Network Working Group
Request for Comments: 4664
Category: Informational

L. Andersson, Ed.
Acreo AB
E. Rosen, Ed.
Cisco Systems, Inc.
September 2006

Framework for Layer 2 Virtual Private Networks (L2VPNs)

Модель виртуальных частных сетей L2 (L2VPN)

Статус документа

В этом документе содержится информация для сообщества Internet. Документ не задаёт каких-либо стандартов Internet. Документ может распространяться без ограничений.

Аннотация

Этот документ определяет модель предоставляемых провайдерами услуг виртуальных частных сетей на канальном уровне (L2VPN[1]). Цель модели заключается в оказании помощи при стандартизации протоколов и механизмов для поддержки взаимодействия L2VPN.

Оглавление

[1]Layer 2 Provider Provisioned Virtual Private Network.

1. Введение

1.1. Уровни требований

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не нужно** (SHALL NOT), **следует** (SHOULD), **не следует** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с RFC 2119 [1].

1.2. Цели и область действия документа

Этот документ определяет модель предоставляемых провайдерами услуг виртуальных частных сетей на канальном уровне (L2VPN). Цель модели заключается в оказании помощи при стандартизации протоколов и механизмов для поддержки интероперабельных L2VPN.

Термин «предоставляемые провайдером VPN» обозначает виртуальные частные сети (VPN[1]), для которых сервис-провайдер (SP[2]) участвует в управлении и обеспечении VPN.

Требования к услугам L2VPN описаны в [RFC4665].

В этом документе представлены эталонные модели и рассмотрены функциональные компоненты L2VPN. В частности, обсуждаются технические вопросы, которые важны для разработки стандартов и механизмов L2VPN, включая требуемые для межсетевого взаимодействия и обеспечения безопасности.

В документе рассматривается множество технических подходов к L2VPN и предпринимается попытка показать их взаимосвязи, а также прояснить вопросы, которые могут возникать при выборе одного подхода вместо другого. Однако документ не пытается выбрать какой-либо конкретный подход.

1.3. Виртуальные частные сети L2

Имеется два фундаментально различающихся типа услуг L2 VPN, предлагаемых сервис-провайдерами своим абонентам - VPWS[3] и VPLS[4]. Возможна также поддержка услуг IPLS[5].

VPWS - это сервис VPN, основанный на предоставлении соединений «точка-точка» на уровне L2. Природа такого сервиса не создаёт проблем с расширяемостью. Проблемы могут быть связаны лишь с числом конечных точек, которые может поддерживать отдельное устройство PE.

VPLS - это сервис канального уровня (L2), основанный на эмуляции услуг ЛВС через распределенную сеть (WAN[6]). Объем данных состояния, который должен храниться на краевых устройствах для поддержки функций пересылки, определяет параметры расширяемости ЛВС. Другие проблемы расширяемости могут быть связаны лишь с числом конечных точек, которые может поддерживать отдельное устройство PE (см. параграф 3.4.4).

Отметим, что в VPLS используются услуги без естественной групповой адресации (multicast) для эмуляции услуг, которые включают естественную групповую адресацию. В результате это создаёт проблемы расширяемости, связанные с обработкой группового трафика в VPLS.

Услуги VPLS могут вносить большие задержки и обеспечивать менее надёжный транспорт по сравнению с естественными ЛВС. Стандартные протоколы управления ЛВС могут оказаться не готовыми к работе в таких средах и это может создавать дополнительные проблемы расширяемости сервиса.

1.4. Терминология

Список технических терминов, используемых при обсуждении L2VPN, приведён в [RFC4026].

2. Модели

2.1. Эталонная модель для VPWS

Эталонная модель VPWS приведена на рисунке 1.

[1]Virtual Private Network.
[2]Service Provider.
[3]Virtual Private Wire Service - услуги виртуального частного провода (соединения).
[4]Virtual Private LAN Service - услуги виртуальной частной ЛВС.
[5]IP-only LAN-like Service - подобный ЛВС сервис, поддерживающий только протокол IP.
[6]Wide Area Network.

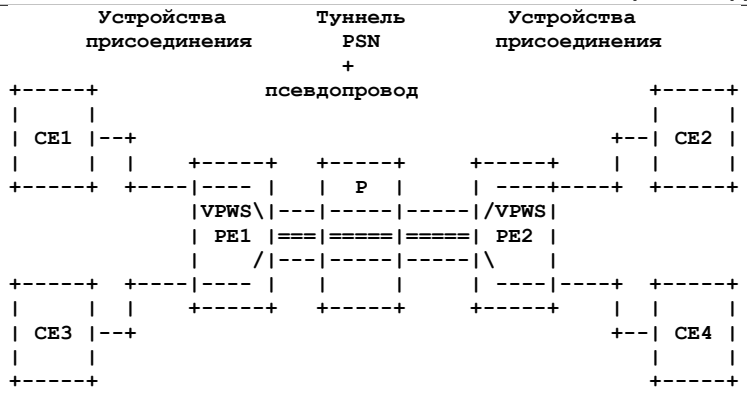


Рисунок 1.

2.1.1. Элементы эталонной модели VPWS

Устройства P, PE (VPWS-PE) и CE, а также туннели PSN определены в [RFC4026]. Устройства присоединения и псевдопровода (PW[1]) описаны в разделе 3. PE просто отображает PW на устройство присоединения в соответствии с локальной информацией, т. е. демультиплексором PW и входным/выходным физическим или логическим портом.

2.2. Эталонная модель для VPLS

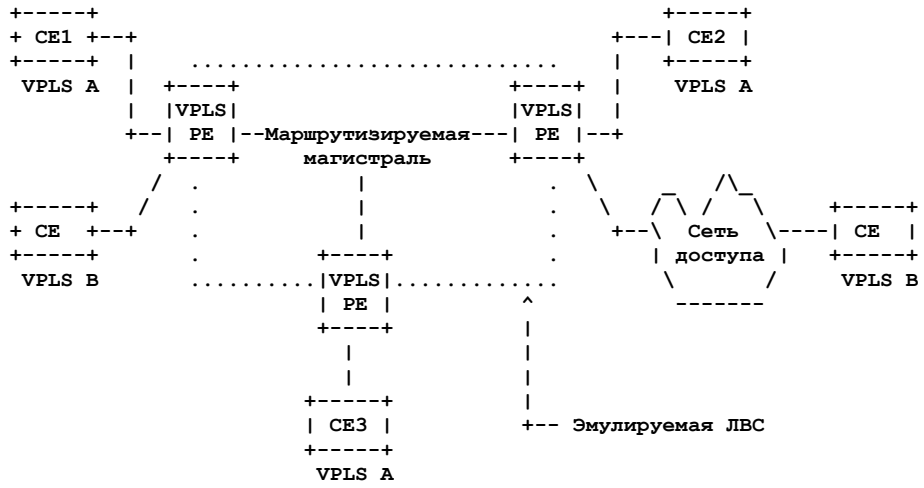


Рисунок 2.

На рисунке 2 показана эталонная модель VPLS, где устройства PE с поддержкой VPLS обеспечивают логические соединения так, что устройства CE, относящиеся к одной VPLS, представляются подключёнными к одной сети на базе мостов Ethernet. VPLS может включать одну виртуальную сеть (VLAN) или множество VLAN с тегами. Логическая модель VPLS представлена на рисунке 3.

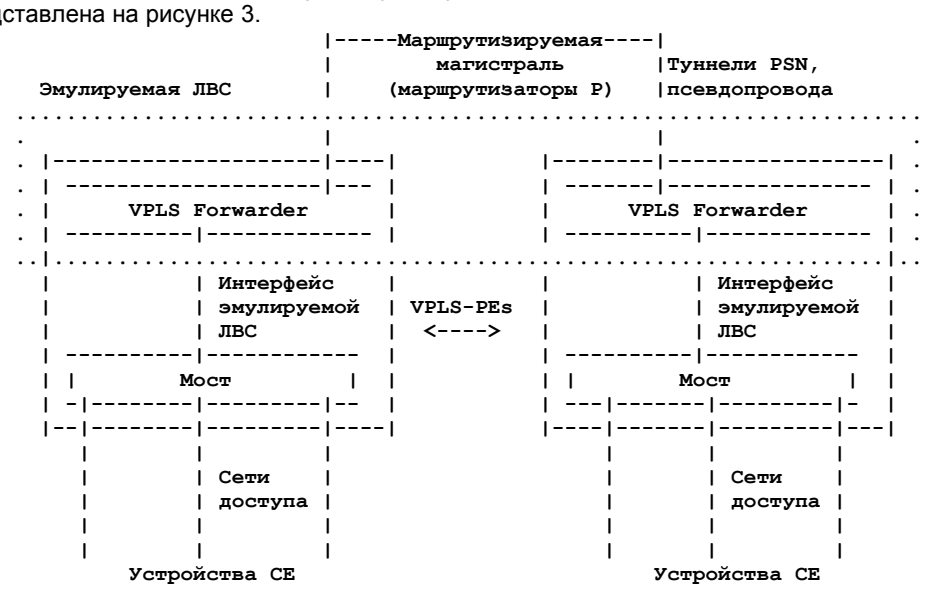


Рисунок 3.

На рисунке 3 показано, что в VPLS устройства CE подключаются (возможно через сеть доступа) к модулю моста в VPLS-PE. Внутри VPLS-PE модуль моста соединяется через интерфейс эмулируемой ЛВС с эмулируемой ЛВС. Для каждого экземпляра VPLS имеется экземпляр эмулируемой ЛВС. На рисунке 3 показана некая внутренняя структура эмулируемой ЛВС, включающая модули пересылки (VPLS Forwarder) соединённые псевдопроводами, которые могут проходить через туннели PSN в маршрутизируемой магистральной сети.

