможность работы на стороне хоста традиционного протокола IPv6 Neighbor Discovery с поддержкой коротких путей пересылки NBMA (перемычек) при доступности динамической сигнализации NBMA.

Канал IPv6 (Link) обобщён до логического канала (Logical Link или LL) по аналогии с логическими подсетями IPv4. Протокол MARS дополнен и применяется для обеспечения сравнительно эффективной групповой адресации в LL для пакетов IPv6 и распространения сообщений Discovery. Поддерживаются короткие пути NBMA на основе обнаружения потоков маршрутизатором или по запросам хостов. протокол ND не изменён для управления внутри LL (включая обнаружение перемычек NBMA). Протокол NHRP применяется между маршрутизаторами для получения сопоставлений адресов IPv6/NBMA с целями коротких соединений вне LL.

7. Вопросы безопасности

Эта архитектура не задаёт новых протоколов, но зависит от существующих (NHRP, IPv6, ND, MARS) и поэтому к ней применимы угрозы, наследуемые от этих протоколов. Архитектуру не следует применять в доменах, где любой из указанных протоколов не считается достаточно безопасным. Однако сама по себе архитектура не добавляет угроз безопасности.

Хотя это предложение не вносит новых механизмов защиты, все имеющиеся механизмы безопасности IPv6 будут без изменений работать с NBMA. Это включает аутентификацию и шифрование для протокола ND и обмена пакетами данных IPv6. Протокол MARS изменён так, что это не влияет на свойства безопасности RFC 2022.

Благодарности

Eric Nordmark подтвердил полезность сообщений ND Redirect в частной переписке по электронной почте в марте 1996 г. Обсуждение с членами рабочей группы ION с июня по декабрь 1996 г. помогло укрепить описанную здесь архитектуру. Первоначальная работа Grenville Armitage по IPv6/NBMA была выполнена в Bellcore. Элементы раздела 5 заимствованы из записки Matt Crawford о работе IPv6 в сетях Ethernet.

Адреса авторов

Grenville Armitage

Bell Laboratories, Lucent Technologies

101 Crawfords Corner Road

Holmdel, NJ 07733

USA

EMail: gja@lucent.com

Peter Schulter

Bright Tiger Technologies

125 Nagog Park

Acton, MA 01720

EMail: paschulter@acm.org

Markus Jork

European Applied Research Center

Digital Equipment GmbH

CEC Karlsruhe

Vincenz-Priessnitz-Str. 1

D-76131 Karlsruhe

Germany

EMail: jork@kar.dec.com

Geraldine Harter

Digital UNIX Networking

Compaq Computer Corporation

110 Spit Brook Road

Nashua, NH 03062

EMail: harter@zk3.dec.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Литература

[1] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460[1], December 1998.

[2] ATM Forum, "ATM User Network Interface (UNI) Specification Version 3.1", ISBN 0-13-393828-X, Prentice Hall, Englewood Cliffs, NJ, June 1995.

[3] Crawford, M., "A Method for the Transmission of IPv6 Packets over Ethernet Networks", RFC 1972, August 1996.

[4] Heinanen, J., "Multiprotocol Encapsulation over ATM Adaptation Layer 5", RFC 1483[2], July 1993.

[5] Armitage, G., "Support for Multicast over UNI 3.1 based ATM Networks", RFC 2022, November 1996.

[6] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[7] Narten, T., Nordmark, E. and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461[3], December 1998.

[8] Luciani, J., Katz, D., Piscitello, D. Cole B and N. Doraswamy, "NBMA Next Hop Resolution Protocol (NHRP)", RFC 2332, April 1998.

[9] Thomson, S. and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC 2462, December 1998.

[10] "64-Bit Global Identifier Format Tutorial", http://standards.ieee.org/db/oui/tutorials/EUI64.html[4].

[11] Katsube, Y., Nagami, K. and H. Esaki, "Toshiba's Router Architecture Extensions for ATM : Overview", RFC 2098, February 1997.

[12] P. Newman, T. Lyon, G. Minshall, "Flow Labeled IP: ATM under IP", Proceedings of INFOCOM'96, San Francisco, March 1996, pp.1251-1260

[13] Piscitello, D. and J. Lawrence, "The Transmission of IP Datagrams over the SMDS Service", RFC 1209, March 1991.

[14] Plummer, D., "An Ethernet Address Resolution Protocol - or - Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, November 1982.

[15] McCann, J., Deering, S. and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981[5], August 1996.

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

[17] Armitage, G., Schulter, P. and M. Jork, "IPv6 over ATM Networks", RFC 2492, January 1999.

[18] C. Perkins, J. Bound, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", Work in Progress[6].

[19] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 2373[7], July 1998.

