

Internet Engineering Task Force (IETF)
Request for Comments: 6325
Category: Standards Track
ISSN: 2070-1721

R. Perlman
Intel Labs
D. Eastlake 3rd
Huawei
D. Dutt
S. Gai
Cisco Systems
A. Ghanwani
Brocade
July 2011

Базовая спецификация протокола маршрутизирующих мостов Routing Bridges (RBridges): Base Protocol Specification

Аннотация

Маршрутизирующие мосты (RBridge¹) обеспечивают оптимальную парную пересылку без настройки конфигурации, сохраняя пересылку даже в моменты возникновения временных петель, и поддерживают множество путей для индивидуального и группового трафика. Это обеспечивается за счет применения маршрутизации IS-IS и инкапсуляции трафика с заголовком, включающим счетчик интервалов пересылки.

Мосты RBridge совместимы с прежними пользовательскими мостами IEEE 802.1, а также с маршрутизаторами и оконечными узлами IPv4 и IPv6. Они невидимы для современных маршрутизаторов IP как мосты и, подобно маршрутизаторам, прерывают действие протокола STP².

Решение поддерживает виртуальные сети VLAN и оптимизацию распространения групповых (multi-destination) кадров на основе VLAN ID и выведенных из IP multicast-групп. Поддерживается также масштабирование таблиц индивидуальной пересылки на транзитных мостах RBridge в соответствии с числом мостов RBridge (а не числом конечных узлов), что позволяет сделать таблицы пересылки существенно меньше, чем в традиционных мостах.

Статус документа

Документ относится к категории Internet Standards Track.

Документ является результатом работы IETF³ и представляет согласованный взгляд сообщества IETF. Документ прошёл открытое обсуждение и был одобрен для публикации IESG⁴. Дополнительную информацию о стандартах Internet можно найти в разделе 2 в RFC 5741.

Информацию о текущем статусе документа, ошибках и способах обратной связи можно найти по ссылке <http://www.rfc-editor.org/info/rfc6325>.

Авторские права

Авторские права (Copyright (c) 2011) принадлежат IETF Trust и лицам, указанным в качестве авторов документа. Все права защищены.

К документу применимы права и ограничения, указанные в BCP 78 и IETF Trust Legal Provisions и относящиеся к документам IETF (<http://trustee.ietf.org/license-info>), на момент публикации данного документа. Прочтите упомянутые документы внимательно. Фрагменты программного кода, включённые в этот документ, распространяются в соответствии с упрощённой лицензией BSD, как указано в параграфе 4.e документа IETF Trust Legal Provisions, без каких-либо гарантий (как указано в Simplified BSD License).

Документ может содержать материалы из IETF Document или IETF Contribution, опубликованных или публично доступных до 10 ноября 2008 года. Лица, контролирующие авторские права на некоторые из таких документов, могли не предоставить IETF Trust права разрешать внесение изменений в такие документы за рамками процессов IETF Standards. Без получения соответствующего разрешения от лиц, контролирующих авторские права этот документ не может быть изменён вне рамок процесса IETF Standards, не могут также создаваться производные документы за рамками процесса IETF Standards за исключением форматирования документа для публикации или перевода с английского языка на другие языки.

Оглавление

1. Введение.....	3
1.1. Алгоритм версии 2 от Ray Perln.....	4
1.2. Нормативное содержимое и приоритеты.....	4
1.3. Терминология и обозначения.....	4
1.4. Категории кадров L2.....	4
1.5. Сокращения.....	5

¹Routing Bridge.

²Spanning tree protocol - протокол связующего дерева.

³Internet Engineering Task Force - комиссия по решению инженерных задач Internet.

⁴Internet Engineering Steering Group - комиссия по инженерным разработкам Internet.

2. Маршрутизирующие мосты.....	5
2.1. Общий обзор.....	5
2.2. Адреса конечных станций.....	6
2.3. Архитектура инкапсуляции RBridge.....	6
2.4. Обзор пересылки.....	7
2.4.1. Заведомо индивидуальные кадры.....	7
2.4.2. Групповые кадры.....	7
2.5. Мосты RBridge и VLAN.....	8
2.5.1. VLAN на канале.....	8
2.6. RBridge и мосты IEEE 802.1.....	8
2.6.1. Порты RBridge и уровни 802.1.....	8
2.6.2. Поэтапное развертывание.....	9
3. Детали заголовка TRILL.....	9
3.1. Формат заголовка TRILL.....	9
3.2. Версия (V).....	9
3.3. Резерв (R).....	9
3.4. Множество получателей (M).....	9
3.5. Op-Length.....	10
3.6. Счетчик интервалов.....	10
3.7. Псевдонимы RBridge.....	10
3.7.1. Псевдоним выходного RBridge.....	10
3.7.2. Псевдоним входного RBridge.....	10
3.7.3. Выбор псевдонима RBridge.....	10
3.8. Опции заголовка TRILL.....	11
4. Другие детали устройства RBridge.....	12
4.1. Инкапсуляция данных Ethernet.....	12
4.1.1. Информация тег VLAN.....	12
4.1.2. Внутренний тег VLAN.....	13
4.1.3. Внешний тег VLAN.....	13
4.1.4. Контрольная сумма FCS.....	13
4.2. Протокол на основе состояния каналов (IS-IS).....	13
4.2.1. Отождествление RBridge для IS-IS.....	14
4.2.2. Экземпляры IS-IS.....	14
4.2.3. Кадры TRILL IS-IS.....	14
4.2.4. Приветствия TRILL, DRB и назначенные устройства пересылки.....	14
4.2.4.1. Каналы с P2P Hello.....	14
4.2.4.2. Назначенный RBridge.....	15
4.2.4.3. Назначенное устройство пересылки VLAN-x.....	15
4.2.4.4. Информация TRILL LSP.....	15
4.2.5. Протокол TRILL ESADI.....	16
4.2.5.1. Участие в TRILL ESADI.....	17
4.2.5.2. Информация TRILL ESADI.....	17
4.2.6. SPF, пересылка и неоднозначные получатели.....	17
4.3. Размер MTU на канале между RBridge.....	18
4.3.1. Определение TRILL IS-IS MTU в масштабе кампуса.....	18
4.3.2. Тестирование размера MTU для канала.....	18
4.4. Протокол TRILL-Hello.....	18
4.4.1. Обоснование TRILL-Hello.....	18
4.4.2. Содержимое TRILL-Hello и синхронизация.....	19
4.4.2.1. Список соседей TRILL.....	19
4.4.3. Теги VLAN для TRILL MTU-Probe и TRILL Hello.....	20
4.4.4. Множество портов на одном канале.....	20
4.4.5. Трансляция VLAN в канале.....	21
4.5. Деревья распространения.....	21
4.5.1. Расчет дерева распространения.....	22
4.5.2. Проверка кадров с множеством получателей.....	22
4.5.3. Сокращение дерева распространения.....	23
4.5.4. Оптимизация дерева распространения.....	23
4.5.5. Пересылка с использованием дерева распространения.....	24
4.6. Поведение обработки пакетов.....	24
4.6.1. Прием естественного кадра.....	24
4.6.1.1. Индивидуальные кадры.....	24
4.6.1.2. Групповые и широковещательные кадры.....	24
4.6.2. Прием кадра TRILL.....	25
4.6.2.1. Кадры управления TRILL.....	25
4.6.2.2. Кадры TRILL ESADI.....	25
4.6.2.3. Кадры данных TRILL.....	25
4.6.2.4. Заведомо индивидуальные кадры данных TRILL.....	25
4.6.2.5. Кадры данных TRILL для множества получателей.....	26
4.6.3. Прием кадров управления L2.....	26
4.7. Изучение IGMP, MLD и MRD.....	26
4.8. Детали адреса конечной станции.....	26
4.8.1. Определение адресов конечных станций.....	27
4.8.2. Обоснование уровня достоверности при обучении.....	27
4.8.3. Забывание адресов конечных станций.....	27
4.8.4. Общее изучение VLAN.....	28
4.9. Порты RBridge.....	28

4.9.1. Конфигурация порта RBridge.....	28
4.9.2. Структура порта RBridge.....	29
4.9.3. Обработка BPDU.....	30
4.9.3.1. Прием BPDU.....	30
4.9.3.2. Смена корневого моста.....	30
4.9.3.3. Передача BPDU.....	30
4.9.4. Динамическая регистрация VLAN.....	30
5. Параметры RBridge.....	30
5.1. Параметры для RBridge в целом.....	30
5.2. Параметры для каждого псевдонима в RBridge.....	31
5.3. Параметры для каждого порта в RBridge.....	31
5.4. Параметры для каждой VLAN в RBridge.....	31
6. Вопросы безопасности.....	31
6.1. Вопросы безопасности VLAN.....	32
6.2. DoS-атаки BPDU/Hello.....	32
7. Вопросы выделения значений.....	32
7.1. Взаимодействие с IANA.....	32
7.2. Регистрация значений в IEEE.....	32
8. Нормативные документы.....	32
9. Дополнительная литература.....	33
Приложение А. Вопросы поэтапного развертывания.....	34
А.1. Определение стоимости канала.....	34
А.2. Назначенные устройства пересылки и ЛВС на базе мостов.....	34
А.3. Топология кабельного шкафа.....	34
А.3.1. Решение на основе RBridge.....	35
А.3.2. Решение на основе VLAN.....	35
А.3.3. Решение на основе STP.....	35
А.3.4. Сравнение решений.....	35
Приложение В. Настройка транкового порта и порта доступа.....	36
Приложение С. Поддержка множества путей.....	36
Приложение D. Определение VLAN и Priority.....	37
Приложение E. Поддержка поправок к стандарту IEEE 802.1Q-2005.....	37
E.1. Завершенные поправки.....	37
E.2. Разрабатываемые поправки.....	38
Приложение F. Благодарности.....	38

Список рисунков

Рисунок 1. Соединение между RBridge.....	7
Рисунок 3. Кадр TRILL, инкапсулированный в Ethernet.....	7
Рисунок 4. Кадр TRILL, инкапсулированный в PPP.....	7
Рисунок 5. Модель порта RBridge.....	9
Рисунок 6. Заголовок TRILL.....	10
Рисунок 7. Начальный октет флагов в области опций.....	12
Рисунок 8. Инкапсуляция TRILL Data для Encapsulation over Ethernet.....	13
Рисунок 9. Информация тега VLAN.....	14
Рисунок 10. Формат кадра TRILL IS-IS.....	15
Рисунок 11. Формат кадра TRILL ESADI.....	18
Рисунок 12. Модель порта RBridge.....	32
Рисунок 13. Стоимость канала с мостами.....	37
Рисунок 14. Топология кабельного шкафа.....	38
Рисунок 15. Множество путей ко множеству адресатов.....	40
Рисунок 16. Множество путей для заведомо индивидуального трафика.....	40

1. Введение

В традиционных сетях IPv4 и IPv6 каждая подсеть имеет свой уникальный префикс. Поэтому узел, входящий в разные подсети имеет множество адресов IP, обычно по одному адресу на интерфейс. Это также означает что при переносе интерфейса в другую подсеть он меняет свой адрес IP. Администрирование сетей IP усложняется в результате того, что маршрутизаторы IP требуют настройки адресов подсети на уровне своиз портов. Требуется аккуратно назначать адреса IP, чтобы не возникало малозаселенных подсетей, занимающих дефицитные адреса.

Мосты IEEE 802.1 не сталкиваются с такими проблемами, прозрачно соединяя множество физических каналов в одну локальную сеть с точки зрения IP [802.1D]. Однако мост 802.1 выполняет пересылку с использованием протокола STP, с которым связан ряд недостатков.

- Протокол STP работает на основе блокировки портов, ограничивая число каналов пересылки, что создает «пробки» за счет концентрации трафика в некоторых каналах.
- Пересылка не происходит по парному кратчайшему пути, вместо этого используются каналы, оставшиеся после удаления протоколом STP избыточных путей.
- В заголовке Ethernet нет счетчика интервалов пересылки (поля TTL¹). Это создает опасность возникновения временных петель (например, при потере сообщений STP или добавлении компонент типа повторителей).
- VLAN могут разбиваться на части при перестройке STP.

¹Time to Live - время жизни.

В этом документе описаны маршрутизирующие мосты RBridge [Rbridges], в которых реализован протокол TRILL, описанные ниже в поэтической форме. RBridge объединяет преимущества мостов и маршрутизаторов, являясь, как указано в этом документе, приложением маршрутизации по состоянию каналов к решению проблемы осведомленных о VLAN пользовательских мостов. За исключением указанных в документе случаев мосты RBridge могут постепенно заменить пользовательские мосты IEEE [802.1Q-2005] и [802.1D].

Хотя мосты RBridge могут применяться для разных протоколов канального уровня, данная спецификация относится к каналам IEEE [802.3]. Предполагается, что другие типы каналов будут описаны в отдельных документах.

Протокол TRILL, как указано в этом документе, разработан в качестве протокола ЛВС и не предназначен для масштабирования за пределы размеров существующих ЛВС на основе мостов. Дополнительное рассмотрение области проблем, решаемых с помощью мостов RBridge, приведено в [RFC5556].

1.1. Алгоритм версии 2 от Ray Perlner

I hope that we shall one day see
A graph more lovely than a tree.

A graph to boost efficiency
While still configuration-free.

A network where RBridges can
Route packets to their target LAN.

The paths they find, to our elation,
Are least cost paths to destination!

With packet hop counts we now see,
The network need not be loop-free!

RBridges work transparently,
Without a common spanning tree.

1.2. Нормативное содержимое и приоритеты

Большая часть нормативного материала приведена в разделах 1 - 4 этого документа. При возникновении конфликтов между этими четырьмя разделами приоритет имеют сведения из раздела с большим номером.

1.3. Терминология и обозначения

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 интерпретируются в соответствии с [RFC2119].

TRILL - это протокол, заданный этим документом, а RBridge - устройство, реализующее этот протокол. Регистр второй буквы в Rbridge значения не имеет, корректны оба варианта Rbridge и RBridge.

В этом документе термин «канал» (если явно не указано иное) означает ЛВС на основе мостов (bridged LAN), т. е. представляет собой комбинацию из одного или множества каналов [802.3], которые могут включать мосты, концентраторы (hub), повторители (repeater) и т. п. Термин «простой канал» означает соединение «точка-точка» или канал с множественным доступом, не включающие мостов или RBridge.

Термин «порт» (если явно не указано иное) обозначает физический, виртуальный [802.1AE] или псевдопорт [802.1X]. Термин «физический порт» обозначает физическую точку подключения между RBridge и каналом.

Термин «кампус» относится к RBridge, а «ЛВС на базе мостов» - к обычным мостам. Кампус RBridge состоит из сети мостов RBridge, обычных мостов, концентраторов, повторителей, простых каналов и т. п., ограниченных конечными станциями и маршрутизаторами.

Термин «связующее дерево» (spanning tree) в этом документе обозначает классический а также ускоренный вариант spanning tree (как в протоколе RSTP¹).

В документе используется 16-ричное представление адресов MAC. Каждый октет (8-битовый байт) указывается двумя шестнадцатеричными цифрами, задающими значение октета в форме целого числа без знака. Октеды разделяются символом дефиса. В документе используется принятый в IETF порядок представления битов, хотя в физической среде IEEE [802.3] биты внутри октета передаются от младшего к старшему, т. е. в обратном порядке.

1.4. Категории кадров L2

В этом документе кадры уровня 2 (L2) делятся на 5 категорий:

- кадры управления L2 (типа BPDU²);
- естественные кадры (кадры данных без инкапсуляции TRILL);
- кадры данных TRILL (кадры данных с инкапсуляцией TRILL);
- кадры управления TRILL;
- прочие кадры TRILL.

Идентификация этих категорий кадров описана ниже.

¹Rapid Spanning Tree Protocol - ускоренный протокол связующего дерева.

²Bridge PDU - блок данных протокола мостов.

- Кадры управления L2 используют групповой адрес получателя из диапазона 01-80-C2-00-00-00 - 01-80-C2-00-00-0F или адрес 01-80-C2-00-00-21. Мостам RBridge **недопустимо** инкапсулировать и пересылать такие кадры, хотя они **могут** (если в документе не указано иное) выполнять для этих кадров функции L2 (типа MAC-security). Кадры с адресом получателя 01-80-C2-00-00-00 (BPDU) или 01-80-C2-00-00-21 (протокол регистрации VLAN) в этом документе называются кадрами управления верхнего уровня (high-level control frames), все остальные управляющие кадры L2 - кадрами управления нижнего уровня.
- Естественными кадрами называются кадры, не относящиеся к управления и имеющие значение Ethertype, отличное от TRILL и L2-IS-IS, а также адресованные получателям, MAC-адреса которых отличаются от 16 зарезервированных для TRILL групповых адресов.
- В кадрах данных TRILL поле Ethertype имеет значение TRILL. Кроме того, в групповых кадрах этой категории используется групповой MAC-адрес All-Rbridges.
- В кадрах управления TRILL поле Ethertype имеет значение L2-IS-IS, а групповые кадры этой категории передаются по MAC-адресу All-IS-IS-RBridges. Отметим, что кадры ESADI снаружи выглядят как кадры данных TRILL и обрабатываются так же, но при декапсуляции имеют Ethertype = L2-IS-IS.
- Остальные кадры TRILL используют в качестве адреса получателя один из 16 зарезервированных для TRILL групповых адресов, но не All-Rbridges или All-IS-IS-RBridges. Мосты RBridge, соответствующие этой спецификации, **должны** отбрасывать такие кадры.

1.5. Сокращения

AllL1ISs - All Level 1 Intermediate Systems - все промежуточные системы физического уровня.

AllL2ISs - All Level 2 Intermediate Systems - все промежуточные системы канального уровня.

BPDU - Bridge PDU - блок данных протокола мостов.

CHbH - Critical Hop-by-Hop - флаг наличия критичной поэтапной опции.

CItE - Critical Ingress-to-Egress - флаг наличия критичной сквозной опции.

CSNP - Complete Sequence Number PDU - блок данных полного порядкового номера.

DA - Destination Address - адрес получателя.

DR - Designated Router - назначенный маршрутизатор.

DRB - Designated RBridge - назначенный RBridge.

EAP - Extensible Authentication Protocol - расширяемый протокол аутентификации.

ECMP - Equal Cost Multipath - множество равноценных путей.

EISS - Extended Internal Sublayer Service - расширенный сервис внутреннего подуровня.

ESADI - End-Station Address Distribution Information - информация о распространении адресов конечных станций.

FCS - Frame Check Sequence - последовательность проверки кадра (контрольная сумма).

GARP - Generic Attribute Registration Protocol - базовый протокол регистрации атрибутов.

GVRP - GARP VLAN Registration Protocol - протокол GARP для регистрации VLAN.

IEEE - Institute of Electrical and Electronics Engineers - институт инженеров по электрике и электронике.

IGMP - Internet Group Management Protocol - протокол управления группами Internet.

IP - Internet Protocol - протокол Internet.

IS-IS - Intermediate System to Intermediate System - протокол маршрутизации между промежуточными системами.

ISS - Internal Sublayer Service - внутренний сервис подуровня.

LAN - Local Area Network - локальная (вычислительная) сеть, ЛВС

LSP - Link State PDU - блок данных состояния канала.

MAC - Media Access Control - контроль доступа к среде.

MLD - Multicast Listener Discovery - обнаружение «слушателей» группового трафика.

MRD - Multicast Router Discovery - обнаружение маршрутизаторов группового трафика.

MTU - Maximum Transmission Unit - максимальный размер передаваемого блока.

MVRP - Multiple VLAN Registration Protocol - протокол регистрации множества VLAN.

NSAP - Network Service Access Point - точка доступа к сетевому сервису.

P2P - Point-to-point - соединение между парой точек, «точка-точка».

PDU - Protocol Data Unit - блок данных протокола.

PPP - Point-to-Point Protocol - протокол соединений «точка-точка».

RBridge - Routing Bridge - маршрутизирующий мост.

RPF - Reverse Path Forwarding - пересылка по обратному пути.

SA - Source Address - адрес отправителя.

SNMP - Simple Network Management Protocol - простой протокол сетевого управления.

SPF - Shortest Path First - сначала кратчайший путь.

TLV - Type, Length, Value - тип, размер, значение.

TRILL - TRansparent Interconnection of Lots of Links - прозрачное соединение множества каналов.

VLAN - Virtual Local Area Network - виртуальная ЛВС.

VRP - VLAN Registration Protocol - протокол регистрации VLAN.

2. Маршрутизирующие мосты

В этом разделе представлен краткий обзор маршрутизирующих мостов RBridge, реализующих протокол TRILL, без учета некоторых деталей. Подробные спецификации представлены в разделах 3 и 4.

TRILL, как описано в этом документе с некоторыми рассмотренными здесь исключениями, обеспечивает сервис [802.1Q-2005] с учетом VLAN. Как указано ниже, TRILL размещается над уровнем портов RBridge.

Маршрутизирующие мосты RBridge, описанные в этом документе не поддерживают услуг провайдерских мостов [802.1ad], включая магистральные [802.1ah], или похожего на них сервиса. Расширение TRILL для поддержки провайдерских служб оставлено для будущей работы и отдельного документа. Однако провайдерские мосты (включая магистральные) могут использоваться для соединения частей кампуса RBridge.

2.1. Общий обзор

Маршрутизирующие мосты RBridge используют между собой протокол на основе состояния каналов. Это дает им достаточно информации для расчета оптимальных парных путей для индивидуального трафика и расчета деревьев распространения для доставки кадров получателям, чье местоположение неизвестно, а также по широковещательным и групповым адресам [RBridges] [RP1999].

Для снижения влияния временных петель RBridge выполняют пересылку на основе заголовка со счетчиком интервалов. Маршрутизирующие мосты RBridge также задают RBridge следующего интервала пересылки как адресата кадра при пересылке индивидуальных кадров через канал с общей средой (shared-media), что позволяет избежать появления дополнительных кадров при наличии временной петли. Выполняется проверка пересылки по обратному пути (RPF) и другие проверки для кадров с множеством получателей с целью дополнительного контроля возможных петель трафика (параграф 4.5.2).

Первый RBridge (RB1), с которым индивидуальный кадр сталкивается в кампусе, инкапсулирует принятый кадр с использованием заголовка TRILL, который задает последний RBridge (RB2), где происходит декапсуляция кадра. RB1 называют входным RBridge, а RB2 - выходным. Для экономии места в заголовке TRILL и упрощения поиска при пересылке, между RBridge поддерживается протокол динамического получения псевдонимов (nickname), выбирающий 2-октетный псевдоним для RBridge, который уникален в масштабе кампуса и служит сокращением для IS-IS ID данного RBridge. Двухоктетные псевдонимы служат для указания входного и выходного RBridge в заголовке TRILL.

Поддерживается распространение кадров со множеством получателей по множеству путей с разными деревьями распространения и ECMP для индивидуальных кадров (Приложение С).

Сети со структурой, более близкой к полносвязной, в большей степени выиграют от поддержки множественных и оптимальных путей, обеспечиваемой TRILL, чем сети с более древовидной структурой соединений.

Маршрутизирующие мосты RBridge используют на канале протокол для выбора DRB. Протокол выбора TRILL-IS-IS на канале несколько отличается от протокола выбора L3 IS-IS [ISO10589], поскольку в TRILL может быть указан только один RBridge в качестве DRB, а в L3 IS-IS можно выбрать множество DR (называемых также also known as Designated Intermediate System). Как и маршрутизатор IS-IS, DRB может давать каналу имя псевдоузла, выдавать LSP от имени псевдоузла (Link State PDU) и выдавать в канал CSNP. В дополнение к этому DRB имеет некоторые специфические для TRILL обязанности, включая указание VLAN, используемой в качестве Designated VLAN при коммуникациях между RBridge на данном канале (параграф 4.2.4.2).

DRB инкапсулирует и декапсулирует весь трафик данных канала или при распределении нагрузки передает эти полномочия для одной или нескольких VLAN другим RBridge на канале. В любой момент на канале должно быть не более одного RBridge, который инкапсулирует и декапсулирует трафик для конкретной VLAN. Будем называть RBridge, назначенный для пересылки трафика VLAN-х от имени канала «назначенным узлом пересылки VLAN-х» (параграф 4.2.4.3). Дополнительное обсуждение VLAN приведено в параграфе 2.5.

Маршрутизирующим мостам **следует** поддерживать протокол SNMPv3 [RFC3411]. База RBridge MIB будет задана в отдельном документе. Если на RBridge доступен сервис IP, ему **следует** поддерживать SNMPv3 по протоколу UDP для IPv4 [RFC3417] и IPv6 [RFC3419], однако в кампусе управление может использоваться даже для RBridge без поддержки IP или другого транспортного стека L3, а также без адреса L3 путем доставки данных SNMP через Ethernet [RFC4789].

2.2. Адреса конечных станций

RBridge RB1, который служит узлом пересылки для VLAN-х, на каждом из своих каналов должен определять местоположение конечных узлов VLAN-х как на каналах, для которых он является узлом пересылки VLAN-х, так и на других каналах в кампусе. RB1 определяет порты, VLAN, и адреса L2 (MAC) конечных узлов на каналах, для которых он является узлом пересылки VLAN-х, по адресам отправителей кадров, для которых он служит мостом (см., например, параграф 8.7 в [802.1Q-2005]), из конфигурации или протокола явной регистрации L2 типа привязки и аутентификации IEEE 802.11. RB1 определяет VLAN и адрес L2 удаленных конечных точек VLAN-х и соответствующего RBridge, к которому они могут быть подключены, путем просмотра псевдонима входного RBridge в заголовке TRILL, а также VLAN и MAC-адреса отправителя во внутренних кадрах данных TRILL, которые он декапсулирует.

Кроме того, RBridge назначенный ретранслятором VLAN-х хотя бы для одного канала, **может** использовать протокол ESADI для анонсирования части или всех конечных узлов VLAN-х на этих каналах.

Протокол ESADI можно использовать для анонсирования конечных узлов, которые были анонсированы явно. Такая информация может быть более достоверной, чем определенная из кадров, декапсулированных в канал. Кроме того, зарегистрированные и распространяемые таким способом адреса могут быть более защищены по двум причинам - (1) регистрация может выполняться с проверкой подлинности (например, с использованием криптографических методов EAP через [802.1X]) и (2) протокол ESADI поддерживает криптографическую аутентификацию своих сообщений [RFC5304] [RFC5310] для более защищенной передачи.

Если конечная станция отключается от одного RBridge и подключается к другому, в зависимости от обстоятельств кадры для этой станции могут попадать в «черную дыру». Т. е. они просто могут быть отправлены прежнему RBridge, который конечная станция использовала, пока кэшированный адрес на некоторых удаленных RBridge не устареет (возможно в течение минут или дольше). При использовании протокола ESADI разрыв соединения при переключении станции может вызвать незамедлительную передачу обновления.