Экземпляр VPLS состоит из набора модулей пересылки (не более одного на PE), соединённых псевдопроводами.

[1]Pseudowire.

Требуемая от моста функциональность зависит от сервиса, который SP предоставляет своим абонентам, а также от устройства сети SP. По меньшей мере мост должен поддерживать стандартное MAC-обучение и старение записей. Однако, если устройства PE имеют «закулисные» соединения между собой через сеть L2, может потребоваться полная функциональность мостов IEEE ([IEEE8021D]) с поддержкой протокола связующего дерева между всеми мостами. Точная спецификация требований к модулям мостов в конкретных обстоятельствах выходит за рамки этого документа.

Эта схема задаёт у каждого модуля моста наличие одного интерфейса эмулируемой ЛВС. Не задается число модулей моста, которые могут присутствовать в VPLS-PE, а также число экземпляров VPLS, которые могут быть подключены к модулю моста через один интерфейс эмулируемой ЛВС.

Эта схема совместима по меньшей мере с перечисленными ниже тремя моделями.

- Модель 1

  VPLS-PE включает один модуль моста и поддерживает один экземпляр VPLS, которым служит эмулируемая ЛВС. Если эта ЛВС поддерживает VLAN, должны применяться теги 802.1Q [IEEE8021Q] для привязки пакетов к VLAN.

- Модель 2

  VPLS-PE включает один модуль моста, но поддерживает множество экземпляров VPLS, каждый из которых считается VLAN (по сути, эмулируемой ЛВС). Множество экземпляров VPLS рассматривается как набор VLAN в одной ЛВС. Поскольку каждая сеть VLAN использует свой набор PW, нет необходимости в тегах 802.1Q.

- Модель 3

  VPLS-PE содержит произвольное число модулей-мостов, каждый из которых подключён к 1 экземпляру VPLS.

Могут быть и другие модели, а некоторые из них могут быть комбинацией перечисленных выше. Различные модели могут иметь разные характеристики и разные сферы применения.

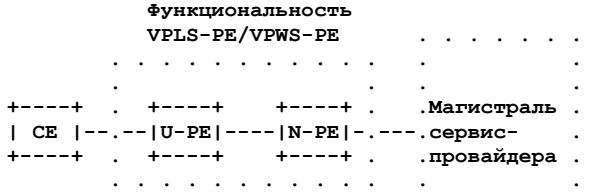
Каждому решению VPLS следует указывать поддерживаемую модель или модели. Каждое решение должно также задавать требуемую от модуля моста функциональность.

Эта схема не задаёт способа использования протоколов управления мостами в эмулируемых ЛВС.

2.2.1. Элементы эталонной модели VPLS

Устройства PE (VPLS-PE) и CE определены в [RFC4026].

2.3. Эталонная модель для распределенных VPLS-PE и VPWS-PE



2.3.1. Элементы эталонной модели распределенного PE

Функциональность VPLS-PE и VPWS-PE может быть распределена между несколькими устройствами. Устройства, расположенные ближе к клиенту, называют U-PE[1], а расположенные ближе к ядру сети - N-PE[2] (параграф 3.4.3).

Определения U-PE и N-PE приведены в [RFC4026].

2.4. VPWS-PE и VPLS-PE

Устройства VPWS-PE и VPLS-PE функционально очень похожи в том, что оба применяют модуль пересылки (forwarder) для отображения устройств присоединения на псевдопровода. Единственное различие заключается в том, что модуль пересылки VPWS-PE сопоставляет одно устройство присоединения с одним псевдопроводом, а модуль пересылки в VPLS-PE является экземпляром виртуального коммутатора VSI[3], который отображает множество устройств присоединения на множество псевдопроводов (раздел 3).

3. Функциональные компоненты L2 VPN

В этом разделе описана функциональная модель L2VPN, которая позволяет разделить архитектуру L2VPN не её функциональные компоненты. Это показывает роли различных протоколов и механизмов, упрощая понимание различий и сходства между разными вариантами архитектуры L2VPN.

В параграфе 3.1 приведён обзор нескольких типов L2VPN, в параграфе 3.2 описаны общие для разных типов компоненты. Затем приводится более подробное описание каждого типа сервиса L2VPN. В заключение рассматриваются функциональные компоненты, которые относятся к отдельным типам L2VPN и зависящие от типа функции компонентов общего назначения.

3.1. Типы L2VPN

Типы L2VPN различаются характеристиками услуг, предоставляемых клиентам сервис-провайдером (SP).

3.1.1. Услуги VPWS

В VPWS каждое устройство CE представлено набором виртуальных устройств (каналов) «точка-точка».

[1]User-Facing PE - обращённое к абоненту устройство PE.
[2]Network-Facing PE - обращённое к сети устройство PE.
[3]Virtual Switching Instance.

На другой стороне каждого виртуального канала размещается другое устройство CE. Кадры, передаваемые CE в такое виртуальное устройство принимаются CE на другой стороне виртуального соединения. Пересылка между устройствами CE не зависит от содержимого кадра и полностью определяется виртуальным устройством, в которое кадр передаётся. Устройства PE выступают в качестве коммутаторов виртуальных каналов.

Этот тип L2VPN был доступен достаточно давно в сетях ATM и Frame Relay. Предоставление таких услуг L2VPN через сети MPLS и/или IP актуально сегодня.

Требования к этому типу сервиса L2VPN указаны в [RFC4665].

3.1.2. Услуги VPLS

В VPLS каждое устройство CE имеет один или несколько интерфейсов ЛВС, ведущих в «виртуальную магистраль».

Два CE подключаются к одной виртуальной магистрали лишь в том случае, когда они относятся к одному экземпляру VPLS (т. е. в одной сети VPN). Когда CE передаёт кадр, принявшее его устройство PE проверяет поле MAC-адреса получателя для решения вопроса о пересылке кадра. Таким образом, PE играет роль моста. Как показано на рисунке 3, если набор устройств PE поддерживает общий экземпляр VPLS, создаётся эмулируемая ЛВС, соответствующая этому экземпляру VPLS, к которой подключён каждый из этих PE (через эмулируемый интерфейс). С точки зрения CE виртуальная магистраль представляет набор мостов PE и эмулируемую ЛВС, к которой они относятся. Таким образом, для устройства CE ЛВС, которая соединяет его с мостом, распространяется через магистральную сеть MPLS и/или IP.

Мост PE рассматривает эмулируемую ЛВС как любую другую ЛВС, в которую у него есть интерфейс. Решение о пересылке принимается как в обычных мостах на основе изучения MAC-адресов отправителей.

Сервис VPLS похож на VPWS в том смысле, что решение о пересылке принимается без привлечения заголовка сетевого уровня (L3). Отличия VPLS от VPWS указаны ниже.

- VPLS позволяет PE использовать адресную информацию из заголовка L2 в кадре для решения вопроса о пересылке кадра.
- VPLS позволяет использовать одно соединение CE-PE для передачи кадров множеству удалённых CE. В этом отношении VPLS больше напоминает L3VPN, нежели VPWS.

Требования к этому типу сервиса L2VPN указаны в [RFC4665].

3.1.3. Услуги IPLS

Сервис IPLS очень поход на VPLS, за исключением перечисленных ниже отличий.

- Предполагается, что устройства CE являются хостами или маршрутизаторами, а не коммутаторами.
- Предполагается, что сервис будет работать только с пакетами IP и протоколами поддержки, такими как ICMP и ARP для IPv4 или Neighbor Discovery для IPv6. Кадры L2, содержащие другие протоколы, не поддерживаются.

Хотя этот сервис является функциональным подмножеством VPLS, он рассматривается отдельно, поскольку для его реализации могут применяться другие механизмы, которые могут работать на некоторых аппаратных платформах, не поддерживающих VPLS полностью.

3.2. Функциональные компоненты базового транспорта L2VPN

Все типы L2VPN должны транспортировать «кадры» через ядро сети, соединяющее устройства PE. Для всех типов L2VPN устройство PE (PE1) получает кадр от CE (CE1), затем передаёт его другому устройству PE (PE2), которое доставляет кадр CE (CE2). В этом параграфе рассматриваются функциональные компоненты, которые нужны для транспортировки кадров L2 во всех типах L2VPN.

3.2.1. Устройства присоединения

Во всех типах L2VPN устройство CE подключается к PE через то или иное устройство или виртуальный канал. Будем называть это устройством присоединения (подключения) или AC[1]. Термин применяется очень широко, устройствами подключения могут служить Frame Relay DLCI, ATM VPI/VCI, порты Ethernet, VLAN, соединения PPP на физических интерфейсах, сессии PPP из туннелей L2TP, MPLS LSP и т. п. Устройством CE может быть маршрутизатор, коммутатор, хост, или что-либо иное, подключённое клиентом к VPN. AC передаёт кадры между CE и PE.

Процедуры организации и поддержки AC выходят за рамки этой архитектуры.

Это процедуры обычно включаются в спецификации конкретной технологии устройств присоединения.

Любой кадр будет проходить через AC от CE к PE на одной стороне, а затем через другое устройство AC от PE к CE.

Первое устройство AC будем называть входным (ingress AC), а второе - выходным (egress AC). Отметим, что устройства AC служат входным и выходным для конкретного кадра и указывают лишь направления в данном AC.

3.2.2. Псевдопровода

Псевдопровод (PW) обеспечивает связь между двумя устройствами PE. Как AC служит для передачи кадров между CE и PE, так PW используется для передачи кадров между PE. Термин PW используется в соответствии с определением [RFC3985].

Организация и поддержка PW являются задачами устройств PE. Информация о состоянии конкретного PW поддерживается на двух соединяемых псевдопроводом PE, но не на других PE и магистральных маршрутизаторах (P).