Приложение A. Описание работы протокола IPv6

Модель IPv6 в сети NBMA, описанная в этом документе, полностью поддерживает семантику протокола IPv6 и не требует изменения сетевого уровня IPv6. Поскольку концепция защищённых связей (security association) не изменялась для NBMA, эта модель полностью поддерживает механизмы и свойства защиты IPv6. Это позволяет узлам IPv6 выбирать отклики на запросы на основе сведений о безопасности, как это делается для других каналов данных, полностью поддерживая семантику ND, поскольку запрошенный узел сам решает, отвечать ли на запрос и что передавать в ответ. Таким образом, сеть NBMA будет прозрачная для сетевого уровня за исключением случаев предоставления других услуг (например, QoS VC).

В оставшейся части этого Приложения описана работа протоколов ядра IPv6 с описанной здесь моделью.

A.1. Операции обнаружения соседей

Перед выполнением любых операций ND узел должен сначала присоединиться к группе all-nodes (все узлы) и своему групповому адресу solicited-node (использование этого адреса для DAD описано в A.1.4). Сетевой уровень IPv6 будет присоединять эти multicast-группы, как описано в параграфе 4.2.

[1] Заменён RFC 8200. *Прим. перев.*
[2] Заменён RFC 2684. *Прим. перев.*
[3] Заменён RFC 4861. *Прим. перев.*
[4] Приведённая ссылка устарела, но документ доступен по другой ссылке. *Прим. перев.*
[5] Заменён RFC 8201. *Прим. перев.*
[6] Документ опубликован в RFC 3315 и позднее заменён RFC 8415. *Прим. перев.*
[7] Заменён RFC 3513, а тот - RFC 4291. *Прим. перев.*

A.1.1. Распознавание адреса

Хост IPv6 распознает адреса путём передачи сообщений Neighbor Solicitation (NS) по групповому адресу solicited-node целевого хоста, как описано в [7]. Сообщение NS содержит опцию Source Link-Layer Address с адресом NBMA запрашивающего узла на канале LL.

Когда драйвер IPv6/NBMA локального узла передаёт сообщение NS от сетевого уровня IPv6, он следует описанной в параграфе 4.4.2 процедуре групповой передачи данных. Сообщение NS получает один или несколько узлов, обрабатывая данные в соответствии с параграфом 4.5 и передавая декапсулированный пакет сетевому уровню IPv6.

Если приёмный узел является целью NS, он обновляет свой кэш соседей, внося в него адрес запрашивающего узла NBMA из опции Source Link-Layer Address в NS, как описано в [7]. Запрошенный хост IPv6 отвечает на NS анонсом Neighbor Advertisement (NA) по индивидуальному адресу IPv6 запрашивающего узла. Сообщение NA включает опцию Target Link-Layer Address Option с адресом NBMA запрашивающего узла на канале LL.

Драйвер IPv6/NBMA запрашиваемого узла передаёт NA и адрес канального уровня запрашивающего узла от сетевого уровня IPv6. Затем он выполняет действия, указанные в параграфе 4.4.1 для передачи сообщения NA запрашивающему узлу. Это создаст pt-pt VC между запрашивающим и запрашиваемым узлом, если его ещё нет.

Запрашивающий узел получает сообщение NA по новому pt-pt VC, декапсулирует его и передаёт сетевому уровню IPv6 для обработки, описанной в параграфе 4.5. Затем запрашивающий узел делает соответствующие записи в своём кэше соседей, включая кэширование адреса канального уровня NBMA запрашиваемого узла, как описано в [7].

С этого момента каждая из систем имеет полную запись в кэше соседей для другой системы. Они могут обмениваться данными через соединение pt-pt VC, созданное запрашиваемым узлом при возврате NA, или создать новое соединение VC.

Хост IPv6 может также передать незапрошенный анонс NA по групповому адресу all-nodes. Когда драйвер IPv6/NBMA локального узла передаёт NA от сетевого уровня IPv6, он выполняет действия, указанные в параграфе 4.4.2, для отправки сообщения NA по групповому адресу all-nodes. Каждый узел обрабатывает входящий пакет, как описано в параграфе 4.5, и передаёт его сетевому уровню IPv6, где сообщение обрабатывается в соответствии с [7].

A.1.2. Обнаружение маршрутизаторов

Механизм Router Discovery описан в [7] и для его поддержки маршрутизатор IPv6 присоединяет групповой адрес IPv6 all-routers. При получении драйвером IPv6/NBMA запроса JoinLocalGroup от сетевого уровня IPv6 Network он будет следовать процедуре, описанной в параграфе 4.2.

Маршрутизаторы IPv6 периодически отправляют незапрошенные анонсы Router Advertisement (RA), указывающие их доступность на канале LL. При отправке маршрутизатором IPv6 незапрошенного RA он передаёт пакет данных по групповому адресу IPv6 all-nodes. Когда драйвер IPv6/NBMA локального узла получает сообщение RA от сетевого уровня IPv6, он передаёт его в соответствии с параграфом 4.4.2. MARS передаёт пакет в ClusterControlVC на канале LL и оно приходит всем узлам LL. Каждый узел на канале LL обрабатывает входящий пакет в соответствии с параграфом 4.5 и передаёт его сетевому уровню IPv6 для дальнейшей обработки.