Даже при использовании протокола ESADI для определения и анонсирования подключенных конечных узлов маршрутизирующие мосты RBridge **должны** продолжать получение информации из принятых обычных кадров и декапсулированных кадров данных TRILL, если в их конфигурации не задано иное. Анонсирование конечных узлов с использованием ESADI является не обязательным, как и извлечение информации (обучение) из этих анонсов.

Дополнительные детали, связанные с адресами конечных станций, приведены в параграфе 4.8.

2.3. Архитектура инкапсуляции RBridge

В качестве технологии L2 для соединения между RBridge может служить IEEE [802.3] или другие технологии канального уровня типа PPP [RFC1661]. Это возможно в результате того, что функция трансляции RBridge работает на основе технологий L2. Однако в этом документе описывается лишь инкапсуляция IEEE 802.3.

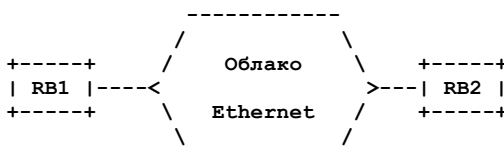


Рисунок 1. Соединение между RBridge.

На рисунке 1 показаны два RBridge RB1 и RB2, соединенных между собой через облако (сеть) Ethernet, которое может включать концентраторы, среды «точка-точка» или с множественным доступом, мосты IEEE 802.1D и 802.1Q.

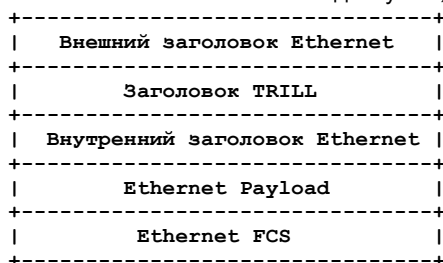


Рисунок 2. Кадр TRILL, инкапсулированный в Ethernet.

На рисунке 2 показан формат кадра данных TRILL или кадра ESADI в облаке Ethernet между RB1 и RB2.

При использовании среды, отличающейся от Ethernet, используется заголовок конкретной среды вместо внешнего заголовка Ethernet. На рисунке 3 показана инкапсуляция TRILL для передачи по протоколу PPP.

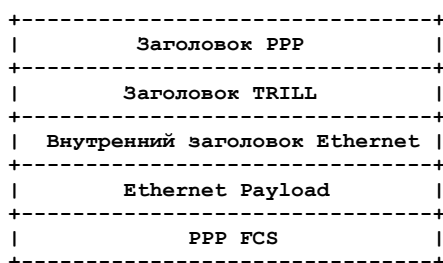


Рисунок 3. Кадр TRILL, инкапсулированный в PPP.

Внешний заголовок определяется типом канала. В данном документе рассматриваются лишь соединения [802.3], но возможны и другие типы.

В любом случае внутренний заголовок Ethernet и данные (Ethernet Payload) берутся из принятого кадра и инкапсулируются с использованием заголовка TRILL для передачи между RBridge. Использование заголовка TRILL обеспечивает ряд преимуществ:

1. снижения влияния петель за счет наличия счетчика интервалов (hop count);
2. отсутствие необходимости определять VLAN и MAC-адрес конечной станции на транзитных RBridge;
3. отправка индивидуальных кадров в направлении выходного RBridge (это позволяет на транзитных RBridge держать таблицы индивидуальной пересылки с размером, определяемым числом RBridge, а не общим числом конечных станций);
4. предоставление отдельного тега VLAN для пересылки трафика между RBridge, независимо от VLAN естественного кадра.

При пересылке индивидуальных кадров между RBridge во внешнем заголовке присутствует в качестве MAC-адреса получателя адрес следующего RBridge, что позволяет предотвратить дублирование кадров в случаях среды с множественным доступом между RBridge. Это также обеспечивает возможность передачи индивидуального трафика по множеству путей, поскольку передающий RBridge может задать следующий интервал. Наличие внешнего заголовка, указывающего передающий RBridge в качестве отправителя гарантирует, что мосты в облаке Ethernet не запутаются при наличии множества путей и будут видеть исходного отправителя или входной RBridge во внешнем заголовке.

2.4. Обзор пересылки

Маршрутизирующие мосты RBridge являются маршрутизаторами в том смысле, что при пересылке кадра транзитным RBridge внешний заголовок L2 заменяется на каждом устройстве пересылки подходящим заголовком L2 для следующего интервала, а значение счетчика интервалов уменьшается. Несмотря на изменение внешнего заголовка L2 и счетчика интервалов в заголовке TRILL, инкапсулированный исходный кадр сохраняется, включая тег VLAN (см. подробности в параграфе 4.6).

С точки зрения пересылки транзитные кадры можно разделить на 2 категории - заведомо индивидуальные (known-unicast) и групповые (multi-destination). Кадры управления L2 и TRILL, а также прочие (не данные) кадры TRILL не являются транзитными, не пересылаются RBridge и не входят в эти категории.

2.4.1. Заведомо индивидуальные кадры

К этой категории относятся кадры, в которых используется индивидуальный внутренний MAC-адрес получателя (Inner.MacDA) и для которых входной RBridge знает выходной RBridge для MAC-адреса получателя в VLAN кадра.

Такие кадры поэтапно передаются через цепочку RBridge до выходного RBridge.

2.4.2. Групповые кадры

Это кадры, которые должны доставляться множеству получателей.

Варианты групповых кадров перечислены ниже.

- 1 Индивидуальные кадры для которых местоположение адресата не известно - Inner.MacDA содержит индивидуальный адрес, но входной RBridge не может определить его местоположение по VLAN в кадре.
- 2 Групповые кадры, для которых адрес получателя L2 выводится из группового адреса IP - Inner.MacDA содержит групповой адрес из набора групповых адресов L2, выведенный из группового адреса IPv4 [RFC1112] или IPv6 [RFC2464]. Эти кадры обрабатываются в зависимости от ситуации:
 - 2.1 отчеты о принадлежности к группам IGMP [RFC3376] и MLD [RFC2710];
 - 2.2 запросы IGMP [RFC3376] и MLD [RFC2710], а также анонсы MRD [RFC4286];
 - 2.3 другие групповые кадры L2, выведенные из IP.
- 3 Групповые кадры, для которых адрес получателя L2 не взят из группового адреса IP, Inner.MacDA содержит групповой адрес, не являющийся групповым адресом L2, полученным из группового адреса IPv4 или IPv6.
- 4 Широковещательные кадры - Inner.MacDA содержит широковещательный адрес FF-FF-FF-FF-FF-FF.

Маршрутизирующие коммутаторы RBridge строят деревья распространения (параграф 4.5) и применяют их для пересылки кадров со множеством получателей. Каждое дерево распространения доходит до всех RBridge в кампусе, совместно используется всеми VLAN и может применяться для распространения естественного кадра, который относится к любой VLAN. Однако распространение любого конкретного кадра по дереву обрезается в зависимости от конкретной ситуации для предотвращения ненужно распространения этого кадра.

2.5. Мосты RBridge и VLAN

Технология VLAN обеспечивает способ разделения конечных узлов в кампусе на группы L2 [802.1Q-2005]. Для использования VLAN нужна настройка. По умолчанию приемный порт определяет VLAN переданного конечной станцией кадра. Конечные станции могут также явно включать эту информацию в кадр.

Мосты IEEE [802.1Q-2005] можно настроить для поддержки множества пользовательских VLAN на одном простом канале путем вставки в кадр (и удаления) тега VLAN. Теги VLAN, используемые TRILL, имеют такой же формат, как теги VLAN, определенные в IEEE [802.1Q-2005]. Как показано на рисунке 2, такие теги могут появляться в двух местах кадров с инкапсуляцией TRILL, передаваемых по каналу IEEE [802.3] - один тег во внешнем заголовке (Outer.VLAN), другой во внутреннем (Inner.VLAN). Внутренние и внешние теги VLAN более подробно рассмотрены в параграфе 4.1.

RBridge обеспечивают доставку естественных кадров из той или иной VLAN лишь в каналы, относящиеся к той же VLAN, однако обработка VLAN в кампусе RBridge и LBC 802.1 на базе мостов различается, как описано ниже.

Работа TRILL IS-IS на канале более подробно описана в параграфе 4.2.4.

2.5.1. VLAN на канале

Некоторые конфигурации мостов могут приводить к разделению VLAN на канале. В таких случаях кадр, переданный RBridge соседу по каналу может не дойти, если он имеет тег VLAN, которая отделена в конфигурации моста.

TRILL требует наличия на канале по меньшей мере одной VLAN, обеспечивающей связь со всеми RBridge на этом канале. По умолчанию это VLAN 1, хотя RBridge могут быть настроены на другую VLAN. DRB указывает другим RBridge, какую VLAN использовать.

Поскольку на каждой VLAN (скажем, VLAN-x) на канале будет лишь один назначенный узел пересылки, при настройке мостов, приводящей к отделению VLAN-x на канале, некоторые конечные узлы VLAN-x могут потерять возможность связи с остальным кампусом.

Настройка моста и порта может приводить к трансляции VLAN на канале (кадры VLAN-х станут кадрами VLAN-y). TRILL детектирует это, помещая копию внешнего тега VLAN в сообщения TRILL-Hello и проверяя такие сообщения на приемной стороне. При обнаружении принимаются меры, обеспечивающие наличие на канале единственного назначенного устройства пересылки, для предотвращения возможных петель и дублирования кадров (параграф 4.4.5).

TRILL ведет себя по возможности консервативно, предотвращая петли, не частичную потерю связности. В результате при ошибочной настройке моста или порта некоторые связи могут теряться.

2.6. RBridge и мосты IEEE 802.1

Порты RBridge за исключением описанных ниже случаев, работают «поверх» портов IEEE [802.1Q-2005].

2.6.1. Порты RBridge и уровни 802.1

Порты RBridge используют функции обработки VLAN и приоритета портов [802.1Q-2005]. В дополнение к этому они могут реализовать другие функции низкого уровня протоколов 802.1, а также протоколов для используемого канала типа PAUSE (Приложение 31В в [802.3]), контроля доступа по портам [802.1X], защиты MAC [802.1AE] или агрегирования каналов [802.1AX].

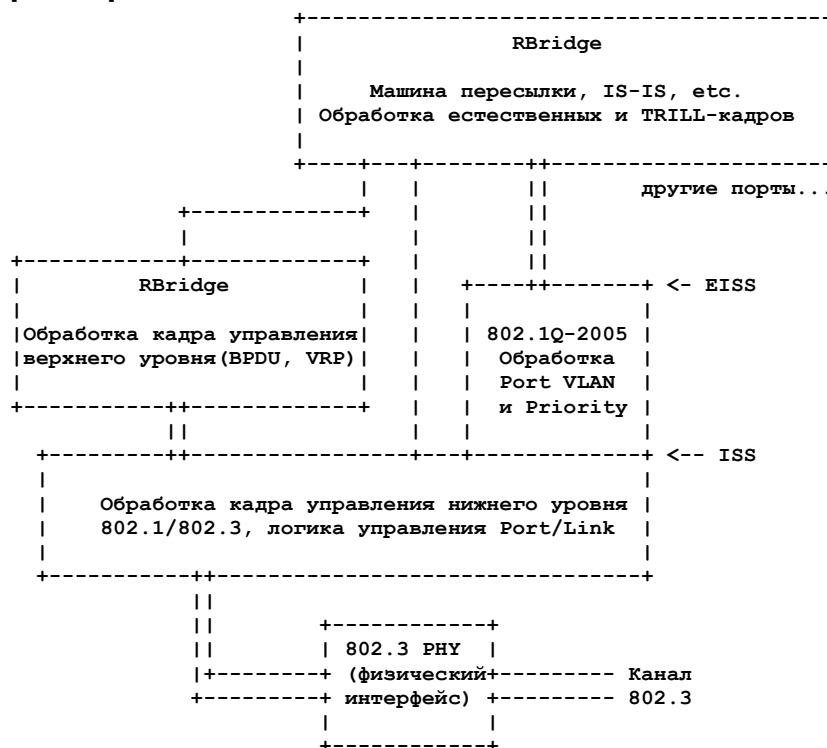


Рисунок 4. Модель порта RBridge.

Однако RBridge не используют связующее дерево и не блокируют порты, как это делает STP. На рисунке 4 показана высокоуровневая диаграмма RBridge с одним портом, подключенным к каналу IEEE 802.3. Одинарные линии представляют поток управляющей информации, двойные - поток кадров и управляющей информации.

Верхний интерфейс к низкоуровневой логике управления портом/каналом соответствует ISS¹ в [802.1Q-2005]. В RBridge кадры управления высокого уровня обрабатываются над интерфейсом ISS.

Верхний интерфейс к обработке VLAN и приоритета соответствует EISS² в [802.1Q-2005]. В RBridge естественные и TRILL-кадры обрабатываются над интерфейсом EISS и относятся к обработке VLAN и приоритета.

2.6.2. поэтапное развертывание

Поскольку маршрутизирующие мосты RBridge совместимы с пользовательскими мостами IEEE [802.1Q-2005] с некоторыми отмеченными в этом документе исключениями, ЛВС на основе мостов можно обновлять постепенно, заменяя традиционные мосты на RBridge. Мосты, которые пока не заменены, будут прозрачны для трафика RBridge. Физические каналы, напрямую соединяющие такие мосты, и сами эти мосты образуют ЛВС. Такая ЛВС на основе мостов с точки зрения RBridge будет выглядеть каналами с множественным доступом.

Если в мостах, заменяемых на RBridge, использовалась принятая по умолчанию конфигурация, устанавливаемые взамен RBridge также не потребуются настраивать.

Поскольку RBridge (как описано в этом документе) обеспечивают лишь обслуживание пользователей, они не могут заменить РВ или РВВ, как пользовательский мост не может заменить провайдерский. Однако такие провайдерские устройства могут быть частью ЛВС на базе мостов между устройствами RBridge. Расширение TRILL для поддержки провайдерских услуг оставлено на будущее и будет описано в отдельных документах.

При замене мостов с настройками на уровне порта протоколов типа управления доступом по портам [802.1X] или защиты на уровне MAC [802.1AE] в устанавливаемых RBridge потребуется аналогичная настройка для таких протоколов. В дополнение к этому устанавливаемые RBridge должны поддерживать тот же тип канала и протоколы канального уровня, какие использовались на заменяемых мостах.

¹Internal Sublayer Service - сервис внутреннего подуровня.

²Extended Internal Sublayer Service - расширенный сервис внутреннего подуровня.

Кампус RBridge будет работать лучше всего при замене всех мостов IEEE [802.1D] и [802.1Q-2005] на RBridge, если маршрутизирующие мосты RBridge имеют такую же скорость и производительность. Однако на промежуточных этапах обновления повышение производительности будет не столь велико по причине замены на RBridge лишь части мостов.

Постепенное обновление более подробно рассмотрено в Приложении А.

3. Детали заголовка TRILL

В этом разделе описан заголовок TRILL, а другие детали устройства RBridge рассмотрены в разделе 4.

3.1. Формат заголовка TRILL

Заголовок TRILL показан на рисунке 5 и не зависит от используемого канального уровня. Для канального уровня IEEE [802.3] используется префикс в виде 16-битового поля TRILL Ethertype [RFC5342], обеспечивающего выравнивание по 64-битовой границе. Если значение Op-Length кратно 64 битам, обеспечивается обычное выравнивание содержимого инкапсулированного кадра.

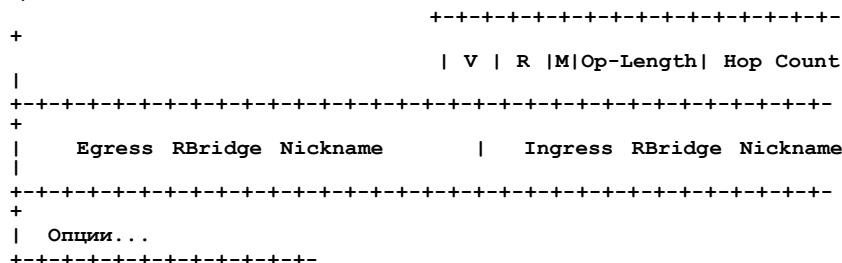


Рисунок 5. Заголовок TRILL.

Поля заголовка перечислены ниже и подробно рассмотрены в последующих параграфах.

- V (Version) - 2-битовое целое число без знака (параграф 3.2).
- R (Reserved) - 2 бита (параграф 3.3).
- M (Multi-destination) - 1 бит (параграф 3.4).
- Op-Length (Options Length) - 5-битовое целое число без знака (параграф 3.5).
- Hop Count - 6-битовое целое число без знака (параграф 3.6).
- Egress RBridge Nickname - 16-битовый идентификатор (параграф 3.7.1).
- Ingress RBridge Nickname: 16-битовый идентификатор (параграф 3.7.2).
- Опции присутствуют при отличном от 0 поле Op-Length (параграф 3.8).

3.2. Версия (V)

Поле версии (V) имеет размер 2 бита и данный документ задает TRILL версии 0. RBridge RB1 **должен** проверять значение поля V в полученных кадрах с инкапсуляцией TRILL. Если поле V имеет не понятное RB1 значение, RB1 **должен** отбрасывать кадр без уведомления отправителя. Выделение новых значений номера версии TRILL происходит по процедуре IETF Standards Action.

3.3. Резерв (R)

Два бита R зарезервированы для использования в будущих расширениях нулевой версии протокола TRILL. Эти поля **должны** устанавливаться в 0 при добавлении заголовка TRILL на входном RBridge, прозрачно копироваться без обработки транзитными RBridge и игнорироваться выходными RBridge. Распределение резервных битов TRILL должно происходить по процедуре IETF Standards Action.

3.4. Множество получателей (M)

Бит M (параграф 2.4.2) указывает, что кадр доставляется некому классу конечных станций через дерево распространения, заданное псевдонимом выходного RBridge.

- M = 0 (FALSE) - поле Egress RBridge Nickname содержит псевдоним выходного RBridge для известного индивидуального MAC-адреса.
- M = 1 (TRUE) - поле Egress RBridge Nickname содержит псевдоним, который задает дерево распространения. Псевдоним выбирается входным RBridge для кадра данных TRILL или исходным RBridge для TRILL ESADI.

3.5. Op-Length

Обеспечивается возможность выразить в заголовке TRILL необязательные возможности (опции) и закодировать связанную с ними информацию

Поле заголовка Op-Length указывает размер опций заголовка TRILL в 4-октетных блоках, что позволяет использовать опции суммарным размером до 124 октетов. Нулевое значение Op-Length говорит об отсутствии опций. При наличии опций они размещаются непосредственно за полем Ingress RBridge Nickname.

Информация о поддерживаемых опциях TRILL приведена в параграфе 3.8.

3.6. Счетчик интервалов

Поле счетчика интервалов Hop Count представляет собой 6-битовое целое число без знака. RBridge отбрасывает кадры с нулевым значением счетчика, а в остальных случаях декрементирует значение поля (это отличается от IPv4 и IPv6 для поддержки последующего добавления функции типа traceroute, позволяющей получить информацию о превышении счетчика интервалов от выходного RBridge).

Для заведомо индивидуальных кадров входному RBridge **следует** устанавливать значение Hop Count, превышающее число интервалов пересылки RBridge, предполагаемых до выходного RBridge, чтобы обеспечить возможность дополнительной маршрутизации дальше в пути.

Для кадров с множеством получателей входному RBridge (или исходному RBridge для кадров TRILL ESADI) **следует** устанавливать в поле Hop Count значение не меньше ожидаемого числа интервалов до самого удаленного RBridge. Для решения этой задачи RBridge Rbп рассчитывает для каждой ветви от Rbп указанного дерева распространения с корнем Rbп число интервалов пересылки в ветви и выбирает максимальное значение.

Кадры с множеством адресатов представляют некоторую опасность, поскольку петля, включающая одно или несколько разветвлений дерева распространения, приводит к быстрой генерации множества копий кадра даже при использовании механизма подсчета интервалов. Именно по этой причине кадры с множеством адресатов подвергаются строгой проверке пересылки по обратному пути и другим проверкам, описанным в параграфе 4.5.2. В качестве дополнительной меры управления трафиком при пересылке кадра с множеством адресатов в ветвь дерева распространения транзитный RBridge Rbм **может** уменьшить значение счетчика интервалов более чем на 1, если это не приводит к недоступности адресатов в данной ветви дерева распространения с корнем Rbп. Использование счетчика интервалов близкого или равного минимальному требуемому для кадров с множеством адресатов обеспечивает дополнительную защиту от проблем пересылки, связанных с временными петлями.

Хотя RBridge **может** уменьшать счетчик интервалов для кадров с множеством адресатов больше чем на 1, при описанных выше обстоятельствах RBridge, пересылающий кадр, **должен** уменьшить счетчик интервалов по меньшей мере на 1 и отбрасывать кадры если он не может сделать это, поскольку счетчик имеет значение 0. Опция декрементирования счетчика интервалов больше чем на 1 при описанных выше обстоятельствах применима лишь к кадрам с множеством получателей и не применима к заведомо индивидуальным кадрам.

3.7. Псевдонимы RBridge

Псевдонимы представляют собой 16-битовые значения, которые назначаются динамически и служат сокращением для идентификаторов RBridge IS-IS ID, обеспечивающим более компактное представление, и могут использоваться для указания потенциально различных деревьев распространения с общим корнем. Такой подход позволяет обозначить до 2^{16} RBridge, однако значение 0x0000 зарезервировано для указания отсутствия псевдонима, диапазон 0xFFC0 - 0xFFFF зарезервирован для будущих спецификаций, а значение 0xFFFF зарезервировано навсегда. RBridge комбинируют протокол получения псевдонимов с протоколом маршрутизации по состоянию канала (параграф 3.7.3) для получения одного или множества псевдонимов, уникальных в масштабе кампуса.

3.7.1. Псевдоним выходного RBridge

Существует два варианта содержимого поля Egress RBridge Nickname в зависимости от бита M (параграф 3.4). Псевдоним задается входным RBridge для кадров данных TRILL и исходным RBridge для кадров TRILL ESADI.

- Для заведомо индивидуальных кадров данных TRILL бит M = 0 и поле Egress RBridge Nickname задает выходной RBridge, т. е. RBridge который должен удалить инкапсуляцию TRILL и переслать естественный кадр. После установки поля Egress RBridge Nickname его **недопустимо** менять на последующих транзитных RBridge.
- Для кадров данных TRILL с множеством получателей и кадров TRILL ESADI бит M = 1. Поле Egress RBridge Nickname содержит псевдоним, задающий дерево распространения, выбранное для пересылки кадра. Этот корневой псевдоним **недопустимо** менять на транзитных RBridge.

3.7.2. Псевдоним входного RBridge

В поле Ingress RBridge Nickname указывается псевдоним входного RBridge для кадров данных TRILL и псевдоним исходного RBridge для кадров TRILL ESADI. Если в конфигурации RBridge входной псевдоним имеет множество значений, **следует** использовать один и тот же псевдоним в поле Ingress RBridge Nickname при инкапсуляции кадров с любыми значениями Inner.MacSA и Inner.VLAN. Это упрощает определение (learning) конечных узлов.

После установки поля Ingress RBridge Nickname его **недопустимо** менять на последующих транзитных RBridge.

3.7.3. Выбор псевдонима RBridge

Протокол выбора псевдонима работает на основе TRILL IS-IS, как показано ниже.

- Псевдоним или псевдонимы для использования RBridge передаются в IS-IS TLV¹ вместе приоритетом [RFC6326]. Каждый RBridge выбирает свои псевдонимы самостоятельно.
- Значения псевдонимов **могут** быть настраиваемыми. RBridge, на котором настроен один или множество псевдонимов, будет иметь приоритет на использование этих значений пред всеми RBridge, у которых нет настроенного псевдонима.
- Значение 0x0000 и диапазон 0xFFC0 - 0xFFFF зарезервированы и их **недопустимо** выбирать для настройки RBridge. Значение 0x0000 служит для индикации того, что псевдоним не известен.
- Приоритет использования поля, сообщаемый с псевдонимом, является 8-битовым целым числом без знака, в котором старший бит (0x80) показывает, что значение псевдонима задано. Для 7 младших битов по умолчанию используется значение 0x40, но **может** быть задано другое значение. Кроме того, RBridge **может** увеличить приоритет после удержания псевдонима в течение некоторого времени. Однако установка старшего бита **недопустима**, если псевдоним не задан в конфигурации.

¹Type-length-value - элемент данных «тип-размер-значение».

- После получения псевдонима RBridge ему **следует** попытаться применить тот же псевдоним после перезагрузки.
- Каждый RBridge отвечает за уникальность каждого из своих псевдонимов. Если RB1 выбрал себе псевдоним x, а затем обнаружил из сообщения LSP для RB2, что тот также выбрал псевдоним x, тогда RBridge или псевдоузел с численно большим приоритетом сохраняет свой псевдоним или его сохраняет RBridge с численно большим IS-IS System ID, если имеется привязка приоритетов, а другой RBridge **должен** выбрать новый псевдоним¹. Это может потребовать замены псевдонима RBridge, заданного в конфигурации.
- Для снижения вероятности совпадения псевдонимов RBridge выбирает псевдоним случайно из числа представляющихся доступными на основе своей копии состояния канала. Этот случайный выбор может происходить путем хэширования RBridge некоторых своих параметров типа SystemID, даты и времени, а также других источников энтропии типа приведенных в [RFC4086] каждый раз или с использованием таких хэш-значений для создания «затравки» (seed) и выбора на основе псевдослучайных значений, созданных с этой затравкой [RFC4086]. Случайным значениям или затравке, а также применяемому алгоритму **следует** обеспечивать равномерное распределение значений, выбираемых из пространства доступных псевдонимов. Схождение бесконфликтного распределения псевдонимов в кампусе может быть ускорено путем выбора новых псевдонимов лишь из числа тех, которые представляются доступными, и выбора при конфликте псевдонима с более высоким приоритетом из числа вовлеченных в конфликт. Не обязательно применять один алгоритм выбора псевдонимов для всех RBridge.
- При слиянии двух кампусов RBridge возможны временные конфликты псевдонимов. Как только каждый RBridge получает LSP от других RBridge, устройство RBridge, которому нужно сменить псевдоним выбирает новые псевдонимы, которые представляются не конфликтующими с имеющимися псевдонимами. Некоторым RBridge может потребоваться неоднократная смена псевдонима в процессе разрешения конфликта.
- Для снижения вероятности узурпирования новым RBridge уже используемого псевдонима устройству RBridge **следует** дожидаться получения от соседа базы данных о состоянии каналов, прежде чем анонсировать какие-либо псевдонимы, не заданные в конфигурации.
- По умолчанию RBridge имеет лишь один псевдоним, но **может** быть настроен на запрос множества псевдонимов. Каждый такой псевдоним будет задавать дерево кратчайшего пути с корнем RBridge, в результате использования номера дерева при разрыве лишних связей (tiebreaking) в случае наличия равноценных путей (параграф 4.5.1), деревья для разных псевдонимов вероятно будут использовать разные каналы. По причине связанной с расчетом деревьев нагрузки запросом множества псевдонимов для RBridge следует пользоваться экономно. Например, это следует использовать на нескольких RBridge, которые в топологии кампуса хорошо подходят в качестве мест для расчета разных деревьев распространения кратчайших путей. Таким деревьям нужны разные псевдонимы, поскольку трафик может идти разными путями.
- Если нужно сделать корнем дерева псевдоузел, DRB **может** запросить один или несколько псевдонимов в LSP псевдоузла.