Псевдопровода могут поддерживать соединения «точка-точка», «один со многими» и «множество с одним». В этой схеме PW типа «точка-точка» всегда считаются двухсторонними, а PW типов «один со многими» и «множество с одним» всегда считаются односторонними. PW типа «множество с одним» могут применяться лишь в тех случаях,

[1]Attachment Circuit.

когда принимающему кадр PE не нужно на основании получившего кадр PW определять входное устройство AC. PW типа «один со многими» могут быть полезны для групповых кадров.

Процедуры организации и поддержки PW типа «один со многими» не рассматриваются в этой версии модели.

Любой конкретный кадр приходит сначала на входное устройство AC, затем в PW, а затем на выходное устройство AC.

Групповые кадры могут реплицироваться PE, поэтому информация, передаваемая в групповых кадрах, может проходить через множество PW и множество выходных AC.

Таким образом, применительно к данному кадру PW можно связать с несколькими AC. Если эти AC используют одну технологию (например, ATM, Ethernet, Frame Relay), говорят, что PW обеспечивает однородны транспорт, в противном случае транспорт будет неоднородным. Для неоднородного транспорта требуется использовать ту или иную функцию межсетевого взаимодействия. Есть по меньшей мере 3 разных подхода к такому взаимодействию, приведённых ниже.

1. Одно из устройств CE может локально выполнять функции межсетевого взаимодействия. Например, если устройство CE1 подключено к PE1 через ATM, а устройство CE2 подключено к PE2 через Ethernet, CE1 может принять решение о приёме и передаче кадров Ethernet через ATM с использованием инкапсуляции RFC 2684. В таком случае устройству PE1 нужно знать, что оно завершает ATM VC локально и принимает/передаёт через PW лишь кадры Ethernet.
2. Функции межсетевого взаимодействия может выполнять одно из устройств PE. Например, если устройство CE1 подключено к PE1 через ATM, а устройство CE2 подключено к PE2 через Frame Relay, PE1 может выполнять функцию «ATM/FR Service Interworking». CE не будут видеть этого и через PW будут передаваться только кадры Frame Relay.
3. Можно использовать IPLS. В этом случае «кадрами» в PW будут дейтаграммы IP и двум PE нужно взаимодействовать для обмана (spoof) различных относящихся к L2 процедур, применяемых IP (параграф 3.5).

При использовании разнородных PW протокол организации должен гарантировать, что каждая из конечных точек знает MTU удалённого AC. Если у двух AC значения различаются, нужно выполнять три перечисленных ниже правила.

- Недопустимо активизировать PW.
- Конечная точка AC с большим MTU должна уменьшить MTU своего AC до значения MTU удалённого AC.
- Конечные точки должны согласовать использование процедуры фрагментации-сборки.

3.2.3. Модули пересылки

Во всех типах L2VPN устройство PE (PE1) получает кадр через AC и пересылает его через PW другому PE (PE2). Затем PE2 пересылает этот кадр другому AC.

Случай, когда PE1 и PE2 являются одним устройством, важен в плане корректности обработки для правильного функционирования сервиса L2VPN. Однако для него не нужен какой-либо протокол, поэтому далее этот случай в документе не рассматривается.

Когда PE1 получает кадр от конкретного AC, нужно определить PW, куда должен пересылаться кадр. В общем случае это выполняется на основе следующей информации:

- входное устройство AC;
- возможно, содержимое заголовка L;
- возможно, некая информация для пересылки, поддерживаемая статически или динамически.

Если рассматривается статическая или динамическая информация, она относится к определённому экземпляру L2VPN (т. е. отдельной VPN).

Аналогично при получении кадра PE2 на конкретном PW нужно определить AC, куда должен пересылаться кадр. Это определяется следующей информацией:

- входной PW;
- возможно, содержимое заголовка L;
- возможно, некая информация для пересылки, поддерживаемая статически или динамически.

Если рассматривается статическая или динамическая информация, она относится к определённому экземпляру L2VPN (т. е. отдельной VPN).

Процедуры, используемые для принятия решения о пересылке, называют «модулем пересылки» (forwarder). Будем считать PW «привязанным» в каждой из конечных точек к модулю пересылки. Эти модули, в свою очередь, «привязывают» псевдопровода PW к устройствам AC. В разных типах L2VPN применяются разные модули пересылки.

Например, forwarder может связывать одно устройство AC с одним PW, игнорируя содержимое кадров и не используя для пересылки другой информации. Другой модуль пересылки может связывать AC с множеством PW и AC, передавая отдельные кадры из AC в PW, из PW в AC или из AC в AC на основе сравнения данных в заголовке L2 с базой данных пересылки. Это более подробно рассматривается при описании разных типов L2VPN.

3.2.4. Туннели

PW организуется в «туннеле» от PE1 до PE2. Предполагается возможность организации в одном туннеле произвольного числа PW[1], если все они завершаются на PE2.

[1]Хотя можно представить методы туннелирования с поддержкой лишь одного PW на туннель, у них будут очевидные проблемы с расширяемостью, поэтому далее такие варианты не рассматриваются.

Не требуется даже начинать все PW в туннеле на устройство PE1, это могут быть туннели типа multipoint-to-point. Не требуется и передача всех PW между парой устройств PE в одном туннеле. Все, что требуется, - это возможность PE2 связать проходящий через туннель кадр с определенным PW.

Для организации туннелей между PE могут применяться различные технологии. Все, что реально требуется от такой технологии, - это поддержка демультиплексирования организованных в туннеле PW. В качестве туннеля могут служить MPLS LSP, туннели L2TP, IPsec, MPLS-in-IP и т. п. В общем случае технология туннелирования будет требовать применения инкапсуляции, содержащей поле демультиплексора, служащее для идентификации PW. Процедуры организации и поддержки туннелей выходя за рамки этой схемы (см. параграф 3.2.6. Сигнализация PW).

При наличии множества туннелей от PE1 до PE2 может оказаться желательной привязка определённого PE1-PE2 PW к конкретному туннелю на основе тех или иных характеристик PW и/или туннеля. Например, можно размещать PW с определёнными требованиями QoS в туннель с подходящими параметрами QoS. Процедуры такой привязки выходят за рамки этой модели.

Хотя PW «точка-точка» являются двухсторонними, туннели, через которые они проходят, не обязаны быть двухсторонними и/или относиться к типу «точка-точка». Например, PW «точка-точка» может проходить через односторонний путь multipoint-to-point MPLS LSP.

3.2.5. Инкапсуляция

Поскольку пакеты L2VPN передаются через псевдопровода, следует применять стандартные для псевдопровода форматы и методы инкапсуляции (заданные рабочей группой IETF PWE3), когда они подходят.

В общем случае инкапсуляция PW затем инкапсулируется в туннель в соответствии с туннельным протоколом.

Может потребоваться определение дополнительных вариантов инкапсуляции PW с учётом потребностей L2VPN, но это может выходить за рамки задач группы PWE3. Примером может служить неоднородный (гетерогенный) транспорт.

3.2.6. Сигнализация PW

Процедуры организации и поддержки PW являются частью этой схема и включают процедуры распространения значений поля демультиплексора, хотя это поле, строго говоря, относится к протоколу туннелирования, а не к PW.

Сигнализация для псевдопровода «точка-точка» должна обеспечивать перечисленные ниже функции.

- Распространение демультиплексора.

  Поскольку в одном туннеле может поддерживаться множество PW, протокол туннелирования должен назначать значение демультиплексора для каждого PW. Эти значения должны быть уникальными в рамках туннеля, а для некоторых технологий туннелирования - в рамках выходного PE. В общем случае устройство PE, служащее выходом туннеля, будет выбирать значения демультиплексора и распространять их входному (входным) PE этого туннеля. Это является важной частью процедуры организации PW.

  Отметим, что как обычно в архитектуре туннелирования поле демультиплексора относится к протоколу туннелирования, а не к туннелируемому протоколу. По этой причине протоколы организации PW могут быть расширениями для организации туннеля.

- Выбор модуля пересылки (Forwarder) на удалённом PE.

  Сигнальный протокол должен содержать достаточно информации, чтобы удалённый маршрутизатор PE мог выбрать подходящий модуль пересылки, к которому будет привязан PW. Эту информацию называют селектором удалённого модуля пересылки (Remote Forwarder Selector). Требуемая для этого информация зависит от типа L2VPN и модели предоставления услуг (см. параграфы 3.3.1 и 3.4.2). Remote Forwarder Selector может однозначно указывать конкретный модуль пересылки или задавать атрибуты таких модулей. В последнем случае выбор конкретного модуля пересылки будет определяться этими атрибутами.

- Поддержка эмуляции псевдопроводов.

  В той степени, с которой конкретный PW должен эмулировать определённую технологию L2, сигнализация PW должна поддерживать требуемые функции.

- Распространение смены состояний.

  Изменения состояний AC могут потребовать смены состояний PW, к которым привязаны эти устройства AC, и наоборот. Спецификация способов отражения изменений обычно относится к задачам рабочей группы PWE3.

- Задание характеристик псевдопровода.

  В той степени, которая требует известности и/или согласования одной или нескольких характеристик PW на обеих сторонах соединения, сигнализация должна обеспечивать требуемое взаимодействие.

Как отмечено выше, сигнализация для PW «точка-точка» должна передавать достаточно информации, чтобы удалённый маршрутизатор PE мог должным образом связать PW с модулем пересылки, а также связать конкретное значение демультиплексора с этим PW. Когда два PE имеют корректные привязки PW-Forwarder и согласовали значения демультиплексоров, можно считать создание PW завершённым. Если нужно согласовать дополнительные характеристики или параметры конкретного PW или передать информацию о состоянии определённого PW, этот псевдопровод должен указываться значением демультиплексора.