Для выполнения Router Discovery хост IPv6 передаёт сообщение Router Solicitation (RS) по групповому адресу all-routers. Когда драйвер IPv6/NBMA локального узла получает от сетевого уровня IPv6 запрос на отправку пакета, он следует процедурам параграфа 4.4.2. Сообщение RS будет передано узлам, присоединившимся к группе all-routers, либо всем узлам. Получившие сообщение RA узлы обрабатывают его в соответствии с параграфом 4.5 и передают сетевому уровню IPv6 для дальнейшей обработки. Лишь маршрутизаторы обрабатывают сообщение и отвечают на него. Маршрутизатор IPv6 отвечает на RS передачей анонса RA по групповому адресу IPv6 all-nodes, если адресом отправителя RS был неуказанный (unspecified) адрес. Если отправитель RS указал иной адрес, маршрутизатор будет возвращать запросившему узлу индивидуальное сообщение RA. При отправке RA по групповому адресу all-nodes маршрутизатор следует процедуре, описанной выше для незапрошенных сообщений RA.

Если анонс RA передан индивидуально запросившему узлу, сетевой уровень IPv6 предоставит драйверу IPv6/NBMA узла сообщение RA и адрес канального уровня запросившего узла (определяется процедурой распознавания адреса при необходимости) и пакет будет передан в соответствии с параграфом 4.4.1. Это приведёт к созданию нового pt-pt VC между маршрутизатором и запросившим узлом, если такого соединения ещё не было.

Запрашивающий узел принимает и обрабатывает анонсы RA в соответствии с параграфом 4.5 и передаёт RA сетевому уровню IPv6. В зависимости от состояния записи в кэше соседей сетевой уровень IPv6 обновляет адрес NBMA для маршрутизатора в записи кэша в соответствии с опцией Source Link-Layer Address в анонсе RA.

Если в процессе обнаружения маршрутизатора создано соединение pt-pt VC, последующие индивидуальные данные IPv6 best effort между маршрутизатором и запросившим узлом будут передаваться через это соединение.

A.1.3. Обнаружение недоступности соседа (NUD)

Обнаружение недоступности соседа (Neighbor Unreachability Detection или NUD) служит хосту IPv6 для определения потери связности с соседом, как описано в [7]. Каждая запись кэша соседей содержит сведения, используемые алгоритмом NUD для обнаружения отказов доступности. Подтверждение доступности соседа приходит от вышележащего уровня (получение тем недавно переданных данных) или в результате приёма анонса NA в ответ на NS.

Отказы связности на уровне драйвера IPv6/NBMA узла, такие как разрыв VC (см. параграф 4.6) и невозможность создать VC с соседом (см. параграф 4.4.1), обнаруживаются и обрабатываются сетевым уровнем IPv6 с помощью NUD. Драйвер IPv6/NBMA на узле не пытается детектировать или корректировать связность.

Сохраняющиеся отказы при создании VC от хоста IPv6 к одному из его соседей IPv6 обнаруживаются и обрабатываются с помощью NUD. При каждой попытке передачи данных от хоста IPv6 соседу драйвер IPv6/NBMA на хосте пытается создать VC с соседом и при отказе отбрасывает пакет. Таймеры подтверждения доступности IPv6 в конечном итоге достигают заданного порога и запись кэше для соседа будет переведена в состояние PROBE. Это заставляет хост IPv6 передать соседу индивидуальный запрос NS, который будет отброшен локальным драйвером

IPv6/NBMA после отказа при организации VC. В результате хост IPv6 не получит запрошенного анонса NA, требуемого для подтверждения доступности, и это вынудит хост удалить из кэше запись для данного соседа. При следующей попытке хоста IPv6 передать данные этому соседу будет выполняться распознавание адреса. В зависимости от причины предшествующего отказа (например, когда прежнее соединение VC было разорвано в результате устаревания адреса канального уровня в кэше соседей) связность с соседом может быть организована заново.

В случае разрыва VC с соседом IPv6 при следующей передаче пакета от хоста IPv6 соседу драйвер IPv6/NBMA на узле увидит отсутствие VC к этому соседу и попытается организовать новое соединение VC с ним. Если при первой или последующих передачах узел не сможет создать VC к соседу, NUD обнаружит и обработает отказ, как описано выше (обработка сохраняющихся отказов при создании VC от хоста IPv6 к одному из его соседей IPv6). Восстановление связности зависит от причин предшествующего отказа.

A.1.4. Обнаружение дубликатов адресов (DAD)

Хост IPv6 проверяет наличие дубликата (Duplicate Address Detection или DAD) для определения возможности использования желаемого адреса на LL (предварительный адрес), как описано в [9] и [7]. Процедура DAD выполняется для всех адресов, которые хочет использовать хост, независимо от способа их настройки.