Каждый используемый в кампусе псевдоним указывает RBridge (или псевдоузел), а каждый псевдоним указывает дерево распространения с корнем в RBridge (или псевдоузле) этого псевдонима. Однако в действительности RBridge кампуса рассчитывают лишь ограниченное число возможных деревьев распространения, как описано в параграфе 4.5.

3.8. Опции заголовка TRILL

Все RBridge **должны** быть способны пропускать заданное полем Op-Length число 4-октетных блоков (параграф 3.5) для нахождения внутреннего кадра, поскольку RBridge должны находить MAC-адрес получателя и тег VLAN во внутреннем кадре. Транзитным RBridge такая информация нужна для фильтрации VLAN, групповой адресации IP и т. п. Выходным RBridge внутренний заголовок нужен для корректной декапсуляции и обработки внутреннего кадра.

Для обеспечения безопасной и совместимой работы при отличном от 0 поле Op-Length, указывающим наличие опций, два старших бита первого октета опций имеют специальное значение, как показано на рисунке 6.

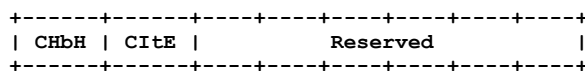


Рисунок 6. Начальный октет флагов в области опций.

Установленный бит CHbH указывает наличие хотя бы одной критической поэтапной (hop-by-hop) опции. Транзитные RBridge, которые не поддерживают присутствие таких опций (например, RBridge без поддержки опций), **должны** отбрасывать такие кадры. Если CHbH=0, кадр является безопасным с точки зрения обработки опций для пересылки транзитными узлами, независимо от поддержки опций данным RBridge. Транзитный RBridge, который не поддерживает присутствующих опций, **должен** просто пересылать область опций в кадре.

Установленный бит CltE указывает наличие хотя бы одной критической сквозной (ingress-to-egress) опции. Сброшенный бит говорит об отсутствии таких опций. Если любой из флагов CHbH и CltE установлен, входные RBridge, которые не поддерживают всех имеющихся критических опций (например, RBridge совсем не поддерживающий опций), **должны** отбросить такой кадр. Если оба флага CHbH и CltE сброшены, кадр является безопасным с точки зрения опций для обработки любым RBridge, независимо от поддержки им тех или иных опций.

Опции, включая значение резервных битов (Рисунок 6), будут описаны в отдельных документах и предполагается включение в них поэтапных и сквозных, а также критических и некритических опций.

Примечание. Предполагается, что большинство реализаций RBridge будет оптимизировано для простейших и массовых случаев пересылки и обработки кадров. Включение опций может, а для сложных и длинных опций будет приводить к разделению обработки кадров между «медленным путем» с ограниченной пропускной способностью и «быстрым путем». Ограниченная пропускная способность медленного пути может вести к отбрасыванию кадров.

¹В оригинале это предложение содержит ошибки, см. <https://www.rfc-editor.org/errata/errata3002>. Прим. перев.

4. Другие детали устройства RBridge

В разделе 3 был описан заголовок TRILL, а в этом разделе рассмотрены другие внутренние детали RBridge.

4.1. Инкапсуляция данных Ethernet

Кадры TRILL с данными и ESADI для передачи по каналам Ethernet инкапсулируются с использованием внешнего заголовка Ethernet (Рисунок 2). Этот внешний заголовок для моста на пути между двумя RBridge выглядит как заголовок обычного кадра Ethernet, поэтому мосты будут пересылать кадры обычным способом. Чтобы RBridge могли отличать такие кадры данных TRILL, во внешнем заголовке используется новое значение TRILL Ethertype (параграф 7.2).

На рисунке 7 показан кадр TRILL Data с внешним тегом VLAN, проходящий через канал Ethernet, как показано в верхней части рисунка, между транзитными RBridge RB3 и RB4. Естественный кадр от конечной станции Esa был инкапсулирован входным RBridge RB1, а затем декапсулирован выходным RBridge RB2 и доставлен адресату Esb. Показанная инкапсуляция будет более эффективна при отсутствии опций TRILL или наличии опций, размер которых кратен 64 битам, за счет выравнивания исходного кадра Ethernet по 64-битовой границе.

При передаче кадра TRILL Data через облако Ethernet, он будет включать три пары адресов, указанные ниже.

- Внешний заголовок Ethernet. Внешние MAC-адреса получателя (Outer.MacDA) и отправителя (Outer.MacSA). Эти адреса служат для указания следующего RBridge и передающего RBridge, соответственно.
- Заголовок TRILL. Egress Nickname и Ingress Nickname указывают псевдонимы выходного и входного RBridge, соответственно, если кадр не имеет множества получателей (в последнем случае Egress Nickname задает дерево распространения, в которое кадр будет передаваться).
- Внутренний заголовок Ethernet. Внутренние MAC-адреса получателя (Inner.MacDA) и отправителя (Inner.MacSA). Эти адреса указываются отправителем исходного кадра.

Кадр данных TRILL может также иметь два тега VLAN, как описано ниже в параграфах 4.1.2 и 4.1.3, для передачи двух разных идентификаторов VLAN и задания приоритета.

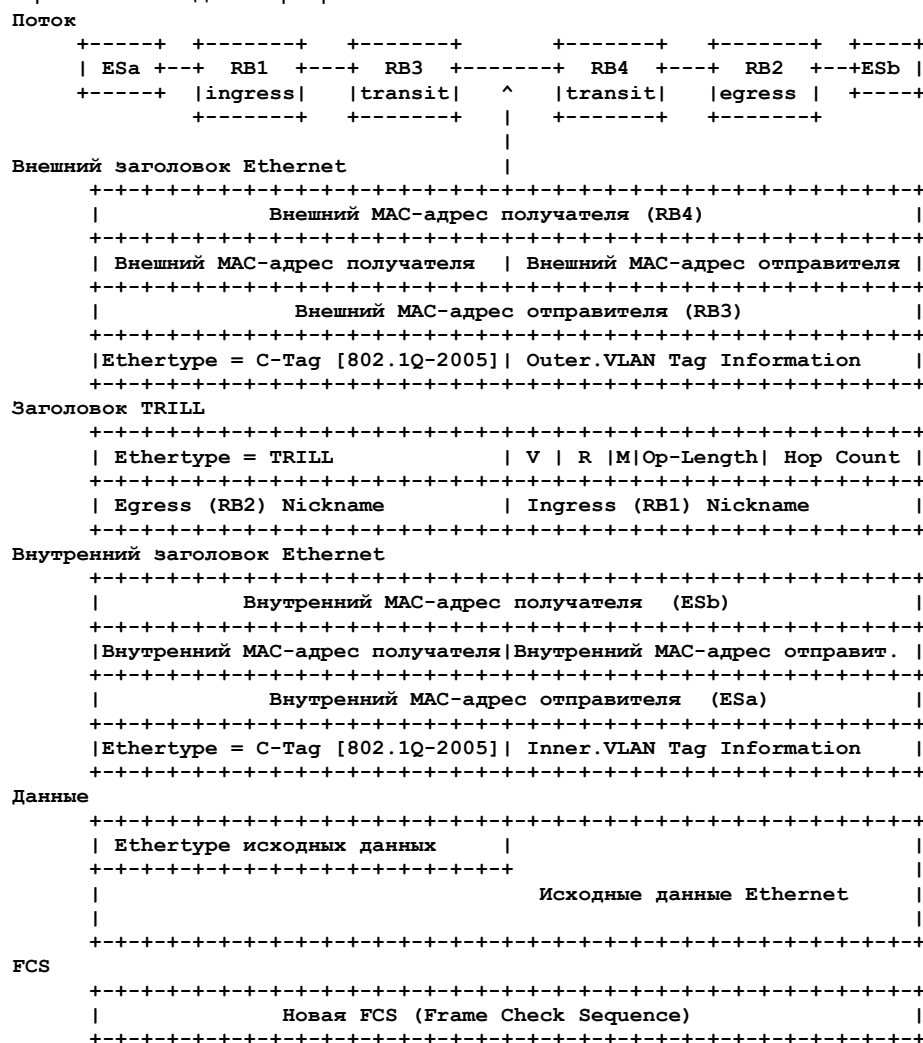


Рисунок 7. Инкапсуляция TRILL Data для Encapsulation over Ethernet.

4.1.1. Информация тега VLAN

Тег VLAN (раньше назывался Q-тегом), известный также как C-тег для пользовательских тегов, включает идентификатор VLAN ID и поле приоритета, как показано на рисунке 8. Поле VLAN ID может иметь значение 0, если VLAN не указана и тег просто задает приоритет, хотя обычно такие кадры называют priority tagged, а не VLAN tagged [802.1Q-2005].

Использование S-тегов [802.1ad], которые называют также тегами сервиса, и стеки тегов выходят за рамки документа.

Как рекомендовано в [802.1Q-2005], RBridge **следует** реализовать так, чтобы обеспечивалась возможность использовать полный диапазон VLAN ID от 0x001 до 0xFFE. Маршрутизирующие мосты RBridge **могут** поддерживать

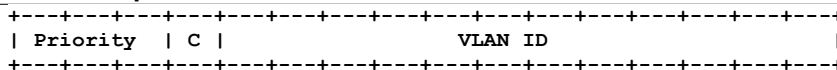


Рисунок 8. Информация тега VLAN.

меньшее число одновременно активных VLAN ID. VLAN ID = 0 указывает, что VLAN не задана, а значение VLAN ID = 0xFFFF зарезервировано.

Использование VLAN ID 0xFFFF **недопустимо**. RBridge **должны** отбрасывать все принятые кадры с Outer.VLAN ID = 0xFFFF. RBridge **должны** отбрасывать все кадры, для которых проверка Inner.VLAN ID дала значение 0xFFFF (такая проверка требуется на всех выходных RBridge, которые декапсулируют кадры).

Бит C, показанный на рисунке 8, не используется в Inner.VLAN протокола TRILL. Он **должен** устанавливаться в 0 входными RBridge, передаваться без изменения транзитными и игнорироваться выходными RBridge.

Как указано в [802.1Q-2005], поле Priority содержит целое число без знака от 0 до 7, где 1 указывает низший приоритет, 7 - высший, а 0 указывает приоритет выше уровня 1, но ниже уровня 2. Поправка [802.1ad] к стандарту [802.1Q-2005] разрешает отображение некоторых смежных пар уровней приоритета на один уровень с правом или без права отбрасывания. Текущая работа IEEE 802.1 (802.1az, Appendix E) предлагает возможность настройки «групп приоритетов», которым предоставляются некие гарантии пропускной способности. Порты RBridge тоже **могут** реализовать такие опции. Маршрутизирующим мостам RBridge не обязательно реализовать любое определенное число различающихся уровней приоритета, но они могут трактовать несколько смежных уровней таким способом.

Полученные на одном порту кадры с совпадающими адресами отправителя и получателя, а также VLAN и приоритета, которые передаются в один порт, **должны** передаваться в порядке их приема, если RBridge не относит их к разным потокам с тонкой настройкой, когда сохранение порядка применяется независимо для каждого потока. Кадры одной VLAN с одинаковым приоритетом, принятые на одном порту, могут передаваться через разные порты, если используется пересылка по множеству путей (см. Приложение C.)

C-Tag Ethertype [RFC5342] имеет значение 0x8100.

4.1.2. Внутренний тег VLAN

Поле внутреннего тега VLAN (Inner.VLAN) содержит информацию VLAN, связанную с естественным кадром, когда он выходит наружу, или с тегом VLAN кадра TRILL ESADI, когда он создается. При прохождении кадра TRILL через транзитный RBridge, поле Inner.VLAN **недопустимо** изменять за исключением преднамеренной трансляции VLAN на RBridge.

При получении RBridge естественных кадров связанные с ним VLAN ID и приоритет определяются в соответствии с [802.1Q-2005] (см. Приложение D и параграф 6.7 в [802.1Q-2005]). Если RBridge является назначенным устройством пересылки для VLAN и доставка кадра требует передачи в один или множество других каналов, входной RBridge создает кадр данных TRILL с VLAN ID и приоритетом на основе Inner.VLAN.

VLAN ID требуется на входном RBridge в качестве одного из элементов, определяющих подходящий выходной RBridge для заведомо индивидуального кадра, а также требуется на входном и каждом транзитном RBridge для кадров со множеством адресатов для корректной обрезки дерева распространения.

4.1.3. Внешний тег VLAN

В кадрах TRILL, передаваемых RBridge, за исключением некоторых кадров TRILL-Hello, используется значение Outer.VLAN ID, заданной назначенным RBridge (DRB) для канала, в который кадры передаются. Этот идентификатор называется Designated VLAN. Для кадров данных TRILL и кадров ESADI в качестве приоритета в теге Outer.VLAN **следует** указывать приоритет из Inner.VLAN tag.

В кадрах TRILL, пересылаемых транзитным RBridge, используется приоритет из Inner.VLAN полученного кадра. Кадры TRILL Data передаются с приоритетом, связанным с соответствующим естественным кадром, который был получен (Приложение D). Кадры TRILL IS-IS **следует** передавать с приоритетом 7.

Реальное присутствие тега Outer.VLAN при передаче кадра TRILL в среду зависит от настройки порта RBridge, через который передается кадр, так же, как наличие тега VLAN при передаче кадра мостом [802.1Q-2005] зависит от настройки порта (параграф 4.9.2).

4.1.4. Контрольная сумма FCS

Каждый кадр Ethernet включает контрольную сумму FCS¹, которая рассчитывается для всего кадра и служит для обнаружения проблем, связанных с битовыми ошибками при передаче в канале. Таким образом, при инкапсуляции кадра исходное значение FCS не включается и отбрасывается. Все кадры, принятые с ошибкой FCS, **следует** отбрасывать (это может быть невозможно при «сквозной» (cut through) коммутации). Поле FCS обычно меняется при инкапсуляции, декапсуляции и на каждом интервале TRILL в результате изменения внешних адресов получателя и отправителя, декремента счетчика интервалов и т. п.

Хотя FCS обычно вычисляется непосредственно перед отправкой, на практике бывает желательно наличие FCS вместе с кадром в RBridge после приема. Это значение FCS может обновляться динамически в процессе обработки кадра в RBridge и использоваться для передачи или сравнения с расчетным значением FCS при отправке кадра. Это необязательное непрерывное использование FCS будет помогать при обнаружении некоторых внутренних отказов RBridge типа ошибок памяти.

4.2. Протокол на основе состояния каналов (IS-IS)

TRILL использует расширение IS-IS [ISO10589] [RFC1195] в качестве протокола маршрутизации. IS-IS обеспечивает два преимущества:

- работа непосредственно на уровне L2, не требующая настройки конфигурации (не нужны адреса IP);

¹Frame Check Sequence - последовательность проверки кадра.

- простота расширения путем определения новых TLV элементов данных и субэлементов для передачи информации TRILL.

В этом параграфе описано использование в TRILL протокола IS-IS без протокола TRILL-Hello, который отдельно описан в параграфе 4.4, а также сообщений MTU-probe и MTU-ack, описанных в параграфе 4.3.

4.2.1. Отождествление RBridge для IS-IS

Каждый RBridge имеет уникальный 48-битовый (6 октетов) идентификатор IS-IS System ID, который может быть создан на основе любого уникального MAC-адреса RBridge.

Псевдоузел получает 7-октетный идентификатор от создавшего его DRB, путем добавления 1 октета к имеющемуся у DRB 6-октетному идентификатору. В качестве 6-октетного идентификатора для создания идентификатора псевдоузла **следует** применять идентификатор DRB, пока этот DRB не создает идентификаторов псевдоузлов больше чем для 255 каналов. Единственным требованием является уникальность 7-октетных идентификаторов в рамках кампуса и отличное от 0 значение седьмого октета. RBridge имеет 7-октетный идентификатор, состоящий из 6-октетного идентификатора системы и седьмого октета со значением 0.

В этом документе IS-IS ID используется для обозначения 7-октетного значения, которое может быть идентификатором RBridge или псевдоузла.

4.2.2. Экземпляры IS-IS

TRILL реализует экземпляр IS-IS, отличающийся от любого, применяемого на уровне L3 (маршрутизаторами). Кадры L3 IS-IS должны отличаться от кадров TRILL IS-IS даже при прохождении этих кадров L3 IS-IS через кампус RBridge.

Естественные кадры L3 IS-IS используют специальные групповые адреса получателей типа AllL1ISs или AllL2ISs. При инкапсуляции TRILL эти групповые адреса отображаются как Inner.MacDA, а Outer.MacDA будет групповым адресом All-RBridges.

В рамках TRILL имеется экземпляр IS-IS для всех RBridge в кампусе, как описано в параграфе 4.2.3. Этот экземпляр использует кадры TRILL IS-IS, которые отличаются по EtherType L2-IS-IS. Кроме того, для групповых кадров TRILL IS-IS имеется специальный групповой адрес All-IS-IS-RBridges. Кадры TRILL IS-IS не имеют заголовка TRILL.

ESADI является отдельным протоколом, отличающимся от экземпляра IS-IS, реализованного на всех RBridge. Отдельный экземпляр ESADI используется для каждой VLAN и кадры ESADI инкапсулируются подобно кадрам данных TRILL. После заголовка TRILL в кадрах ESADI размещается внутренний заголовок Ethernet с Inner.MacDA = All-ESADI-Rbridges и EtherType = L2-IS-IS, за которым следует кадр ESADI.

4.2.3. Кадры TRILL IS-IS

Все RBridge **должны** принимать участие в работе экземпляра TRILL IS-IS, что создает единую область L1 IS-IS с нулевым адресом. Кадры TRILL IS-IS никогда не пересылаются RBridge и при получении обрабатываются локально (такая обработка может заставить RBridge передать дополнительные кадры TRILL IS-IS).

Кадр TRILL IS-IS на канале 802.3 имеет показанную ниже структуру. Все такие кадры кодируются в EtherType. Порт RBridge, из которого передается такой кадр, будет вырезать внешний теги VLAN, если это задано в его конфигурации.

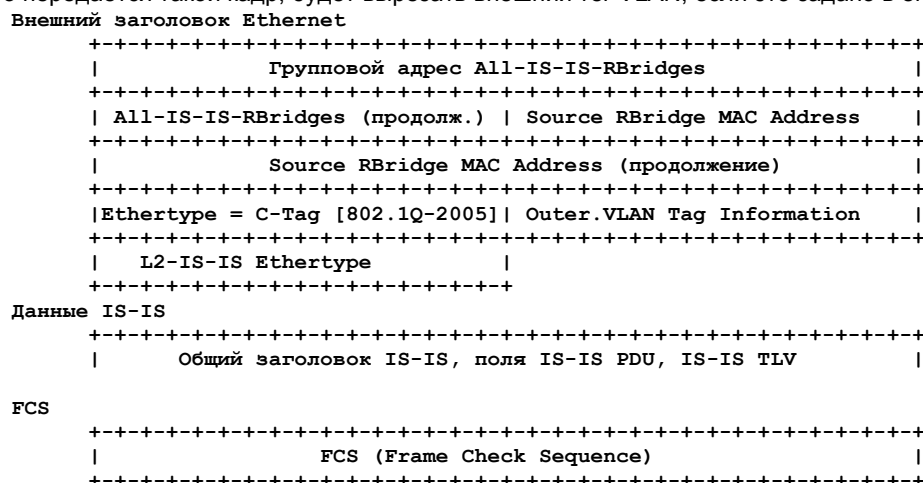


Рисунок 9. Формат кадра TRILL IS-IS.

VLAN, указанная в Outer.VLAN, будет назначенной (Designated) VLAN для канала, в который передается кадр, за исключением некоторых случаев TRILL Hello.

4.2.4. Приветствия TRILL, DRB и назначенные устройства пересылки

RBridge используют TRILL Hello, если только на уровне портов не настроено использование P2P Hello. Кадры TRILL-Hello описаны в параграфе 4.4.

RBridge обычно настраиваются на использование P2P Hello лишь в тех случаях, когда на канале имеется лишь два устройства. Однако могут возникать случаи когда настройка типа приветствий в RBridge ошибочна. Это не представляет опасности, но может снижать уровень связности между RBridge. Порт RBridge, настроенный на использование P2P Hello, игнорирует TRILL Hello и наоборот.

Если любой из портов RBridge на канале настроен на использование TRILL Hello, один из использующих TRILL Hello портов RBridge выбирается DRB (назначенный RBridge) для канала. Этот выбор основан на заданном в конфигурации приоритете (старшее поле) и MAC-адресе отправителя в кадрах TRILL-Hello. DRB, как описано в параграфе 4.2.4.2,

назначает VLAN для использования на канале в коммуникациях между RBridge через порты, не являющиеся P2P RBridge, и назначает себя или другой RBridge на канале узлом пересылки (параграф 4.2.4.3) для VLAN на этом канале.

4.2.4.1. Каналы с P2P Hello

Порты RBridge можно настроить на использование IS-IS P2P Hello. Это предполагает, что порт использует для соединения с другим RBridge соединение «точка-точка». RBridge **недопустимо** обслуживать какие-либо конечные станции (естественные кадры) на портах, настроенных на использование P2P Hello.

Как и в случае L3 IS-IS, такие порты P2P не участвуют в выборе DRB. Они передают все кадры с тегами VLAN как относящиеся к желаемой назначенной (Desired Designated) VLAN, настроенной для порта, хотя тег может быть вырезан при соответствующей настройке порта. Поскольку весь трафик через порт должен представлять кадры TRILL или кадры управления L2, такой порт не может быть назначенным узлом пересылки. Порты RBridge P2P **должны** использовать трехэтапное согласование IS-IS [RFC5303], чтобы с каналом были связаны расширенные идентификаторы устройств для целей прерывания (параграф 4.5.2).

Если некоторые узлы являются мостами, сеть на базе мостов, включающая эти мосты, будет выглядеть каналами с множественным доступом к подключенным устройствам RBridge, даже когда все простые каналы в сети являются физическими соединениями «точка-точка». Это во многих случаях требует использования TRILL Hello для корректной работы.

Хотя ошибочная настройка портов как P2P не создает опасностей, она может приводить к потере связности.

4.2.4.2. Назначенный RBridge

TRILL IS-IS выбирает один RBridge для каждого канала ЛВС в качестве назначенного RBridge (DRB), выполняющего особые функции, перечисленные ниже.

- Выбор для канала и анонсирование в TRILL Hello идентификатора назначенной VLAN ID для использования в коммуникациях между RBridge. Эта VLAN используется для всех кадров данных с инкапсуляцией TRILL, кадров ESADI и кадров TRILL IS-IS за исключением некоторых кадров TRILL-Hello.
- Если канал представлен в топологии IS-IS как псевдоузел, выбор идентификатора этого псевдоузла и анонсирование его в TRILL Hello, а также выпуск LSP от имени псевдоузла.
- Выпуск CSNP.
- Для каждой VLAN-х на канале выбор RBridge в качестве назначенного устройства пересылки VLAN-х (DRB **может** выбрать себя в качестве назначенного узла пересылки VLAN-х для всех или некоторых VLAN).
- Перед назначением узла пересылки VLAN-х (включая самого себя) ожидание не менее Holding Time (чтобы убедиться в том, что реально является DRB).
- Если настроена передача кадров TRILL-Hello, продолжение их отправки во все разрешенные VLAN, которые были указаны в наборе Announcing VLANs этого DRB (по умолчанию все разрешенные VLAN).

4.2.4.3. Назначенное устройство пересылки VLAN-х

Назначенный узел пересылки VLAN-х для канала отвечает за перечисленные ниже вопросы. В канале с точками предотвращения петель узел пересылки для порта, находящийся в состоянии inhibited (заблокирован), отбрасывает получаемые естественные кадры, не пересылая их, но все естественные кадры VLAN, для которой он назначен, узел не отбрасывает, а декапсулирует.

- Предотвращение петель.
 - Блокировка самого себя на задаваемое в конфигурации порта время от 0 до 30 секунд (по умолчанию) после обнаружения смены корневого моста на канале (параграф 4.9.3.2).
 - Блокировка самого себя для VLAN-х при получении кадра Hello, в котором отправитель заявляет себя назначенным узлом пересылки и который:
 - принят в VLAN-х (имеет VLAN-х в качестве Outer.VLAN) или
 - был изначально передан в VLAN-х, как указано в теле Hello.
 - Необязательный отказ от декапсуляции кадра от входного RBridge Rbm, если у него LSP этого RBm, а корневой мост на канале, в который следует пересылать не указан в списке корневых мостов RBm для VLAN-х. Это называется «проверкой декапсуляции» (decapsulation check) или «проверкой конфликта корневых мостов» (root bridge collision check).
- Пока нет блокировки (см. выше), прием естественного трафика VLAN-х и соответствующая пересылка его.
- Прием трафика VLAN-х для канала и (пока нет блокировки), передача его в естественной форме после декапсуляции.
- Определение (learning) MAC-адреса локальных узлов VLAN-х путем просмотра адреса отправителя в кадрах VLAN-х из канала.
- Необязательное определение (learning) порта локальных узлов VLAN-х на основе любых протоколов регистрации L2 типа ассоциаций и аутентификации IEEE 802.11.
- Отслеживание комбинаций {входной RBridge, VLAN, MAC-адрес} удаленных конечных узлов VLAN-х, определенных путем просмотра полей {входной RBridge, Inner.VLAN ID, Inner.MacSA} в кадрах VLAN-х принимаемых для декапсуляции в канал.
- Необязательное наблюдение естественных кадров IGMP [RFC3376], MLD [RFC2710] и MRD [RFC4286] для определения наличия локальных получателей группового трафика и групповых маршрутизаторов.

- Необязательное прослушивание сообщений TRILL ESADI для VLAN-х с целью определения триплетов {входной RBridge, VLAN-х, MAC-адрес} и уровня доверия к таким явно анонсируемым конечным узлам.
- Необязательное анонсирование в сообщениях ESADI конечных узлов VLAN-х на каналах, для которых выполняются функции назначенного узла пересылки.
- Передача кадров TRILL-Hello в VLAN-х, если набор Announcing VLANs, заданный для порта, не запрещает это.
- Прослушивание BPDU на общем связующем дереве для определения корневого моста (если он есть) на канале и указания в своем LSP полного набора корневых мостов, видимых на всех каналах, для которых выполняются функции назначенного узла пересылки для VLAN-х.

Когда назначенный узел пересылки видит смену DRB на канале, он больше не считает себя назначенным узлом пересылки, пока новый DRB не назначит его.