Сигнальные процедуры для псевдопроводов «точка-точка» чаще всего представляют обычные процедуры «точка-точка», применяемые двумя конечными точками PW. Однако имеются предложения использовать сигнализацию point-to-multipoint для организации псевдопроводов «точка-точка», поэтому это включено в описание модели. Когда PW сами по себе организуют соединения «один со многими» (point-to-multipoint), для их организации можно применять сигнализацию point-to-point или point-to-multipoint. Эти вопросы рассматриваются ниже.

3.2.6.1. Сигнализация «точка-точка»

Существует несколько способов сигнализации «точка-точка», включая перечисленные ниже.

- LDP

  Можно определить расширения протокола LDP [RFC3036] для сигнализации псевдопроводов. Эта форма сигнализации может применяться для псевдопроводов, организуемых в туннелях MPLS или MPLS с инкапсуляцией в другие протоколы.

- L2TP

  Протокол L2TP [RFC2661] можно применять для сигнализации псевдопроводов, организуемых как «сессии» в туннелях L2TP. Для этого могут потребоваться специальные расширения L2TP.

Возможны и другие методы сигнализации.

Одно управляющее соединение между парой PE может применяться для управления множеством PW.

Использование сигнализации «точка-точка» для организации PW «точка-точка» достаточно просто. PW типа «множество с одним» также можно организовывать с помощью сигнализации «точка-точка», поскольку удаленным PE не требуется знать, какой из PW относится к типу «точка-точка», а какой является многоточечным. В некоторых сигнальных процедурах одно значение демультиплексора может применяться для всех удалённых PE.

3.2.6.2. Многоточечная сигнализация

Рассмотрим следующие условия:

- нужно организовать множество PW с одинаковыми характеристиками;
- не требуется использовать сигнальный протокол PW для передачи изменения состояний псевдопроводов;
- для всех PW можно использовать одно значение Remote Forwarder Selector.

Назовём эти условия «условиями среды» (Environmental Conditions).

Предположим, что имеется механизм, с помощью которого каждое из множества устройств PE может делать уникальный и детерминированный выбор из заданного набора значений демультиплексора. Назовём это «условием демультиплексора» (Demultiplexor Condition). В качестве альтернативы предположим, что кто-то пытается организовать PW типа «множество с одним» вместо PW «точка-точка». Назовём это «многоточечным условием» (Multipoint Condition).

Если

- выполняются «условия среды»;
- и выполняется какое-либо из двух условий:
  - Demultiplexor Condition;
  - Multipoint Condition.

Тогда для данного набора PW, завершающихся на выходном PE1, информация, которую устройству PE1 нужно передать входным (входному) PE для каждого псевдопровода будет совпадать и все входные PE получат одно значение Forwarder Selector, а также получать одинаковые наборы параметров PW (если они имеются). Они также получат одно значение демультиплексора (для PW типа «точка-точка») или общий набор значений демультиплексора из которого каждое устройство может выбрать своё уникальное значение.

Вместо соединения с каждым входным PE и репликации одних и тех же данных имеет смысл передать эту информацию по групповому адресу или отправить её один раз «рефлектору», который распространить её другим PE.

Этот метод сигнализации называется «один со многими» (point-to-multipoint). Он может применять, например, протокол BGP [RFC1771] для устройств PE, не являющихся партнёрами BGP, но имеющими в качестве партнёра один или несколько рефлекторов маршрутов BGP [RFC2796].

3.2.6.3. Работа через несколько AS

Псевдопровода могут потребоваться между PE в сети одного SP и PE в сети другого SP. Это имеет ряд последствий, перечисленных ниже.

- Протокол сигнализации, применяемый для организации PW, должен работать через границы сетей. Такая возможность обеспечивается всеми протоколами на основе IP.
- Два PE в конечных точках PW должны иметь маршрутизируемые и доступные один для другого адреса.
- Протокол сигнализации должен обеспечивать каждой конечной точке PW возможность проверки подлинности другой стороны. Для использования аутентификации потребуется также тот или иной метод распространения ключей, приемлемый на обеих сторонах.

3.2.7. Качество обслуживания

Качество обслуживания характеризует способность сети обеспечить заданный уровень сервиса (SLS[1]) для таких атрибутов обслуживания, как защита, безопасность и QoS[2]. Качество обслуживания зависит от требований абонента и может характеризоваться множеством параметров.

Требуемое качество обслуживания должно обеспечиваться на устройствах присоединения (AC), а также на псевдопроводах (PW). Механизмы обеспечения качеств обслуживания на PW могут быть обусловлены PW или туннелем. В последнем случае привязка PW к туннелю может определяться требованиями качества сервиса.

[1]Service level Specification - спецификация уровня обслуживания.
[2]Quality of Service - качество обслуживания.

3.2.7.1. QoS

QoS описывает поведение очередей применительно к определённому «потоку» для достижения заданных целей в плане предпочтений, пропускной способности, задержки и её вариаций и т. п.

На основе соглашения об уровне обслуживания абонента (SLA[1]) трафик от абонента может приоритизироваться и формоваться для выполнения требований QoS. Правила постановки в очередь и пересылки могут сохранять порядок пакетов и параметры QoS для абонентского трафика. Классы обслуживания могут выбираться на основе информации в пользовательских кадрах или независимо от содержимого кадров.

Функции QoS перечислены ниже.

- Приоритизация абонентского трафика. Для услуг L2VPN может применяться доставка «по мере возможностей» (best effort) или гарантия QoS. Для трафика абонента может потребоваться изменение приоритета по отношению к другому трафику при распределении общих ресурсов сети. Это требует поддержки решением L2VPN классификации и маркировки приоритета для обеспечения требований абонента к QoS.

  На входном устройстве AC, а возможно и в магистральной сети, требуется обеспечить подобающее поведение очередей. Если требуется контроль очередей в магистральной сети, он должен выполняться на основе информации о классе обслуживания CoS[2] в заголовке MPLS или IP, а может также быть обеспечен на основе встраивания определённых туннелей в соответствующие туннели организации трафика.

- Правила служат для того, чтобы пользователь L2VPN применял ресурсы сети, не выходя за пределы, заданные SLA. Любой избыточный трафик L2VPN может быть отвергнут или обработан в соответствии с особыми правилами провайдера.

  Правила обычно применяются на входном устройстве AC.

- Формование (Shaping). В некоторых случаях неоднородный (случайный) трафик L2VPN может приводить к неоптимальному использованию ресурсов сети. С помощью механизмов управления очередями и пересылкой трафик можно «формовать» без нарушения порядка доставки пакетов.

  Формование трафика обычно выполняется на входном устройстве AC.

3.2.7.2. Отказоустойчивость

Отказоустойчивость описывает способность инфраструктуры L2VPN защищать потоки данных при отказах в сети с сохранением доступности сервиса.

L2VPN, как и другие типы сервиса, может сталкиваться с отказами в сети, такими как отключение каналов, транков или узлов как в инфраструктуре ядра SP, так и в устройствах AC.

Желательно обнаруживать отказы «немедленно» и позволять механизмам защиты быстро восстанавливать работу, чтобы сбои практически не отражались на сервисе L2VPN. Восстановление следует выполнять до того, как устройства CE смогут отреагировать на отказ. Важные аспекты обеспечения отказоустойчивости приведены ниже.

- Детектирование отказов узлов и каналов. Сервису L2VPN следует поддерживать механизмы незамедлительного обнаружения отказов узлов и каналов, влияющих на работу сервиса.

- Политика отказоустойчивости. Способ обработки обнаруженного отказа будет зависеть от политики восстановления в SLA, связанном со сервисом L2VPN. Соглашение может требовать незамедлительной обработки отказа, обрабатывать отказ лишь при отсутствии других критических отказов, требующих использования ресурсов защиты, или просто игнорировать отказ, если он не выходит за рамки допустимого простоя для сервиса L2VPN.

- Механизмы восстановления. Решения L2VPN могут поддерживать защиту на физическом, логическом или обоих уровнях. Например, за счёт подключения абонентов через избыточные и физически разделённые устройства AC к разным провайдерам можно сделать одно устройство AC активным, а другое - резервным с переключением на него в случае отказа основного AC.

Отказоустойчивость в значительной степени зависит от наличия соответствующих механизмов детектирования и восстановления в ядре сети, включая «обычную» адаптивную маршрутизацию, а также возможности «быстрой перемаршрутизации» Возможность поддержки избыточных AC между CE и PE также играет важную роль.

3.2.8. Управление

Решение L2VPN может включать механизмы управления и мониторинга для разных компонентов L2VPN. С точки зрения SLA решение L2VPN может разрешать мониторинг характеристик сервиса L2VPN и предоставлять механизмы, используемые SP для информирования о собранных данных. Поиск неисправностей и контроль действий при работе и обслуживании L2VPN являются важными требованиями для SP.

3.3. VPWS

VPWS - сервис L2VPN, в котором каждый модуль пересылки связывает одно устройство AC с одним PW. Кадры, полученные устройством AC, передаются в PW, а кадры, принятые из PW, передаются в AC. Содержимое заголовков L2 в кадрах не играет роли при решении о пересылке за исключением того, что содержимое заголовка L2 используется для связывания кадра с определенным AC (например, поле DLCI в кадре Frame Relay указывает AC).

Комбинация <AC, PW, AC> определяет «виртуальное устройство» (канал) между двумя устройствами CE.

Конкретная сеть VPN (экземпляр VPWS) может рассматриваться как набор таких виртуальных каналов или «наложение» (overlay) псевдопроводов PW на магистраль MPLS или IP. Это создаёт наложенную топологию, которая служит «виртуальной магистралью» отдельной сети VPN.

[1]Service Level Agreement - соглашение об уровне обслуживания.
[2]Class of Service - класс обслуживания. Прим. перев.

Вопрос принадлежности двух виртуальных устройств (каналов) к одной или разным VPN является административным и решается на основе соглашения между SP и абонентом. Решение этого вопроса может влиять на модель предоставления услуг (см. ниже), а также на связывание конкретных PW с туннелями, способ назначения QoS конкретным AC и PW и т. п.