До выполнения DAD хост присоединяет групповые адреса all-nodes и solicited-node, соответствующие предварительному адресу хоста (см. параграф 4.2. Подключение к Multicast-группе). Хост IPv6 запускает DAD, отправляя сообщение NS по групповому адресу solicited-node, соответствующему предварительному адресу хоста, с указанием предварительного адреса как цели. Когда драйвер IPv6/NBMA локального узла получает сообщение NS от сетевого уровня IPv6, он следует процедурам параграфа 4.4.2. Сообщение NS передаётся узлам, подключившимся к группе solicited-node, или всем узлам. Сообщение DAD NS получает один или несколько узлов на канале LL и каждый узел обрабатывает его в соответствии с параграфом 4.5. Отметим, что клиент MARS на передавшем узле фильтрует сообщение, поэтому уровень IPv6 на этом узле не получит сообщения. Сетевой уровень IPv6 на узлах, не входящих в целевую группу solicited-node, будет отбрасывать сообщение NS.

Если ни один хост не присоединился к групповому адресу solicited-node, соответствующему предварительному адресу, хост не получит сообщения NA с предварительным адресов как целью. Хост повторит передачу в соответствии с логикой [9], прервёт DAD и назначит предварительный адрес интерфейсу NBMA. В противном случае, когда другие хосты на LL присоединили групповой адрес solicited-node, соответствующий предварительному адресу, они будут обрабатывать запрос NS. Результат будет зависеть от того, считает ли принявший сообщение хост IPv6 целевой адрес предварительным. Если для принявшего хоста IPv6 адрес не является предварительным, хост будет передавать анонс NA с целевым адресом. Поскольку источником NS является неуказанный адрес, сообщение NA передаётся по групповому адресу all-nodes в соответствии с параграфом 4.4.2. Сообщение DAD NA будет принято и обработано клиентом MARS на всех узлах канала LL, как описано в параграфе 4.5. Отметим, что передающий узел отфильтрует входящее сообщение, поскольку CMI в его заголовке совпадает с идентификатором узла. Все другие узлы декапсулируют сообщение и передают его сетевому уровню IPv6. Хост, запустивший DAD, увидит, что адрес является целью в NA, и поймёт, что этот адрес не является уникальным на канале и не может быть назначен интерфейсу NBMA.

Если адрес принимающего хоста IPv6 является предварительным, это говорит о совпадении предварительных адресов этого хоста и инициатора DAD. Принимающий хост понимает, что предварительный адрес неуникален и не может назначаться его интерфейсу NBMA.

A.1.5. Обработка Redirect

Маршрутизатор IPv6 использует сообщения Redirect для информирования хоста IPv6 о лучшем первом маршрутизаторе (first-hop) для доступа к конкретному адресату, как описано в [7]. Это можно использовать для направления хоста к лучшему первому маршрутизатору, другому хосту на том же канале LL или временному соседу на другом LL. Маршрутизатор IPv6 передаёт индивидуальное сообщение Redirect по адресу отправителя IPv6, вызвавшего Redirect. Драйвер IPv6/NBMA в маршрутизаторе передаёт сообщение Redirect в соответствии с параграфом 4.4.1 и это ведёт к созданию VC между маршрутизатором и перенаправляемым хостом, если его не было.

Драйвер IPv6/NBMA на хосте IPv6, вызвавшем Redirect получит инкапсулированное сообщение Redirect через одно из своих соединений pt-pt VC. Он декапсулирует пакет и передаёт Redirect сетевому уровню IPv6 в соответствии с параграфом 4.5. Последующие данные от этого хоста IPv6 к адресату передаются по адресу next-hop из сообщения Redirect. Для сетей NBMA в Redirect следует включать опцию Link-Layer Address, описанную в [7] и параграфе 5.2, что позволяет перенаправленному узлу не использовать NS для определения адреса канального уровня, на который он перенаправлен. Таким образом, перенаправление может быть задано для любого узла сети NBMA, независимо от LL целевого узла. Это позволяет перенаправлять хосты NBMA за пределы их LL через перемычки с использованием стандартных протоколов IPv6.

После перенаправления сетевой уровень IPv6 предоставляет драйверу IPv6/NBMA пакет IPv6 и адрес канального уровня next-hop, которому он отправляет данные для перенаправленного адресата. Драйвер IPv6/NBMA на узле определяет наличие VC к next-hop. При наличии pt-pt VC драйвер IPv6/NBMA помещает пакет данных в очередь и запускает организацию VC к адресату. Когда соединение VC уже имеется или создано, узел инкапсулирует исходящий пакет и передаёт его в VC.