4.2.4.4. Информация TRILL LSP

Ниже перечислены информационные элементы TRILL IS-IS LSP, упоминаемые в этом документе. Элементы, не указанные в списке как необязательные, **должны** включаться всегда. Остальные элементы **можно** включать, если это явно не запрещено в данном документе. Реальное представление этой информации, а также значения типа и подтипа IS-IS для всех новых элементов данных TLV заданы в отдельных документах [RFC6165] [RFC6326].

- 1 Идентификаторы соседей IS-IS (псевдоузлы и RBridge) маршрутизирующего моста RBridge RBn и стоимость каналов до каждого из этих соседей. RBridge **должны** использовать Extended IS Reachability TLV (#22 или «метрика ширины» [RFC5305]), а использование IS Reachability TLV (#2 или «метрика узости») **недопустимо**. Для упрощения эффективной работы без настройки конфигурации и в соответствии с [802.1D] RBridge **следует** устанавливать по умолчанию в качестве стоимости канала целую часть результата деления числа 20 000 000 000 000 (двадцать триллионов) на скорость порта RBridge, но не более $2^{24}-2$ (16 777 214). Например, стоимость канала в порт 1 Гбит/с составит 20 000 (отметим, что число $2^{24}-1$ имеет в IS-IS специальное значение и будет исключать канал из маршрутов SPF). Принятая по умолчанию стоимость канала **может** быть уменьшена для агрегированных соединений и/или увеличена до значения, не превышающего $2^{24}-2$, если канал является ЛВС на базе мостов. Проверенное значение MTU для канала (параграф 4.3) **может** включаться в суб-TLV.
- 2 Указанная ниже информация в сочетании с каждым псевдонимом RBridge RBn:
 - 2.1 значение псевдонима (2 октета);
 - 2.2 8-битовое целое число без знака, указывающее приоритет RBn для данного псевдонима (параграф 3.7.3);
 - 2.3 16-битовое целое число без знака, указывающее приоритет данного псевдонима при выборе корня дерева распространения.
- 3 Максимальный номер версии заголовка TRILL, поддерживаемой RBridge RBn.
- 4 Приведенная ниже информация в дополнение к приоритету корня дерева на уровне псевдонима вкпе с определением и анонсированием дерева распространения (использование этой информации рассмотрено в параграфе 4.5).
 - 4.1 16-битовое целое число без знака, указывающее количество деревьев всех RBridge в кампусе, которое вычисляется в случае когда RBn имеет корень дерева с высшим приоритетом.
 - 4.2 16-битовое целое число без знака, указывающее количество деревьев, которые будет использовать RBn.
 - 4.3 16-битовое целое число без знака, указывающее максимальное количество деревьев, которые RBn способен рассчитать.
 - 4.4 Список псевдонимов, предназначенных для расчета деревьев распространения на всех RBridge в кампусе.
 - 4.5 Список псевдонимов, являющихся деревьями распространения, которые RBn может использовать, выступая в качестве входа для кадров со множеством получателей.
- 5 Список идентификаторов VLAN, подключенных непосредственно к Rbn для каналов, на которых RBn является назначенным узлом пересылки для этих VLAN¹. RBridge могут связывать анонсируемые соединения для разных групп VLAN с конкретными псевдонимами, которыми они владеют. В дополнение к этому LSP содержит указанную ниже информацию для каждой VLAN.
 - 5.1 Флаги подключенных групповых маршрутизаторов для VLAN. Это два бита информации, которые указывают наличие подключенного к RBridge группового маршрутизатора IPv4 и/или IPv6 для данной VLAN. RBridge, который не отслеживает управление групповой адресацией IP, **должен** устанавливать оба эти флага (параграф 4.5.4). Эта информация используется, поскольку отчеты о принадлежности к группам IGMP [RFC3376] и MLD [RFC2710] **должны** передаваться во все каналы с групповыми маршрутизаторами IP, а в каналы без таких маршрутизаторов передавать их **не следует**. Кроме того, все кадры для полученных от IP групповых адресов **должны** передаваться во все каналы с групповыми маршрутизаторами IP (в данной VLAN) в дополнение к каналам, откуда узел IP явно запросил включение в группу, для которой предназначен кадр, за исключением некоторых групповых адресов IP, которые **должны** трактоваться как широковещательные.
 - 5.2 Для каждой VLAN обязательный анонс набора идентификаторов корневых мостов для всех каналов RBn, на которых RBn назначен узлом пересылки для данной VLAN. При использовании на канале протокола MSTP это будет корневой мост CIST². Это нужно для быстрого обнаружения ситуаций, когда два облака L2

¹RBridge может анонсировать свое подключение к другим VLAN для того, чтобы получать дополнительные кадры с целью поддержки неких функций на базе VLAN, выходящих за рамки этой спецификации, как описано в параграфе 4.8.4 и в отдельном документе, посвященном отображению VLAN внутри RBridge.

²Common and Internal Spanning Tree - общее и внутреннее связующее дерево.

случайно сливаются, поскольку без этого временно могут существовать два DRB для одной VLAN на одном канале (параграф 4.2.4.3.).

- 5.3 (Необязательно) для каждой VLAN групповые адреса L2, выведенные из уведомлений IPv4 IGMP и IPv6 MLD, полученных от присоединенных конечных узлов данной VLAN, с указанием местоположения получателей для этих групповых адресов (параграф 4.5.5).
- 5.4 Для каждой VLAN флаг участия в протоколе ESADI, приоритет и время удержания. Установленный флаг показывает желание RBridge получать такие кадры TRILL ESADI (параграф 4.2.5.1).
- 5.5 Для каждой VLAN счетчик потерь статуса назначенного узла пересылки (параграф 4.8.3).
- 6 (Необязательно) размер наибольшего кадра TRILL IS-IS, который RBridge может обрабатывать с использованием `originatingLSPBufferSize TLV #14` (параграф 4.3).
- 7 (Необязательно) список групп VLAN, где определение адресов выполняется совместно во всей группе VLAN (параграф 4.8.4). Каждая группа VLAN является списком идентификаторов VLAN, где первый идентификатор является «основным», а все прочие - «вторичными». Это служит для обнаружения некорректных конфигураций и выходит за рамки данного документа. RBridge, не поддерживающие совместное определение адресов VLAN, игнорируют это поле.
- 8 (Необязательно) Authentication TLV #10 (раздел 6).

4.2.5. Протокол TRILL ESADI

Маршрутизирующие мосты RBridge, которые назначены узлами пересылки VLAN-х для канала, **могут** принимать участие в работе протокола TRILL ESADI для этой VLAN. Все транзитные RBridge **должны** корректно пересылать кадры TRILL ESADI, как будто они являются групповыми кадрами данных TRILL. Кадры TRILL ESADI структурно похожи на кадры IS-IS, но всегда инкапсулируются в кадры TRILL при передаче к линии, как будто это кадры данных TRILL.

В результате аткой пересылки протоколу ESADI на RBridge кажется, будто он напрямую подключен к общему виртуальному каналу, соединяющему со всеми RBridge кампуса, но каоторых работает ESADI для этой VLAN. RBridge, которые не реализуют протокол ESADI или не назначены узлами пересылки для данной VLAN, не будут декапсулировать или обрабатывать локально какие-либо кадры TRILL ESADI, полученные ими для данной VLAN. Иными словами, эти кадры будут просто туннелироваться через транзитные RBridge, которые будут трактовать их в точности так же, как групповые кадры данных TRILL, не выполняя при пересылке какой-либо специальной обработки.

Структура кадров TRILL ESADI, передаваемых в канал IEEE 802.3, показана ниже. Внешний тег VLAN не будет присутствовать, если он был вырезан портом, через который кадр был передан.

```

Внешний заголовок Ethernet
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Next Hop Destination Address                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Next Hop Destination Address | Sending RBridge MAC Address |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Sending RBridge Port MAC Address                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Ethertype = C-Tag [802.1Q-2005] | Outer.VLAN Tag Information |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
Заголовок TRILL
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Ethertype = TRILL | V | R | M | Op-Length | Hop Count |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Egress (Dist. Tree) Nickname | Ingress (Origin) Nickname |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
Внутренний заголовок Ethernet
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     All-ESADI-RBridges Multicast Address                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| All-ESADI-RBridges continued | Origin RBridge MAC Address |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Origin RBridge MAC Address continued                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Ethertype = C-Tag [802.1Q-2005] | Inner.VLAN Tag Information |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Ethertype = L2-IS-IS |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
Данные ESADI (в формате IS-IS)
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| IS-IS Common Header, IS-IS PDU Specific Fields, IS-IS TLVs |
FCS
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     FCS (Frame Check Sequence)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

Рисунок 10. Формат кадра TRILL ESADI.

Поля Next Hop Destination Address или Outer.MacDA содержит групповой адрес All-Rbridges. VLAN в поле Outer.VLAN всегда будет назначенной (Designated) VLAN для канала, через который кадр передается. Поля V и R будут иметь значение 0, а M - 1. VLAN в поле Inner.VLAN будет указывать VLAN, к которой применяется кадр ESADI. В поле Origin RBridge MAC Address или Inner.MacSA **должен** указываться уникальный в глобальном масштабе MAC-адрес, принадлежащий RBridge, создавшему кадр ESADI (например, MAC-адрес любого из портов) и каждый RBridge **должен** использовать одно значение Inner.MacSA для всех создаваемых им кадров ESADI.

4.2.5.1. Участие в TRILL ESADI

RBridge не передает никаких Hello в результате участия в протоколе ESADI. Информации, доступной в базе данных о состояниях каналов TRILL IS-IS, достаточно для определения ESADI DRB на виртуальном канале протокола ESADI для каждой VLAN. В частности, база состояний каналов для каждого RBridge включает VLAN (если они есть), для которых RBridge участвует в протоколе ESADI, приоритет выбора в качестве DRB для протокола ESADI в каждой из этих VLAN, время удержания, а также идентификатор системы IS-IS для исключения лишнего (breaking tie) при выборе приоритета.

RBridge не нужно выполнять каких-либо маршрутных расчетов в результате участия в протоколе ESADI. Поскольку все RBridge, участвующие в ESADI для конкретной VLAN, представляются соединенными общим виртуальным каналом, маршрутные решения не требуются. Участвующий в протоколе RBridge просто передает создаваемые им кадры ESADI в этот виртуальный канал.

ESADI DRB передает кадры TRILL-ESADI-CSNP в виртуальный канал ESADI. Для отказоустойчивости принимающий участие в протоколе RBridge, который определил, что ESADI DRB следует быть некому другому RBridge на этом виртуальном канале, но тот не принял и не передал TRILL-ESADI-CSNP в течение времени удержания ESADI DRB, **может** сам передать TRILL-ESADI-CSNP в виртуальный канал. Участвующему в протоколе RBridge, который определил отсутствие других RBridge в протоколе ESADI для конкретной VLAN, **не следует** передавать информацию ESADI или кадры TRILL-ESADI-CSNP в виртуальный канал для этой VLAN.

4.2.5.2. Информация TRILL ESADI

Информация, распространяемая протоколом ESADI, представляет собой список MAC-адресов локальных конечных станций, известных исходному RBridge, с 1-октетным уровнем «доверия» (0 - 254) для каждого адреса (параграф 4.8).

Это необязательное представление для трансляции VLAN ID в RBridge, как указано в [VLAN-MAPPING], включая трансляцию кадров TRILL ESADI. Если кадры TRILL ESADI могут содержать идентификаторы VLAN в произвольном месте внутри себя, такая трансляция становится непрактичной. Поэтому в кадры TRILL ESADI **недопустимо** включать идентификаторы VLAN для VLAN, к которой они применяются, после тега Inner.VLAN.

4.2.6. SPF, пересылка и неоднозначные получатели

В этом параграфе описан желаемый логический результат. Для его достижения могут использоваться разные методы реализации, если они обеспечивают такое же поведение пересылки кадров.

При построении таблицы пересылки RBridge RB1 рассчитывает кратчайшие пути от самого себя, как описано в Приложении С.1 к [RFC1195]. Псевдонимы добавляются в расчет кратчайшего пути на финальном этапе, как с конечными узлами. Если один и тот же псевдоним заявляет несколько RBridge (скажем, RbA и RbB), это будет временным состоянием и RbA или RbB выберет другой псевдоним. Однако RB1 просто добавит этот псевдоним, как будто он связан сразу с RbA и RbB, при этом он будет использовать свой стандартный расчет кратчайшего пути для выбора следующего этапа пересылки.

Входной RBridge RB2 отображает заведомо индивидуальные MAC-адреса получателей и VLAN в естественных кадрах на псевдоним выходного RBridge. Если RB2 узнает адреса лишь из просмотра полученных и декапсулированных кадров, тогда эти MAC-не могут дублироваться внутри VLAN в таблицах RB2, поскольку свежая информация с уровнем доверия не меньше имеющегося у прежней будет переписывать имеющуюся информацию, а информация с меньшим уровнем доверия будет игнорироваться. Однако дубликаты MAC внутри VLAN могут появляться в данных ESADI, а также возможны совпадения адресов между данными ESADI и адресами, полученными из просмотра полученных и декапсулированных кадров, заданными в конфигурации вручную или узнанными от протокола регистрации L2. При возникновении дубликатов MAC-адресов внутри VLAN устройство RB2 будет передавать кадр по MAC с более высоким уровнем доверия. Если уровни доверия для дубликатов совпадают, для согласованности предполагается, что RB2 будет отправлять все такие кадры (или все такие кадры одного потока ECMP) в направлении одного выходного RBridge. Однако и при использовании других правил не будет возникать проблем с сетью, поскольку транзитные RBridge не проверяют Inner.MacDA для заведомо индивидуальных кадров.

4.3. Размер MTU на канале между RBridge

Есть две причины, по которым важно знать размер кадров, который может поддерживать каждый канал между RBridge в кампусе.

1. RBridge RB1 должен знать размер сообщений с информацией о состоянии канала, которые он может генерировать так, чтобы они гарантировано пересылались через все каналы между RBridge в кампусе.
2. Если средства организации трафика знают какие каналы поддерживают размер кадров сверх минимального, можно будет рассчитать пути передачи больших пакетов данных.

4.3.1. Определение TRILL IS-IS MTU в масштабе кампуса

В стабильном кампусе в конечном итоге должно быть достигнуто соглашение между всеми RBridge в части значения Sz - минимального приемлемого размера на каналах между RBridge в кампусе для корректной работы TRILL IS-IS. Все RBridge **должны** форматировать свои сообщения с информацией о состоянии канала в блоки, размер которых не превышает согласованного Sz. Каждому RBridge RB1 **следует** проверить соседние RBridge (скажем, RB2), чтобы убедиться в возможности пересылки через канал RB1-RB2 пакетов, размер которых составляет по меньшей мере Sz.

Значение Sz не оказывает прямого влияния на конечные станции и не связан напрямую с какими-либо значениями path MTU на пути между конечными станциями. Способы использования Sz или каких-либо данных об MTU на каналах, собранных TRILL IS-IS при организации трафика на маршрутах или определении MTU для любого пути, выходят за рамки этого документа. Естественные кадры, размер которых после инкапсуляции TRILL превысит MTU для канала, обычно будут отбрасываться.

Sz определяется за счет того, что каждый RBridge (необязательно) анонсирует в своих LSP свое представление о значении Sz в кампусе. Этот элемент LSP называется в IS-IS originatingLSPBufferSize, TLV #14. Принятое по умолчанию и минимальное значение для Sz, а также неявно анонсируемое значение Sz в отсутствие TLV составляет 1470 октетов. Этот размер (который также является минимальным размером TRILL-Hello) был выбран для того, чтобы сделать

крайне маловероятным возникновение проблем MTU для управляющих кадров TRILL с разумными дополнительными заголовками, тегами и/или инкапсуляцией на канале между RBridge.

Значением Sz в масштабе кампуса выбирается минимальное из значений Sz анонсируемых RBridge.

4.3.2. Тестирование размера MTU для канала

Имеется два новых типа сообщений TRILL IS-IS для обмена между парой соседних RBridge с целью проверки размера пакетов, передаваемых через соединение в обоих направлениях:

- MTU-probe;
- MTU-ack.

Оба типа сообщений заполняются до проверяемого размера.

Передача сообщения MTU-probe не обязательна, однако RBridge RB2, получивший MTU-probe от RB1, **должен** ответить сообщением MTU-ack, дополненным до размера полученного MTU-probe. Сообщение MTU-probe **можно** передать по групповому адресу All-IS-IS-Rbridges¹ или индивидуальному адресу конкретного RBridge. Сообщения MTU-ack обычно передаются по индивидуальному адресу отправителя MTU-probe, но **могут** передаваться и по групповому адресу All-IS-IS-Rbridges¹.

Если RB1 не получает MTU-ack на пробный пакет размера X от RB2 после k попыток (k - конфигурационный параметр, по умолчанию равный 3), RB1 предполагает, что канал RB1-RB2 не поддерживает размер X. Если X не превышает Sz, RB1 устанавливает флаг failed minimum MTU test (отказ при проверке минимального MTU) для RB2 в своем сообщении Hello. Если кадры размера X проходят и X > Sz, RB1 анонсирует наибольшее из проверенных значений X для каждого отношения смежности в сообщениях TRILL Hello, передаваемых в соответствующий канал, и **может** анонсировать X как атрибут канала устройству RB2 в своих LSP.

4.4. Протокол TRILL-Hello

Протокол TRILL-Hello несколько отличается от протокола L3 IS-IS LAN Hello и использует новый тип сообщений IS-IS - TRILL-Hello.

4.4.1. Обоснование TRILL-Hello

Причина определения нового типа каналов в TRILL заключается в том, в L3 IS-IS протокол LAN Hello может приводить к выбору нескольких назначенных маршрутизаторов (DR²), поскольку при выборе DR маршрутизаторы игнорируют другие маршрутизаторы, с которыми нет двухсторонней связи. Кроме того, сообщения L3 IS-IS LAN Hello дополняются, чтобы предотвратить организацию отношений смежности между соседями, которые не могут обмениваться между собой пакетами максимального размера. Это означает для L3 IS-IS, что соседи, которые соединены между собой, но MTU в этом соединении меньше значения, которое они считают максимальным размером пакетов, не будут видеть сообщений Hello от другой стороны. В результате маршрутизаторы могут формировать «клик» и канал превращается в несколько псевдоузлов.

Такое поведение хорошо для уровня L3, но не подходит для L2, где могут возникать петли при наличии множества DRB. Поэтому протокол TRILL-Hello несколько отличается от протокола LAN Hello в L3 IS-IS.

Еще одна связанная с TRILL-Hello проблема заключается в гарантии того, что подмножества информации, которые могут появляться в любом отдельном сообщении, будут обрабатываться в духе IS-IS LSP и CSNP. Кадры TRILL-Hello могут быть очень большими, хотя они и не дополняются. Примером этого могут быть ситуации когда некая магистральная технология используется для соединения сотен сайтов TRILL, в результате чего TRILL становится гигантской сетью Ethernet, где RBridge, подключенные к этому облаку, будут воспринимать эту магистраль как один канал с сотнями соседей.

В TRILL (в отличие от L3 IS-IS) DRB выбирается исключительно на основе приоритета и MAC-адреса. Иными словами, если RB2 получает TRILL-Hello от RB1 с более высокими (приоритет, MAC), RB2 выбирает RB1 в качестве DRB, независимо от того, указывает ли RB1 устройство RB2 в своем TRILL-Hello.

Хотя список соседей в TRILL-Hello не влияет на выбор DRB, он определяет, что анонсируется в LSP. RB1 указывает только каналы к RBridge, с которыми у него есть двухсторонняя связь. Если RB1 является DRB на канале и по какой-либо причине (несоответствие MTU или односторонняя связь) RB1 и RB2 не имеют двухсторонней связи, RB2 не указывает свой канал к RB1 (или псевдоузлу), а RB1 (или RB1 от имени псевдоузла) не указывает канал к RB2.

4.4.2. Содержимое TRILL-Hello и синхронизация

TRILL-Hello - это новый тип кадров IS-IS. Сообщение начинается с обычного фиксированного заголовка (как в IS-IS LAN Hello), который включает 7-битовое значение приоритета для выбора RBridge в качестве DRB на данном канале. Кадры TRILL-Hello передаются с такой же временной координацией, как IS-IS LAN Hello.

Кадрам TRILL-Hello, включая Outer.MacDA и Outer.MacSA, но без учета Outer.VLAN и других тегов, **недопустимо** превышать размер 1470 октетов и дополнять эти сообщения **не следует**. Приведенная ниже информация **должна** присутствовать в каждом кадре TRILL-Hello, TLV в действительности могут быть sub-TLV, как описано в [RFC6165] [RFC6326].

1. VLAN ID назначенной VLAN для канала.
2. Копия Outer.VLAN ID которая была в кадре Hello при передаче.
3. 16-битовый идентификатор порта, назначенный передающим RBridge порту, куда передается TRILL-Hello, так, чтобы у двух портов данного RBridge идентификаторы не совпадали.
4. Псевдоним передающего RBridge.

¹В оригинале ошибочно сказано All-Rbridges, см. <https://www.rfc-editor.org/errata/eid3004>. Прим. перев.

²Designated Router.

5. Перечисленные ниже флаги:
 - 5.a. флаг, показывающий что отправитель обнаружил трансляцию VLAN на канале в течение двух последних интервалов Holding Time;
 - 5.b. флаг, показывающий что отправитель считает себя назначенным узлом пересылки для VLAN и порта, в который передан TRILL-Hello.

В кадрах **может** появляться также указанная ниже информация.

1. Набор VLAN, для которых на порту разрешено обслуживание конечных станций.
2. Флаги, описанные ниже:
 - 2.a. флаг, установка которого показывает, что порт отправителя был настроен как порт доступа;
 - 2.b. флаг, установка которого показывает, что порт отправителя был настроен как транковый;
 - 2.c. флаг обхода псевдоузла описанный ниже в этом параграфе.
3. Если отправитель является DRB, могут присутствовать мосты Rbridge (исключая самого себя), назначенные пересылающими для данного канала и VLAN, к которым они подключены. Как описано ниже, этот блок TLV рассчитан на то, что не всю информацию о назначениях следует включать в каждое сообщение Hello. Его отсутствие означает, что назначенным пересылающим устройствам, следует продолжать работать, как было назначено ранее.
4. Список соседей TRILL. Это новый блок TLV, отличающийся от IS-IS Neighbor TLV для согласования с фрагментацией и отчетов об MTU на канале (параграф 4.4.2.1).

Блок Appointed Forwarders TLV задает диапазон VLAN, а в этом диапазоне - Rbridge (если таковой имеется), отличный от DRB и назначенный пересылающим устройством для VLAN этого диапазона [RFC6326]. Назначение RBridge пересылающим на порту для VLAN, которая не разрешена на этом порту, не оказывает влияния.

Предполагается, что многие каналы между RBridge будут относиться в типу «точка-точка» и в этом случае использование псевдоузлов явно увеличивает сложность. Если DRB задает бит обхода псевдоузлов в своих TRILL-Hello, мосты RBridge на канале просто укажут свою смежность типа «точка-точка». Это не оказывает влияния на лавинную рассылку LSP в канал и влияет лишь на генерацию LSP.

Например, если кроме RB1 и RB2 на канале нет других RBridge и RB1 является DRB, тогда при создании RB1 псевдоузла, который он будет использовать, будет три 3 LSP - скажем, RB1.25 (псевдоузел), RB1 и RB2, где RB1.25 будет сообщать о связности с RB1 и RB2, а RB1 и RB2 просто сообщать, что они соединены с RB1.25. Если DRB RB1 установит бит обхода псевдоузлов в своих сообщениях Hello, будет только два 2 LSP - RB1 и RB2 сообщат лишь о соединении в другом узлом.

DRB **следует** устанавливать бит обхода псевдоузлов для своих каналов, если только он не видит для конкретного канала одновременных смежностей с той или иной точкой в связи с ее последней перезагрузкой.

4.4.2.1. Список соседей TRILL

Новый блок TRILL Neighbor TLV включает для каждого соседа из списка перечисленную ниже информацию.

1. MAC-адрес соседа.
2. Размер MTU для этого соседа в виде 2-октетного целого числа без знака, указывающего количество 4-октетных блоков. Нулевое значение говорит, что величина MTU не проверялась.
3. Флаг отказа при проверке минимального MTU (failed minimum MTU test).

Чтобы разрешить частичные отчеты о соседях, идентификаторы соседей **должны** быть отсортированы. Если в сообщении Hello от RB1 указан набор соседей { X1, X2, X3, ... Xn }, то X1 < X2 < X3, ... < Xn. Если идентификатор RBridge RB2 находится между X1 и Xn и не присутствует в RB1 Hello, RB2 знает, что RB1 не слышит Hello от RB2.

В дополнение к этому имеется два общих флага TRILL Neighbor List TLV для наименьшего и наибольшего идентификатора, указанных в этом сообщении Hello. Если все соседи попадают в TRILL-Hello от RB1, устанавливаются оба флага.

Если RB1 указывает { X1, ... Xn } в своем сообщении Hello с установленным флагом smallest, а RB2 ID меньше X1, RB2 знает, что RB1 не слышит Hello от RB2. Аналогично, если RB2 ID больше Xn и установлен флаг largest, RB2 знает, что RB1 не слышит Hello от RB2.

Чтобы любой RBridge RB2 мог достоверно определить, может ли RB1 слышать RB2, список соседей RB1 **должен** в конечном итоге охватывать все возможные диапазоны идентификаторов, т. е. в течение периода, определяемого политикой RB1 и не обязательно в какой-либо конкретный интервал типа времени ожидания. Иными словами, если X1 является наименьшим идентификатором, указанным в обном из списков соседей RB1 и флаг smallest не установлен, идентификатор X1 **должен** также появляться как наибольший идентификатор в другом списке соседей TRILL-Hello. Или фрагменты могут перекрываться, пока нет зазоров, при которых некий диапазон между Xi Xj не появляется ни в одном из фрагментов.

4.4.3. Тегу VLAN для TRILL MTU-Probe и TRILL Hello

Механизм проб MTU предназначен для определения MTU между мостами RBridge. Пробы MTU и подтверждения передаются лишь на Designated VLAN.

RBridge RBn поддерживает для каждого порта одну и ту же информацию VLAN как клиент моста IEEE [802.1Q-2005], включая набор VLAN, для которых разрешен вывод через порт (параграф 4.9.2). В дополнение к этому RBn поддерживает относящиеся к TRILL параметры VLAN на уровне порта, перечисленные ниже.

- a) Желаемая Designated VLAN - виртуальная ЛВС, которую RBN, будучи DRB, станет указывать в своих TRILL-Hello как the VLAN для использования всеми RBridge на канале для обмена всеми кадрами TRILL за исключением некоторых TRILL-Hello. Это **должна** быть VLAN, разрешенная на порту RBN. По умолчанию это VLAN с наименьшим из числа разрешенных номером, в качестве которой по умолчанию в RBridge служит VLAN 1.
- b) Designated VLAN - виртуальная ЛВС, которая будет использоваться на канале для всех кадров TRILL за исключением некоторых TRILL Hello. Это желаемая RBN Desired Designated VLAN, если RBN считает себя DRB или Designated VLAN DRB Hello, если RBN не является DRB.
- c) Набор анонсирующих VLAN. По умолчанию это набор разрешенных VLAN на порту, но в конфигурации может задаваться и подмножество разрешенных VLAN.
- d) Набор пересылающих VLAN - множество VLAN, для которых порт RBridge назначен рассылающим для VLAN на порту. В этом наборе **должны** содержаться лишь разрешенные для порта VLAN (возможно все).

На каждом из своих портов, где не настроена передача P2P Hello, RBridge передает TRILL-Hello с Outer.VLAN каждой VLAN из набора VLAN. Этот набор зависит от статуса DRB в RBridge и указанных выше параметров VLAN. RBridge передают TRILL Hello с Outer.VLAN назначенной сети (Designated VLAN), пока эта VLAN не разрешена на данном порту. В дополнение к этому DRB передает TRILL Hello с тем же Outer.VLAN каждой включенной VLAN в своем наборе Announcing VLANs. Все RBridge, не являющиеся DRB, передают TRILL-Hello с тем же Outer.VLAN tagged каждой включенной VLAN, которая входит в пересечение набора Forwarding VLAN и Announcing VLAN. При передаче кадров TRILL-Hello выполняются приведенные ниже условия.