Отметим, что VPWS использует исключительно PW «точка-точка».

3.3.1. Предоставление и автоматическое обнаружение сервиса

Предоставление услуг VPWS включает:

1. предоставление устройств (каналов) AC;
2. предоставление устройствам PE информации, позволяющей им организовать PW между устройствами AC для создания наложенной топологии;
3. настройку нужных характеристик PW.

3.3.1.1. Предоставление AC

Во многих случаях устройства AC должны предоставляться индивидуально для PE и/или CE. Это безусловно будет возникать в случаях, когда для соединения CE - PE используется коммутируемая сеть, такая как ATM или FR и в качестве VC применяются PVC[1], а не SVC[2]. Это также возникает в случаях, когда для отдельных AC требуются определённые параметры (например, QoS, гарантированная пропускная способность) различающиеся у разных устройств.

Имеются также случаи, когда специально предоставлять AC не нужно. Например, если в качестве AC используется MPLS LSP между CE и PE, соединение может быть организовано в **результате** создания PW и его не требуется организовывать **до** создания PW. То же самое применимо в случае, когда AC является коммутируемым виртуальным устройством любого типа, хоты в таких случаях могут потребоваться другие правила управления предоставлением канала (например, ограничение числа AC, которые могут быть организованы между данной парой CE - PE).

Вопросы индивидуального предоставления AC, применения коммутируемых или постоянных VC, тип управления правилами могут решаться на уровне реализации или развёртывания и выходят за рамки данной модели.

3.3.1.2. Предоставление PW для произвольной наложенной топологии

Для поддержки произвольных наложенных топологий нужно разрешить предоставление отдельных PW. В этой модели при предоставлении PW на устройстве PE, псевдопровод локально связывается с конкретным AC. Предоставляется также информация, указывающая конкретное соединение AC на удалённом PE.

Существуют два базовых варианта этой модели предоставления, рассмотренных ниже.

- Двухстороннее предоставление, когда PE на каждой стороне PW предоставляется приведённая ниже информация.
  - Идентификатор локального AC, с которым связан PW.
  - Тип и параметры PW.
  - IP-адрес удалённого PE (PE на удалённом конце PW).
  - Значимый для удалённого PE идентификатор, который может быть передан протоколом сигнализации PW, чтобы позволить удалённому PE связать PW с нужным AC. Это может быть идентификатор PW или удалённого AC. При использовании идентификатора PW он должен быть уникальным на каждом из двух PE. Если используется идентификатор AC, он должен быть уникальным на удалённом PE.

  Этот идентификатор применяется в качестве Remote Forwarder Selector по завершении сигнализации (параграф 3.2.6.1).
- Одностороннее предоставление, когда PE на одном конце PW предоставляется приведённая ниже информация.
  - Идентификатор локального AC, с которым связан PW.
  - Тип и параметры PW.
  - Уникальный в глобальном масштабе идентификатор удалённого AC.

  Этот идентификатор применяется в качестве Forwarder Selector по завершении сигнализации (параграф 3.2.6.1).

  В этой модели IP-адрес удалённого PE не предоставляется, а вместо этого будет применяться схема автоматического обнаружения для отображения уникального в глобальном масштабе идентификатора на IP-адрес удалённого PE вместе с идентификатором (возможно уникальным лишь в этом PE) для AC на этом PE. Затем сигнальный протокол PW организует соединение с удаленным PE, передавая идентификатор AC, чтобы удалённый маршрутизатор PE связал PW с нужным AC.

  Эта схема требует предоставления PW лишь на одном PE, но не избавляет от необходимости (если она имеется) предоставления AC на обоих PE.

Эти модели предоставления хорошо подходят для сигнализации «точка-точка». Когда каждый псевдопровод PW предоставляется индивидуально, не возникает необходимости применять сигнализацию point-to-multipoint.

3.3.1.3. Модель предоставления PW с цветными группами

[1]Permanent Virtual Circuit - постоянное виртуальное устройство (канал). *Прим. перев.*
[2]Switched Virtual Circuit - коммутируемое виртуальное устройство (канал). *Прим. перев.*

Предположим, что каждый маршрутизатор PE собирает AC в группу (pool) и с каждой группой связывает определённый цвет (например, группа может содержать все AC от данного PE к отдельному CE и не включать других соединений). Далее предположим, что вводится правило, в соответствии с которым при наличии у PE1 и PE2 групп одного цвета между PE1 и PE2 будет псевдопровод PW, привязанный на PE1 и PE2 к произвольно выбранному AC из данной группы (не исключается случай наличия в PE нескольких групп одного цвета).

Например, каждая группа в определённом PE может представлять определённое устройство CE, для которого AC из этой группы являются AC, соединяющими данный CE с данным PE. В качестве «цвета» может использоваться VPN-id. Применение этой модели предоставления приведёт к полносвязным соединениям между всеми CE в сети VPN, где каждое устройство CE в сети VPN имеет виртуальный канал (устройство) к каждому другому CE в данной VPN.

Более конкретно, для предоставления VPWS в соответствии с этой моделью нужно обеспечить набор групп и настроить для каждой группы приведённые ниже параметры.

- Набор AC, относящихся к группе (ни одно устройство AC не входит в несколько групп).
- «Цвет».
- Идентификатор группы, уникальный по меньшей мере в части «цвета».

  Затем используется процедура автоматического обнаружения с целью отобразить каждый цвет на список упорядоченных пар <IP-адрес PE, pool id>. Наличие пары <X, Y> в таком списке означает, что PE с IP-адресом X имеет группу заданного цвета с идентификатором (pool id) Y.

Эта информация может применяться для поддержки нескольких различных методов сигнализации, один из которых описан ниже.

- PE определяет наличие у себя группы с цветом C.
- С помощью автоматического обнаружения PE получает набор упорядоченных пар <X,Y> для цвета C.
- Для каждой такой пары <X,Y> устройство PE:
  - удаляет AC из группы;
  - привязывает AC к определённому PW;
  - сообщает PE X с помощью сигнализации «точка-точка» о привязке PW к AC из группы Y.

Другой возможный метод сигнализации описан ниже.

- PE определяет наличие у себя группы с цветом C, содержащей n каналов AC.
- PE связывает каждый канал AC с PW, создавая набор PW, который организуется в форме последовательности (например, с каждым PW может быть связано значение демультиплексора и PW упорядочиваются по числовым значениям полей демультиплексора).
- С помощью автоматического обнаружения PE находит список маршрутизаторов PE, имеющих одну или несколько групп с цветом C.
- PE сообщает каждому такому маршрутизатору PE последовательность Q из псевдопроводов PW.
- Если PE получает такой сигнал и имеет группу Y с заданным цветом, этот PE:
  - удаляет AC из группы;
  - привязывает AC к определённому PW, являющемуся Y-м в последовательности Q.

При этом предполагается, идентификаторы групп отображаются или могут быть отображены однозначно на небольшие упорядоченные числа, поэтому назначение идентификаторов групп становится требованием для системы обеспечения.

Отметим, что по причине передачи некой информации всем удаленным PE этот метод может поддерживаться с помощью сигнализации point-to-multipoint.

Эта модель предоставления применима при выполнении всех перечисленных ниже условий.

- Не требуется обеспечение различных характеристик для разных PW.
- Не важно, какая пара устройств AC связывается PW, если оба AC в паре относятся группам одного цвета.
- Можно создавать желаемую наложенную топологию, просто назначая цвета для групп (это просто для полносвязных соединений или конфигурации с концентратором и лучами, а создание произвольных топологий сложнее и возможно не во всех случаях).

3.3.2. Требования к процедурам автоматического обнаружения

Некоторые из требований к автоматическому обнаружению могут быть выведены из написанного выше.

Для поддержки модели одностороннего предоставления автоматическое обнаружение должно быть способно отобразить уникальный в глобальном масштабе идентификатор (PW или устройства AC) на IP-адрес PE.

Для поддержки модели с цветными группами автоматическое обнаружение должно позволять PE определить набор других PE, содержащих группы с таким же цветом.

Эти требования позволяют схеме автоматического обнаружения предоставить информацию, которая нужна устройствам PE для организации псевдопроводов PW.

Ниже перечислен ряд требований к автоматическому обнаружению, которые не могут быть просто выведены из модели предоставления.

- Конкретным схемам сигнализации может требоваться дополнительная информация перед началом работы и этом вносит дополнительные требования к процедурам автоматического обнаружения.
- Данный SP может поддерживать несколько разных типов сигнальных процедур и устройствам PE может потребоваться определение используемых процедур с помощью автоматического обнаружения.
- Изменения конфигурации PE должны своевременно отражаться процедурами автоматического обнаружения без необходимости явного изменения конфигурации других PE.
- Процедуры автоматической настройки должны работать через границы SP. Это исключает, например, применение схем, объединяющих данные автоматического обнаружения на магистральных IGP.

3.3.3. Разнородные PW

В некоторых случаях могут быть желательны PW, связывающие пару AC, которые используют разные технологии (например, ATM и Ethernet). Имеется много способов решения этой задачи в зависимости от типа AC.