Отметим, что сообщения Redirect являются однонаправленными. Перенаправленный хост создаёт VC к next-hop в соответствии с сообщением Redirect, но next-hop не перенаправляется на хост-источник. Поскольку ND не применяется, у next-hop может не быть способа определить вызывающего при получении нового VC. По этой же причине next-hop не знает о событиях, которые могут обновлять кэш соседей и адресатов. Пакеты для перенаправленного хоста будут передаваться по прежнему пути. Узлу next-hop следует иметь возможность использовать новое соединение VC для перенаправленного адресата, если он тоже получает Redirect для перенаправленного узла. Это поведение согласуется с [7].

A.2. Настройка адреса

Адреса IPv6 настраиваются автоматически с использованием механизмов настройки с учётом или без учёта состояния, как описано в [9] и [18]. Процесс автоматической настройки IPv6 включает создание и проверку уникальности адреса

Link-Local на канале LL, определение возможности механизмов настройки с учётом или без учёта состояния для получения адреса, а также определение другой информации (кроме адреса) для настройки. Адреса IPv6 можно настраивать вручную, если, например, при автоматической настройке возникает отказ в результате совпадения адресов Link-Local. Администратор LL задаёт тип используемой автоматической настройки и хосты на канале LL получают эти сведения в анонсах RA.

В следующих параграфах рассматривается автоматическая настройка с учётом и без учёта состояния, а также настройка адресов вручную для работы в среде IPv6/NBMA.

A.2.1. Настройка адреса без учёта состояния

Настройка адреса IPv6 без учёта состояния позволяет хосту IPv6 автоматически получить адреса для своих интерфейсов в соответствии с [IPV6-ADDRCONF]. При первом запуске хоста IPv6 он генерирует адрес Link-Local для интерфейса, подключённого к LL. Затем проверяется уникальность этого адреса с использованием DAD. Если хост IPv6 обнаруживает неуникальность адреса Link-Local, процесс автоматической настройки прерывается и хост IPv6 нужно настраивать вручную.

Если выбранный хостом IPv6 адрес Link-Local оказался уникальным, он назначается интерфейсу на канале LL и хост выполняет поиск маршрутизатора (Router Discovery) для получения данных автоматической настройки. Хост IPv6 передаёт сообщение RS и получает анонс RA или ждёт незапрошенного анонса RA. Хост IPv6 обрабатывает биты M и O из RA, как описано в [9] и по результату может вызвать автоматическую настройку с учётом состояния.

Если на канале LL нет маршрутизаторов, хост IPv6 может взаимодействовать с другими хостами IPv6 на LL, используя адрес Link-Local. Хост IPv6 получает адрес канального уровня для соседей с помощью распознавания адресов (Address Resolution). Хост IPv6 также пытается вызвать автоматическую настройку с учётом состояния, если это явно не отключено.

A.2.2. Настройка адреса с учётом состояния (DHCP)

Хосты IPv6 используют протокол динамической настройки (Dynamic Host Configuration Protocol или DHCPv6) для настройки адресов с учётом состояния, как описано в [18]. Сервер DHCPv6 или ретранслятор присутствует на канале LL, настроенном вручную или автоматически с учётом состояния. Сервер или ретранслятор DHCPv6 присоединяется к группе IPv6 DHCPv6 Server/Relay-Agent на канале LL. Когда драйвер IPv6/NBMA на узле получает запрос JoinLocalGroup от сетевого уровня IPv6, он следует процедуре, описанной в параграфе 4.2.

Хост IPv6 вызывает автоматическую настройку с учётом состояния, если биты M и O в анонсе RA указывают это, а также может вызывать такую настройку при отсутствии маршрутизаторов на канале LL. Хост IPv6, получающий параметры конфигурации от механизма с учётом состояния далее называется клиентом DHCPv6.

Клиент DHCPv6 передаёт сообщение DHCPv6 Solicit по групповому адресу DHCPv6 Server/Relay-Agent для нахождения агента DHCPv6. Когда драйвер IPv6/NBMA запрашивающего узла получает от сетевого уровня IPv6 запрос на передачу пакета, он следует процедуре из параграфа 4.4.2. Это ведёт к получения сообщения одним или несколькими узлами на канале LL. Каждый узел, получивший запрос, обрабатывает его в соответствии с параграфом 4.5. Однако воспринимают и полностью обрабатывают запрос лишь сетевые уровни IPv6 серверов или агентов DHCPv6.

Сервер или ретранслятор DHCPv6 на канале LL передаёт индивидуальное сообщение DHCPv6 Advertisement клиенту DHCPv6. Сетевой уровень предоставит драйверу IPv6/NBMA на узле пакет и адрес канального уровня клиента DHCPv6 (при необходимости определяется с помощью ND). Затем драйвер IPv6/NBMA передаёт пакет, как указано в параграфе 4.4.1. Это ведёт к созданию pt-pt VC между сервером и клиентом, если соединения ещё не было.