```
Если отправитель является DRB
пересечение ( включенные VLAN, объединение ( Designated VLAN, Announcing VLANs ) )
Если отправитель не является DRB
пересечение ( включенные VLAN, объединение ( Designated VLAN,
пересечение ( Forwarding VLAN, Announcing VLANs ) ) )
```

Установка для Announcing VLANs пустого значения минимизирует число TRILL-Hello. В этом случае TRILL-Hello помечаются лишь Designated VLAN. Следует проявлять осторожность при настройке RBridge без отправки TRILL Hello в любую сеть VLAN, где этот RBridge назначен пересылающим, поскольку при некоторых обстоятельствах отказ передавать такие сообщения Hello может приводить к возникновению петель.

В этой спецификации число TRILL-Hellos максимизируется путем установки по умолчанию в Announcing VLANs полного набора идентификаторов включенных VLAN ID. В этом случае DRB будет передавать кадры TRILL-Hello, помеченные всеми тегами Enabled VLAN, кроме того, не являющийся DRB мост RBridge RBN будет передавать кадры TRILL-Hello с тем же Designated VLAN (если она включена) и с тегами всех VLAN, для которых RBN назначен пересылающим. Возможна передача даже большего числа TRILL-Hello. В частности, мосты RBridge, не являющиеся DRB, могут передавать TRILL-Hello в разрешенные VLAN для которых они не назначены пересылающими и которые не являются Designated VLAN. Это не причинит никакого вреда, но увеличит нагрузку на каналы связи и повысит объем обработки.

При активизации (up) порта RBridge он будет считать себя DRB для этого порта, пока не получит TRILL-Hello от RBridge с большим приоритетом, и на основании этого будет передавать пакеты TRILL-Hello. Даже если мост распознал другой RBridge на канале в качестве DRB, при отсутствии на этом порту TRILL-Hello от RBridge с более высоким приоритетом, как DRB в течение достаточно долгого времени (как указано в IS-IS), он продолжит считать себя DRB.

4.4.4. Множество портов на одном канале

RBridge RB1 может иметь несколько портов на одном канале и в этом случае для RB1 важно распознать такие порты. Например, если RB1 назначен пересылающим для VLAN A, он знает лишь об одном своем порту как о назначенном пересылающем для VLAN A на этом канале.

RB1 обнаруживает такие ситуации на основе приема сообщений с одним идентификатором псевдоузла IS-IS через несколько портов. RB1 может иметь набор портов (скажем, { p1, p2, p3 }) на одном канале, другой (например, { p4, p5 }) - на втором, а остальные порты (скажем, p6, p7, p8), подключены. Назовем множество портов, подключенных к одному каналу «группой портов» (port group).

Если RB1 видит, что набор портов (например, { p1, p2, p3 }) является группой на канале, тогда он **должен** обеспечить предотвращение петель при инкапсуляции и декапсуляции пакетов, связанных с этим каналом. Если RB1 назначен пересылающим для VLAN A на канале Ethernet, он **должен** инкапсулировать и декапсулировать VLAN A лишь на одном из портов. Однако, будучи назначенным пересылающим для нескольких VLAN, RB1 **может** распределять нагрузку между портами, используя один порт для одного набора VLAN, другой - для другого (не связанного с первым).

Если RB1 видит трансляцию VLAN (параграф 4.4.5), ему **недопустимо** распределять нагрузку, будучи назначенным пересылающим, и вместо этого он **должен** работать как назначенный для VLAN пересылающий на этом канале лишь через один из портов в группе.

При пересылке групповых кадров с инкапсуляцией TRILL в канал (или из него), где RB1 имеет группу портов, RB1 **может** распределять нагрузку между портами, не дублируя кадры и сохраняя кадры одного потока в одном порту. Если сосед RB1 на этом канале (RB2) воспринимает групповые кадры на этом дереве данного канала от RB1, RB2 **должен** воспринимать кадр из любой смежности между RB2 и RB1 на этом канале.

Если несколько портов RBridge подключено к каналу и эти порты имеют один MAC-адрес, их можно различать по идентификаторам порта в TRILL-Hello.

4.4.5. Трансляция VLAN в канале

IEEE [802.1Q-2005] не предусматривает для мостов изменения C-tag VLAN ID в принятых кадрах с тегами, т. е. трансляцию VLAN. Тем не менее, некоторые мосты поддерживают такую возможность и в таких случаях ЛВС на основе мостов можно настроить для отображения такого поведения. Например, порт моста можно настроить на вырезание тегов VLAN на выходе и передачу полученных кадров без тега в канал, ведущий к порту другого моста, настроенному помечать такие кадры тегом другой VLAN. Хотя настройки каждого порта корректны по отношению к [802.1Q-2005], в совокупности выполняются действия, не разрешенные для отдельного клиентского моста [802.1Q-2005]. Поскольку

порты RBridge имеют такие же функции VLAN, как у пользовательских мостов [802.1Q-2005], это может происходить даже при отсутствии мостов (отображение VLAN называется в IEEE 802.1 трансляцией - VLAN ID translation).

RBridge включают идентификатор Outer.VLAN ID в каждое сообщение TRILL-Hello. При получении TRILL-Hello RBridge сравнивает сохраненную копию с Outer.VLAN ID из принятого кадра. Если значения различаются и VLAN ID в пакете Hello имеет значение X, а Outer.VLAN - Y, можно предположить трансляцию VLAN ID X в VLAN ID Y.

Когда не являющийся DRB мост RB2 обнаруживает отображение VLAN на основе принятого TRILL-Hello, где тег VLAN в теле Hello отличается от тега во внешнем заголовке, он устанавливает флаг во всех своих TRILL-Hello на время двух интервалов Holding Time с момента последнего обнаружения им трансляции VLAN. Когда DRB RB1 знает о трансляции VLAN из принятого кадра TRILL-Hello с отображением или из флага VLAN mapping detected в TRILL-Hello от соседа на канале, RB1 заново назначает пересылающие мосты для VLAN, чтобы на канале был лишь один пересылающий для всех VLAN.

4.5. Деревья распространения

RBridge используют деревья распространения для пересылки групповых кадров (параграф 2.4.2). Деревья распространения являются двунаправленными (не ориентированными). Хотя одного дерева логически достаточно для всего кампуса, расчет дополнительных деревьев оправдан по нескольким причинам. Это позволяет использовать несколько путей для групповых кадров и выбрать корень дерева ближе (в пределах - входной RBridge). Более близкий корень повышает эффективность доставки групповых кадров, которые будут передаваться в подмножество каналов кампуса, и снижает вероятность нарушения порядка доставки когда индивидуальный адрес меняет состояние между неизвестным и известным. Если используется приложение, где случайное нарушение порядка доставки вызывает проблемы, кампус RBridge следует организовать так, чтобы вероятность нарушения порядка была максимально низкой, например, с помощью протокола ESADI, настраивая адреса для устранения кадров неизвестным индивидуальным получателем или используя кадры keep alive.

Дополнительный уровень гибкости обеспечивает возможность RBridge иметь несколько псевдонимов и, следовательно, несколько деревьев с корнем в одном RBridge. Поскольку номер дерева используется для устранения конфликтов равноценных путей, разные деревья даже с корнем в одном RBridge будут давать разную стоимость путей.

Способ выбора мостом RBridge одного или нескольких деревьев распространения для использования с групповыми кадрами выходит за рамки этого документа. Однако по указанным выше причинам при отсутствии других факторов хорошим выбором является дерево корень которого обеспечивает наименьшую стоимость от входного RBridge и оно применяется по умолчанию для входного RBridge при использовании для распространения групповых кадров одного дерева.

RBridge заранее рассчитывают все деревья, которые могут использоваться, и сохраняют состояние для фильтров проверки пересылки по обратному пути (Reverse Path Forwarding Check, параграф 4.5.2). Поскольку номер дерева служит для устранения конфликтов, всем RBridge важно знать:

- сколько деревьев нужно рассчитывать;
- какие деревья рассчитывать;
- какой номер имеет каждое дерево;
- какое из деревьев каждого входного RBridge можно выбрать (для фильтров Reverse Path Forwarding Check).

Каждый RBridge анонсирует в своих LSP приоритет корня дерева для своих псевдонимов - 16-битовое целое число без знака, которое для ненастроенного RBridge имеет значение 0x8000. Корни деревьев упорядочиваются в соответствии с приоритетом (большее значение - выше приоритет), затем по идентификатору системы, а при их совпадении - по большему значению псевдонима, трактуемого как целое число без знака.

Каждый RBridge анонсирует в своих LSP максимальное число деревьев распространения, которые мост может рассчитать, и число деревьев, которые он хочет, чтобы рассчитывалось на всех RBridge кампуса. Число деревьев (k), рассчитываемых для кампуса, - это число, желаемое RBridge RB1, у которого псевдоним имеет высший приоритет корня дерева, но не более наименьшего числа деревьев, поддерживаемых RBridge в кампусе. Если RB1 не задает конкретные корни деревьев распространения, как описано ниже, все RBridge в кампусе будут рассчитывать k деревьев с высшим приоритетом. Отметим, что некоторые из этих k деревьев могут оказаться в одном RBridge, имеющем множество псевдонимов.

Если RBridge указывает число рассчитываемых им или желаемым для кампуса деревьев 0, это считается указанием одного дерева. Т. е. по умолчанию k = 1.

В дополнение к этому RBridge RB1 с высшим приоритетом корня может явно анонсировать набор из s деревьев, предоставив список псевдонимов s. В этом случае будут рассчитываться первые k из этих s деревьев. Если s < k или любой из s псевдонимов, связанных с деревьями, анонсируемыми RB1, отсутствует в базе данных LSP, устройства RBridge по-прежнему рассчитывают k деревьев и выбранные деревья будут иметь наивысший приоритет.

Для приведенных правил имеется 2 исключения, которые могут снижать число рассчитываемых деревьев.

- Псевдоним с нулевым приоритетом корня не выбирается по приоритету, но может быть выбран путем указания устройством RBridge, владеющим псевдонимом с наивысшим приоритетом. Единственным исключением из этого правила является случай, когда все псевдонимы имеют приоритет 0, то в качестве корня используется наивысший приоритет, определенный с помощью устранения конфликтов (tiebreaker), поэтому во всех случаях гарантируется наличие хотя бы одного дерева распространения.
- Временно в списке рассчитываемых деревьев может появиться несколько идентичных псевдонимов. В таком случае бесполезно рассчитывать деревья для псевдонимов, которые будут утрачены содержащими их RBridge. Поэтому дерево распространения рассчитывается только для экземпляра псевдонима, у которого приоритет сохранения псевдонима является наивысшим, с целью снижения числа рассчитываемых деревьев. Также не было бы большого смысла идти вниз по порядку приоритета для поддержки числа деревьев, поскольку дополнительные корни, найденные таким способом, будут действительны лишь в течение короткого времени.

Рассчитанные для кампуса k деревьев упорядочиваются по номерам от 1 до k . В дополнение к анонсированию значения k устройство RB1 может явно анонсировать множество из s деревьев, предоставляя список из s псевдонимов.

- Если $s == k$, деревья нумеруются в порядке их анонсирования.
- Если $s == 0$, деревья нумеруются в порядке снижения приоритета. Например, если RB1 анонсирует лишь $k=2$, дерево с высшим приоритетом будет иметь номер 1, а со следующим - номер 2.
- Если $s < k$, анонсированные RB1 деревья нумеруются с 1 в порядке анонсирования. Затем среди оставшихся выбираются деревья с наивысшим приоритетом и нумерация продолжается. Например, если RB1 анонсирует $k=4$ и $\{T_x, T_y\}$ как псевдонимы корней деревьев, а приоритеты в кампусе упорядочены как $T_y > T_a > T_c > T_b > T_x$, T_x получит номер 1, а T_y - 2, поскольку в этом порядке они объявлены RB1. Затем T_a получит номер 3, а T_c - 4 в соответствии с приоритетом деревьев, которые еще не пронумерованы.

4.5.1. Расчет дерева распространения

RBridge не применяют связующее дерево для расчета деревьев распространения, используя в этих расчетах данные о состоянии каналов и выбирая псевдоним конкретного RBridge в качестве корня. Каждый RBridge RBn независимо рассчитывает дерево с корнем Rbi, выполняя расчет кратчайшего пути SPF с Rbi в качестве корня, для чего не требуется дополнительного обмена информацией.

При расчете дерева распространения важно, чтобы все RBridge выбирали для дерева одни и те же каналы. Поэтому при наличии равноценных путей для конкретного дерева все RBridge должны использовать одни и те же средства устранения конфликтов. Желательно также разрешить распределение трафика через как можно большее число каналов. По этой причине простые средства устранения конфликтов (tiebreaker), такие как «выбор всегда родителя с меньшим идентификатором», применять нежелательно. Взамен TRILL использует номер дерева в качестве параметра алгоритма устранения конфликтов.

При расчете дерева с номером j запоминаются все возможные равноценные родители для узла N . После расчета всего «дерева» (в реальности ориентированного графа) для каждого узла N , если он имеет p родителей, они упорядочиваются по возрастанию в соответствии с 7-октетными IS-IS ID, рассматриваемыми как целые числа без знака и нумерация начинается с 0. Для дерева j выбирается родитель с $N = j \bmod p$.

Отметим, что может быть множество равноценных каналов между N и возможным родителем P , не имеющих псевдоузлов, поскольку эти каналы имеют тип «точка-точка» или подавляют псевдоузлы. Такие каналы считаются одним каналом при построении дерева, поскольку все они имеют одного родителя P с идентификатором IS-IS ID = $P.0$.

Иными словами, набор возможных родителей N для дерева с корнем R состоит из узлов с равноценными путями от R к N^1 и разными идентификаторами IS-IS ID, полученными из LSP.

4.5.2. Проверка кадров с множеством получателей

При получении RBridge группового кадра с инкапсуляцией TRILL выполняются 4 проверки, которые могут приводить к отбрасыванию кадра.

1. **Проверка смежности в дереве.** Каждый RBridge RBn хранит свои смежности (пары { port, neighbor }) для каждого рассчитанного им дерева. Одна из этих смежностей ведет к корневому Rbi, а остальные - к листьям. После выбора смежностей неважно, какая из них ведет к корневому Rbi, а какие - от него. RBridge **должны** отбрасывать групповые кадры, приходящие в порт от RBridge, не являющегося смежным для дерева, по которому распространяется кадр. Предположим, что RBn считает, что смежности a , c и f находятся в дереве Rbi. Групповой кадр для дерева распространения Rbi, полученный от любого из смежных узлов a , c , f пересылается двум другим смежным узлам, а прочие кадры отбрасываются. Если RBn имеет несколько портов на канале, групповые кадры, передаваемые им в один из таких портов, будут получены на другой стороне, но отброшены, поскольку RBridge не является смежным для себя.
2. **Проверка RPF.** Другим методом, используемым в RBridge для предотвращения петель групповой пересылки при изменении топологии является проверка пересылки по обратному пути (RPF). Здесь проверяется, что групповой кадр прибыл из ожидаемого канала на основе дерева и входного RBridge. RBridge **должны** отбрасывать групповые кадры, не прошедшие проверку RPF.

Для ограничения числа состояний, требуемых при проверке RPF, каждый RBridge RB2 **должен** анонсировать деревья, которые он может выбрать в качестве входных для групповых пакетов. Когда любой из RBridge (например, RB3) рассчитывает дерево по псевдониму X , он вычисляет для каждого RBridge RB2, который может быть входным для дерева X , канал, откуда RB3 следует получать пакеты от RB2 по дереву X , и отмечает для этого канала, что RB2 является действительным входным RBridge для дерева X .

Информация о том, какие деревья может выбрать RB2, включается в LSP от RB2. Подобно тому, как RBridge RB1 указывает k деревьев, которые будут рассчитываться всеми RBridge, RB2 задает число j , которое указывает общее количество разных деревьев, которые может задать RB2, а конкретные деревья, которые может указать RB2, представляют собой комбинацию указанных деревьев и деревьев, выбранных из числа имеющих наиболее высокий приоритет.

j возможных входных деревьев для RB2 - это деревья с псевдонимами, явно указанными RB2 в specified ingress tree picknames (и включенные в k деревьев, выбранных в масштабе кампуса RBridge RB1 с наибольшим приоритетом), а остальные (вплоть до k , если $j = 0$ и или $\text{minimum}(j, k)$ в противном случае ²) - деревья с наибольшим приоритетом из числа выбранных в масштабе кампуса k деревьев.

По умолчанию $j = 1$. Значение 0 для j является особым и означает, что RB2 может выбрать любое k деревьев, рассчитанных для кампуса.

¹В оригинале ошибочно сказано от N к R , см. <https://www.rfc-editor.org/errata/eid3508>. Прим. перев.

²В оригинале обычно сказано «до максимального из $\{j, k\}$ », см. <https://www.rfc-editor.org/errata/eid3052>. Прим. перев.

3. *Проверка параллельных каналов.* Если при построении дерева и устранении конфликтов для конкретного дерева группового распространения выбирается канал без псевдоузлов между RB1 и RB2, этот канал может фактически состоять из нескольких. Эти параллельные каналы будут видимы для RB1 и RB2, но не для остального кампуса (поскольку каналы не представлены псевдоузлами). Если эти параллельные каналы включены в дерево, для RB1 и RB2 важно решить, какой из каналов использовать, но для остальных RBridge это не имеет значения, поэтому алгоритму устранения петель (tiebreaking) не требуется быть видимым для всех RBridge, кроме RB1 и RB2. В этом случае смежности RB1-RB2 упорядочиваются как описано ниже, причем более предпочтительной будет та, через которую RB1 и RB2 передают групповые кадры между собой.
 - a) Наиболее предпочтительны каналы, организованные P2P Hello. Алгоритм устранения конфликтов между ними основан на предпочтении смежности с наибольшим Extended Circuit ID устройства RBridge с большим System ID.
 - b) Затем рассматриваются каналы, организованные через кадры TRILL-Hello с подавленными псевдоузлами. Отметим, что псевдоузлы подавляются в LSP, но остаются в TRILL-Hello, поэтому они доступны для устранения конфликтов. Среди этих каналов тот, у которого больше идентификатор псевдоузла.
4. *Проверка группы портов.* Если RBridge имеет несколько портов на одном канале, групповые пакеты будут приходить из каждого такого канала. Все такие копии, кроме одной, **должны** отбрасываться. Все кадры, являющиеся частью одного потока должны приниматься через один порт во избежание разупорядочения.

При изменении топологии (включая изменения при активизации) RBridge **должен** уточнить фильтры входного дерева распространения не позднее, чем он уточняет свою выходную пересылку.

4.5.3. Сокращение дерева распространения

Каждое дерево распространения **следует** сокращать по VLAN, сокращая нисходящие ветви без потенциальных получателей. Групповые кадры TRILL Data **следует** пересылать лишь в ветви, которые не сокращаются.

В двух случаях **следует** выполнять дополнительное сокращение - (1) сообщения IGMP [RFC3376], MLD [RFC2710] MRD [RFC4286], которые доставляются лишь в каналы с групповыми маршрутизаторами IP и (2) другие групповые кадры, полученные из групповых адресов IP, которые следует доставлять лишь в каналы с зарегистрированными получателями и каналы с групповыми маршрутизаторами IP, за исключением групповых адресов IP, которые должны быть ширококестельными. Каждый из этих случаев обрабатывается на уровне VLAN.

Пусть RBridge RBn знает, что смежности (a, c, f) находятся в дереве распространения nickname1. RBn отмечает данные сокращения для каждой смежности в дереве nickname1:

- набор VLAN, доступных в нисходящем направлении;
- для каждой из таких VLAN устанавливаются флаги, указывающие, есть ли в них групповые маршрутизаторы нисходящей рассылки IPv4 или IPv6;
- множество групповых адресов L2, выведенных из multicast-групп IP, для которых есть нисходящие адресаты.

4.5.4. Оптимизация дерева распространения

RBridge **должны** определять VLAN, связанную со всеми получаемыми естественными кадрами и корректно применять правила VLAN при отправке естественных кадров через выходные порты RBridge в соответствии с настройками этих портов и их назначением пересылающими (forwarder). Узлам RBridge **следует** также сокращать дерево распространения групповых кадров в соответствии с VLAN. Поскольку такое сокращение не обязательно, узлы могут получать кадры данных TRILL или ESADI, которые следовало сократить ранее в дереве распространения. RBridge отбрасывает такие кадры без уведомления. Некоторые RBridge в кампусе могут сокращать деревья распространения для VLAN.

Для групповых пакетов ситуация сложнее. RBridge **следует** анализировать полученные из IP естественные групповые кадры, определяя и анонсируя получателей и групповые маршрутизаторы IP для таких кадров, как описано в параграфе 4.7. **Следует** сокращать распространение полученных из IP групповых кадров на основе такого изучения и анонсов, но не требуется сокращения на основе групповых получателей IP и статуса подключения маршрутизатора. В отличие от VLAN, где статус подключения VLAN на порту **должен** поддерживаться и соблюдаться, от RBridge не требуется поддерживать статус групповых получателей IP и подключения маршрутизатора.

RBridge, который не проверяет поступающие естественные кадры IGMP [RFC3376], MLD [RFC2710], MRD [RFC4286], **должен** сообщить, что к нему подключены групповые маршрутизаторы IPv4 и IPv6 IP для всех VLAN, где он назначен пересылающим. Не требуется анонсировать полученных из IP групповых получателей. Это приведет к тому, что весь полученный из IP групповой трафик будет передаваться данному RBridge для этих VLAN. Трафик будет потом пересылаться в каналы, для которых RBridge назначен пересылающим, при совпадении VLAN со значением VLAN для которой узел назначен пересылающим на этом канале (это может вести к подавлению некоторых сообщений IGMP о принадлежности к группе, но оно не существенно, поскольку весь групповой трафик, запрашиваемый такими сообщениями будет передаваться конечным станциям).

Кампус может включать комбинацию RBridge с разными уровнями оптимизации полученного из IP группового трафика. RBridge может получать извлеченные из IP групповые кадры, которые следовало бы сократить ранее в дереве распространения. Такие кадры отбрасываются без уведомления.

Дополнительная информация приведена в документе Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches [RFC4541].

4.5.5. Пересылка с использованием дерева распространения

Пересылка групповых кадров при полной оптимизации описана ниже (ссылки на смежности в описании не включают канал, из которого был получен кадр).

- RBridge RBn принимает групповой кадр TRILL Data с внутренним VLAN-х и заголовком TRILL, указывающим выбор дерева с псевдонимом nickname1.
- Если источник принятого кадра не находится в одной из смежностей дерева nickname1 для указанного входного RBridge, кадр отбрасывается (параграф 4.5.1).
- Если кадр содержит анонс IGMP или MLD или запрос MRD, инкапсулированный в нем кадр пересылается в смежности дерева nickname1, которые указывают наличие нисходящих групповых маршрутизаторов IPv4 или IPv6 (что подходит) для VLAN-х.
- В остальных случаях, если кадр направлен по групповому адресу L2, полученному из multicast-группы IP, но адрес IP не находится в диапазоне групповых адресов IP, которые должны считаться ширококестельными, кадр пересылается в смежности дерева nickname1, которые указывают наличие нисходящих групповых маршрутизаторов IPv4 или IPv6 (что подходит) для VLAN-х, а также в смежности, указывающие наличие получателей VLAN-х для данного группового адреса.
- В противном случае (внутренний кадр предназначен для группового адреса L2, не выведенного из multicast-группы IP, неизвестного адресата, является ширококестельным или является групповым адресом IP, который нужно считать ширококестельным) кадр пересылается в смежность тогда и только тогда, когда эта смежность присутствует в дереве nickname1 и помечена как ведущая к каналам VLAN-х.

Для каждого канала, где RBn назначен пересылающим, RBn дополнительно проверяет, следует ли деинкапсулировать кадр и передавать его в естественной форме или обрабатывать кадр локально.

Кадры TRILL ESADI будут доставляться лишь мостам RBridge, назначенным пересылающими для их VLAN. Такие кадры будут групповыми для всего кампуса, подобно другим групповым кадрам, не полученным из групповых кадров данных IP, в дереве распространения, выбранном RBridge, который создал кадр TRILL ESADI и сокращаются в соответствии с внутренним Inner.VLAN ID. Таким образом, все RBridge, назначенные пересылающими для канала в эту VLAN, получат данный кадр.

4.6. Поведение обработки пакетов

В этом параграфе описано поведение RBridge для всех вариантов принимаемых пакетов, включая пакеты, которые будут пересылаться. В параграфе 4.6.1 рассмотрены естественные кадры, в 4.6.2 - кадры TRILL, а в 4.6.3 - кадры управления L2. Обработку можно организовать или упорядочить не так, как описано здесь, если обеспечивается описанный здесь результат. Определения кадров приведены в параграфе 1.4.

Поврежденные кадры, например, кадры с некратным 8 битам размером, слишком короткие для используемого протокола или оборудования, а также кадры с некорректным значением FCS отбрасываются при получении портом RBridge, как это делается портами мостов IEEE 802.1.

Информация об адреса получателя ({ VLAN, Outer.MacSA, port }) по умолчанию изучается для всех кадров с индивидуальным адресом отправителя (параграф 4.8).

4.6.1. Прием естественного кадра