- Если одно устройство AC использует ATM, а другое - FR, можно применить стандартное решении ATM/FR Network Interworking. В этом случае для сигнализации PW может применяться ATM, а межсетевое взаимодействие происходить между PW и FR AC.
- Может применяться общая для обоих AC инкапсуляция. Например, если одно устройство AC использует Ethernet, а другое FR, на втором устройстве можно использовать инкапсуляцию Ethernet over FR. В этом случае для сигнализации PW используется Ethernet с локальной обработкой инкапсуляции Ethernet over FR на PE с FR AC.
- Если известно, что два AC подключены к маршрутизаторам или хостам IP и передают лишь трафик IP, можно применять PW, передающий пакеты IP и выбор инкапсуляции L2 будет локальной задачей для обоих PE. Однако, если одно из устройств AC является ЛВС, а другое - каналом «точка-точка», нужно будет обеспечить корректную работу таких процедур, как ARP и Inverse ARP, а для этого может потребоваться та или иная сигнализация и функции прокси. Кроме того, при использовании в CE алгоритма маршрутизации с разными процедурами для ЛВС и каналов «точка-точка» могут потребоваться дополнительные механизмы межсетевого взаимодействия.

3.4. Эмулируемые ЛВС в VPLS

VPLS - это сервис L2VPN, в котором:

- AC соединяют устройства CE с модулями мостов в PE;
- каждый модуль моста в PE подключается через «интерфейс эмулируемой ЛВС» к «эмулируемой ЛВС».

Логическая модель VPLS представлена на рисунке 3.

Ниже приведена функциональная декомпозиция эмулируемой ЛВС сервиса VPLS. Устройствами AC эмулируемой ЛВС являются «интерфейсы эмулируемой ЛВС», соединяющие модуль моста в PE с модулем пересылки (VPLS Forwarder), как показано на рисунке 3. Данными (payload) AC являются кадры Ethernet с тегами VLAN или без них.

Данный модуль VPLS Forwarder с данном PE будет иметь множество AC лишь в том случае, когда в PE имеется множество модулей мостов, подключённых к модулю пересылки. Этот вариант рассматривается в данной схеме, хотя рассмотрение его реальной применимости выходит за рамки документа.

Множество модулей VPLS Forwarder в одном экземпляре VPLS соединяется через PW. Два VPLS Forwarder будут соединены PW лишь в том случае, когда эти модули относятся к одному экземпляру VPLS (могут быть и другие ограничения, например, наличие PW между парой VPLS Forwarder может быть обусловлено их принадлежностью к одной VLAN внутри одной сети VPN). Набор соединённых между собой модулей VPLS Forwarder образует эмулируемую ЛВС (VPLS Emulated LAN).

В реальной ЛВС любой кадр, переданный одним элементом, получают все остальные. Однако в VPLS Emulated LAN ситуация иная. Когда модуль VPLS Forwarder принимает индивидуальный кадр через один из своих интерфейсов Emulated LAN, он не обязан пересылать этот кадр всем другим модулям Forwarder в данной эмулируемой ЛВС. Индивидуальный кадр нужно переслать лишь одному модулю Forwarder, чтобы тот мог доставить его по MAC-адресу получателя. Если передающий модуль Forwarder знает какому модулю пересылки нужно отправить конкретный индивидуальный кадр, он будет пересылать кадр лишь этому модулю Forwarder. Такая оптимизация пересылки является важной частью любой попытки обеспечить сервис VPLS через распределенную или городскую сеть.

По сути, каждый модуль пересылки ведёт себя как экземпляр виртуального коммутатора (VSI[1]), поддерживающий таблицу пересылки с отображением MAC-адресов на PW. Таблица пересылки VSI заполняется почти так же, как обычный мост заполняет таблицу пересылки. VPLS Forwarder изучает MAC-адреса отправителей (SA[2]) в кадрах, полученных на PW от модулей пересылки, а также должен выполнять набор процедур, таких как контроль старения адресных записей. Кадры с неизвестными или групповыми DA[3], должны широковещательно рассылаться одним модулем пересылки всем другим модулям Forwarder той же эмулируемой ЛВС. Однако имеется ряд существенных различий между VPLS Forwarder VSI и функцией пересылки стандартного моста, отмеченных ниже.

- VPLS Forwarder никогда не изучает MAC SA в кадрах, полученных на его AC, изучая лишь MAC SA а кадрах, полученных на PW от других модулей VPLS.
- Модули VPLS Forwarder конкретной эмулируемой ЛВС не участвуют в протоколе STP[4] с другими модулями пересылки. Для предотвращения петель при пересылке применяется метод «расщепления горизонта.

Эти различия более подробно рассматриваются ниже.

[1]Virtual Switch Instance.
[2]Source Address - адрес отправителя.
[3]Destination Address - адрес получателя.
[4]Spanning tree protocol - протокол связующего дерева.

Отметим, что модули моста PE в одной эмулируемой ЛВС могут (но не обязаны) участвовать в протоколе STP эмулируемой ЛВС (вопрос участи в протоколе выходит за рамки спецификации VPLS). Модули моста PE будут изучать MAC-адреса на AC, а также на интерфейсах Emulated LAN, но не будут изучать MAC-адреса на PW, поскольку PW «спрятаны» за интерфейсом Emulated LAN. Поэтому таблицы пересылки модуля моста PE и VSI модуля VPLS Forwarder будут разными (конкретная реализация может объединять эти таблицы, но этот вопрос выходит за рамки документа).

Другой вопрос связан с использованием мостами PE протоколов управления мостами эмулируемой ЛВС. Обычно протоколы управления мостами предназначены для работы в реальных ЛВС и для их работы могут быть важны некоторые характеристики ЛВС, такие как малые задержки. Если эмулируемая ЛВС не обеспечивает нужных характеристик, протоколы управления могут не выполнять некоторых функций, пока не будут применяться специальные механизмы для доставки кадров управления.

Следует отметить, что в результате изменения связующего дерева в сети абонента (если оно есть) или на мостах PE (если оно есть), некоторые MAC-адреса могут менять своё местоположение, переходя с одного PE на другой. Эти изменения могут быть не видимы для модулей VPLS Forwarder, что может привести к недоступности таких MAC-адресов, пока они не устареют в VSI прежнего PE. Если такое поведение не приемлемо, потребуется тот или иной механизм оповещения модулей VPLS Forwarder о произошедших изменениях.

3.4.1. Топология и пересылка наложенного сервиса VPLS

Внутри одной сети VPLS модули VPLS Forwarder соединяются псевдопроводами PW, формирующими «наложенную топологию».

Таблицы VSI в модулях VPLS Forwarder заполняются при изучении MAC-адресов, т. е. VSI хранит адреса MAC SA с привязкой к PW, через который адрес был получен. Если определённый MAC-адрес присутствовал в качестве SA в кадре, принятом через определённый PW, кадры с таким MAC в поле DA, следует передавать в таблицу VSI, расположенную на удалённом конце этого PW. Для выполнения этого на удаленной стороне PW требуется наличие уникального VSI, что означает невозможность соединения VSI с помощью многоточечных (multipoint-to-point) PW. PW должны быть типа «точка-точка» или, возможно, - point-to-multipoint.

Изучение MAC от PW «точка-точка» выполняется стандартными методами IEEE, где PW трактуется модулем VPLS Forwarder как «порт моста». Естественно, при изучении MAC-адресов из point-to-multipoint PW, экземпляр VSI должен указывать, что пакеты для этого адреса будут передаваться в PW «точка-точка» к корню point-to-multipoint PW.

Решения VSI о пересылке должны координироваться для обеспечения беспетлевой пересылки в наложенной топологии.

Имеется несколько возможных типов наложенной топологии.

- Полносвязные соединения (Full mesh).

  При полносвязных соединениях каждый экземпляр VSI в данной сети VPLS имеет в точности один PW «точка-точка» к каждому другому VSI в той же VPLS.

  В такой топологии пересылка без петель обеспечивается простыми правилами.

  Если VSI получает через PW кадр от другого VSI, **недопустимо** пересылать этот кадр **любому** другому PW для любого другого VSI. Это гарантирует пересылку кадра за пределы эмулируемой ЛВС после прохождения через неё.

  Если VSI получает на одном из интерфейсов Emulated LAN индивидуальный кадр с известным DA, кадр передаётся. в точности одному PW.

  Если VSI получает на одном из интерфейсов Emulated LAN групповой кадр с неизвестным DA, копия этого кадра передаётся. каждому VSI в той же эмулируемой ЛВС. Это может быть реализовано путём репликации кадра и передачи копии через каждый PW «точка-точка». Псевдопровода «точка-точка» могут быть дополнены в полносвязной сети псевдопроводами point-to-multipoint PW, где каждый экземпляр VSI в VPLS служит передатчиком в один псевдопровод point-to-multipoint, а получателями на этом PW являются все другие VSI в этой сети VPLS.

- Структурированное дерево (Tree structured)

  В топологии структурированного дерева каждый виртуальный коммутатор VSI в конкретной сети VPLS предоставляется на определённом уровне дерева и имеет не более одного псевдопровода на вышележащий уровень. Верхним уровнем является корень дерева.

  В этой топологии отсутствие петель обеспечивается простым правилом - если кадр получен из псевдопровода от вышележащего уровня, он не может быть передан через псевдопровод, ведущих вверх.

- Дерево с полной связностью на верхнем уровне (Tree with Meshed Highest Level)

  В этом варианте топологии со структурированным деревом на верхнем уровне может быть более одного коммутатора VSI и все VSI верхнего уровня должны образовывать полносвязную сеть. Для предотвращения петель нужно внести правило пересылки кадра в тот же или вышележащий уровень лишь в том случае, когда он получен из псевдопровода от нижележащего уровня. Кадры, принятые из PW того же уровня, не могут пересылаться в PW этого уровня.

Возможны и другие наложенные топологии, например частичная полносвязность PW между VSI сети VPLS. Для предотвращения петель может применяться, например, протокол STP для наложенной топологии. Эти варианты топологии далее не рассматриваются в модели.

Отметим, что отсутствие петель в наложенной топологии не гарантирует их отсутствия в ЛВС, содержащей VPLS. Не гарантируется даже отсутствие петель между модулями мостов PE. Обеспечивается лишь гарантия того, что кадр, переданный в эмулируемую ЛВС не окажется в петле пересылки до того, как он покинет Emulated LAN.