Драйвер IPv6/NBMA клиента DHCP получает инкапсулированный пакет от сервера или ретранслятора DHCP, как описано в параграфе 4.5. Узел декапсулирует групповой пакет и затем передаёт его сетевому уровню IPv6 для обработки. Сетевой уровень IPv6 доставляет сообщение DHCPv6 Advertise клиенту DHCPv6.

Другие сообщения DHCPv6 (Request, Reply, Release, Reconfigure) передаются по индивидуальным адресам между клиентом и сервером DHCPv6. В зависимости от доступности адреса клиента DHCPv6 сообщения между клиентом DHCPv6 и сервером DHCPv6 на другом канале LL передаются через маршрутизатор или ретранслятор (DHCPv6 Relay-Agent). Перед отправкой сообщения DHCPv6 сетевой уровень IPv6 выполняет процедуру ND (при необходимости) для определения адреса канального уровня next-hop для пакета. Организуется соединение pt-pt VC между отправителем и next-hop, через которое передаётся инкапсулированный пакет, как описано в параграфе 4.4. Передача данных.

A.2.3. Настройка адреса вручную

Хост IPv6 можно настроить вручную, если он с помощью DAD определяет, что его адрес Link-Local не уникален. После установки на хосте IPv6 уникального маркера интерфейса можно вызвать механизм автоматической настройки.

A.3. Протокол управления группами Internet (IGMP)

Групповые маршрутизаторы IPv6 используют протокол IGMPv6 для периодического определения принадлежности локальных хостов к группам. В описанной здесь схеме протоколы IGMPv6 можно применять без изменения NBMA. Хотя эти протоколы могут быть не самыми эффективными в данной среде, они работают, как описано ниже. Однако multicast-маршрутизаторы IPv6, подключённые к NBMA LL, могут оптимизировать функции IGMP, передавая сообщения MARS_GROUPLIST_REQUEST обслуживающему LL серверу MARS и определяя принадлежность к группам их сообщений MARS_GROUPLIST_REPLY. Запросы к MARS о принадлежности к multicast-группам являются необязательным расширением и не требуются маршрутизаторам для определения членства в группах IPv6 на LL.

Имеется 3 типа сообщений ICMPv6, передающий сведения о принадлежности к multicast-группам, - Group Membership Query, Group Membership Report, Group Membership Reduction. IGMPv6 работает без изменений в описанной здесь архитектуре IPv6/NBMA.

Групповой маршрутизатор IPv6 получает все multicast-пакеты IPv6 на канале LL, присоединяясь ко всем multicast-группам в неразборчивом режиме (promiscuous) [5]. Сервер MARS заставляет multicast-маршрутизатор добавляться во

все имеющиеся и будущие групповые VC, после чего маршрутизатор IPv6 становится получателем всех групповых пакетов IPv6, переданных внутри LL.

Групповой маршрутизатор IPv6 обнаруживает группы, имеющие получателей на канале LL, периодически передавая сообщения Group Membership Query по групповому адресу IPv6 all-nodes. Когда драйвер IPv6/NBMA локального узла получает от сетевого уровня IPv6 запрос на передачу пакета Group Membership Query, он следует параграфу 4.4.2. Узел определяет, что пакет направлен по групповому адресу all-nodes и передаёт пакет клиенту MARS на узле, который инкапсулирует пакет и передаёт напрямую в MARS, где пакет ретранслируется всем узлам LL. Драйверы IPv6/NBMA на каждом узле получают пакет, декапсулируют его и передают сетевому уровню IPv6. Исходный узел фильтрует такие пакеты с помощью клиента MARS, поскольку его Cluster Member ID совпадает с CMI в заголовке инкапсуляции пакета MARS.

Хосты IPv6 на канале LL отвечают на Group Membership Query сообщением Group Membership Report для каждой группы IPv6, к которой данных хост присоединён. Хосты IPv6 передают Group Membership Report также при вхождении в группу IPv6 по адресу данной multicast-группы. Когда драйвер IPv6/NBMA локального узла получает от сетевого уровня IPv6 запрос на передачу пакета, он выполняет процедуру из параграфа 4.4.2. Узел понимает, что пакет передаётся по групповому адресу и пересылает его локальному клиенту MARS для отправки в подходящий канал VC.

Пакеты Group Membership Report будут поступать на каждый узел, входящий в соответствующую группу, через один из каналов VC, подключённых к каждому клиенту MARS на узле. Клиент MARS декапсулирует входящий пакет и передаёт его сетевому уровню IPv6 для обработки. Клиент MARS на передающем узле отфильтрует свой пакет при получении.

При выходе из группы хост IPv6 передаёт сообщение Group Membership Reduction по адресу покидаемой группы. Передача и приём сообщений Group Membership Reduction соответствуют операциям для Group Membership Report.