Если порт настроен как отключенный (disabled), отключена служба конечной станции на порту (end-station service) настройкой порта в качестве транкового или порт настроен на использование P2P Hello, кадр отбрасывается.

Входной RBridge RB1 определяет VLAN ID для естественного кадра по тем же правилам, что и мосты IEEE [802.1Q-2005] (Приложение D и параграф 4.9.2). После определения VLAN если RB1 не назначен пересылающим для этой VLAN на принявшем кадр порту или заблокирован, кадр отбрасывается. В случае назначения пересылающим для VLAN и отсутствия блокировки (параграф 4.2.4.3), естественный кадр пересылается в соответствии с параграфом 4.6.1.1 (индивидуальный) или 4.6.1.2 (групповой или ширококестельный).

4.6.1.1. Индивидуальные кадры

Если MAC-адрес получателя естественного кадра является индивидуальным (unicast) выполняются указанные ниже действия. Для адрес получателя L2 и VLAN выполняется в базе данных входного RBridge для MAC-адресов и VLAN с целью найти выходной RBridge Rbm, локальный выходной порт или определить, что адресат недоступен. Возможны 4 случая.

1. Адресатом является принявший кадр RBridge, кадр обрабатывается локально.
2. Если известно, что адресат размещен на том же канале, откуда был принят естественный кадр, но не является данным RBridge, этот RBridge отбрасывает кадр без уведомления, поскольку адресат уже получил его.
3. Если известно, что адресат размещен на другом локальном канале, где RBm назначен пересылающим, RB1 преобразует кадр в кадр TRILL Data с Outer.MacDA следующего RBridge в направлении RBm, заголовком TRILL с M = 0, и входным псевдонимом для RB1 и выходным псевдонимом для RBm. Если входной RB1 имеет множество псевдонимов, **следует** всякий раз использовать один псевдоним во входном поле nickname при каждой инкапсуляции естественного кадра с любым MAC-адресом отправителя и VLAN. Это упрощает обучение конечных узлов. Если RBm является RB1, обработка выполняется в соответствии с параграфом 4.6.2.4, в противном случае для Outer.MacSA устанавливается MAC-адрес порта RB1 на пути к следующему RBridge в направлении RBm и кадр помещается в очередь для передачи через этот порт.
4. Если индивидуальный MAC-адрес получателя не известен в указанной в кадре VLAN, RB1 обрабатывает кадр в соответствии с параграфом 4.6.1.2 для ширококестельной передачи, но в поле Inner.MacDA указывается исходный индивидуальный адрес получателя.

4.6.1.2. Групповые и ширококестельные кадры

Если несколько портов RBridge подключено к одному каналу, все принятые копии естественных групповых и ширококестельных кадров, кроме одной, отбрасываются для предотвращения дубликатов. Относящиеся к одному потоку кадры всегда должны восприниматься на одном порту во избежание нарушения порядка доставки.

Если кадр является естественным кадром IGMP [RFC3376], MLD [RFC2710] или MRD [RFC4286], RB1 **следует** анализировать его, определяя принадлежность к группе или индикацию группового маршрутизатора IP и анонсируя эти данные в подходящие VLAN в своих LSP (параграф 4.7).

Естественные групповые кадры RB1 пересылает в естественной форме в свои каналы, где размещаются подходящие назначенные узлы пересылки для VLAN кадра с учетом дополнительной отсечки и блокировки. В дополнение к этому естественный кадр преобразуется в кадр TRILL Data с групповым адресом All-RBridges в поле Outer.MacDA, заголовком TRILL с $M = 1$ (групповой), входным псевдонимом для RB1 и выходным псевдонимом для выбранного дерева распространения. Затем кадр пересылается в сокращенное дерево распространения (параграф 4.5) с установкой в Outer.MacSA каждой переданной копии MAC-адреса порта RB1, куда кадр передается.

По умолчанию RB1 записывает в выходное поле nickname псевдоним дерева распространения и набора анонсированных для использования, корень которого имеет наименьшую стоимость из RB1. Однако RB1 **может** выбирать другие деревья распространения, если RB1 настроен на распределение группового трафика по разным путям. В этом RB1 **должен** выбрать дерево, указав псевдоним, являющийся корнем дерева распространения (параграф 4.5). Кроме того, RB1 **должен** выбрать псевдоним, объявленный им (в своем LSP), как один из тех, которые RB1 может использовать. Используемая RB1 стратегия выбора для распространения группового трафика по нескольким путям выходит за рамки этого документа.

4.6.2. Прием кадра TRILL

Кадр TRILL имеет тип (EtherType) TRILL или L2-IS-IS и групповой адрес Outer.MacDA, выделенный для TRILL (параграф 7.2). Приведенные ниже проверки выполняются последовательно и первое совпадение определяет обработку кадра.

1. Если EtherType == L2-IS-IS, а Outer.MacDA содержит All-IS-IS-RBridges или индивидуальный MAC-адрес принявшего кадр порта RBridge, кадр обрабатывается в соответствии с параграфом 4.6.2.1¹.
2. Если в Outer.MacDA указан групповой адрес, выделенный для TRILL, но не All-RBridges, кадр отбрасывается.
3. Если Outer.MacDA содержит индивидуальный адрес, не совпадающий с MAC-адресом принявшего кадр порта RBridge, кадр отбрасывается (такие кадры скорее всего адресованы другим RBridge на канале с множественным доступом и эти RBridge обработают их).
4. EtherType отличается от TRILL, кадр отбрасывается.
5. Если поле Version в заголовке TRILL больше 0, кадр отбрасывается.
6. Если счетчик интервалов равен 0, кадр отбрасывается.
7. Если Outer.MacDA содержит групповой адрес, а $M == 0$ или адрес в Outer.MacDA является индивидуальным, а $M == 1$, кадр отбрасывается.
8. По умолчанию RBridge **недопустимо** пересылать кадры с инкапсуляцией TRILL от соседа, с которым нет смежности TRILL IS-IS. RBridge **можно** настроить на уровне порта на восприятие таких кадров для пересылки в ключах, когда известно, что не являющееся партнером устройство (например, конечная станция) настроено на передачу кадров с инкапсуляцией TRILL, которые можно без опаски пересылать.
9. Затем проверяется Inner.MacDA. Если это групповой адрес All-ESADI-Rbridges и RBn реализует протокол ESADI, выполняется обработка в соответствии с параграфом 4.6.2.2. если адрес иной или RBn не поддерживает ESADI, выполняется обработка, описанная в параграфе 4.6.2.3.

4.6.2.1. Кадры управления TRILL

Кадры обрабатываются экземпляром TRILL IS-IS на RBn и не пересылаются.

4.6.2.2. Кадры TRILL ESADI

Если $M == 0$, кадр отбрасывается без уведомления.

Выходной псевдоним указывает дерево распространения. Кадр пересылается в соответствии с параграфом 4.6.2.5. Кроме того, если пересылающий RBridge назначен пересылающим (appointed forwarder) для канала в указанную VLAN, реализует протокол TRILL ESADI и этот протокол включен на пересылающем RBridge для этой VLAN, внутренний кадр декапсулируется и передается локальному протоколу ESADI.

4.6.2.3. Кадры данных TRILL

Затем проверяется флаг M. При нулевом значении обработка продолжается в соответствии с параграфом 4.6.2.4, в противном случае - в соответствии с параграфом 4.6.2.5.

4.6.2.4. Заведомо индивидуальные кадры данных TRILL

Проверяется выходной псевдоним в заголовке TRILL и при неизвестном или зарезервированном имени кадр отбрасывается.

Если RBn является транзитным RBridge, счетчик интервалов уменьшается на 1 и кадр передается следующему RBridge в направлении выходного RBridge (возможность уменьшать счетчик больше чем на 1 в некоторых обстоятельствах (параграф 3.6) применима лишь для групповых кадров, а здесь рассматриваются лишь заведомо индивидуальные). Поле Inner.VLAN не проверяется транзитным RBridge при пересылке заведомо индивидуальных кадров TRILL Data. Для пересылаемого кадра Outer.MacSA является MAC-адресом передающего порта, Outer.MacDA — индивидуальным адресом следующего RBridge, а VLAN - Designated VLAN на канале, куда кадр передается.

Если RBn не является транзитным RBridge, т. е. указанный выходной RBridge является выполняющим обработку, Inner.MacSA и Inner.VLAN ID (по умолчанию) определяются как связанные со входным псевдонимом, если он известен и не является зарезервированным, или Inner.MacSA не является индивидуальным. Тогда пересылаемый кадр декапсулируется и выполняются указанные ниже действия.

¹В оригинале это предложение содержит ошибки, см. <https://www.rfc-editor.org/errata/eid3003>. Прим. перев.

- Проверяется Inner.MacDA и кадра с неиндивидуальным получателем отбрасываются.
- Если Inner.MacDA соответствует RBridge, выполняющему обработку, кадр доставляется локально.
- Проверяется Inner.VLAN ID и при значении 0x0 или 0xFFFF кадр отбрасывается.
- Выполняется поиск Inner.MacDA и Inner.VLAN ID в локальном кэше адресов RBr и кадр пересылается в один из каналов, соответствующих адресату, если RBridge назначен пересылающим для данного канала в VLAN кадра и не заблокирован (отбрасывается при блокировке), или обрабатывается, как описано ниже.

Заведомо индивидуальный кадр TRILL Data может прийти на выходной RBridge лишь для того, чтобы обнаружить, что RBridge реально не знает комбинации Inner.MacDA и Inner.VLAN. Это может произойти, например, в результате завершения срока действия кэшированных в таблице выходного RBridge адресов MAC. В этом случае выходной RBridge передает естественный кадр во все каналы, которые относятся к VLAN кадра, где RBridge назначен пересылающим и не заблокирован. Исключение состоит в том, что RBridge **может** воздержаться от передачи кадра в каналы, где ему известно об отсутствии конечной станции с MAC-адресом получателя, например, в канал, настроенный как транк (параграф 4.9.1).

Если в результате настройки или изучения регистраций L2 значения MAC-адреса получателя и VLAN появляются в локальном кэше адресов RBr для двух или более каналов, где RBr является не заблокированным назначенным пересылающим для VLAN кадра, RBr передает кадр во все такие каналы.

4.6.2.5. Кадры данных TRILL для множества получателей

Проверяются выходной и входной псевдоним в заголовке TRILL и кадр отбрасывается, если любой из псевдонимов не известен или зарезервирован.

Проверяется Outer.MacSA и кадр отбрасывается при отсутствии смежности для дерева, указанного псевдонимом выходного RBridge на порту, где был принят кадр. Выполняется проверка RPF для входного и выходного псевдонимов и кадр отбрасывается при отрицательном результате. При наличии нескольких параллельных каналов с подавлением TRILL-Hello от псевдоузлов к RBridge предыдущего интервала кадр отбрасывается, если он получен на неверном канале. Если несколько портов RBridge подключены к одному каналу, кадр отбрасывается при его получении из неверного канала. Дополнительные сведения о таких проверках даны в параграфе 4.5.2.

Если Inner.VLAN ID в кадре имеет значение 0x0 или 0xFFFF, кадр отбрасывается.

Если RBridge назначен пересылающим для Inner.VLAN ID из кадра, значения Inner.MacSA и Inner.VLAN ID по умолчанию определяются как связанные со входным псевдонимом, если он не является неизвестным или зарезервированным или Inner.MacSA не является индивидуальным адресом. Тогда копия кадра декапсулируется и передается в естественной форме в те каналы для данной VLAN, где RBridge назначен пересылающим, с учетом дополнительного сокращения и блокировки (параграф 4.2.4.3) и/или обрабатывается локально, когда это применимо.

Счетчик интервалов уменьшается (возможно, более чем на, как указано в параграфе 3.6) и кадр пересылается вниз по дереву, заданному псевдонимом выходного RBridge, как указано в параграфе 4.5.

Для пересылаемого кадра в поле Outer.MacSA указывается адрес порта, куда кадр передается, в Outer.MacDA - групповой адрес All-RBridges, а в VLAN - Designated VLAN для канала, куда передается кадр.

4.6.3. Прием кадров управления L2

Низкоуровневые кадры управления, принимаемые RBridge, обрабатываются на приемном порту, как описано в параграфе 4.9. Имеется два типа таких кадров, различающихся адресами получателя. Описание их обработки приведено в параграфах, указанных ниже.

Имя	Описание	Адрес получателя
BPDU	параграф 4.9.3	01-80-C2-00-00-00
VRP	параграф 4.9.4	01-80-C2-00-00-21

4.7. Изучение IGMP, MLD и MRD

Входным RBridge **следует** на основе естественных кадров IGMP [RFC3376], MLD [RFC2710] и MRD [RFC4286], принимаемых на портах в VLAN, где они назначены пересылающими (и не заблокированы) узнавать, какие выведенные из групповых сообщений IP кадры следует пересылать в какие каналы. Такие кадры (в общем случае) являются инкапсулированными в TRILL Data и распространяются в соответствии с параграфом 4.5.

Сообщения о членстве IGMP или MLD, полученные в естественной форме из канала, указывают групповых получателей для данной группы на данном канале. Запросы IGMP или MLD, а также анонсы MRD в естественной форме из канала указывают присутствие на канале группового маршрутизатора IP.

Сообщения о принадлежности к multicast-группам IP, должны передаваться в масштабе кампуса и доставляться всем групповым маршрутизаторам IP (различая IPv4 и IPv6). Весь полученный из IP групповой трафик также должен передаваться всем групповым маршрутизаторам IP с той же версией протокола IP.

Групповые данные IP **следует** передавать лишь в каналы, где имеется групповой маршрутизатор для данной версии IP (IPv4 или IPv6) или групповой получатель IP для выведенного из IP адреса MAC, если групповой адрес IP не относится к диапазону, который должен считаться ширококестельным.

RBridge не обязаны анонсировать себя в качестве получателей группы IPv4 All-Snoopers (применяется для отчетов MRD [RFC4286]), поскольку групповой адрес IPv4 для этой группы относится к диапазону, куда передаются все групповые кадры IP, т. е. является ширококестельным (см. параграф 2.1.2 в [RFC4541]). Однако RBridge, выполняющие оптимизацию выведенных из IPv6 групп, **должны** анонсировать себя в качестве получателей группы IPv6 All-Snoopers.

Дополнительная информация приведена в параграфе Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches [RFC4541].

4.8. Детали адреса конечной станции

RBridge узнают MAC-адреса и VLAN локально подключенных конечных станций для пар канал-VLAN, где они назначены пересылающими. Эта информация позволяет:

- пересылать естественную форму заведомо индивидуальных кадров TRILL в корректный канал;
- решать для входящих естественных индивидуальных кадров на канале, где RBridge назначен пересылающим для VLAN, является ли кадр:
 - заведомо адресованным другой конечной станции на том же канале (не нужно делать ничего);
 - нужно преобразовать кадр TRILL Data и переслать его.

RBridge должны изучать MAC-адреса, VLAN и удаленные RBridge удаленно подключенных станций для VLAN, где данный и удаленный RBridge являются назначенными пересылающими, чтобы можно было эффективно направлять полученные естественные кадры для индивидуальных адресатов по соответствующим адресам и VLAN.

4.8.1. Определение адресов конечных станций

Для RBridge имеется 5 независимых способов узнать адреса конечных станций.

1. Просмотр в кадрах VLAN-х, полученных на портах, где он назначен пересылающим для VLAN-х, триплетов { MAC-адрес отправителя, VLAN, порт }.
2. Триплеты { MAC-адрес отправителя, VLAN, псевдоним входного RBridge } в любом декапсулируемом им естественном кадре.
3. Изучение в протоколах регистрации L2 { MAC-адрес отправителя, VLAN, порт } для конечных станций, зарегистрированных на локальном порту.
4. Использование протокола TRILL ESADI для одной или нескольких VLAN и получение за счет этого информации об удаленных адресах и/или передача локальной информации об адресах.
5. Настройка в системе управления.

RBridge **должны** реализовать пп. 1 и 1 и 2. RBridge используют эти возможности, пока для одной или нескольких VLAN и/или портов не отключено обучение по принятым кадрам или/и декапсуляции передаваемых кадров.

RBridge **могут** реализовать пп. 3 и 4. При реализации п. 4 протокол ESADI работает лишь в том случае, когда RBridge настроен делать это на уровне VLAN.

RBridge **следует** реализовать п. 5.

Записи в таблице полученных при обучении адресов MAC и VLAN, а также связанная с ними информация также используют однооктетный (без знака) уровень доверия, обоснование которого приведено ниже. Информация, полученная в результате наблюдения за данными, имеет уровень достоверности 0x20, если в конфигурации не задано другое значение. Этот уровень можно настраивать на уровне RBridge отдельно для информации из естественных кадров и данных из инкапсулированных кадров удаленного происхождения. Информация, полученная от протокола TRILL ESADI имеет уровень достоверности от 0 до 254. Информация, заданная системой управления, по умолчанию имеет уровень достоверности 255, но можно задать в конфигурации иное значение.

Таблица полученных при обучении MAC-адресов включает квартеты (1) { достоверность, VLAN, MAC-адрес, локальный порт } для адресов, полученных из локальных естественных кадров и локальных протоколов регистрации, квартеты (2) { достоверность, VLAN, MAC-адрес, псевдоним выходного RBridge } для адресов, полученных из инкапсулированных удаленных кадров и баз данных ESADI о состоянии каналов, (3) дополнительную информацию для реализации таймаутов изученных и статически заданных адресов и т. п.

При получении нового адреса и связанной информации из наблюдений за кадрами данных эта информация вводится в локальную базу данных. При этом возможны три описанных ниже варианта.

- A. Для новой пары { адрес, VLAN } данные добавляются вместе с уровнем достоверности.
- B. Если запись для пары { адрес, VLAN } с такой же информацией о доставке уже имеется в базе, для уровня достоверности устанавливается большее из значений имеющейся и полученной путем обучения достоверности. Если полученная при обучении информация имеет более высокий уровень достоверности, таймер для записи сбрасывается.
- C. Если в базе уже есть запись для пары { адрес, VLAN } с другой информацией, новая запись включается взамен имеющейся лишь в том случае, когда она получена с тем же или более высоким уровнем достоверности. При замене записи ее таймер сбрасывается.

4.8.2. Обоснование уровня достоверности при обучении

Механизм уровней достоверности позволяет администраторам кампуса RBridge установить предпочтения для некоторых источников получения информации по отношению к другим. В принятой по умолчанию конфигурации без дополнительного протокола ESADI изучение адресов происходит лишь путем наблюдения за естественными локальными кадрами и декапсуляции принятых кадров TRILL Data. Оба эти источника по умолчанию имеют уровень достоверности 0x20 и, поскольку получение данных с таким же или более высоким уровнем достоверности ведет к переопределению имеющихся записей, такое поведение имитирует принятое по умолчанию обучение в мостах 802.1.

Хотя правила управления кампусом RBridge выходят за рамки документа, ниже приведено два примера реализации таких правил с использованием уровней достоверности данных.

- Локально полученные естественные кадры можно считать более надежными, нежели декапсулированные кадры из удаленных частей кампуса. Чтобы полученные из таких локальных кадров MAC-адреса не могли быть

захвачены удаленными фиктивными кадрами, можно повысить уровень достоверности для адресов из локальных кадров или снизить его для адресов, полученных при декапсуляции удаленных кадров.

- MAC-адреса, полученные от протоколов регистрации L2 с криптографической защитой, такой как протокол 802.1X с криптографией EAP, можно считать более надежными, чем адреса, полученные из простого наблюдения за кадрами данных. При передаче такой защищенной информации по протоколу ESADI использование аутентификации в кадрах TRILL ESADI LSP может существенно осложнить подделку кадров в процессе передачи. В результате может оказаться разумным повышение уровня достоверности для данных, полученных по протоколу ESADI, и переопределение этими данными записей, полученных иным путем.

Заданные вручную адресные данные обычно считаются статически и по умолчанию имеют уровень достоверности 0xFF, тогда как для всех остальных источников информации этот уровень не может превышать 0xFE. Однако в некоторых случаях, например, при использовании заданной вручную информации лишь в качестве отправной точки, когда администратор кампуса RBridge желает потом динамически переопределять адреса, можно задать для этих данных меньший уровень достоверности.

4.8.3. Забывание адресов конечных станций

Хотя RBridge должны изучать адреса конечных станций, как описано выше, не менее важно забывать адресную информацию. Иначе при перемещении конечной станции в другую часть кампуса она окажется на неопределенное время заблокирована («черная дыра») устаревшей информацией о канале, к которому станция была подключена.

Для адресов конечных станций, полученных из принятых локально кадров, время ожидания следующего принятого или декапсулированного естественного кадра соответствует рекомендациям [802.1D]. Оно называется временем старения (Ageing Time), настраивается на уровне RBridge в диапазоне от 10 до 1000000 и по умолчанию составляет 300 секунд.

Для адресов, полученных по протоколу TRILL ESADI ситуация отличается. Вопрос удаления информации из ESADI LSP (или таймаута протокола ESADI, если RBridge становится недоступным) решает RBridge-источник.

Когда RBridge перестает быть назначенным пересылающим для VLAN-х на порту, он забывает все адреса конечных станций, полученные из наблюдения естественных кадров VLAN-х, принятых на этом порту. Он также увеличивает на уровне VLAN значения счетчиков фактов утраты статуса назначенного пересылающего для данной VLAN.

Когда RBridge RBn перестает быть назначенным пересылающим для VLAN-х на всех своих портах, он забывает все адреса конечных станций, полученные из декапсуляции естественных кадров VLAN-х. Если RBn участвует в работе протокола TRILL ESADI для VLAN-х, он прекращает это участие после передачи финального LSP, обнуляющего информацию об адресах конечных станций для VLAN. Кроме того, все другие RBridge, пересылающие кадры для VLAN-х хотя бы на одном из портов, отмечают, что данные о состоянии канала для RBn изменились и говорят об утрате канала в VLAN-х. В результате они забывают адресную информацию, полученную при декапсуляции кадров VLAN-х, которая указывает, что RBn является входным RBridge.

Когда отмечается увеличение значения счетчика потерь статуса назначенного пересылающего RBridge RBn для VLAN-х в базе данных TRILL IS-IS о состоянии каналов, но RBn продолжает быть назначенным пересылающим для VLAN-х хотя бы на одном порту, все остальные RBridge, являющиеся назначенными пересылающими для VLAN-х, меняют информацию о старении всех адресов, полученных при декапсуляции естественных кадров VLAN-х от входного RBridge RBn. В каждой записи оставшееся время корректируется так, чтобы оно не превышало конфигурационный параметр на уровне RBridge, называемый задержкой пересылки (Forward Delay) в соответствии с [802.1D]. Этот параметр имеет значения от 4 до 30 и по умолчанию установлено 15 секунд.

4.8.4. Общее изучение VLAN

RBridge могут транслировать VLAN ID в меньшее число идентификаторов для изучения адресов, как это делают мосты [802.1Q-2005]. Эти идентификаторы применяются для сопоставления при поиске адресов взамен VLAN ID. Если идентификатор VLAN, для которой адрес был получен, не был сохранен, возникают два последствия.

- RBridge больше не имеет информации, требуемой для участия в протоколе TRILL ESADI для VLAN, чьи идентификаторы не были сохранены.
- В случаях, где параграф 4.8.3 требует отбрасывания полученной адресной информации для конкретной VLAN, когда VLAN ID недоступен для записей с идентификатором VLAN общего пользования, оставшееся время для каждой записи с этим идентификатором VLAN устанавливается не больше Forward Delay для RBridge.

Несмотря на то, что это выходит за рамки данной спецификации, имеются некоторые функции L2, в которых множество VLAN использует общее обучение, где одна из VLAN является первичной, а другие VLAN в группе - вторичными. Примером этого является разделение трафика разных сообществ по тегам VLAN с предоставлением некоторых ресурсов (например, маршрутизатор IP или сервер DHCP) всем сообществам. Реализация этой функции возможна путем предоставления некоего тега VLAN (например, Z) каналу к общему ресурсу, а другие теги (скажем, A, C, D) включить в группу вторичных { primary = Z, secondaries = A, C, D }. RBridge, знающий об этой группировке и подключенный к одной из вторичных VLAN в группе, заявляет также свою принадлежность к первичной VLAN. Таким образом, RBridge, подключенный к A, будет также заявлять свое подключение к Z. RBridge, подключенный к первичной VLAN, будет также заявлять свое подключение ко всем VLAN в группе.

Этот документ не задает способ использования групп VLAN. Информация о группе VLAN включается лишь в RBridge, участвующие в группе. Однако для обнаружения конфигурационных ошибок RBridge, знающим о группе VLAN, **следует** указывать группу в своих LSP.

4.9. Порты RBridge

В параграфе 4.9.1 описано несколько битов конфигурации порта RBridge, в параграфе 4.9.2 приведена логическая структура порта, а параграфы 4.9.3 и 4.9.4 посвящены обработке высокоуровневых кадров управления.

4.9.1. Конфигурация порта RBridge

Имеется 4 бита, задающие конфигурацию порта¹.

- Бит отключения порта, при установке которого все кадры, принимаемые или предназначенные для передачи портом, отбрасываются. Возможны исключения лишь для некоторых кадров управления L2 (параграф 1.4), которые могут генерироваться и передаваться или приниматься и обрабатываться внутри порта. По умолчанию порты включены.
- Бит запрета обслуживания конечных станций (транковый порт), при установке которого отбрасываются все естественные кадры принятые и предназначенные для передачи портом (см. Приложение В). Отметим, что в этом документе к естественным кадрам не относятся кадры управления L2. По умолчанию порты не являются транковыми.