Некорректная конфигурация абонентской ЛВС или модулей мостов PE может создавать петли и попавшие в такую петлю кадры могут многократно проходить через наложенную топологию. Можно рекомендовать процедуры и для обнаружения и/или предотвращения таких петель.

3.4.2. Предоставление и автоматическое обнаружение сервиса

Каждой сети VPLS должен быть назначен уникальный в глобальном масштабе идентификатор VPN-id.

Устройства AC, соединяющие CE с PE, должны предоставляться как на PE, так и на CE. VSI для сети VPLS должны предоставляться на PE и локальные AC этой сети VPLS должны связываться с соответствующими VSI. Для VSI должны предоставляться идентификаторы VPLS, к которым относятся виртуальные коммутаторы.

Схема автоматического обнаружения может применяться устройствами PE для отображения идентификатора VPLS на множество удалённых PE, имеющих виртуальные коммутаторы VSI в данной сети VPLS. После этого PE может использовать сигнализацию псевдопроводов для организации PW к каждому из этих коммутаторов VSI. Идентификатор VPLS будет служить в качестве Forwarder Selector для протокола сигнализации. Это приведёт к организации полной связности PW между VSI в отдельной сети VPLS.

Если одна сеть VPLS содержит множество VLAN, может оказаться желательным связывание между собой лишь VSI, относящихся к одной виртуальной сети VLAN.

В этом случае для каждого виртуального коммутатора VSI будет требоваться один или несколько идентификаторов VLAN и схема автоматического обнаружения должна будет отобразить идентификатор VPLS на пары <PE, VLAN id>.

Если полносвязная топология VSI нежелательна, для каждого коммутатора VSI нужно предоставить дополнительную информацию, задающую место коммутатора в топологии. Эту информацию также должна предоставлять схема автоматического обнаружения.

Другим вариантом является метод одностороннего предоставления, описанный в параграфе 3.3.1.2. Поскольку этот метод более сложен, его применяют лишь при необходимости связать с конкретными PW определённые характеристики. Например, если для разных пар сайтов в VPLS нужна разная пропускная способность, PW следует предоставлять индивидуально.

3.4.3. Распределенное устройство PE

При организации сервиса VPLS устройства CE зачастую подключаются к управляемым провайдером устройствам CPE. Одно устройство CPE может применяться для подключения CE нескольких абонентов, особенно в случаях размещения множества абонентов в одном здании. Однако эти устройства реально являются частью сети SP, поэтому они могут рассматриваться как устройства PE.

В некоторых вариантах развёртывания VPLS устройства CE подключаются к управляемому провайдером промежуточному устройству. Это устройство также может подключать CE разных абонентов. Такая ситуация часто возникает при размещении множества абонентов в одном здании. Промежуточное устройство является частью сети SP и поэтому может считаться устройством PE, однако в простейшем случае оно лишь выполняет функцию агрегирования, а не функции, связанные с VPLS.

Применительно к функциям VPLS устройства с таким положением в сети провайдера могут играть 2 разных роли.

- Агрегирование и поддержка только услуг L2, без выполнения функций устройства PE для сервиса VPLS. В этом случае промежуточная система должна подключаться к устройствам, выполняющим функции VPLS PE, а сама не будет частью архитектуры VPLS.
- Выполнение функций PE, относящихся к VPLS. В этом случае устройство называется VPLS-PE [RFC4026]. Такие устройства соединяются с маршрутизаторами ядра (P).

Функциональность PE для VPLS может быть распределена между двумя устройствами, одно из которых (low-end) размещается ближе к абонентам и выполняет, например, функции изучения MAC-адресов и принимает решение о пересылке, а другое (high-end) выполняет функции управления, например, организацию туннелей, PW и VC. Первый тип устройств называют U-PE, а второй - N-PE.

Возможно, что U-PE может быть размещено очень близко к клиенту, например в здании с несколькими абонентами. Предполагается, что N-PE будет размещаться на площадке SP.

Распределенный вариант PE интересен по нескольким причинам, рассмотренным ниже.

- N-PE может быть устройством, на котором сложено реализовать функции VSI, описанные выше. Например, N-PE может быть маршрутизатором, не способным к быстрому изучению MAC-адресов, которое требуется для реализации модуля пересылки VSI. В то же время, U-PE может быть недорогим устройством, которое не может реализовать все функции VPLS.

  Это стимулирует дополнительные исследования эффективного распределения функциональности VPLS PE между U-PE и N-PE.

- Обычно в архитектуре L2VPN предполагается, что устройства PE выступают партнёрами в протоколе маршрутизации магистральной сети. Поскольку число устройств U-PE может быть очень велико, по сравнению с числом N-PE, разделение функций может быть оправдано с точки зрения расширяемости маршрутизации.
- U-PE может быть сравнительно недорогим устройством, которое не способно выполнить все процедуры сигнализации и автоматического обнаружения, которые требуются для сервиса VPLS.

Функциональность VPLS можно распределить между U-PE и N-PE разными способами и для этого предложено множество решений. Все они предполагают, что U-PE будет поддерживать модуль пересылки (VSI forwarder), подключенный с помощью PW к удаленным VSI и N-PE не потребуется выполнять функции пересылки VSI. Предложения обычно различаются по перечисленным ниже вопросам.

- Поддерживать в U-PE все сигнальные функции для организации PW к удаленным VSI или отдать это N-PE?

Поскольку U-PE требуется передавать пакеты в PW для удалённых VSI и получать пакеты из PW от удалённых VSI, при поддержке сигнализации PW на устройствах N-PE потребуется тот или иной (облегченный) вариант сигнализации между N-PE и U-PE, позволяющий «продлить» PW от N-PE до U-PE.

- Поддерживать в U-PE функции автоматического обнаружения или отдать это N-PE?

  При отказе от автоматического обнаружения устройствам U-PE может потребоваться тот или иной способ уведомить N-PE об интересующих VPLS, а N-PE должны будут поддерживать способ передачи результатов автоматического обнаружения в U-PE.

  Выбор точки реализации автоматического обнаружения может зависеть от конкретного протокола обнаружения. Например, не следует ожидать участия U-PE в автоматическом обнаружении на основе BGP, но вполне возможно их участие в автоматическом обнаружении на основе RADIUS.

- Если устройство U-PE не участвует в резервировании маршрутизации, но подключено к двум разным N-PE, может ли U-PE выполнять интеллектуальный выбор лучшего N-PE в качестве next hop для трафика, адресованного определённому VSI? Если нет, можно ли сделать этот выбор на основе того или иного взаимодействия между N-PE и U-PE или этот выбор должен быть предоставлен?

- Если U-PE не участвует в маршрутизации, но полностью участвует в сигнализации PW и применяется MPLS, как N-PE может передать U-PE метки, которые нужны U-PE для передачи трафика своим партнёрам по сигнализации (если U-PE участвует в маршрутизации, это происходит автоматически)?

- Выполнять репликацию групповых кадров в N-PE или U-PE?

Эти вопросы взаимосвязаны и ответ на один из них может влиять на другие.

3.4.4. Проблемы расширяемости VPLS

В общем случае PSN поддерживает решение VPLS с туннелями между каждой парой устройств, участвующих в одном экземпляре VPLS. Строго говоря, устройствам VPLS-PE, участвующим в нескольких экземплярах VPLS, в общем случае нужен лишь один туннель, но в целях выделения ресурсов может потребоваться организация нескольких туннелей. Для каждого экземпляра VPLS на данном VPLS-PE требуется организовать псевдопровод к каждому другому VPLS-PE этого экземпляра VPLS. В результате между n маршрутизаторами VPLS-PE необходимо организовать n*(n-1) псевдопроводов. В больших системах это обычно ограничивает возможности расширения. Одним из способов решения проблемы является использование иерархии.

3.5. IPLS

Если вместо VPLS общего назначения достаточно создать сеть VPLS, служащую лишь для соединения маршрутизаторов и хостов IP (т. е. все устройства CE являются маршрутизаторами или хостами IP), это позволяет воспользоваться некоторыми упрощениями.

В такой среде все кадры Ethernet, переданные от данного CE конкретному устройству PE через данный канал AC, будут иметь один MAC-адрес отправителя. Поэтому вместо изучения MAC-адресов на уровне данных PE может использовать для изучения MAC-адресов уровень управления. Это позволяет реализовать PE на устройствах, не способных выполнять изучение MAC-адресов на уровне данных.

Для избавления от необходимости изучать MAC-адреса на PW и AC сигнальный протокол псевдопроводов будет переносить MAC-адреса от одного PW к другому. В случае IPv4 каждое устройство PE будет служить proxy ARP для подключённых напрямую устройств CE. В случае IPv6 каждое устройство PE будет передавать proxy Neighbor и/или анонсы маршрутизаторов (Router Advertisement).

Отказ от изучения MAC-адресов на PW избавляет от необходимости использовать PW типа «точка-точка» и позволяет применять вместо них многоточечные (multipoint-to-point) PW.

В отличие от VPLS все AC в IPLS не обязаны передавать кадры Ethernet и требуется лишь передача через сеть пакетов IP, а не их инкапсуляции в L2. Однако при наличии связанных с L2 протоколов, которые предоставляют услуги для L3 (например, преобразование адресов), может потребоваться «трансляция» или иное преобразование одних протоколов L2 в другие. Например, если экземпляр IPLS имеет AC типов Ethernet и Frame Relay, на которых работает IPv4, может потребоваться взаимодействие между ARP и Inverse ARP.

Набор протоколов маршрутизации, которые могут передаваться через IPLS, также может быть ограничен.

Экземпляр IPLS должен иметь своё значение MTU и при наличии разных AC с разными MTU для всех должно применяться одно значение. Если AC не может менять MTU, ему не будет разрешено входить в экземпляр IPLS.