Приложение B. Модели поддержки Intra-LL ND в MARS

B.1. Упрощённый подход к использованию MARS

Драйвер IPv6/NBMA использует стандартный протокол MARS для организации петей пересылки VC с интерфейса, на котором он может передавать все групповые пакеты IPv6, включая ICMPv6. Пакеты IPv6 передаются, а затем принимаются предусмотренным набором адресатов и использованием отдельного pt-mpt VC для каждой целевой группы.

В этой модели все протокольные элементы [5] используются «как есть». Однако следует учитывать расход ресурсов для SVC. К сожалению, ND предполагает, что ресурсы multicast лучше всего сохранять за счёт генерации распределенных (неплотных) наборов групповых адресов solicited-nodes (по которым изначально передаются запросы обнаружения). Исходная цель состояла в минимизации числа непричастных устройств, которые одновременно получают сообщения, предназначенные кому-либо другому.

Однако в ориентированной на соединения среде NBMA также (или более) важно минимизировать число независимых VC, которые нужно начинать или завершать на данном интерфейсе NBMA. Если рассматривать службу MARS как «чёрный ящик», неплотное пространство адресов solicited-nodes ведёт к большому числу коротко используемых, но долго живущих pt-mpt VC (создаются при каждой передаче узлом запросов NS). Ещё хуже то, что эти VC полезны лишь для дополнительных пакетов, передаваемых по связанным с ними адресам solicited-nodes. Для фактической передачи индивидуального трафика IPv6, вызвавшего передачу NS, требуется новый канал pt-pt VC.

Неэффективность, связанная с неплотностью адресного пространства solicited-nodes ортогональна компромиссу между mesh-сетью VC и Multicast Server. Обычно сервер групповой адресации агрегирует потоки трафика одной multicast-группы в один канал VC. Для снижения числа VC, потребляемых ND, нужно агрегирование в пространстве адресов solicited-nodes, выполняемое по предназначению пакета, а не явному адресату IPv6. Компромисс здесь состоит в том, что агрегирование устраняет проблему рассеянности узлов в неплотном пространстве адресов solicited-nodes. Это плата за несоответствие ND ориентированным на соединения сетям.

B.2. MARS как Link (Multicast) Server

Одним из возможных механизмов агрегирования является захват каждым драйвером IPv6/NBMA групповых пакетов ICMPv6 с групповыми сообщениями ND или RD и логическое сопоставление их адресатов с группой all-nodes (на локальном канале). С учётом поддержки группы all-nodes в MCS, число VC внутри LL существенно снижается.

Дополнительной оптимизацией является захват драйвером IPv6/NBMA каждого узла групповых пакетов ICMPv6 с групповыми сообщениями ND или RD и передача их самому MARS для ретрансляции в ClusterControlVC (с тривиальным расширением для MARS). Этот подход основан на том, что в любом LL с поддержкой IPv6 multicast:

- узлы уже имеют pt-pt VC к своему MARS;
- MARS имеет pt-mpt VC (ClusterControlVC) ко всем членам кластера (члены LL, зарегистрировавшие поддержку multicast).

Поскольку каналы VC между MARS и клиентами MARS передают инкапсулированные пакет LLC/SNAP, пакеты ICMP можно мультиплексировать с обычными управляющими сообщениями MARS. По сути, MARS ведёт себя как multicast-сервер, для чужих пакетов (не MARS), получаемых по каналу LL.

Поскольку от клиента MARS не требуется воспринимать лишь управляющие сообщения MARS на канале ClusterControlVC, полученные таким путём пакеты ICMP могут передаваться уровню IP каждого узла без дополнительных комментариев. Внутри уровня IP происходит фильтрация на основе IP-адреса фактического получателя и отвечать будет лишь целевой узел.

К сожалению такой подход ведёт к тому, что все члены кластера получают разные сообщения ICMPv6, которые в большинстве случаев отбрасываются.

Приложение C. Обнаружение потоков

Связи между потоками пакетов IPv6, гарантиями QoS и оптимальным использованием ресурсов базовой сети IP и NBMA все ещё остаются предметом исследований IETF (в частности, рабочих групп ISSLL, RSVP, IPNG, ION). В этом документе описано лишь обнаружение потоков как средство оптимизации ресурсов сети NBMA за счёт организации коротких соединений между LL.

C.1. Использование ненулевого FlowID для подавления обнаружения потоков

Для описываемой архитектуры IPv6/NBMA потоком считается связанная последовательность пакетов IPv6, для которой первому маршрутизатору (first-hop) разрешено выполнять обнаружения потока для запуска обнаружения коротких соединений. Сопоставление пакетов с потоком (например, по общим полям заголовков IPv6, таким как адрес получателя) определяется локальной конфигурацией.

Правила обнаружения задают учёт лишь потоков с FlowID=0 при выборе потоков для запуска обнаружения перемычек. Обоснование этого приведено ниже.