Если в DRB RB1 этот бит не установлен, но у соседа по каналу RB2 установлен, RB1 не будет назначать RB2 пересылающим для какой-либо VLAN и ни один из RBridge (включая псевдоузлы) не будет указывать RB2 как соседа в своих LSP.

Если порт с отключенным обслуживанием конечных станций сообщает в переданном через этот порт кадре TRILL-Hello, для каких VLAN он поддерживает конечные станции, он указывает отсутствие таковых.

- Бит запрета трафика TRILL (порт доступа) устанавливается для предотвращения отправки через порт любых кадров TRILL, за исключением TRILL-Hello, поскольку порт предназначен лишь для естественного трафика конечных станций. По умолчанию этот запрет не установлен. Этот бит указывается в кадрах TRILL-Hello. Если RB1 является DRB и установил этот бит в своем TRILL-Hello, DRB по-прежнему назначает пересылающие мосты для VLAN. Однако обычно в LSP не сообщается о псевдоузлах и каких-либо каналах между RBridge, связанных с этим каналом.

В некоторых случаях даже при установленном DRB флаге порта доступа этот DRB **может** создать псевдоузел для порта доступа. Тогда другие RBridge сообщают о связности с псевдоузлом в своих LSP, но DRB устанавливает флаг перегрузки (overload) в LSP псевдоузла.

- Бит использования P2P Hello, при установке которого пакеты Hello, передаваемые этим портом, являются IS-IS P2P Hello. По умолчанию используются TRILL-Hello (см. параграф 4.2.4.1).

Приоритеты (доминирование) этих битов показаны ниже - биты, указанные левее, доминируют над битами справа. То есть, когда утверждается, что какая-либо пара битов несовместима между собой, поведение определяется значением бита, указанного левее.

Отключение > P2P > Доступ > Транк

4.9.2. Структура порта RBridge

Порт RBridge можно смоделировать как низкоуровневую структуру, похожую на структуру порта в мосту [802.1Q-2005], как показано на рисунке 11. Двойные линии на рисунке представляют общий поток кадров и информации, а одинарные - только потоки информации. Прерывистый пунктир канала с VRP (GVRP/MVRP) показывает необязательность поддержки VRP. Реальная реализация порта RBridge может иметь иную структуру, обеспечивающую такое же поведение.

Низкоуровневые кадры управления обрабатываются управляющей логикой порта (канала) как это происходит в портах моста [802.1Q-2005]. Это может дополнительно включать разные протоколы 802.1 или связанные с каналом протоколы, такие как PAUSE (Annex 31B of [802.3]), обнаружение канального уровня [802.1AB], агрегирование каналов [802.1AX], защита MAC [802.1AE] или контроль доступа на уровне порта [802.1X]. Эти кадры обрабатываются на низком уровне, но могут влиять на высокоуровневую обработку. Например, протокол регистрации L2 может влиять на достоверность полученных при обучении адресов. Верхний интерфейс к этой низкоуровневой логике управления в порту соответствует службе внутреннего подуровня ISS (Internal Sublayer Service) [802.1Q-2005].

¹В оригинале этот параграф содержал ошибки, см. <https://www.rfc-editor.org/errata/eid4573>. Прим. перев.

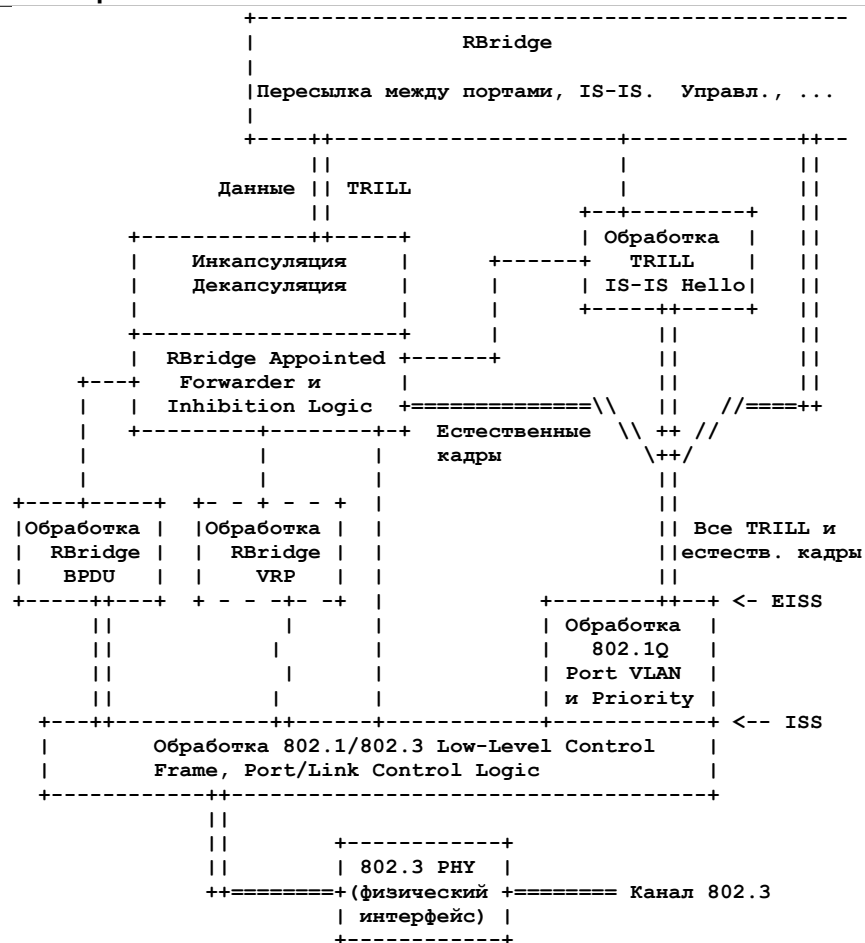


Рисунок 11. Модель порта RBridge.

Высокоуровневые кадры управления (BPDU и VRP, при его поддержке) не имеют тегов VLAN. Хотя эти кадры передаются через интерфейс ISS, они не подвергаются обработке VLAN на порту. Поведение при получении BPDU или VRP с тегом VLAN не определено. Если VRP не реализован, все кадры VRP отбрасываются. Обработка BPDU описана в параграфе 4.9.3, а обработка кадров VRP - в параграфе 4.9.4.

Для кадров, отличающихся от кадров управления L2 (т. е. все кадры TRILL и естественные кадры), выполняется обработка VLAN и приоритета, как в мостах [802.1Q-2005]. Верхний интерфейс к обработке в порту VLAN и приоритета соответствует расширенной службе внутреннего подуровня EISS (Extended Internal Sublayer Service) [802.1Q-2005].

В этой модели операции порта RBridge ниже уровня EISS идентичны операциям в мостах [802.1Q-2005] за исключением (1) обработки высокоуровневых кадров и (2) того, что отбрасывание кадров с превышением максимальной задержки при передаче (MTD - Maximum Transit Delay) не обязательно, но **возможно**.

Как описано в этом документе, входящие естественные кадры воспринимаются лишь в том случае, когда RBridge является незаблокированным назначенным пересылающим для VLAN кадра, после чего они обычно инкапсулируются и передаются. Исходящие естественные кадры обычно получаются путем декапсуляции и передаются лишь в том случае, когда RBridge является незаблокированным назначенным пересылающим для VLAN кадра.

TRILL-Hello, зонды MTU и подтверждения MTU обрабатываются на уровне порта и, подобно кадрам TRILL IS-IS, никогда не пересылаются. Они могут влиять на назначенного пересылающего и логику блокировки, а также на LSP RBridge.

За исключением TRILL-Hello, зондов и подтверждений MTU, все управляющие кадры TRILL, а также кадры данных TRILL и кадры ESAD передаются для высокоуровневой обработки RBridge при получении и передаются вниз для отправки или пересылки. Отметим, что эти кадры никогда не блокируются логикой назначенного пересылающего и блокировки, которая влияет лишь на естественные кадры, но имеются дополнительные фильтры, такие как RPF.

4.9.3. Обработка BPDU

При статичной топологии кампуса RBridge узлы RBridge были бы с точки зрения моста просто конечными станциями, завершающими каналы, но не взаимодействующими со связующим деревом. Однако у RBridge имеются причины слушать, а иной раз и передавать BPDU, как описано ниже. Даже прием и передача RBridge пакетов BPDU остается локальной активностью порта RBridge и порты RBridge никогда не взаимодействуют как узлы связующего дерева.

4.9.3.1. Прием BPDU

RBridge **должны** «слушать» BPDU с конфигурацией связующего дерева, принимаемые на порту, и отслеживать корневой мост на канале, если он имеется. Если на канале применяется протокол MSTP, это корень CIST. Такая информация сообщается на уровне VLAN мостом RBridge в его LSP и может применяться, как указано в параграфе 4.2.4.3. Кроме того, получение BPDU с конфигурацией связующего дерева используется в качестве индикации того, что канал относится к ЛВС на основе мостов, что может влиять на передачу RBridge пакетов BPDU.

RBridge **недопустимо** инкапсулировать или пересылать получаемые кадры BPDU.

RBridge отбрасывает BPDU с изменениями топологии, учетом параграфа 4.9.3.3.

4.9.3.2. Смена корневого моста

Смена корневого моста, видимая в BPDU связующего дерева, принимаемых на порту RBridge, может указывать изменение топологии ЛВС на основе мостов, включая возможное слияние двух ЛВС и т. п., без какой-либо физической индикации на порту. Во время изменения топологии мосты могут переходить в состояние pre-forwarding, где блокируются кадры TRILL-Hello. По этой причине RBridge, видя смену корневого моста на порту, для которого он назначен пересылающим в одной или нескольких VLAN, блокируется на время до 30 секунд. Блокировка заставляет назначенного для пересылки отбрасывать все естественные кадры принятые или предназначенные для отправки через канал. Интервал блокировки настраивается в конфигурации и по умолчанию составляет 30 секунд.

Рассмотрим в качестве примера две ЛВС на основе мостов, поддерживающие множество VLAN, со своими RBridge в качестве назначенных пересылающих. Если сети будут объединены путем подключения кабеля или иным способом, RBridge, подключенные к исходной ЛВС на основе мостов с более низким приоритетом корня, увидят смену корневого моста, а в другой ЛВС этого не произойдет. Таким образом, все назначенные пересылающие в сети с более низким приоритетом будут заблокированы на время, пока все не будет рассортировано с помощью BPDU в объединенной сети и кадров TRILL-Hello между вовлеченными RBridge.

4.9.3.3. Передача BPDU

Когда RBridge перестает быть назначенным пересылающим для одной или нескольких VLAN на определенном порту, ему **следует** отправлять BPDU с изменениями топологии, пока он продолжает получать BPDU протокола связующего дерева на этом порту и не видит в нем отражения данного изменения.

RBridge **могут** поддерживать передачу spanning tree BPDU в целях принудительного разделения ЛВС на основе мостов в соответствии с Приложением А.3.3.

4.9.4. Динамическая регистрация VLAN

Динамическая регистрация VLAN позволяет мостам (реже, конечным станциям) запрашивать включение или отключение VLAN на портах, ведущих к запрашивающей стороне. Это делается с помощью кадров протокола регистрации VLAN (VRP¹) - GVRP или MVRP. RBridge **могут** реализовать GVRP и/или MVRP, как описано ниже.

Кадры VRP никогда не инкапсулируются в кадры TRILL между RBridge и не пересылаются RBridge в естественной форме. Если RBridge не реализует VRP, он отбрасывает все принятые кадры VRP и не передает их.

На портах RBridge можно включить динамические VLAN. Если RBridge поддерживает VRP, реальное включение динамических VLAN определяется кадрами GVRP/MVRP, полученными на порту, как на мостах [802.1Q-2005] и [802.1ak].

RBridge, поддерживающий VRP, передает кадры GVRP/MVRP как мост [802.1Q-2005] или [802.1ak] на каждом порту, не настроенном как транковый порт RBridge или порт P2P. С этой целью он передает кадры VRP для запроса трафика в тех VLAN, где он назначен пересылающим и в Designated VLAN, если Designated VLAN не отключена на порту. Трафик других VLAN не запрашивается.

5. Параметры RBridge

В этом разделе приведен список параметров RBridge. Предполагается, что TRILL MIB будет включать многие из указанных здесь элементов, а также другие данные о состоянии RBridge, а также сведения о трафике и ошибках.

TRILL добавляет принятые по умолчанию значения и диапазоны. Для параметров в форме 16-битовых целых чисел без знака при отсутствии указанного максимума принимается максимальное значение $2^{16}-1$. Для параметров, импортированных из [802.1Q-2005], [802.1D] и IS-IS [ISO10589] [RFC1195] принятые по умолчанию значения и диапазоны указаны в соответствующих стандартах.

5.1. Параметры для RBridge в целом

Ниже перечислены параметры, встречающиеся в RBridge.

- Число псевдонимов, которое по умолчанию составляет 1 и может быть задано в диапазоне от 1 до 256.
- Желаемое число деревьев распространения для расчета каждым RBridge в кампусе и желаемое число деревьев распространения, анонсируемых RBridge для использования. Задаются 16-битовыми целыми числами без знака, по умолчанию 1 (параграф 4.5).
- Максимальное число деревьев распространения, которые RBridge может рассчитать. Выражается 16-битовым целым числом без знака, которое зависит от реализации и окружения, но не может быть изменено в конфигурации.
- Два списка псевдонимов деревьев распространения - рассчитываемые и используемые, как описано в параграфе 4.5. По умолчанию списки пусты.
- Параметры Ageing Timer и Forward Delay, их значения по умолчанию и диапазоны описаны в [802.1Q-2005].
- Два октета без знака, представляющие достоверность триплетов { MAC, VLAN, локальный порт }, полученных из принятых локально естественных кадров, и триплетов { MAC, VLAN, удаленный RBridge }, полученных из декапсулированных кадров. По умолчанию октеты имеют значение 0x20, а диапазон составляет 0x00 - 0xFE.
- Желаемое значение минимального подходящего MTU на каналах между RBridge в кампусе, т. е. originatingLSPBufferSize. Это 16-битовое целое число без знака, задающее размер в октетах и по умолчанию равное 1470, которое является минимальным приемлемым MTU. При анонсировании меньших значений они игнорируются.
- Число неудачных попыток проверки MTU, после которого RBridge считает, что конкретное значение MTU не поддерживается. По умолчанию заданы 3 попытки, допустимы значения от 1 до 255.

¹VLAN registration protocol.

Данные о статическом адресе конечной станции и достоверность таких статически настроенных данных также можно задать в конфигурации. По умолчанию принято значение 0xFF, а диапазон составляет 0x00 - 0xFF. По умолчанию статических адресных данных нет. Объем статически задаваемой информации зависит от реализации.

5.2. Параметры для каждого псевдонима в RBridge

RBridge имеет перечисленные ниже параметры конфигурации для каждого псевдонима.

- Приоритет удержания псевдонима со значением по умолчанию 0x40, если конкретного псевдонима не задано и 0xC0 в противном случае (параграф 3.7.3).
- Приоритет псевдонима при выборе корня дерева распространения в виде 16-битового целого числа без знака (по умолчанию 0x8000).
- Значение псевдонима в виде 16-битового целого числа без знака. По умолчанию принимается заданное в конфигурации значение, а при его отсутствии - последнее использованное значение, если RBridge инициализируется после перезагрузки и случайное значение в остальных случаях. Однако в любом случае резервные значения 0x0000 и 0xFFC0 - 0xFFFF не используются (параграф 3.7.3).

5.3. Параметры для каждого порта в RBridge

RBridge имеет перечисленные ниже параметры конфигурации для каждого порта.

- Такой же набор параметров, как у порта [802.1Q-2005] в терминах C-VLAN ID. В дополнение к этому имеется набор Announcing VLANs с разрешенными по умолчанию VLAN на порту (параграф 4.4.3) и диапазоном от пустого набора до всех разрешенных VLAN.
- Такой же набор параметров, как у порта [802.1Q-2005] в терминах отображения кода приоритета для кадра (см. [802.1Q-2005]).
- Время запрета для порта когда он фиксирует смену корневого моста в подключенной ЛВС на базе мостов. Время указывается в секундах и может принимать значения от 0 до 30 (принято по умолчанию).
- Desired Designated VLAN, которую RBridge будет анонсировать в своих TRILL Hello, если он является DRB для канала через этот порт. По умолчанию используется меньший идентификатор VLAN, разрешенных на порту, и в конфигурации может быть задан идентификатор любой действительной VLAN, разрешенной на этом порту (от 0x001 до 0xFFE).
- 4 конфигурационных бита на порт - отключен (по умолчанию 0 - включен), отключено обслуживание конечных станций (транк, по умолчанию 0 - не транк), порт доступа (по умолчанию 0 - не только порт доступа) и использование P2P Hello (по умолчанию 0 - используются TRILL Hello), см. параграф 4.9.1.
- Один бит на порт, установка которого запрещает определение триплетов {MAC-адрес, VLAN, порт} из локально принятых портом естественных кадров (по умолчанию 0 - обучение разрешено).
- Приоритет RBridge при выборе DRB и время удержания (Holding Time) для этого порта с использованием по умолчанию значения, заданных в IS-IS [ISO10589] [RFC1195].
- Бит, установка которого разрешает прием кадров с инкапсуляцией TRILL от Outer.MacSA, с которым RBridge не имеют смежности IS-IS (по умолчанию 0 - запрещено).
- Необязательная конфигурация передачи BPDU для решения проблемы кабельного шкафа, описанной в Приложении A.3.3. По умолчанию адресом моста является System ID устройства RBridge с меньшим значением System ID. Если RB1 и RB2 являются частью топологии кабельного шкафа, оба устройства должны быть настроены так, чтобы знать об этом и знать идентификатор, который следует использовать в протоколе STP на заданном порту.

5.4. Параметры для каждой VLAN в RBridge

RBridge имеет перечисленные ниже параметры конфигурации для каждой VLAN.

- Флаг участия в протоколе ESADI, 7-битовое значение приоритета и время удержания (Holding Time). По умолчанию флаг сброшен, используемые по умолчанию значения приоритета и времени удержания заданы в IS-IS [ISO10589] [RFC1195].
- Один бит для каждой VLAN, указывающий запрет определения триплетов {MAC-адрес, VLAN, удаленный RBridge} из кадров, декапсулированных в VLAN. По умолчанию бит сброшен (0) и обучение разрешено.

6. Вопросы безопасности

Мосты L2 не защищены по своей природе. Например, они могут быть подвержены атакам с обманными адресами отправителя или управляющими сообщениями. Одна из целей TRILL заключается в том, чтобы RBridge не добавляли проблем безопасности к имеющимся в современной технологии мостов.

К числу мер противодействия относятся настройка для протоколов TRILL IS-IS и ESADI защиты IS-IS [RFC5304] [RFC5310], а также игнорирование неаутентифицированных кадров управления TRILL и кадров ESADI. При использовании защиты IS-IS требуется настройка конфигурации RBridge.

Могут использоваться также механизмы допуска портов и защиты каналов IEEE 802.1 типа [802.1X] и [802.1AE]. Их лучше всего рассматривать как реализованные ниже TRILL (параграф 4.9.2) и не входящие в сферу действия (просто потому, что они, как правило, выходят за рамки стандартов для мостов [802.1D] и 802.1Q). Однако TRILL может использовать защищенную регистрацию через уровень доверия дополнительного протокола TRILL ESADI (параграф 4.8).

TRILL инкапсулирует естественные кадры в кампусе RBridge, пока они находятся в пути между входным и выходным(и) RBridge, как описано в параграфах 2.3 и 4.1. Таким образом, не понимающие TRILL устройства с функциями межсетевых экранов, которые не могут быть обнаружены RBridge как конечные станции, в общем случае будут не способны проверять содержимое таких кадров для целей обеспечения безопасности. Это может сделать такие устройства неэффективными. Маршрутизаторы L3 и хосты представляются RBridge конечными станциями и естественные кадры будут декапсулироваться перед отправкой таким устройствам. Таким образом, они не будут видеть TRILL EtherType. Межсетевые экраны, которые не кажутся RBridge конечными станциями (например, мосты с функциями межсетевых экранирования), следует обновить для понимания инкапсуляции TRILL.

RBridge не мешают узлу выдавать себя за другой узел, например, путем отправки ложных откликов ARP/ND. Однако RBriges не конфликтуют с какими-либо схемами, защищающими обнаружение соседей.

6.1. Вопросы безопасности VLAN

Протокол TRILL поддерживает VLAN. Это обеспечивает логическое разделение трафика, но требует внимания к вопросам безопасности использования VLAN. Для обеспечения гарантий разделения все RBridge и каналы (включая ЛВС на базе мостов) в кампусе должны быть защищены и настроены так, чтобы блокировать использование кадров динамической регистрации VLAN конечными станциями или получение ими доступа к каким-либо VLAN, передающим трафик, для которого эти станции не имеют полномочий.

Кроме того, если VLAN используются для изоляции некоей информации от каналов, где ее можно просматривать в ЛВС на базе мостов, это может больше не работать в общем случае при замене мостов на RBridge. Инкапсуляция и другой внешний тег VLAN будут приводить к передаче кадров по пути с меньшей стоимостью независимо от (внутренней) VLAN. Подходящей мерой защиты будет использование сквозного шифрования и подходящих опций защиты TRILL.

6.2. DoS-атаки BPDU/Hello

Протокол TRILL требует, чтобы назначенный узел пересылки на порту RBridge временно блокировался, если он видит кадр TRILL-Hello от другого RBridge, заявляющий себя назначенным узлом пересылки для той же VLAN, или корневой мост меняет порт. Таким образом, может показаться, что обманные BPDU, указывающие повторяющуюся смену корневого моста, или обманные кадры TRILL-Hello с установленным флагом Appointed Forwarder позволяют организовать серьезную атаку на отказ служб. Однако на деле ситуация не столь плоха.

Лучшей мерой против обманных кадров TRILL-Hello и других сообщений IS-IS является использование защиты IS-IS [RFC5304] [RFC5310]. Злонамеренные конечные станции обычно не смогут получить доступа к нужному ключевому материалу IS-IS, а без этого не смогут создать обманных сообщений.

Аутентификация, похожая на защиту IS-IS, обычно недоступна для BPDU. Однако в большинстве современных проводных ЛВС используется только соединения «точка-точка». Если у вас кампус RBridge с соединениями «точка-точка», худшее, что может сделать конечная точка с помощью обманных BPDU или кадров TRILL-Hello, это предотвращение обслуживания самой себя. Это можно сделать путем блокировки пересылки естественных кадров на RBridge, к которому станция подключена, или путем ложно активации проверки декапсуляции (параграф 4.2.4.3).

Однако в кампусах RBridge с ЛВС на базе мостов эти ЛВС с точки зрения соединенных с ними RBridge являются каналами с множественным доступом. Обманные BPDU от конечных станций, подключенных к таким ЛВС, могут оказывать влияние на обслуживание других станций той же ЛВС на базе мостов. Отметим, что мосты обрабатывают BPDU, но не пересылают их, хотя обработка может приводить к передаче других BPDU. Таким образом, конечной станции для использования обманных BPDU с целью вызвать повторяющуюся смену корневого моста (с точки зрения RBridge) через промежуточные мосты обычно требуется воздействие на корневой мост через ЛВС, что представляет угрозу даже при отсутствии RBridge.

Некоторые мосты можно настроить так, чтобы они не передавали BPDU и/или игнорировали BPDU на отдельных портах, а RBridge можно настроить на блокировку назначенной пересылки на порту в результате смены корневого моста. Однако использовать такую конфигурацию следует с осторожностью, поскольку она не безопасна.

7. Вопросы выделения значений

В этом разделе рассмотрено выделение значений через IANA и IEEE 802 (см. [RFC5226]).

7.1. Взаимодействие с IANA

Создан новый реестр IANA для параметров TRILL, в котором организовано два описанных ниже субреестра.

Начальное содержимое субреестра TRILL Nicknames приведено ниже.

0x0000 - резерв для индикации отсутствия псевдонима
0x0001-0xFFBF - динамически назначаются устройствами RBridge в каждом кампусе RBridge
0xFFC0-0xFFFE - доступны для выделения по процедуре RFC Required (1 значение) или IETF Review (1 или множество значений)
0xFFFF - постоянный резерв.

Начальное содержимое субреестра TRILL Multicast Address приведено ниже.

01-80-C2-00-00-40 - All-RBridges
01-80-C2-00-00-41 - All-IS-IS-RBridges
01-80-C2-00-00-42 - All-ESADI-RBridges
01-80-C2-00-00-43 to 01-80-C2-00-00-4F - доступны для выделения по процедуре IETF Review.

7.2. Регистрация значений в IEEE

Значение EtherType 0x22F3 назначено IEEE Registration Authority для протокола TRILL.

Значение EtherType 0x22F4 назначено IEEE Registration Authority для L2-IS-IS.

Блок из 16 групповых MAC-адресов <01-80-C2-00-00-40> - <01-80-C2-00-00-4F> назначен IEEE Registration Authority для использования протоколом IETF TRILL.

8. Нормативные документы

- [802.1ak] "IEEE Standard for Local and metropolitan area networks / Virtual Bridged Local Area Networks / Multiple Registration Protocol", IEEE Standard 802.1ak-2007, 22 June 2007.
- [802.1D] "IEEE Standard for Local and metropolitan area networks / Media Access Control (MAC) Bridges", 802.1D-2004, 9 June 2004.
- [802.1Q-2005] "IEEE Standard for Local and metropolitan area networks / Virtual Bridged Local Area Networks", 802.1Q-2005, 19 May 2006.
- [802.3] "IEEE Standard for Information technology / Telecommunications and information exchange between systems / Local and metropolitan area networks / Specific requirements Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications", 802.3-2008, 26 December 2008.
- [ISO10589] ISO/IEC, "Intermediate system to Intermediate system routing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002.
- [RFC1112] Deering, S., "Host extensions for IP multicasting", STD 5, [RFC 1112](#), August 1989.
- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", [RFC 1195](#), December 1990.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, [RFC 2119](#), March 1997.
- [RFC2464] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", RFC 2464, December 1998.
- [RFC2710] Deering, S., Fenner, W., and B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, October 1999.
- [RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", [RFC 3376](#), October 2002.
- [RFC3417] Presuhn, R., Ed., "Transport Mappings for the Simple Network Management Protocol (SNMP)", STD 62, RFC 3417, December 2002.
- [RFC3419] Daniele, M. and J. Schoenwaelder, "Textual Conventions for Transport Addresses", RFC 3419, December 2002.
- [RFC4286] Haberman, B. and J. Martin, "Multicast Router Discovery", RFC 4286, December 2005.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, [RFC 5226](#), May 2008.
- [RFC5303] Katz, D., Saluja, R., and D. Eastlake 3rd, "Three-Way Handshake for IS-IS Point-to-Point Adjacencies", RFC 5303, October 2008.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.
- [RFC6165] Banerjee, A. and D. Ward, "Extensions to IS-IS for Layer-2 Systems", RFC 6165, April 2011.
- [RFC6326] Eastlake, D., Banerjee, A., Dutt, D., Perlman, R., and A. Ghanwani, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC 6326, July 2011.

9. Дополнительная литература

- [802.1AB] "IEEE Standard for Local and Metropolitan Networks / Station and Media Access Control Connectivity Discovery", 802.1AB-2009, 17 September 2009.
- [802.1ad] "IEEE Standard for and metropolitan area networks / Virtual Bridged Local Area Networks / Provider Bridges", 802.1ad-2005, 26 May 2005.
- [802.1ah] "IEEE Standard for Local and Metropolitan Area Networks / Virtual Bridged Local Area Networks / Provider Backbone Bridges", 802.1ah-2008, 1 January 2008.
- [802.1aj] Virtual Bridged Local Area Networks / Two-Port Media Access Control (MAC) Relay, 802.1aj Draft 4.2, 24 September 2009.
- [802.1AE] "IEEE Standard for Local and metropolitan area networks / Media Access Control (MAC) Security", 802.1AE-2006, 18 August 2006.
- [802.1AX] "IEEE Standard for Local and metropolitan area networks / Link Aggregation", 802.1AX-2008, 1 January 2008.
- [802.1X] "IEEE Standard for Local and metropolitan area networks / Port Based Network Access Control", 802.1X-REV Draft 2.9, 3 September 2008.
- [RBridges] Perlman, R., "RBridges: Transparent Routing", Proc. Infocom 2005, March 2004.