4. Вопросы безопасности

Раздел вопросов безопасности в документе с требованиями к L2VPN [RFC4665] рассматривает множество потенциально небезопасных аспектов L2VPN. Это связано с вопросами безопасности на уровнях данных и управления, которые могут возникать в разных местах:

- сеть провайдера;
- сеть абонента;
- интерфейс между сетями провайдера и абонента.

Эти три области рассматриваются ниже.

4.1. Вопросы безопасности сети провайдера

В этом параграфе рассматриваются вопросы безопасности, связанные исключительно с оборудованием SP.

Ряд проблем безопасности связан с управляющими соединениями между устройствами PE, служащими для создания и поддержки псевдопроводов.

Устройству PE не следует взаимодействовать через управляющее соединение с другим PE, пока у него нет уверенности в том, что это именно то устройство PE, с которым следует организовать PW. В противном случае трафик L2PVN может попасть туда, куда не следует. Если пакеты управления злонамеренно и незаметно подменяются в процессе передачи, это может приводить отказам в обслуживании или снижению качества сервиса.

Если применяется динамическое обнаружение партнёров (с помощью той или иной процедуры), процедура обнаружения должна обеспечивать достаточный уровень доверия.

Устройствам PE не следует воспринимать соединения от произвольных элементов. На PE следует указывать партнёров в конфигурации или находить их с помощью доверенной процедуры автоматического обнаружения. Если партнёр должен находиться в сети того же SP, следует применять списки контроля доступа на границах этой сети для предотвращения подмены адресов отправителей у партнёров. Если партнёр находится в сети другого SP, установка таких фильтров осложняется, а в некоторых случаях становится невозможной (в зависимости от способа соединения между двумя SP). Даже при установке фильтров они не смогут обеспечить полной гарантии.

Поэтому для управляющих соединений между SP рекомендуется применять ту или иную криптографическую процедуру проверки подлинности. Протоколы управления на базе TCP могут использовать опцию TCP MD5 для взаимной аутентификации устройств PE, для чего потребуется хотя один общий для SP секрет. Возможно использовать между PE защиту IPsec, обеспечивающую надёжные гарантии хотя и за счёт роста стоимости.

Любые средства защиты, пригодные для протокола управления, обычно подойдут и для организации PW. Если протокол управления использует сообщения UDP, может иметь смысл применение той или иной защиты от обманных сообщений, которые представляются сообщениями от легитимного партнёра, но этот вопрос требует изучения.

Для снижения влияния DoS-атак[1] на PE могут применяться меры ограничения скорости обработки трафика управления.

В отличие от целостности и подлинности, конфиденциальность сигнальных сообщений обычно не считается важной. При необходимости можно защитить конфиденциальность управления с помощью IPsec.

Если PE не может обслуживать большие объёмы группового трафика в установившемся состоянии, это открывает возможность для атак на отказ служб VPLS путём передачи PE большого числа кадров с групповыми адресами получателей в поле MAC DA. Похожие атаки можно организовать путём передачи PE большого числа кадров с обманными адресами MAC SA. Обманные адреса могут заполнить таблицы MAC в PE и в результате кадры для реальных MAC-адресов будут рассылаться в лавинном режиме (т. е. по групповым адресам). Отметим, что такая лавинная рассылка может нарушать конфиденциальной этой или иной сети на базе мостов.

4.2. Вопросы безопасности на границе сетей

Множество проблем безопасности связано с сетями доступа между провайдером и абонентом. Обычно в таких сетях сложно обеспечить защиту на физическом уровне.

Типичные проблемы на стыке между сетями абонента и провайдера включают:

- корректность настройки абонентского интерфейса;
- предотвращение несанкционированного доступа к PE;
- предотвращение несанкционированного доступа к конкретному порту PE;
- корректность полей разграничения сервиса (VLAN, DLCI и т. п.).

Поскольку сетями доступа для сервиса L2VPN являются сети L2, предположительно использование механизмов проверки подлинности, не предполагающих возможностей IP на устройствах CE.

Имеются протоколы L2 и позитивный опыт их применения для защиты сетей. Например, протокол IEEE 802.1x определяет аутентификацию на уровне канала доступа к мосту Ethernet, расширение Frame Relay Forum (FRF.17) определяет расширения LMI для проверки подлинности.

4.3. Вопросы безопасности сети абонента

Даже если все устройства CE должным образом уполномочены подключаться к их PE, ошибки в конфигурации PE могут привести к подключению устройств CE, которые не могут работать в одной сети L2VPN.

В VPWS на устройствах CE можно применять IPsec для взаимной аутентификации. Другие протоколы L3 или L4 могут иметь свои методы проверки подлинности.

В VPLS применение IPsec между устройствами CE более проблематично, поскольку IPsec не поддерживает многоточечных конфигураций, обеспечиваемых сервисом VPLS.

Имеется много других методов для взаимной проверки подлинности устройств CE, если сигнальный протокол может передавать между ними неанализируемые (opaque) объекты в основной полосе L2VPN или по отдельному каналу сигнализации. Этот вопрос требует дополнительного изучения.

Процедуры L2VPN не обеспечивают аутентификации, защиты целостности и конфиденциальности абонентского трафика. Если такая защита нужна, ответственность за неё. ложится на абонента. Для абонентов, которым реально требуются такие услуги или у которых нет доверия к сервис-провайдеру в части обеспечения защиты рассмотренная здесь схема L2VPN может не подойти. Такие абоненты могут рассмотреть другие схемы L2VPN, основанные не на наложенных псевдопроводах, а на наложенных туннелях IPsec, конечные точки которых размещаются на сайтах абонента. Однако эти вопросы выходят за рамки документа.

[1] Denial of Service - отказ в обслуживании.

При наличии управляющего трафика CE-CE (например, BPDU) от целостности которого зависит работа абонентской сети L2, может оказаться целесообразной передача трафика управления с использованием более защищённого механизма, нежели применяется для трафика данных.

В общем случае все способы организации атак на службы в сетях на базе мостов применимы и для атак на сервис VPLS конкретного абонента. Здесь рассмотрены лишь атаки, основанные на специфике сервиса VPLS, которые не применимы к сетям на основе мостов в целом.

5. Благодарности

Этот документ является результатом дискуссий в команде по организации L2 VPN, все члены которой могут считаться соавторами документа. Особенно следует отметить Loa Andersson, Waldemar Augustyn, Marty Borden, Hamid Ould-Brahim, Juha Heinanen, Kireeti Kompella, Vach Kompella, Marc Lasserre, Pascal Menezes, Vasile Radoaca, Eric Rosen и Tissa Senevirathne.

Авторы благодарны Marco Carugi за поддержку в организации контекста, направлений работы и время, потраченное на обсуждения, Tove Madsen и Pekka Savola за ценный вклад и отзывы, а также Norm Finn, Matt Squires и Ali Sajassi за полезное обсуждение вопросов VPLS.

6. Нормативные документы

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

[RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.

[RFC4026] Andersson, L. and T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", RFC 4026, March 2005.

[RFC4665] Augustyn, W., Ed. and Y. Serbest, Ed., "Service Requirements for Layer 2 Provider-Provisioned Virtual Private Networks (L2VPNs)", RFC 4665, September 2006.

7. Дополнительная литература

[IEEE8021D] IEEE 802.1D-2003, "IEEE Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges"

[IEEE8021Q] IEEE 802.1Q-1998, "IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks"

[RFC1771] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, March 1995[1].

[RFC2661] Townsley, W., Valencia, A., Rubens, A., Pall, G., Zorn, G., and B. Palter, "Layer Two Tunneling Protocol "L2TP"", RFC 2661, August 1999.

[RFC2796] Bates, T., Chandra, R., and E. Chen, "BGP Route Reflection - An Alternative to Full Mesh IBGP", RFC 2796, April 2000[2].

[RFC3036] Andersson, L., Doolan, P., Feldman, N., Fredette, A., and B. Thomas, "LDP Specification", RFC 3036[3], January 2001.

Адреса авторов

Loa Andersson
Acreo AB
EMail: loa@pi.se

Eric C. Rosen
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, MA 01719
EMail: erosen@cisco.com

Waldemar Augustyn
EMail: waldemar@wdmsys.com

Marty Borden
EMail: mborden@acm.org

Juha Heinanen
Song Networks, Inc.
Hallituskatu 16
33200 Tampere, Finland
EMail: jh@song.fi

Kireeti Kompella
Juniper Networks, Inc.
1194 N. Mathilda Ave
Sunnyvale, CA 94089
EMail: kireeti@juniper.net

Vach Kompella
TiMetra Networks
274 Ferguson Dr.
Mountain View, CA 94043
EMail: vach.kompella@alcatel.com

Marc Lasserre
Riverstone Networks
5200 Great America Pkwy
Santa Clara, CA 95054
EMail: mlasserre@lucent.com

Pascal Menezies
EMail: pascalm1@yahoo.com

Hamid Ould-Brahim
Nortel Networks
P O Box 3511 Station C
Ottawa, ON K1Y 4H7, Canada
EMail: hbrahim@nortelnetworks.com

Vasile Radoaca
Nortel Networks
600 Technology Park
Billerica, MA 01821
EMail: radoaca@hotmail.com

Tissa Senevirathne

[1]Ссылка ошибочна, поскольку RFC 1771 был отменен RFC 4271, см. https://www.rfc-editor.org/errata/eid902. *Прим. перев.*
[2]Ссылка ошибочна, поскольку RFC 2796 был отменен RFC 4456, см. https://www.rfc-editor.org/errata/eid902. *Прим. перев.*
[3]Документ заменён RFC 5036. *Прим. перев.*

1567 Belleville Way
Sunnyvale CA 94087

EMail: tsenevir@hotmail.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

Интеллектуальная собственность

Подтверждение

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