- Перемычки NBMA служат для оптимизации пересылки пакетов IPv6 в отсутствие других указаний от хоста.
- Для хостов IPv6/NBMA желательно наличие механизма, позволяющего переопределить попытки «сети» оптимизировать свои пути пересылки.
- FlowID=0 означает для IPv6, что источник разрешает сети пересылать пакеты best-effort по своему усмотрению.
- Семантика IPv6 для FlowID=0 согласуется с правилом обнаружения потоков в этом документе, позволяющим оптимизировать для пакетов с FlowID=0 пути пересылки с использованием коротких соединений.
- Отличное от 0 значение FlowID означает в IPv6, что источник установил предпочтительное поведение пересылки для пакетов с таким FlowID.
- Семантика IPv6 для FlowID≠0 совместима с правилом обнаружения потоков в этом документе, запрещающим создание коротких соединений для пакетов с ненулевым FlowID.

Отличное от 0 значение FlowID может быть задано хостом-источником после согласования предпочтительного механизма пересылки с «сетью» (например, динамическими средствами, такими как RSVP, или административно). Кроме того, FlowID может выбираться источником случайно и сеть будет обеспечивать принятую по умолчанию пересылку best effort (по умолчанию маршрутизатор IPv6 обеспечивает пересылку best-effort для пакетов, где пара «FlowID - адрес отправителя» не распознана). Таким образом, этот документ поддерживает два режима работы.

FlowID=0
: Пересылка best effort с необязательным созданием коротких путей по обнаружению потока.

FlowID≠0
: Пересылка best effort, если на маршрутизаторах в пути не были заданы иные правила обработки для пары «FlowID - адрес отправителя». Обнаружение потоков для создания коротких путей не применяется. Если на маршрутизаторах заданы правила обработки для пары «FlowID - адрес отправителя», поток обслуживается по этим правилам без обнаружения потоков для создания коротких путей.

Механизмы установки правил поэтапной обработки пакетов с отличным от 0 FlowID не задаются и не подразумеваются в этом документе.

C.2. Будущие направления для обнаружения потоков

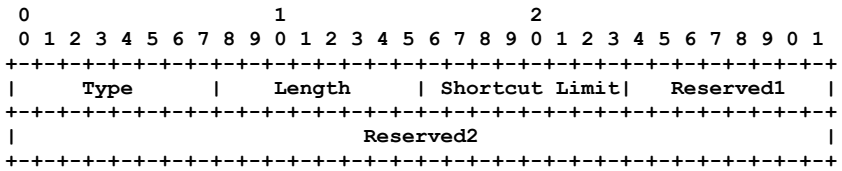
В будущем для точного сопоставления потоков IPv6 с NBMA VC может потребоваться обмен дополнительной информацией в процессе ND по сравнению с доступной сейчас в пакетах ND. В таких случаях протоколы IPv6 ND могут быть расширены путём включения новых опций TLV (см. параграф 4.6 в RFC 1970 [7]), однако такие опции должны быть согласованы с рабочей группой IPNG. Поскольку RFC 1970 задаёт для узлов игнорирование непонятных опций, такие опции можно добавлять в любой момент без нарушения совместимости с имеющимися реализациями.

В NHRP имеются механизмы добавления необязательных TLV в запросы и отклики NHRP. Развитие описанной в этом документе архитектуры потребует согласованных расширений QoS для ND и NHRP, чтобы обеспечить их семантическую эквивалентность (синтаксические различия нежелательны, но допустимы).

Поддержка QoS для индивидуальных потоков IPv6 не потребует расширения имеющегося протокола MARS, однако будущая поддержка QoS для групповых потоков IPv6 может потребовать расширений. Управляющие сообщения MARS используют такой же механизм расширения TLV, как и NHRP, что позволяет создавать расширения QoS при необходимости.

Приложение D. Опция Shortcut Limit

Для сообщений NS, переданных для создания перемычки, нужен новый тип опции ND для передачи сведений о максимальном числе пересылок потока данных между хостом и маршрутизатором. Использование этой опции ND определено в параграфе 3.2.2 этой спецификации. Её двоичное представление следует параграфу 4.6 в RFC 1970.



Type
: 6

Length
: 1

Shortcut Limit

8-битовое целое число без знака, ограничивающее число интервалов пересылки для попытки создать перемычку.

Reserved1

Неиспользуемое поле. При передаче **должно** устанавливаться в 0, а при получении **должно** игнорироваться.

Reserved2

Неиспользуемое поле. При передаче **должно** устанавливаться в 0, а при получении **должно** игнорироваться.

Опция Shortcut Limit применяется хостом в сообщении NS, передаваемом как триггер перемычки принятому по умолчанию маршрутизатору. Она ограничивает число интервалов пересылки (hop) при организации короткого соединения с целью. Число пересылок задаётся относительно запрашивающего перемычку хоста. Сообщения NS с предельным числом пересылок 0 или 1 игнорируются.