- [RFC1661] Simpson, W., Ed., "The Point-to-Point Protocol (PPP)", STD 51, [RFC 1661](#), July 1994.
- [RFC3411] Harrington, D., Presuhn, R., and B. Wijnen, "An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks", STD 62, RFC 3411, December 2002.
- [RFC4086] Eastlake 3rd, D., Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, [RFC 4086](#), June 2005.
- [RFC4541] Christensen, M., Kimball, K., and F. Solensky, "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches", RFC 4541, May 2006.

- [RFC4789] Schoenwaelder, J. and T. Jeffree, "Simple Network Management Protocol (SNMP) over IEEE 802 Networks", [RFC 4789](#), November 2006.
- [RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, October 2008.
- [RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, February 2009.
- [RFC5342] Eastlake 3rd, D., "IANA Considerations and IETF Protocol Usage for IEEE 802 Parameters", BCP 141, [RFC 5342](#), September 2008.
- [RFC5556] Touch, J. and R. Perlman, "Transparent Interconnection of Lots of Links (TRILL): Problem and Applicability Statement", RFC 5556, May 2009.
- [RP1999] Perlman, R., "Interconnection: Bridges, Routers, Switches, and Internetworking Protocols, 2nd Edition", Addison Wesley Longman, Chapter 3, 1999.
- [VLAN-MAPPING] Perlman, R., Dutt, D., Banerjee, A., Rijhsinghani, A., and D. Eastlake 3rd, "Rbridges: Campus VLAN and Priority Regions", Work in Progress, April 2011.

Приложение А. Вопросы поэтапного развертывания

Некоторые аспекты частичного развертывания RBridge описаны ниже для определения стоимости каналов (Приложение А.1) и возможных перегрузок в результате пробок на назначенном устройстве пересылки (Приложение А.2). В Приложении А.3 рассмотрен конкретный пример проблемы, связанной с использованием TRILL одного назначенного устройства пересылки на канал и VLAN.

А.1. Определение стоимости канала

Когда в кампусе RBridge нет мостов и повторителей на каналах между RBridge, устройства RBridge могут точно определить число физических интервалов пути и скорость в линии на каждом интервале при условии, что эту информацию сообщают их порты. При наличии промежуточных устройств это становится невозможным. Например, как показано на рисунке 12, два моста B1 и B2 могут полностью скрыть медленный канал между ними и оба RBridge RB1 и RB2 будут некорректно считать канал быстрым.

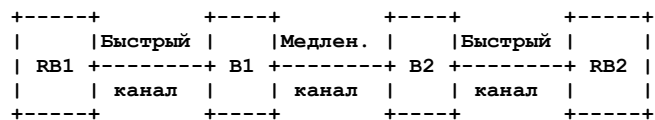


Рисунок 12. Стоимость канала с мостами.

Даже при наличии одного промежуточного моста два RBridge могут значить о соединении между ними, но каждый будет видеть для этого соединения свою скорость.

Однако эта проблема присуща не только RBridge. Мосты могут сталкиваться с подобными ситуациями в результате сокрытия каналов повторителями, а маршрутизаторы - в результате сокрытия мостами, повторителями и RBridge.

А.2. Назначенные устройства пересылки и ЛВС на базе мостов

При частичном развертывании RBridge эти устройства могут разбивать ЛВС на базе мостов на относительно небольшое число ЛВС сравнительно большого размера или не разбивать ЛВС совсем. Такая конфигурация может создавать описанную ниже проблему.

Требование входа (и выхода) естественных кадров в канал через назначенный для канала узел пересылки для VLAN, к которой относится кадр, может приводить к насыщению и неоптимальной маршрутизации (похожие проблемы могут возникать в ЛВС на базе мостов в результате использования алгоритма STP). Область влияния этой проблемы сильно зависит от топологии сети. Например, если ЛВС использует топологию «звезда» с мостами ядра, к которым подключены лишь другие мосты, и соединенными с ядром периферийными мостами, к которым подключены конечные станции, замена всех мостов в ядре на RBridge без замены периферийных мостов обычно будет повышать производительность без возникновения перегрузки устройств пересылки.

Решения этой проблемы описаны ниже, а в Приложении А.3 рассмотрен конкретный пример.

Установка RBridge с сохранением соединений между всеми частями ЛВС на базе мостов и наличием у каждой части множества соединений с RBridge в общем случае является наименее эффективным решением.

Имеется 4 метода решения описанной выше проблемы, которые в той или иной степени можно комбинировать.

1. Замена больше части пользовательских мостов IEEE 802.1 на RBridge для минимизации числа остающихся ЛВС на базе мостов между RBridge. Это не требует настройки RBridge, если заменяемые ими мосты не требовали настройки конфигурации.
2. Изменение топологии сети для минимизации проблемы. Если вовлеченные в процесс мосты и RBridge были настроены, может потребоваться изменение настроек.
3. Настройка RBridge и мостов так, чтобы оставшиеся станции ЛВС на базе мостов были разделены по разным VLAN с разными назначенными узлами пересылки. Если конечные станции уже находятся в разных VLAN, сделать это просто (параграф 4.2.4.2). если конечные станции находятся в одной VLAN и должны быть распределены по разным, этот метод может привести к проблемам связности между конечными станциями.
4. Настройка RBridge так, чтобы их порты, подключенные к ЛВС на базе мостов, передавали BPDU настройки STP (Приложение А.3.3) для форсированного разделения ЛВС на базе мостов¹. Для использования этого метода RBridge должны поддерживать эту необязательную функцию и для этого потребуется их настройка, но

¹Связующее дерево никогда не организуется через RBridge, но всегда завершается на портах RBridge.

вовлеченные в процесс мосты обычно настраивать не требуется. Этот метод делает сеть на базе мостов недоступной для трафика TRILL по причине ее разделения.

В отличие от п. 3 могут существовать ЛВС на базе мостов, использующие VLAN, может применяться больше VLAN, чем требуется для поддержки протокола MSTP¹ или могут использоваться иные меры снижения перегрузок, вызываемых одним деревом STP. Замена мостов IEEE 802.1 в таких ЛВС на RBridge может привести к снижению числа VLAN или отказу от них с соответствующим упрощением настройки.

А.3. Топология кабельного шкафа

Если имеются мосты 802.1, а RBridge не настроены корректно, связующее дерево моста или DRB могут принимать некорректные решения. Ниже приведен конкретный пример более общей проблемы, которая может возникать при подключении ЛВС на основе мостов по множеству RBridge.

При наличии двух (и более) групп конечных узлов, каждая из которых подключена к мосту (скажем, В1 и В2), а каждый мост подключен к RBridge (скажем, RB1 и RB2, соответственно) и имеется дополнительный канал между В1 и В2 (см. Рисунок 13) может оказаться желательным использовать канал В1-В2 лишь в качестве резервного на случай отказа одного из устройств RB1 и RB2 или каналов В1-RB1 и В2-RB2.

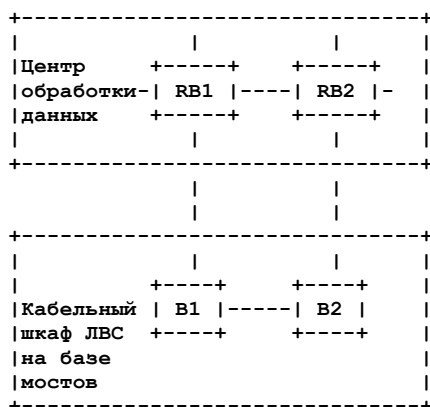


Рисунок 13. Топология кабельного шкафа.

Например, В1 и В2 могут размещаться в кабельном шкафу и для них легко можно организовать короткое, высокоскоростное и недорогое соединения, а RB1 и RB2 - в удаленном центре обработки данных, поэтому каналы RB1-В1 и RB2-В2 будут более медленными и дорогими.

В используемом по умолчанию варианте поведения один из RB1 и RB2 (скажем, RB1) может стать DRB для ЛВС на базе мостов, включающей В1 и В2, и назначит себя устройством пересылки для VLAN этой ЛВС. В результате RB1 будет пересылать весь трафик в канал (из канала) и конечные узлы, подключенные к В2, будут соединены с кампусной сетью по пути В2-В1-RB1, а не по желаемому каналу В2-RB2. Это будет вести к расходу пропускной способности пути В2-RB2 и сокращать пропускную способность между конечными станциями и центром обработки данных наполовину. Желательным поведением будет использование обоих каналов RB1-В1 и RB2-В2.

Ниже описаны три варианта решения этой проблемы.

А.3.1. Решение на основе RBridge

Если В1 и В2 заменить на RBridge, все будет работать без настройки (за исключением поддержки VLAN), но это может быть неразумно при постепенном переходе на RBridge.

А.3.2. Решение на основе VLAN

Если конечные станции, подключенные к В1 и В2 уже разделены между множеством VLAN, RB1 и RB2 могут оказаться настроенными так, что в зависимости от выбора одного из них в качестве DRB для этого канала он будет назначать себя устройством пересылки для некоторых VLAN, а другой RBridge - для оставшихся VLAN. При отказе или отключении любого из RBridge оставшийся будет назначать себя устройством пересылки для всех VLAN.

Если все конечные станции относятся к одной VLAN, это решение требует разделить их по меньшей мере между двумя VLAN. Это может создавать проблемы связности, для решения которых потребуются те или иные меры.

А.3.3. Решение на основе STP

Другим решением является настройка соответствующих портов RB1 и RB2 в составе «группы кабельного шкафа» с заданием для каждого RBridge порта Bridge Address Vx (которые может быть идентификатором системы RB1 RB2). Оба устройства RB1 и RB2 передают BPDU для своих настроенных портов, указывая их как корень Vx с высшим приоритетом. Это разделяет ЛВС на базе мостов путем блокировки канала В1-В2 на той или другой стороне (если один из мостов не настроен с высшим приоритетом и меньшим ID, что будет считаться ошибкой конфигурации). При заблокированном канале В1-В2 устройства RB1 и RB2 не смогут видеть сообщений TRILL-Hello друг от друга через этот канал и каждый будет действовать как DRB и назначенное устройство пересылки в соответствующей части дерева. При таком разделении трафик TRILL не будет проходить по пути RB1-В1-В2-RB2.

В BPDU настройки связующего дерева в качестве Root указывается Vx с высшим приоритетом, стоимость для Root равна 0, идентификатором DRB является RB1 (когда кадр передает RB1) или RB2 (когда передает RB2), а значение идентификатора порта выбирается RB1 и RB2 независимо, чтобы различать свои порты. Флаг изменения топологии сброшен, а флаг подтверждения изменения топологии установлен, если на порту был принят BPDU изменения топологии после того, как этот порт передал последний конфигурационный BPDU. Если RB1 и RB2 являются мостами

¹Multiple Spanning Tree Protocol - множество деревьев STP.

одной сети с разделяемой средой (т. е. между ними нет других мостов), мост с большим идентификатором увидит «лучшие» BPDU (по причине tiebreaker в третьем поле - идентификатора передающего моста) и выключит свой порт.

Если возникнет отказ RB1 или канала RB1-B1, либо RB2 или канала RB2-B2, алгоритм связующего дерева перестанет видеть один из корней RBx и разблокирует канал B1-B2, поддерживая связность конечных станций с центром обработки данных.

Если канал RB1-B1-B2-RB2 находится на срезе кампуса, а RB2 и RB1 настроены как часть группы кабельного шкафа, кампус разделится на две части при блокировке канала.

А.3.4. Сравнение решений

Замена всех пользовательских мостов 802.1 на RBridge обычно является наилучшим решением, когда нужна минимальная настройка конфигурации или ее нет совсем.

Решение на базе VLAN хорошо работает с небольшим объемом настроек конфигурации, если конечные станции уже поделены между разными VLAN. Если этого нет, решение становится более сложным и проблематичным.

В данном конкретном случае решение с связующим деревом достаточно хорошо. Но оно зависит от наличия на RB1 и RB2 дополнительной возможности настройки порта на передачу STP BPDU в соответствии с Приложением А.3.3. Это также делает ЛВС на базе мостов, разделение которой было форсировано, недоступной для сквозного трафика. Наконец, хотя в этом конкретном примере аккуратно разрывается канал между двумя мостами B1 и B2, при наличии более сложной ЛВС на базе мостов, а не просто пары мостов, не будет гарантии разделения на две равные части. В таких случаях может возникнуть сильно разбалансированная нагрузка на каналы RB1-B1 и RB2-B2, хотя это лучше, чем явное использование лишь одного из этих каналов.

Приложение В. Настройка транкового порта и порта доступа

Во многих современных ЛВС на базе мостов используется модель с уровнями ядра и доступа. Мосты уровня ядра имеют с другими мостами лишь соединения «точка-точка», а мосты уровня доступа соединены с конечными станциями, мостами ядра и возможно с другими мостами доступа. Представляется очевидной организацией некоторых кампусов RBridge по аналогичной схеме.

Порт RBridge можно настроить в качестве транкового, т. е. соединяющего с другим или другими RBridge, путем запрета на нем поддержки конечных станций. Это не является причиной поддержки на таком порту множества VLAN и их включения в Announcing Set. Конечно другие RBridge, которым подключен данный, должны поддерживать ту же VLAN. Нет причин использовать для этого VLAN, отличную от принятой по умолчанию VLAN 1, если канал не проходит через операторскую сеть Ethernet или другие среды, которые потребуют работать с другим значением VLAN. Такая конфигурация минимизирует издержки, связанные с сообщениями TRILL-Hello и предотвращает ненужную декапсуляцию, а также передачу кадров для множества получателей в естественной форме (параграфы 4.2.4 и 4.9.1).

Предполагается, что порт доступа RBridge ведет в канал, соединяющий с конечными станциями и возможно с одним или несколькими мостами. К такому каналу может быть также подключен один или несколько RBridge для повышения надежности обслуживания конечных станций. Имеет смысл минимизировать или исключить транзитный трафик в таких каналах, поскольку они предназначены для естественного трафика конечных станций. Это можно обеспечить настройкой конфигурации, как описано в параграфе 4.9.1.

При разработке модели конфигурации пользовательских интерфейсов RBridge следует обращать внимание на удобство его настройки в качестве транкового порта или порта доступа.

Приложение С. Поддержка множества путей

Маршрутизирующие мосты RBridge поддерживают множество путей для заведомо индивидуального и группового трафика. Реализация такой поддержки не обязательна.

Трафик с множеством получателей может передаваться по множеству путей за счет использования разных корней дерева распространения для разных кадров. Например, предположим, что на рисунке 14 конечные станции, подключенные к RBy являются источниками разных групповых потоков, каждый из которых имеет множество «слушателей», подключенных к RB1 - RB9. В предположении каналов с одинаковой пропускной способностью дерево распространения с корнем RBy будет использовать преимущественно вертикальные каналы между RB1 - RB9, а дерево с корнем RBz будет использовать преимущественно горизонтальные каналы. Если RBy выберет свой псевдоним в качестве корня дерева для половины этого трафика, а RBz сделает то же для другой половины, это позволит существенно увеличить суммарную пропускную способность за счет использования вертикальных и горизонтальных каналов между RB1 - RB9.

Поскольку деревья распространения, которые должен рассчитывать RBridge, одинаковы для всех RBridge и транзитные RBridge **должны** принимать во внимание корень дерева, заданный RBridge, кампус будет корректно работать с RBridge, часть которых использует разные корни для разных входящих кадров с множеством получателей, а другая часть использует один корень для всех таких кадров.

ЕСМР для заведомо индивидуального трафика может возникать на RBridge, если вместо использования критерия «удаления лишнего» (tiebreaker) при построении SPF будет сохраняться информация о портах, через которые доступны равноценные пути. Разные индивидуальные кадры могут тогда передаваться через разные порты и пересылаться по равноценным путям. Например, на рисунке 15, где показаны лишь RBridge и опущены все имеющиеся мосты, существует три равноценных пути между RB1 и RB2, а также два равноценных пути между RB2 и RB5. Таким образом, для трафика проходящего через эту часть кампуса слева направо, RB1 сможет выбирать из трех вариантов ЕСМР, а RB2 - из двух.

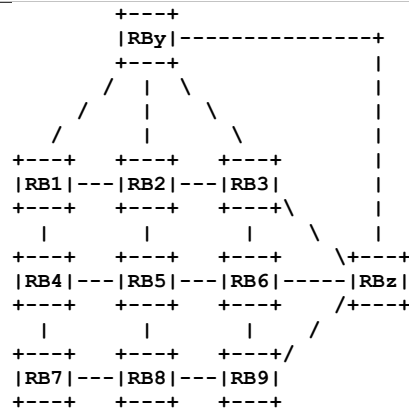


Рисунок 14. Множество путей ко множеству адресатов.

Транзитный RBridge, получая заведомо индивидуальные кадры, пересылает их в направлении выходного RBridge и не занимается вопросом своего присутствия на каком-либо определенном пути от входного или предыдущего транзитного RBridge. Таким образом кампус может корректно работать с RBridge, часть из которых поддерживает ECMP, а остальные не поддерживают.

Существует три варианта использования параллельных путей между RB1 и RB2, как описано ниже.

1. Если два или три из этих путей включают псевдоузлы, все 3 пути будут четко видны в состояниях каналов кампуса и применимо описанное выше использование ECMP.
2. Если пути используют сообщения P2P Hello или (в противном случае) не включают псевдоузлов, эти три пути в состоянии каналов будут представляться единой смежностью (adjacency). В этом случае использование множества путей между ними полностью определяется локально в RB1 и RB2. Множество путей можно свободно использовать для заведомо индивидуальных кадров, но не для кадров с множеством получателей (параграф 4.5.2).
3. Если все три пути между RB1 и RB2 являются одноэтапными (single hop) и имеют одинаковую пропускную способность, предпочтительным решением будет их объединение в один логический канал с помощью функции агрегирования [802.1AX]. Устройства RBridge **могут** реализовать объединение каналов, поскольку эта функция реализована ниже TRILL (параграф 4.9.2).

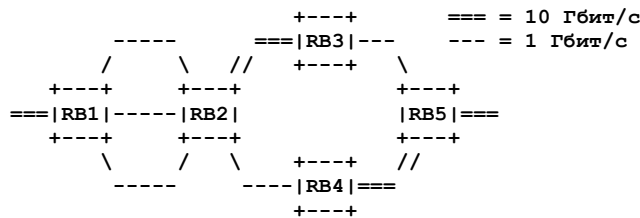


Рисунок 15. Множество путей для заведомо индивидуального трафика.

При использовании множества путей кадры, проходящие по разным путям, могут сталкиваться с разными задержками, что может приводить к нарушению порядка доставки. Хотя часть трафика не чувствительна к таким нарушениям, большая часть трафика связана с потоками кадров, в которых нарушение порядка играет роль. Способы определения потоков и уровень их детализации выходят за рамки документа (эти вопросы рассмотрены в [802.1AX]).

Приложение D. Определение VLAN и Priority

Ниже приведена краткая информация о способах определения VLAN ID и приоритета во входящих кадрах. Более подробные сведения можно получить из [802.1Q-2005].

- При получении естественного кадра без тегов не настроенный RBridge связывает с ним принятый по умолчанию приоритет 0 и VLAN ID 1. Фактически он устанавливает для непомиченного кадра в качестве идентификатора VLAN связанный с портом идентификатор «port VLAN ID». Этот идентификатор по умолчанию соответствует VLAN ID 1, но в конфигурации может быть задано другое значение VLAN ID. RBridge можно также настроить на уровне отдельного порта на отбрасывание таких кадров или связывание с ними жругого уровня приоритета. Определение VLAN ID для не помеченного входного кадра, не относящегося к управлению, возможно также на основе Ethertype или NSAP (в 802.1 называется Protocol) в принятом кадре, MAC-адреса отправителя и других локальных правил.
- При получении естественного кадра с тегом приоритета не настроенный RBridge связывает с ним VLAN ID для порта (по умолчанию 1) и код приоритета, представленный тегом приоритета в кадре. RBridge можно настроить на уровне порта на отбрасывание таких кадров или связывание с ними другого VLAN ID, как указано выше. Можно также настроить отображение на код приоритета путем задания для каждого из 8 возможных значений в кадре нужного кода приоритета, который RBridge свяжет с этим кадром.
- Естественные кадры с C-тегом (раньше назывался Q-тегом) на не настроенном RBridge связываются с VLAN ID и приоритетом из C-тега. RBridge можно настроить на уровне порта и VLAN на отбрасывание таких кадров. Можно также на уровне порта настроить отображение приоритета, как описано выше для кадров с тегом приоритета.

В 802.1 процесс назначения кода приоритета с кадром, включая отображение приоритета из кадра на другой код, называется регенерацией приоритета.

Приложение E. Поддержка поправок к стандарту IEEE 802.1Q-2005

В этом информационном приложении кратко описана поддержка в RBridge завершенных и разрабатываемых поправок к стандарту IEEE [802.1Q-2005]. Нет никаких гарантий, что текущие спецификации протокола RBridge и имеющиеся мосты будут поддерживать еще не завершённые поправки к [802.1Q-2005] точно так же, как нет гарантий поддержки существующими протоколами мостов или RBridge еще не разработанных поправок для TRILL.

Приведенная ниже информация отражает состояние на 25 октября 2009 г. Более свежую информацию можно найти на странице рабочей группы IEEE 802.1 (<http://grouper.ieee.org/groups/802/1/>).

E.1. Завершённые поправки

802.1ad-2005 Provider Bridges

Иногда называется Q-in-Q, поскольку теги VLAN называют Q-тегами. Стандарт 802.1ad задает мосты PB, которые туннелируют трафик пользовательских мостов с использованием сервисных VLAN (S-теги). Если пользовательская ЛВС является кампусом RBridge, трафик будет передаваться мостами PB. Элементы пользовательских мостов, осведомленные о PB (типа возможности настройки порта в пользовательском мосту на добавление в кадр S-тега до передачи кадра мосту PB), размещаются ниже уровня EISS и могут поддерживаться мостами RBridge без изменения протокола TRILL.

802.1ag-2007 Connectivity Fault Management (CFM)

Это свойство 802.1 по крайней мере частично зависит от симметрии пути и других характеристик связующего дерева. Комментарии, представленные группе IETF TRILL рабочей группой IEEE 802.1, говорят, что «TRILL снижает применимость CFM».

802.1ak-2007 Multiple Registration Protocol

Поддерживается для расширения, описанного в параграфе 4.9.4.

802.1ah-2008 Provider Backbone Bridges

Иногда называется MAC-in-MAC. Стандарт 802.1ah обеспечивает для PBB, которые туннелируют трафик пользовательских мостов с использованием других внешних MAC-адресов и тега (I-tag) для сохранения исходных MAC-адресов и передачи других сигналов. Если пользовательская ЛВС является кампусом RBridge, трафик будет передаваться мостами PBB. Функции пользовательских мостов, включающие осведомленность о PBB, типа способности настраивать порт пользовательского моста на добавление в кадр I-тега до отправки мосту PBB, размещаются ниже уровня EISS и могут поддерживаться мостами RBridge без изменения протокола TRILL.

802.1Qav-2009 Management of Data-Driven and Data-Dependent

Отказы соединений. Поправка на основе 802.1ag. См. Комментарий к 802.1ag-2007 выше.

802.1Qay-2009 Provider Backbone Bridge Traffic Engineering

Поправка на основе 802.1ah для настройки маршрутизации при построении трафика. См. комментарий к 802.1ah-2008 выше.

E.2. Разрабатываемые поправки

Ниже перечислены поправки к стандарту IEEE [802.1Q-2005], находящиеся в разработке. Приведенные краткие описания основаны на предварительных вариантах стандартов и могут оказаться некорректными для более поздних или окончательных версий.

802.1aj Two-port MAC Relay [802.1aj]

Эта поправка задает ретранслятор MAC, который будет прозрачен для RBridge. Маршрутизирующие мосты RBridge совместимы в настоящее время с устройствами IEEE 802.1aj в том же смысле, что и мосты IEEE 802.1Q-2005.

802.1aq Shortest Path Bridging

Эта поправка обеспечивает улучшенную маршрутизацию в ЛВС на основе мостов.

802.1Qat Stream Reservation Protocol

Изменение 802.1Q для поддержки синхронизации 802.1. Этот протокол резервирует ресурсы на поддерживающих мостах.

802.1Qau Congestion Notification

В настоящее время представляется, что для поддержки этой поправки придется менять поведение RBridge выше уровня EISS. Такие изменения выходят за рамки этого документа.

802.1Qav Forwarding and Queuing Enhancements for Time-Sensitive

Потоки - изменение 802.1Q для поддержки протокола синхронизации 802.1 (Timing and Synchronization). Эта поправка задает методы поддержки резервирования ресурсов с помощью протокола 802.1Qat (см. выше).

802.1Qaz Enhanced Transmission Selection

Предполагается, что эта поправка будет работать ниже уровня EISS и сможет поддерживаться на портах RBridge без изменения протокола TRILL.

802.1Qbb Priority-based Flow Control

Обычно называется reg-priority pause (пауза для приоритета). Похоже, что эта поправка будет ниже уровня EISS и сможет поддерживаться на портах RBridge без изменения протокола TRILL.

802.1bc Remote Customer Service Interfaces

Расширение 802.1Q PBB. См. 802.1ad-2005 выше.

802.1Qbe Multiple Backbone Service Instance Identifier (I-SID)

Протокол регистрации (MIRP). Расширение для 802.1Q PBB. См. 802.1ah-2008 выше.

802.1Qbf Provider Backbone Bridge Traffic Engineering (PBB-TE)

Защита сегмента инфраструктуры. Расширяет стандарт 802.1Q для поддержки некоторых типов отказоустойчивости между PBB. См. 802.1ah-2008 выше.

Приложение F. Благодарности

В этой работе приняло участие множество людей, часть которых перечислена ниже в алфавитном порядке.

Bernard Aboba, Alia Atlas, Ayan Banerjee, Caitlin Bestler, Suresh Boddapati, David Michael Bond, Stewart Bryant, Ross Callon, James Carlson, Pasi Eronen, Dino Farinacci, Adrian Farrell, Don Fedyk, Bill Fenner, Eric Gray, Sujay Gupta, Joel Halpern, Andrew Lange, Acee Lindem, Vishwas Manral, Peter McCann, Israel Meilik, David Melman, Nandakumar Natarajan, Erik Nordmark, Jeff Pickering, Tim Polk, Dan Romascanu, Sanjay Sane, Pekka Savola, Matthew R. Thomas, Joe Touch, Mark Townsley, Kate Zebrose.

Адреса авторов**Radia Perlman**

Intel Labs
2200 Mission College Blvd.
Santa Clara, CA 95054-1549 USA
Phone: +1-408-765-8080
EMail: Radia@alum.mit.edu

Donald E. Eastlake, 3rd

Huawei Technologies
155 Beaver Street
Milford, MA 01757 USA
Phone: +1-508-333-2270
EMail: d3e3e3@gmail.com

Dinesh G. Dutt

Cisco Systems
170 Tasman Drive
San Jose, CA 95134-1706 USA
Phone: +1-408-527-0955
EMail: ddutt@cisco.com

Silvano Gai

Cisco Systems
170 Tasman Drive
San Jose, CA 95134-1706 USA
EMail: silvano@ip6.com

Анооп Ghanwani

Brocade
130 Holger Way
San Jose, CA 95134 USA
Phone: +1-408-333-7149
EMail: anoop@alumni.duke.edu

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru