Network Working Group
Request for Comments: 3985
Category: Informational

S. Bryant, Ed.
Cisco Systems
P. Pate, Ed.
Overture Networks, Inc.

March 2005

Архитектура сквозной эмуляции псевдопровода (PWE3)

Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture

Статус документа

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

Аннотация

Этот документ описывает архитектуру сквозной эмуляции псевдопровода (PWE3[1]). Обсуждается эмуляция таких служб, как Frame Relay, ATM, Ethernet, TDM, SONET/SDH в сетях с коммутацией пакетов (PSN), использующих IP или MPLS. Документ представляет архитектурную схему псевдопроводов (PW), определяет терминологию, задаёт различные протокольные элементы и их функции.

Оглавление

[1]Pseudo Wire Emulation Edge-to-Edge.

1. Введение

В этом документе описана архитектура сквозной эмуляции псевдопровода (PWE3[1]) в соответствии с [RFC3916]. В документе рассматривается эмуляция таких служб, как Frame Relay, ATM, Ethernet, TDM и SONET/SDH в сетях с коммутацией пакетов (PSN[2]), использующих IP или MPLS. Документ представляет архитектурную схему псевдопроводов (PW[3]), определяет используемые термины, а также описывает различные протокольные элементы и их функции.

1.1. Определение псевдопровода

PWE3 представляет собой механизм, который эмулирует существенные атрибуты телекоммуникационного сервиса (такого, как выделенные линии T1 или Frame Relay) в сетях PSN. Задачей PWE3 является лишь обеспечение минимальной требуемой функциональности для эмуляции провода с требуемой степенью достоверности для данного типа сервиса. За реализацию всех функций коммутации отвечает механизм пересылки (FWRD). Все преобразования и другие операции, требующие знания семантики передаваемой информации, осуществляются элементами естественного сервиса (NSP[4]). Функциональные определения каких-либо элементов FWRD или NSP выходят за пределы PWE3.

В число обязательных функций PW входит инкапсуляция связанных с сервисом битовых потоков, ячеек или PDU, прибывающих на входной (ingress) порт или проходящих через туннель IP или MPLS. В некоторых случаях требуется выполнение иных операций типа управления синхронизацией и порядком следования для эмуляции поведения и характеристик эмулируемого сервиса с требуемым уровнем достоверности.

С точки зрения оконечного пользовательского оборудования (CE[5]) PW представляется как выделенный (unshared) канал или устройство соответствующей службы. В некоторых случаях эмуляции PW присущи недостатки, оказывающие влияние на проходящий через PW трафик и, следовательно, ограничивающие применимость этой технологии. Такие ограничения должны быть полностью описаны в соответствующей документации для сервиса.

Для каждого типа сервиса имеется один принятый по умолчанию режим работы, который должны поддерживать все PE, предлагающие такой сервис. Однако для повышения уровня достоверности эмуляции могут поддерживаться дополнительные режимы, если дополнительные сложности, связанные с этими режимами, компенсируются дополнительными возможностями, предоставляемыми пользователям PW.

1.2. Функциональность PW

Для обеспечение эмуляции поведения и характеристик естественного сервиса PW обеспечивают следующие функции:

- инкапсуляция используемых сервисом PDU или данных устройства, прибывающих на граничный порт PE (логический или физический);
- доставка инкапсулированных данных через туннель PSN;
- организация PW, включая обмен и/или распределение идентификаторов PW, используемых конечными точками туннеля PSN;

[1]Pseudo Wire Emulation Edge-to-Edge.
[2]Packet Switched Network.
[3]Pseudo Wire.
[4]Native Service Processing
[5]Customer Edge Equipment

- управление сигнализацией, синхронизацией, порядком следования и другими параметрами сервиса на границе PW;
- поддержка присущих сервису сигналов о состоянии и тревожных ситуациях.

1.3. Что не рассматривается в этом документе и не относится к PW

Перечисленные ниже аспекты не входят в сферу внимания данного документа:

- спецификация инкапсуляции PW;
- детальная спецификация протоколов, вовлечённых в организацию и поддержку PW.

Перечисленные ниже аспекты лежат за пределами PWE3:

- любые службы с групповой адресацией, не являющиеся естественными для эмулируемого сервиса; таким образом, передача кадров Ethernet по групповому адресу IEEE-48 относится к PWE3, а службы с групповой адресацией типа MARS [RFC2022] не относятся;
- методы сигнализации и контроля для сети PSN, в которой используется инкапсуляция.

1.4. Терминология

Ниже приведены определения используемых в этом документе терминов. Иллюстрация к их использованию имеется на рисунке 2.

Attachment Circuit (AC) - устройство подключения
Физическое или виртуальное устройство, обеспечивающее подключение CE к PE. В качестве AC может выступать Frame Relay DLCI, ATM VPI/VCI, порт Ethernet, VLAN, канал HDLC, соединение PPP на физическом интерфейсе, сессия PPP через туннель L2TP, MPLS LSP и т. п. Если физические или виртуальные AC используют одну и ту же технологию (например, оба устройства ATM, Ethernet или Frame Relay), говорят, что PW обеспечивает «однородный транспорт»; в противном случае транспорт является разнородным (гетерогенным).

CE-bound - к абоненту
Направление трафика, при котором PW-PDU принимаются на PW через сеть PSN, обрабатываются и передаются устройству CE (адресату).

CE Signaling - сигнализация CE
Сообщения, принимаемые и передаваемые в управляющей плоскости CE. Такие сообщения могут быть желательными и даже необходимыми для PE, чтобы обеспечить участие в сигнальных процессах и мониторинг сигнализации с целью эффективной эмуляции сервиса.

Control Word (CW) - управляющее слово
Четырехоктетный заголовок, используемый в некоторых схемах инкапсуляции для передачи относящейся к отдельному пакету информации в тех случаях, когда PSN работает на основе MPLS.

Customer Edge (CE) - пользовательский край
Устройство, один край которого является источником или завершением сервиса. Устройство CE не знает какой сервис используется - естественный или эмулируемый.

Forwarder (FWRD) - система пересылки
Подсистема PE, которая выбирает PW для передачи информации, полученной в AC.

Fragmentation - фрагментация
Операция деления одного PDU на множество PDU перед передачей в предположении, что исходный блок PDU будет заново собран где-либо в сети. Пакеты могут фрагментироваться, если их размер превышает значение MTU для сети, через которую пакет нужно передать.

Maximum Transmission unit (MTU) - максимальный передаваемый блок
Размер максимального пакета (без учёта заголовка канального уровня), который интерфейс может передать без фрагментации.

Native Service Processing (NSP)
Обработка данных, полученных PE от устройства CE до их представления PW для передачи в сеть, или обработка данных, полученных от PW устройством PE до их передачи устройству AC. Функционально NSP определяется не IETF, а комитетами по стандартизации (такими, как ITU-T, ANSI, ATMF)

Packet Switched Network (PSN) - сеть с коммутацией пакетов
В контексте PWE3 это сеть, использующая IP или MPLS в качестве механизма пересылки пакетов.

PE-Bound - в сторону сети
Направление трафика, при котором информация от CE адаптируется для PW и блоки PW-PDU передаются в сеть PSN.

PE/PW Maintenance - обслуживание PE/PW
Используется устройствами PE для организации, обслуживания и разрыва соединений PW. Может быть связано с сигнализацией CE для обеспечения эффективного управления PW.

Protocol Data Unit (PDU) - модуль данных протокола
Единица данных, передаваемая в сеть лили принимаемая из неё на уровне протокола.

Provider Edge (PE) - провайдерский край
Устройство, обеспечивающее PWE3 для CE.

Pseudo Wire (PW) - псевдопровод
Механизм, обеспечивающий передачу существенных элементов эмулируемого сервиса из устройства PE в одно или множество других устройств PE через сеть PSN.

Pseudo Wire Emulation Edge to Edge (PWE3) - сквозная эмуляция псевдопровода
Механизм, эмулирующий существенные атрибуты сервиса (такого, как выделенная линия T1 или Frame Relay) через сеть PSN.

Pseudo Wire PDU (PW-PDU) - PDU псевдопровода
Блок PDU, передаваемый в PW и содержащий все данные и управляющую информацию, требуемые для эмуляции нужного сервиса.

PSN Tunnel - туннель PSN
Туннель через сеть PSN, внутри которого могут передаваться один или множество PW.

PSN Tunnel Signaling - сигнализация туннеля PSN
Используется организации, поддержки и разрыва туннелей PSN.

PW Demultiplexer - демультиплексор PW
Работающий в плоскости данных метод идентификации PW, завершающихся на устройстве PE.

Time Domain Multiplexing (TDM) - мультиплексирование с разделением по времени
Мультиплексирование TDM. Этот термин часто используется для обозначения синхронных битовых потоков со скоростями, определёнными в стандарте G.702.

Tunnel - туннель
Метод прозрачной передачи информации через сеть.

2. Применимость PWE3

В сети PSN используемой для PW может происходить потеря пакетов, их задержка, которая может варьироваться в широких пределах, и нарушение порядка доставки пакетов. В периоды неустойчивости сети могут возникать достаточно продолжительные интервалы существенного снижения качества сервиса и даже его полной неработоспособности. Применимость PWE3 для того или иного сервиса зависит от чувствительности данного сервиса (или реализации CE) к перечисленным воздействиям, а также от возможностей уровня адаптации по маскированию вредных воздействий. Некоторые службы (например, IP over FR over PWE3) могут быть достаточно устойчивы к характеристикам сетей IP и MPLS. Другие типы сервиса (например, соединение PBX через PWE3) будут требовать более осторожного подхода к характеристикам PSN и уровня адаптации. В некоторых случаях требуется использование средств организации трафика PSN, а иногда ограничения будут делать невозможными требуемые для сервиса гарантии.

3. Многоуровневая модель

Многоуровневая модель PWE3 предназначена для минимизации влияния различий между PW, работающими в различных типах PSN. При разработке многоуровневой модели ставилась задача сделать каждое определение PW независимым от нижележащей сети PSN и по максимуму использовать определения протоколов IETF и реализации протоколов.

3.1. Протокольные уровни

Логическая структура многоуровневой модели для поддержки PW показана на рисунке 1.

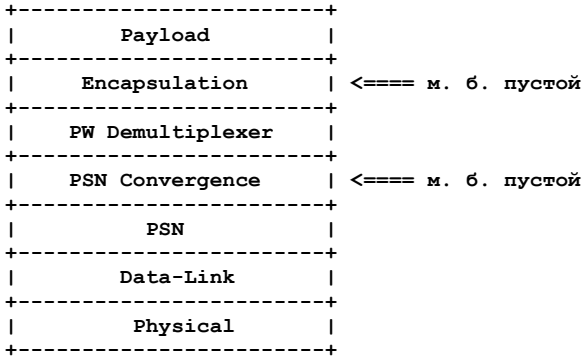

Рисунок 1. Модель логических уровней протокола.

Информация (payload) транспортируется через уровень инкапсуляции (Encapsulation Layer). Этот уровень передаёт всю информацию (не только собственно данные - payload), которая требуется интерфейсу PW CE-bound PE для передачи данных устройству CE через физический интерфейс. Если кроме данных не передаётся никакой информации, этот уровень остаётся пустым.

Уровень инкапсуляции обеспечивает также поддержку обработки в реальном масштабе времени, если это требуется для упорядочения.

Уровень демультиплексирования (PW Demultiplexer) обеспечивает возможность доставки множества PW через один туннель PSN. Значение PW Demultiplexer, используемое для идентификации PW в плоскости данных, может быть уникальным для каждого PE, но это требование не является обязательным для PWE3. Однако идентификаторы должны быть уникальными в масштабах конечной точки туннеля. Если требуется обеспечить идентификацию отдельных туннелей, ответственность за это ложится на уровень PSN.

Уровень конвергенции (PSN Convergence) обеспечивает расширение, требуемое для того, чтобы обеспечить соответствие PSN требованиям к сервису PSN. Следовательно, этот уровень обеспечивает согласованный интерфейс с PW, который делает PW независимым от конкретного типа PSN. Если PSN уже соответствует всем требованиям, уровень остаётся пустым.

Определения заголовков PSN, канального (MAC/Data-Link) и физического уровня выходят за пределы настоящего документа. PSN может быть IPv4, IPv6 или MPLS.

3.2. Сфера действия PWE3

PWE3 определяет уровень инкапсуляции - метод передачи различных типов информации (payload) и интерфейс с уровнем PW Demultiplexer. Предполагается, что другие уровни будут обеспечиваться с помощью методов туннелирования типа L2TP или MPLS через PSN.

3.3. Типы информационных полей

Данные классифицируются на несколько типов блоков данных естественного сервиса:

- пакет;
- ячейка;
- битовый поток;
- структурированный битовый поток.

В рамках этих базовых типов имеются связанные с сервисом подтипы:

Базовый тип	***Сервис PW***
пакет	Ethernet (все типы), кадрирование HDLC, Frame Relay, ATM AAL5 PDU
ячейка	ATM
битовый поток	Неструктурированные потоки E1, T1, E3, T3
структурированный битовый поток	SONET/SDH (например, SPE, VT, NxDS0)

3.3.1. Пакетные данные

Пакеты (packet payload) представляют собой блоки данных переменной длины, доставляемые PE через AC. Размер пакетов может превышать PSN MTU. Границы пакетов определяются типом инкапсуляции. Примерами пакетных данных могут служить HDLC или Ethernet. Обычно из пакетов удаляется избыточная служебная информация (типа флагов HDLC и битов заполнения) перед тем, как пакеты будет переданы через PW.

Пакетные данные обычно передаются через PW в виде одного блока. Однако возможны случаи, когда размер пакетных данных в сумме с заголовками PWE3 и PSN будет превышать значение MTU на пути через PSN MTU. В таких случаях используется тот или иной механизм фрагментации. Это может быть, например, в случаях, когда пользователь обеспечивает сервис и соединяется с провайдером через Ethernet или когда используются “вложенные” псевдопровода. Фрагментация более подробно рассматривается в параграфе 5.3.

Пакетные данные могут требовать упорядочения и поддержки работы в реальном масштабе времени.

В некоторых случаях пакетные данные могут выбираться из присутствующих в эмулируемом проводе пакетов на основе того или иного метода субмультиплексирования. Например, один или несколько блоков Frame Relay PDU могут быть выбраны для транспортировки через определённый псевдопровод на основе Frame Relay DLCI[1] или, для случая Ethernet, может использоваться подходящий фильтр MAC-уровня. Эта функция относится к системе пересылки и выбор, следовательно, будет происходить до того, как пакет будет представлен уровню инкапсуляции (PW Encapsulation).

3.3.2. Данные в ячейках

Данные в ячейках (cell payload) создаются путём сбора, доставки и восстановления (replaying) групп октетов, представленных в проводе в формате с фиксированным размером. Границы группы битов, составляющих ячейку, определяются типом инкапсуляции. Двумя наиболее распространёнными примерами могут служить ячейки ATM размером 53 октета и пакеты MPEG Transport Stream [DVB] размером 188 октетов.

Для снижения объёма служебной информации, передаваемой через PSN, множество ячеек может объединяться в один блок данных (payload). Уровень инкапсуляции может завершать формирование такого блока по таймеру, количеству ячеек или при получении специальной ячейки (например, ATM OAM). Преимущества объединения множества ячеек следует оценивать с учётом возможного роста вариации задержек и более серьёзных потерь в случаях потери пакетов. В некоторых ситуациях на уровне инкапсуляции целесообразно использовать ту или иную компрессию (например, подавление пауз - silence suppression или сжатие голосовых данных).

Базовая поддержка данных в ячейках обычно требует упорядоченной доставки и может также потребовать поддержки работы в реальном масштабе времени. Базовый сервис работы с данными в ячейках обычно не требует фрагментации.

Уровень инкапсуляции может применять ту или иную компрессию для некоторых субтипов (например, могут подавляться пустые ячейки - idle cell).

В некоторых случаях ячейки, помещаемые в поле данных (payload), могут выбираться путём фильтрации потока ячеек, присутствующих в проводе. Например, сервис ATM PWE3 может фильтровать ячейки по значениям полей VCI или VPI. Эта функция реализуется системой пересылки и, следовательно, выбор происходит до представления пакета уровню инкапсуляции.

3.3.3. Битовый поток

Данные в форме битовых потоков (bit stream payload) создаются путём захвата, доставки и воспроизведения битовых потоков в проводе без учёта структуры потока, которая на поверку может быть видна в транслируемом трафике (к примеру, внутренняя структура не оказывает влияния на фрагментацию пакетов).

В некоторых случаях по отношению к битовым потокам могут те или иные алгоритмы подавления. Например, E1 и T1 используют последовательность, состоящую только из 1 (all-ones) для индикации сбоев. Такие ситуации можно обнаруживать без использования сведений о структуре потока и при пакетной передаче такие последовательности могут подавляться.

Этот тип сервиса будет требовать поддержки упорядоченной доставки и работы в реальном масштабе времени.

3.3.4. Структурированный битовый поток

Данные в виде структурированного битового потока создаются с использованием той или иной информации о структуре этого потока для захвата, доставки и воспроизведения битовой последовательности в эмулируемом проводе.

Между структурированными и неструктурированными битовыми потоками существует два важных отличия:

[1]Data-Link Connection Identifier - идентификатор соединения на канальном уровне.

- Некоторые части исходного битового потока могут вырезаться при передаче в направлении PSN-bound в блоке NSP. Например, раздел Structured SONET и служебная строка (возможно и более) могут фильтроваться. Для обеспечения возможности такого вырезания требуется формирователь кадров (framer). Требуется также выравнивание кадров/данных для дробных потоков T1/E1.

- От PW требуется сохранение структуры при передаче через PSN, поэтому блок CE-bound NSP может вставлять информацию в восстановленный неструктурированный поток битов. Вырезанная информация (например, выравнивание указателей SONET) может появляться на уровне инкапсуляции для упрощения этой реконструкции.

В качестве опции уровень инкапсуляции может также использовать подавление пауз (silence/idle suppression) или аналогичные механизмы компрессии по отношению к структурированному потоку.

Структурированные битовые потоки отличаются от ячеек тем, что период структуры может оказаться слишком большим для включения в один пакет. Отметим, что структуры с коротким периодом неотличимы от ячеек и для них могут использоваться преимущества методов, описанных в параграфе 3.3.2.

Этот тип сервиса требует поддержки упорядоченной доставки и работы в реальном масштабе времени.

3.3.5. Принцип минимального вмешательства

Для минимизации видимости данных и повышения эффективности потока данных через уровень инкапсуляции содержимое (payload) следует транспортировать в том виде, как оно было получено с минимальными изменениями [RFC1958].

Принцип минимального вмешательства позволяет отделить изменение данных (payload) от изменения PW и требует меньшего числа преобразований в NSP систем с одинаковыми интерфейсами CE на обеих сторонах. Также предотвращаются нежелательные побочные эффекты, связанные с ошибочной интерпретацией данных в промежуточном формате.

Варианты с большим вмешательством могут быть более эффективными для передачи и в некоторых случаях способны снижать число преобразований NSP для различных интерфейсов CE по разные стороны. Все промежуточные форматы по сути становятся новыми типами кадрирования, требующими документирования и обеспечения взаимодействия. Это увеличивает объем работы по поддержке протокола, который использует промежуточный формат и, по этой причине, нежелательно.

4. Архитектура псевдопроводов

В этой главе описана архитектурная модель PWE3.

4.1. Эталонная модель

На рисунке 2 показана эталонная модель PW типа “точка-точка”.

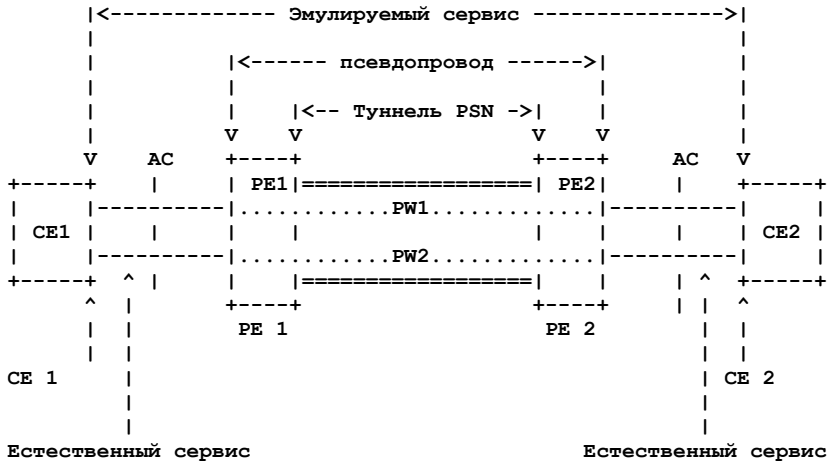

Рисунок 2. Эталонная модель PWE3.

Два устройства PE (PE1 и PE2) обеспечивают один или более PW от имени клиентских устройств CE (CE1 и CE2) чтобы обеспечить взаимодействие клиентских CE через сеть PSN. Чтобы обеспечить путь передачи данных для PW создаётся туннель PSN. Трафик PW остаётся невидимым для промежуточной сети, которая прозрачна для CE. Естественные блоки данных (биты, ячейки или пакеты), поступающие через AC, инкапсулируются в PW-PDU и передаются через сеть с использованием туннеля PSN. Устройства PE выполняют необходимую инкапсуляцию и декапсуляцию PW-PDU, а также выполняют другие функции, требуемые сервисом PW (такие, как упорядочение или синхронизация).

4.2. Предварительная обработка PWE3

Некоторые приложения выполняют те или иные операции по отношению к естественным блокам данных, полученным от CE (включая информацию и сигнальный трафик), до передачи через PW. Примерами могут служить мосты Ethernet, кросс-соединения SONET, трансляция идентификаторов локальной значимости (таких, как VCI/VPI) или преобразование к другому типу сервиса. Эти операции могут выполняться внешним оборудованием с передачей обработанных данных устройству PE через один или множество физических интерфейсов. Во многих случаях выполнение таких операций в PE обеспечивает экономические и эксплуатационные преимущества. Обработанные данные в таких случаях передаются в PW через виртуальный интерфейс внутри PE. Такие операции предварительной обработки включены в эталонную модель PWE3 для обеспечения единой контрольной точки (reference point), но детальное описание этих операций выходит за пределы рассматриваемых здесь определений PW.

Рисунок 3. Предварительная обработка в эталонной модели PWE3.

На рисунке 3 показано межсетевое взаимодействие устройства PE с предварительной обработкой (PREP) с другим устройством, которое не поддерживает таких функций. Контрольная точка подчёркивает, что функциональный интерфейс между PREP и PW представляется физическим интерфейсом, обеспечивающим сервис. Это эффективно определяет требуемую спецификацию межсетевого взаимодействия.

Работа систем, в которых оба устройства PE включают PREP, также поддерживается.

Требуемую предварительную обработку можно поделить на две части:

- пересылка (FWRD);
- естественная для сервиса обработка (NSP).

4.2.1. Системы пересылки

Некоторые приложения селективно пересылают блоки данных от одного или множества AC одному или многим PW. В таких случаях требуется также выполнять обратные операции для PWE3-PDU, полученных PE из сети PSN. Это относится к функциям системы пересылки.

Система пересылки (forwarder) выбирает PW на основе AC-отправителя, содержимого поля данных или неких статически или динамически задаваемых параметров.

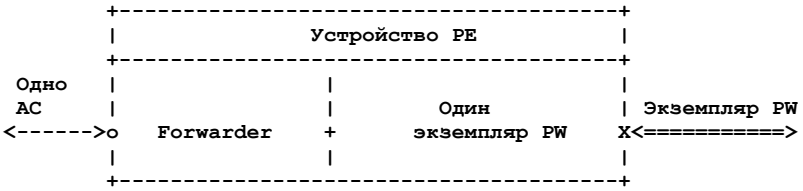

Рисунок 4a. Простой сервис "точка-точка".

На рисунке 4a показана простая система пересылки, выполняющая некоторые операции по фильтрации. Поскольку система пересылки имеет один входной интерфейс и один выходной, фильтрация является единственным типом операций, выполняемых при пересылке. На рисунке 4b показана более распространённая ситуация, когда данные от одного или множества AC направляются одному или множеству PW. В таких случаях к данным могут применяться операции пересылки, фильтрации или их комбинация. Например, если AC использует Frame Relay, система пересылки может выполнять коммутацию Frame Relay, а экземпляры PW могут быть соединениями между коммутаторами.

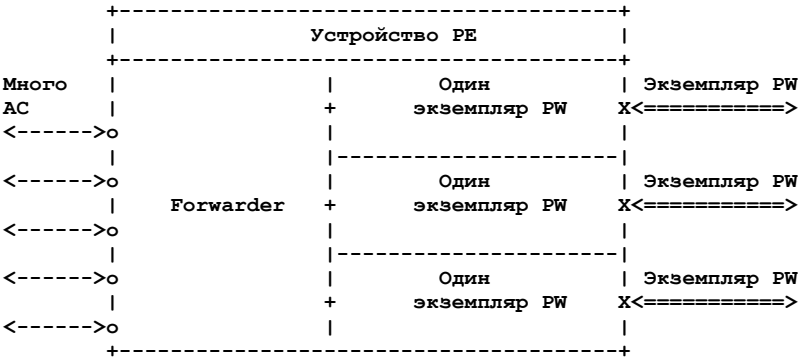

Рисунок 4b. Пересылка от множества AC во множество PW.

4.2.2. Естественная обработка

Некоторым приложениям требуется та или иная форма трансляции (преобразования0 данных или адресов или иные операции, требующие знания семантики данных. Эта функция выполняется процессором естественной обработки (NSP[1]).

[1]Native Service Processor.

Использование NSP упрощает структуру PW, ограничивая работу PW гомогенными операциями. NSP включён в эталонную модель для обеспечения интерфейса к таким функциям. Спецификация различных типов NSP выходит за пределы PWE3.

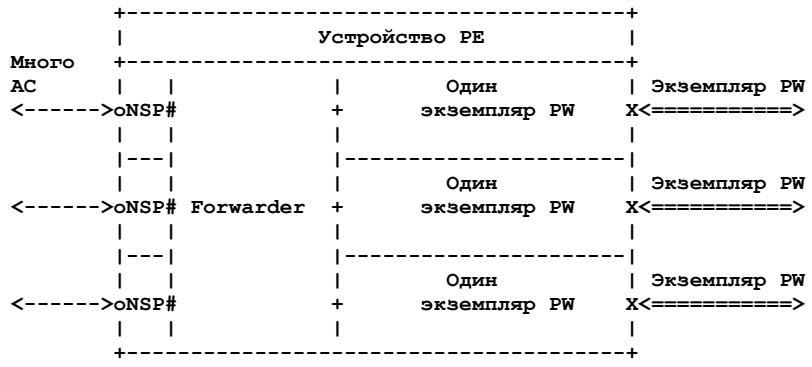

Рисунок 5. Пересылка от множества AC во множество PW.

На рисунке 5 показаны связи между NSP, системой пересылки и PW внутри PE. Функции NSP могут выполнять любые преобразования (изменение, вставку и т. п.) по отношению к данным, проходящим через физический интерфейс к CE и через виртуальный интерфейс к системе пересылки. Эти преобразования будут, естественно, ограничены теми операциями, которые могут быть выполнены на пути передачи данных и разрешены конфигурацией PE. Устройство PE может включать более одной системы пересылки.

Эта модель также поддерживает работу систем, в которых функциональность NSP включает завершение каналов данных, и приложение сетевого уровня обрабатывает информацию (payload).

4.3. Эталонная модель поддержки

На рисунке 6 показана эталонная модель поддержки для PW.

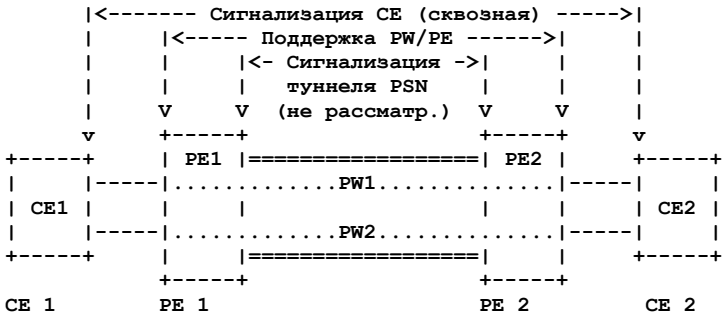

Рисунок 6. Эталонная модель поддержки PWE3.

Требуются перечисленные ниже сигнальные механизмы:

- Сквозная сигнализация между устройствами CE, в качестве которой может использоваться Frame Relay PVC status, ATM SVC, TDM CAS и т. п.
- Поддержка PW/PE используется между устройствами PE (или NSP) для организации, поддержки и разрыва PW, включая всю требуемую координацию параметров.
- Сигнализация туннеля PSN управляет мультиплексированием PW и некоторыми элементами PSN. Примерами могут служить протокол управления L2TP, MPLS LDP и RSVP-TE. Определение информации, которую нужно передавать в качестве сигналов PWE3 входит в задачи PWE3, но сами сигнальные протоколы не входят сюда.

4.4. Эталонная модель стека протоколов

На рисунке 7 показана эталонная модель стека протоколов для PW.

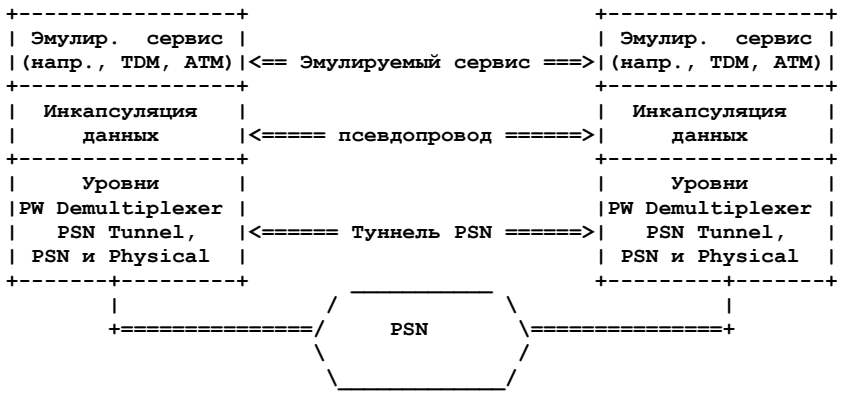

Рисунок 7. Эталонная модель стека протоколов PWE3 .

PW обеспечивает CE эмулируемым физическим или виртуальным соединением с удаленным партнёром. PDU естественного сервиса от CE проходят через уровень инкапсуляции на передающем устройстве PE и пересылаются после этого через сеть PSN. Принимающее устройство PE выполняет декапсуляцию и восстанавливает информацию в её естественном формате для передачи CE-адресату.

4.5. Предварительная обработка для эталонной модели стека протоколов

На рисунке 8 показано как эталонная модель стека протоколов расширяется для выполнения предварительной обработки (пересылка и NSP). Приведённая схема показывает расположение физического интерфейса относительно CE.

Рисунок 8. Эталонная модель стека протоколов с предварительной обработкой.

5. Инкапсуляция PW

Уровень инкапсуляции PW обеспечивает требуемую инфраструктуру для адаптации конкретного типа данных (payload), которые будут передаваться через PW, для уровня демультиплексирования PW, служащего для передачи PW через PSN.

Уровень инкапсуляции PW включает три подуровня:

- Payload Convergence - конвергенция данных;
- Timing - синхронизация;
- Sequencing - упорядочение.

Подуровни уровня инкапсуляции PW в контексте стека протоколов показаны на рисунке 9.

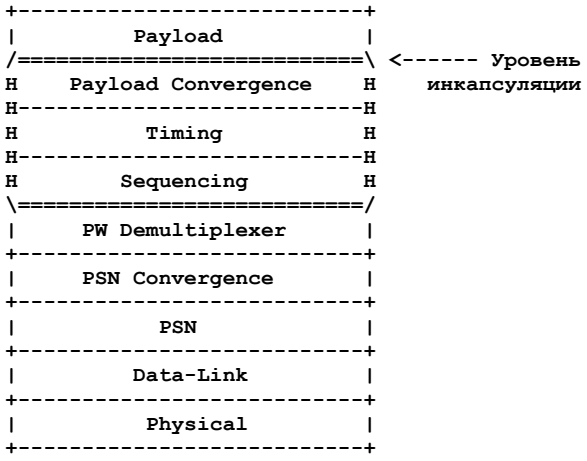

Рисунок 9. Уровень инкапсуляции PWE3.

Подуровень конвергенции данных жёстко связан с конкретным типом данных. Однако группировка множества типов данных в базовый класс и реализация для этого класса общего подуровня конвергенции снижает число типов этого подуровня, что ведёт к снижению сложности реализации. Обеспечение попакетной сигнализации и другой информации за пределами основного канала (out-of-band), не относящейся к синхронизации и обеспечению упорядоченной доставки, гарантируется этим подуровнем.

Подуровни синхронизации и обеспечения упорядоченной доставки обеспечивают подуровню конвергенции базовые функции для всех типов данных.

5.1. Подуровень конвергенции данных

5.1.1. Инкапсуляция

Основной задачей подуровня конвергенции является инкапсуляция данных в PW-PDU. Инкапсулируемые естественные блоки данных могут содержать служебную информацию в канального (L2) и физического (L1) уровня. Эта информация зависит от типа сервиса. Заголовок подуровня конвергенции содержит дополнительную информацию, требуемую для восстановления естественных блоков данных на физическом интерфейсе в направлении CE-bound (к абоненту). Заголовок уровня демультиплексирования не рассматривается как часть заголовка PW.

Не вся дополнительная информация, требуемая для восстановления естественных блоков данных, передаётся в PW-заголовке PW PDU. Часть информации (например, тип сервиса для PW) может сохраняться на приёмной стороне (destination PE) во время организации PW.

5.1.2. Типы каналов PWE3

Уровень инкапсуляции PW и связанная с ним сигнализация требуют одного или нескольких каналов (канал типа 1 + один или несколько каналов типа 2 - 4) из числа перечисленных ниже от нижележащих уровней демультиплексирования PW и PSN:

1. Канал управления с гарантированной доставкой для сигнализации, индикации состояний и, в исключительных случаях, событий CE-CE, которые должны транслироваться и гарантированно передаваться между PE. PWE3 может потребоваться этот тип канала для обеспечения достоверной эмуляции сложных протоколов канального уровня.
2. Канал с высоким приоритетом и соблюдением порядка доставки, но без гарантии доставки пакетов. Такие каналы обычно используются для сигнализации между устройствами CE. “Высокий приоритет” может просто задаваться битами DSCP для протокола IP или EXP для MPLS, обеспечивая пакетам приоритетную доставку. Этот тип каналов может также использовать бит в заголовке самого туннеля для индикации того, что полученные PE пакеты следует обрабатывать с более высоким приоритетом [RFC2474].
3. Канал с упорядоченной доставкой для передачи данных, чувствительных к нарушению порядка следования пакетов (одним из вариантов такого трафика могут быть данные протоколов, отличных от IP).
4. Канал с неупорядоченной доставкой для данных, не чувствительных к нарушению порядка пакетов.

Каналы данных (2, 3, 4) следует передавать в одной полосе (in band) с другими для эффективного использования возможностей PSN.

Там, где сквозная связность может нарушаться системами трансляции адресов [RFC3022], списками контроля доступа, межсетевыми экранами и т. п., могут возникать ситуации, когда канал управления сможет передавать трафик и создавать PW, а трафик данных PW будет блокироваться одним или несколькими упомянутыми механизмами. В этих случаях, если канал управления не передаётся в той же полосе (in band), сигнализация при организации PW не будет подтверждать существование сквозного пути передачи данных. В некоторых случаях требуется синхронизация событий CE с данными, передаваемыми через PW. В частности, такая ситуация возникает при использовании устройств TDM (например, информация о поднятии/опускании телефонной трубки в коммутаторах PSTN может передаваться через сигнальный канал с гарантированной доставкой, тогда как связанные с телефонным разговором данные будут передаваться через упорядоченный канал данных).

Типы каналов PWE3, не требуемые для поддерживаемых PW, могут не включаться в реализацию.

5.1.3. Качество обслуживания

По возможности следует реализовать для PW поддержку механизма QoS[1] через PSN.

5.2. Независимые от типа данных подуровни инкапсуляции PW

Подуровни упорядочения (Sequencing) и синхронизации (Timing) уровня инкапсуляции PWE3 обеспечивают общий сервис для всех типов данных. Этот сервис является необязательным и используется только в случае необходимости для конкретного экземпляра PW. Если сервер не нужен, соответствующий заголовок может быть опущен в целях экономии ресурсов при обработке и передаче через сеть.

Тот или иной конкретный тип данных может требовать транспорта с упорядоченной доставкой или без таковой и/или поддержки работы в реальном масштабе времени. Например, одной из характеристик транспорта Frame Relay является упорядоченная доставка. Некоторые приложения Frame Relay ожидают доставки с сохранением порядка и могут не принять кадры, полученные с нарушением порядка. Однако при использовании сервиса Frame Relay только в качестве транспорта для IP может оказаться разумным отказ от соблюдения порядка, позволяющий снизить издержки на обработку пакетов.

Для обеспечения такого сервиса следует, по возможности, использовать существующие протоколы IETF. Когда подходящего протокола нет, следует расширять или изменять существующие протоколы в соответствии с требованиями PWE3, сохраняя возможность использования этого протокола для других задач. Для синхронизации может потребоваться более одного метода общего назначения, чтобы выполнить все требования по синхронизации данных.

5.2.1. Упорядоченная доставка

Функция упорядочения выполняет три задачи: упорядочение кадров, детектирование дубликатов и детектирование потери кадров. Это позволяет обеспечивать эмуляцию соответствующих свойств проводного соединения. Поддержка упорядочения зависит от типа данных и может быть отключена, если в ней нет необходимости.

[1] Quality of Service - качество обслуживания.

Размер пространства порядковых номеров зависит от скорости эмулируемого сервиса и максимального времени нахождения пакета в сети PSN. Следовательно, требуются порядковые номера более 2^16, чтобы избавиться от проблем, связанных с переходом через максимум в течение срока передачи пакета через сеть.

5.2.1.1. Упорядочение кадров

Когда пакеты, содержащие PW-PDU, проходят через PSN, они могут поменять порядок следования и прийти в PE с нарушением порядка. Для некоторых случаев кадры (управление, данные или оба типа) должны доставляться с сохранением порядка. Для таких типов сервиса должен поддерживаться тот или иной механизм, гарантирующий соблюдение порядка доставки. Задание порядковых номеров на подуровне упорядочения является одним из вариантов решения задачи. Отметим, что упорядоченная доставка является частным случаем задачи синхронизированной доставки и задача доставки с соблюдением порядка может быть решена вместе с задачей синхронизированной доставки с помощью одного комбинированного механизма (например, [RFC3550]).

Существует два варианта стратегии сохранения порядка доставки:

- отбрасывание PW PDU, доставленных с нарушением порядка;
- попытка сортировки PW PDU с целью восстановления порядка.

Выбор одного из этих вариантов будет зависеть от перечисленных ниже условий:

- критичность потери пакетов для работы PW (например, допустимое число битовых ошибок в единицу времени);
- скорость работы PW и PSN;
- допустимая задержка (при попытке восстановления порядка);
- предполагаемая частота нарушения порядка доставки.

5.2.1.2. Детектирование дубликатов

В редких случаях при передаче пакетов PW через PSN могут возникать дубликаты этих пакетов. Для некоторых типов сервиса появление таких дубликатов неприемлемо. Для такого сервиса должен обеспечиваться тот или иной механизм предотвращения доставки дубликатов CE-адресату. Возможно использование того же механизма, который служит для обеспечения упорядоченной доставки.

5.2.1.3. Детектирование потери кадров

PE-адресат может обнаружить потерю кадра, контролируя порядковые номера полученных PW PDU.

В некоторых системах, если PW PDU не поступает в течение некоторого времени, PE-адресат будет предполагать потерю пакета. Если такой PW-PDU впоследствии прибывает целевое устройство PE должно отбросить его.

5.2.2. Синхронизация

Многие естественные типы сервиса предъявляют определённые требования к синхронизации на основе характеристик сетей, для которых этот сервис разрабатывался. Эмулируемый сервис может воспроизводить такие характеристики с достаточной точностью (например, трафик может доставляться с такой же скоростью, задержками, их вариациями и т. п., какие наблюдаются в направлении приёма на передающем устройстве PE).

В таких ситуациях принимающее устройство PE “воспроизводит” естественный трафик, как он был получен на стороне передающего PE. При этом используются синхросигналы, передаваемые между двумя PE или (в некоторых случаях) полученные от внешнего источника.

Следовательно, подуровень синхронизации должен поддерживать две функции: восстановление синхросигналов и синхронизированную доставку данных[1]. Те или иные типы сервиса могут использовать одну из этих функций или обе.

5.2.2.1. Восстановление синхросигналов

Восстановление синхросигналов представляет собой извлечение битов синхронизации из потока доставленных пакетов и требует для этого тот или иной механизм. В физических проводах синхросигналы передаются естественным способом, а восстановление синхросигналов из потока данных со значительными вариациями задержки представляет собой непростую задачу. Следовательно, использовать для этих целей существующие протоколы, работающие в реальном масштабе времени (например, [RFC3550]), если для этого нет тех или иных препятствий применительно к конкретному типу данных.

5.2.2.2. Синхронизированная доставка

Синхронизированная доставка (Timed delivery) представляет собой доставку дискретных PW PDU в выходной интерфейс PW с постоянным фазовым сдвигом относительно входного интерфейса. Синхронизация доставки может осуществляться по сигналам, восстановленным из потока пакетов, полученных через PSN, или по сигналам от внешнего источника.

5.3. Фрагментация

В идеальном случае данные будут транслироваться через PW в виде одного блока. Однако возникают ситуации, когда суммарный размер данных и связанных с ними заголовков PWE3 и PSN будет превышать значение MTU для пути через PSN. Когда размер пакета превышает MTU данной сети, в процессе доставки пакетов используются операции фрагментации и сборки. Поскольку фрагментация и сборка пакетов требуют больших ресурсов, нежели простая передача пакетов, уровень фрагментации (и связанной с ней последующей сборки) следует снижать, насколько это возможно. Фрагментация и сборка пакетов могут стать существенной проблемой для узлов, в которых обрабатывается множество PW (например, в точках PE).

[1] В оригинале - timed payload delivery. *Прим. перев.*

В идеальном оборудование, создающее трафик, который передаётся через PW, будет использовать механизмы адаптации (например, [RFC1191], [RFC1981]), которые обеспечат передачу пакетов, не требующих фрагментации. Когда фрагментации избежать не удаётся, точке, ближайшей к передающему хосту и поддерживающей возможность фрагментации и сборки, следует предпринять попытку снижения размера пакетов до требуемой величины (PSN MTU). Таким образом, в эталонной модели PWE3 (рисунок 3), фрагментацию следует пытаться выполнить в устройстве CE. Если CE не может обеспечить соответствующий MTU размер пакетов, PW следует предпринимать свои методы фрагментации.

В тех случаях, когда системе управления MTU не удаётся снизить размер пакетов до уровня, подходящего для передачи PW, устройство PE может вернуться к базовому методу фрагментации PW или использовать (если это возможно) сервис фрагментации PSN.

Реализация PE может не поддерживать фрагментацию. В этом случае PE будет отбрасывать пакеты, размер которых превышает PSN MTU, и система управления инкапсуляцией PE может получить уведомление об этом.

Если размер кадра L2/L1, восстановленного из PW PDU, превышает значение MTU для AC-адресата, такой кадр должен быть отброшен. В этом случае система управления целевого PE может получить уведомление.

5.4. Конкретизация уровней протокола

В этом документе не рассматривается детальное отображение многоуровневой модели протокола на существующие или разрабатываемые стандарты IETF. Конкретизация логической многоуровневой модели протокола показана на рисунке 9.

5.4.1. PWE3 в сетях IP PSN

Определению протокола работы PWE3 через IP PSN следует использовать существующие протоколы IETF там, где это возможно.

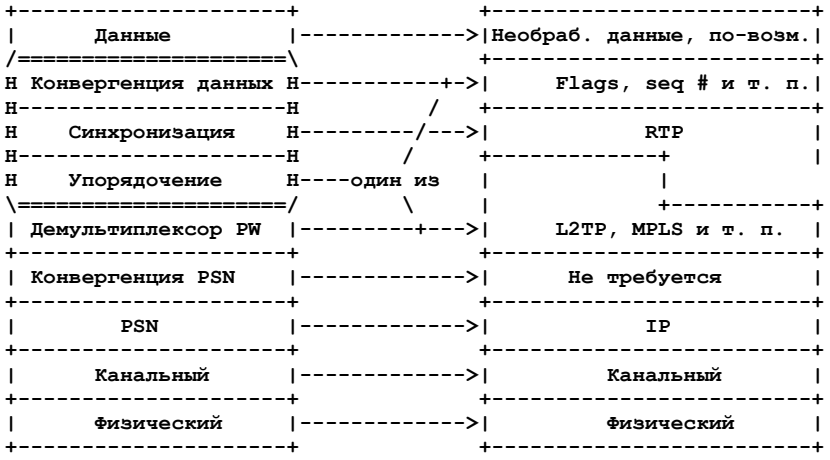

Рисунок 10. PWE3 через IP PSN.

На рисунке 10 показаны уровни протокола для работы PWE3 через IP PSN. Как правило данные следует передавать в том виде, в котором они были получены от NSP с использованием при необходимости подуровня конвергенции данных. Однако в некоторых ситуациях может оказаться предпочтительной передача данных в трансформированном виде. Причина преобразования должна быть документирована в определении уровня инкапсуляции для этого типа данных.

В тех случаях, когда это применимо, синхронизация обеспечивается с помощью протокола RTP [RFC3550], который (при его использовании) обеспечивает также сервис упорядочения. Когда сеть PSN работает на базе UDP/IP, заголовок RTP следует за заголовком UDP и предшествует полю управления PW. Во всех остальных случаях заголовок RTP следует за управляющим заголовком PW.

Уровень инкапсуляции может также передавать порядковый номер. Упорядочение обеспечивается RTP или уровнем инкапсуляции PW (но не обоими).

Демультиплексирование PW обеспечивается метками PW, которые могут иметь форму, заданную для множества протоколов IETF - например, метка MPLS [MPLSIP], идентификатор сессии L2TP [RFC3931] или номер порта UDP [RFC768]. Когда PW передаются через IP, уровень конвергенции PSN не требуется.

В том случае, когда демультиплексирование происходит по меткам MPLS, вместо описанной здесь архитектуры может использоваться архитектура, рассмотренная в параграфе 5.4.2.

5.4.2. PWE3 в сетях MPLS PSN

Специфика MPLS весьма хороша для обеспечения эффективности “проводов”. За счёт использования управляющего слова некоторые компоненты протоколов PWE3 могут быть сжаты для дополнительного повышения эффективности.

На рисунке 11 показана многоуровневая модель работы PWE3 через MPLS PSN. Метки MPLS служат для демультиплексирования PW. Управляющее слово служит для передачи большей части информации, требуемой уровню инкапсуляции PWE3 и конвергенции PSN, в компактном формате. Флаги в управляющем слове обеспечивают требуемую конвергенцию данных. Поле порядкового номера позволяет поддерживать как упорядоченную доставку, так и сервис фрагментации PSN на уровне конвергенции PSN (поддерживается с помощью метода управления фрагментацией). Ethernet дополняет все кадры до минимального размера 64 байта. Заголовок MPLS не включает поле размера. Следовательно, при передаче PWE3 через MPLS для корректного прохождения через каналы Ethernet требуется поле коррекции размера в управляющем слове. Как и для случаях IP PSN, синхронизацию (по возможности) следует обеспечивать с помощью RTP [RFC3550].

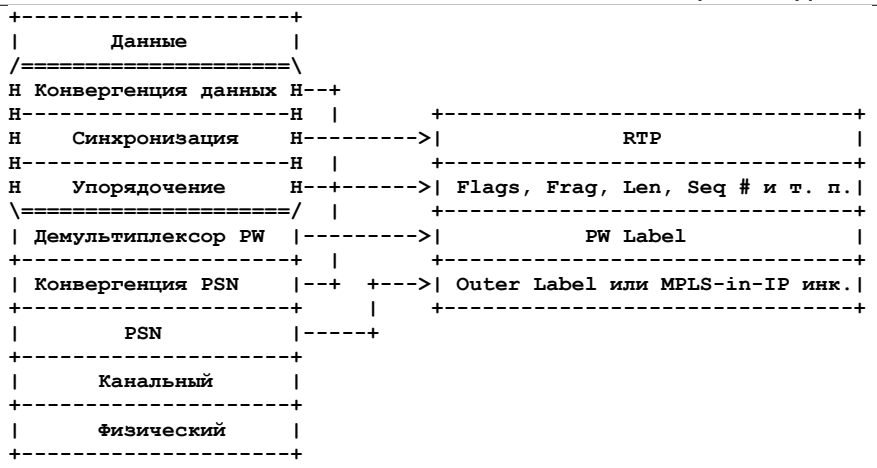

Рисунок 11. PWE3 через MPLS PSN - использование Control Word.

В некоторых сетях может потребоваться передача PWE3 через MPLS, который, в свою очередь, передаётся через IP. В таких случаях PW инкапсулируются для передачи через MPLS, как описано выше, а после этого к полученным PW-PDU применяются методы доставки MPLS через IP PSN (такие, как GRE [RFC2784], [RFC2890]).

5.4.3. Дискриминация пакетов PW-IP

Для MPLS PSN существует дополнительное ограничение на формат пакетов PW. Некоторые маршрутизаторы с коммутацией по меткам детектируют пакеты IP по 4 первым битам содержимого пакета[1]. Для обеспечения корректной работы эти биты в пакетах PW не должны совпадать с текущим номером версии протокола IP.

6. Уровень демультиплексирования PW и требования к PSN

PWE3 предъявляет три требования к протокольным уровням, используемым для передачи через PSN:

- мультиплексирование;
- фрагментация;
- размер и доставка.

6.1. Мультиплексирование

Задача уровня PW Demultiplexer состоит в том, чтобы обеспечить возможность передачи множества PW через один туннель в целях снижения уровня сложности и экономии ресурсов.

Некоторые типы естественного сервиса способны группировать множество устройств в транк (например, множество ATM VC в одном VP, множество Ethernet VLAN в одном физическом канале, множество DS0 в T1 или E1). PW может использоваться для соединения пары таких транков. Транк в таких случаях будет иметь один идентификатор мультиплексирования.

При использовании меток MPLS для мультиплексирования установка значения TTL [RFC3032] в поле PW определяется приложением.

6.2. Фрагментация

Если PSN обеспечивает достаточную производительность фрагментации и сборки пакетов, эти функции можно использовать для разбиения крупных PW PDU в соответствии с MTU. Более подробное описание вопросов фрагментации и сборки пакетов приведено в параграфе 5.3.

6.3. Размер и доставка

Доставка PDU на принимающее устройство PE является функцией уровня PSN.

Если нижележащая PSN не предоставляет информации, достаточной для определения размера PW-PDU, этот размер должен задаваться уровнем инкапсуляции (Encapsulation Layer).

6.4. Проверка PW-PDU

Общепринятым является использование механизмов детектирования ошибок (например, CRC) для контроля целостности доставляемых кадров. Для конкретных типов сервиса должны задаваться механизмы, которые будут определять, следует ли передавать контрольную сумму через PW или её нужно удалять из PDU, передаваемых в направлении PE-bound и заново рассчитывать для вставки в пакеты, передаваемые в направлении CE-bound.

Первый вариант обеспечивает экономию вычислительных ресурсов, а второй - экономию полосы. Для конкретной реализации выбор может диктоваться аппаратными ограничениями, которые могут не позволить сохранение контрольной суммы.

Для таких протоколов, как ATM и FR, сфера действия контрольной суммы ограничена одним соединением. Это связано с тем, что идентификаторы (например, FR DLCI или ATM VPI/VCI) имеют лишь локальную значимость и изменяются на каждом отрезке пути. Если идентификатор устройства (и, следовательно, контрольная сумма) изменяются в процессе эмуляции PW, разумно будет исключить и заново рассчитать контрольную сумму.

Схему проверки корректности пакетов должны задавать документы для соответствующего протокола.

[1]Номер версии протокола IP. *Прим. перев.*

6.5. Перегрузки

В PSN, используемой для организации PW могут наблюдаться перегрузки (насыщение). Характеристики насыщения зависят от типа PSN, архитектуры сети, конфигурационных параметров и уровня загрузки PSN.

Если известно, что трафик, передаваемый через PW, может использовать TCP (например, контроль пакетов), отбрасывание пакетов PSN будет инициировать ограничение уровня трафика и дополнительных действий по предотвращению перегрузки не требуется.

Если PW работает через сеть PSN, обеспечивающую расширенные возможности доставки, устройствам PE следует вести мониторинг потери пакетов, чтобы гарантировать реальную доставку запрошенного сервиса. Если этого нет, PE следует предполагать, что PSN не обеспечивает гарантии доставки (best-effort) и использовать механизмы предотвращения перегрузок, описанные ниже.

Если гарантий доставки не обеспечивается и нет уверенности в том, что передаваемый трафик может использовать TCP, устройствам PE следует вести мониторинг потери пакетов, чтобы убедиться в том, что потеря не превышает допустимого уровня. Уровень потери считается допустимым, если поток TCP по тому же пути через сеть и при тех же условиях будет достигать средней пропускной способности (измеренной в течение разумного интервала) не меньше той, которая достигнута для потока PW. Это условие можно выполнить путём ограничения скорости в NSP или путём отключения одного или нескольких PW. Выбор одного из этих вариантов будет определяться типом передаваемого трафика. Когда предотвращение перегрузки выполняется путём отключения PW, должен обеспечиваться подходящий механизм предотвращения незамедлительного восстановления сервиса, которое может приводить к возникновению пульсаций загрузки сети.

Сравнение с TCP не может быть выполнено в точности, но оно предназначено лишь для оценки и сравнения порядка величин. Период времени, в течение которого измеряется пропускная способность TCP составляет время кругового обхода для соединения. По сути, это требование говорит о недопустимости развёртывания приложений (использующих PWE3 или другой транспортный протокол) в сети Internet без гарантий (best-effort), которые потребляют произвольную полосу и отличаются по порядку значений от TCP. Один из методов определения допустимой полосы PW описан в [RFC3448].

7. Управление

В этой главе рассматриваются службы управления PWE3.

7.1. Организация и разрыв псевдопроводов

PW должны организовываться до того, как будет активизирован эмулируемый сервис и разрываться после того, как этот сервис станет ненужным.

Организация и разрыв PW могут выполняться по команде оператора через систему управления (management plane) PE, с помощью сигналов организации и разрыва от AC (например, ATM SVC) или с помощью механизма автоматического детектирования.

В процессе организации устройства PE обмениваются информацией (например, определяют возможности партнёра). Сигнальный протокол туннеля может быть расширен для обеспечения механизма, который позволит устройствам PE обмениваться необходимой информацией от имени PW.

Ручная настройка PW вполне допустима и может рассматриваться как специальный случай сигнализации.

7.2. Мониторинг состояния

Некоторые естественные типы сервиса обеспечивают мониторинг состояния. Например, ATM поддерживает для этих целей OAM. Для такого сервиса соответствующий эмулируемый сервис также должен обеспечивать способ мониторинга состояний.

7.3. Уведомления о смене состояния псевдопровода

7.3.1. Уведомления об организации и разрыве PW

Если естественный сервер требует двухсторонней связности, соответствующий эмулируемый сервис может лишь сигнализировать о своей активности, когда PW и туннели PSN (если они используются) функционируют в обоих направлениях.

Поскольку два устройства CE эмулируемого сервиса не являются смежными, могут возникать отказы, при которых одно или оба физических соединения между CE и PE сохранят работоспособность. Например, если на рисунке 2 физическое соединение между CE1 и PE1 будет разорвано, это не окажет влияния физическое соединение между CE2 и PE2, которое продолжит работать. Если устройству CE2 не сообщить об удалённой аварии, оно будет продолжать передачу трафика через эмулируемый сервис устройству CE1 и этот трафик будет отбрасываться PE1. Некоторые типы естественного сервиса поддерживают индикацию сбоев, поэтому при отказе сервиса оба CE получат уведомление. Для такого естественного сервиса соответствующий сервис PWE3 должен обеспечивать механизм уведомления об отказах.

Подобно этому, если естественный сервис имеет механизм уведомления, позволяющим всем затрагиваемым службам менять состояние Down на состояние Up при устранении неполадок в сети, соответствующий эмулируемый сервис должен обеспечивать механизм поддержки таких уведомлений.

Такие механизмы могут уже быть встроены в протокол туннелирования. Например, протокол управления L2TP [RFC2661] [RFC3931] поддерживает такую возможность, а LDP имеет возможность отзывать соответствующие метки MPLS.

7.3.2. Ошибочные соединения и несоответствие типов данных

При использовании PWE3 возможны ошибочные соединения и несоответствие типов данных. Ошибочные соединения могут нарушать целостность системы. Несоответствие типов данных может нарушить работу сети пользователя. В обоих случаях могут возникать проблемы с безопасностью и функционированием.

Для предотвращения проблем могут использоваться службы нижележащего механизма туннелирования и связанного с ним протокола управления. В процессе организации PW происходит обмен идентификаторами PW-TYPE, которые впоследствии используются системой пересылки и NSP для проверки совместимости AC.

7.3.3. Потеря и повреждение пакетов, доставка с нарушением порядка

PW может сталкиваться с потерей пакетов, их повреждением и нарушением порядка доставки на пути через PSN между устройствами PE. Это может оказывать влияние на работу эмулируемого сервиса. Для некоторых типов данных потеря или повреждение пакетов и нарушение порядка доставки могут отображаться на выброс числа битовых ошибок или потерю сигнала несущей в PW. Если естественный сервис имеет механизмы обработки битовых ошибок, соответствующему сервису PWE3 следует поддерживать аналогичные механизмы.

7.3.4. Другие уведомления о состояниях

PWE3 может обеспечивать другие механизмы уведомления о состояниях, если в таких механизмах есть необходимость.

7.3.5. Коллективные уведомления о состоянии

Один инцидент в сети может оказывать влияние на состояния группы эмулируемых служб. Например, при повреждении физического канала (или подсети) между CE и PE все проходящие через это соединение (подсеть) эмулируемые службы также перестанут работать. Возможно, что вся группа таких эмулируемых служб заканчивается на одном устройстве CE. Однако в некоторых ситуациях последствия отказа будут влиять на работу множества CE. Следовательно, желательно обеспечить индикацию об отказе для группы устройств с помощью одного уведомления.

PWE3 может обеспечивать механизм для уведомления об изменении состояния группы эмулируемых устройств. Один из возможных методов связывает с каждым эмулируемым устройством идентификатор группы в процессе организации PW для эмулируемого сервиса. После этого устройства объединяются в группы по значениям таких идентификаторов. В уведомлениях о состоянии идентификатор группы может использоваться для указания всех эмулируемых устройств, относящихся к данной группе. В качестве механизма поддержки идентификаторов группы следует использовать механизм, обеспечиваемый нижележащим сигнальным протоколом туннеля.

7.4. Механизм Keep-Alive

Если естественный сервис поддерживает механизм keep-alive, соответствующий эмулируемый сервис должен обеспечивать для него механизм передачи информации через PW. Прозрачная передача сообщений keep-alive через PW соответствует принципу минимального вмешательства. Однако для точного воспроизведения семантики естественного механизма некоторые PW могут требовать дополнительного решения (например, piggy-backing в сигнальном механизме PW).

7.5. Обслуживание управляющих сообщений естественного сервиса

Некоторые типы естественного сервиса используют управляющие сообщение для обслуживания устройств. Такие сообщения могут передаваться в основной полосе (например, управление потоком данных в Ethernet, управление производительностью ATM, тональная сигнализация TDM) или по отдельному каналу (например, VC-сигнализация в ATM VP, или сигнализация TDM CCS).

Следуя принципу минимального вмешательства, желательно обеспечить минимальное участие устройств PE в сигнализации и поддержке естественного сервиса. Однако принцип минимального вмешательства не отменяет необходимости обеспечения удовлетворительной эмуляции естественного сервиса.

При сквозной передаче управляющих сообщений может оказаться желательной их передача с использованием канала с высоким приоритетом или гарантированной доставкой, обеспечиваемого уровнем де мультиплексирования PW (см. параграф 5.1.2).

8. Управление и мониторинг

В этой главе описана архитектура управления и мониторинга для PWE3.

8.1. Состояния и статистика

Устройствам PE следует сообщать о состоянии интерфейсов и таблицы статистики, чтобы помочь в мониторинге сети и контроле сервисных соглашений (SLA[1]). Обычно счётчики используются для следующих параметров:

- принятые и переданные PW-PDU с учётом ошибок или без такового;
- число потерянных один за другим PW-PDU;
- число сервисных PDU, принятых и переданных через PSN с учётом ошибок или без него (кроме TDM);
- связанные с сервисом счётчики для интерфейсов;
- задержка в одном направлении и её вариации.

Значения этих счётчиков содержатся в связанных с PW базах MIB и в них не следует дублировать существующие в MIB счётчики.

[1]Service-level greements - соглашение об уровне обслуживания.

8.2. Архитектура PW SNMP MIB

В этом параграфе описана общая архитектура SNMP MIB, используемых для управления PW и нижележащей PSN. Задача состоит в создании чёткой картины объединения всех имеющих отношение к делу MIB в одну согласованную схему управления для развёртывания служб PWE3. Отметим, что приведённые ниже имена MIB являются лишь предложенным вариантом и реальные модули, используемые для реализации компонент, совсем не обязаны иметь именно такие имена.

8.2.1. Уровни MIB

Базам SNMP MIB, создаваемым для PWE3, следует соответствовать архитектуре, показанной на рисунке 12. Эта архитектура обеспечивает многоуровневую модульную схему, в которой любой поддерживаемый эмулируемый сервис может быть соединён с любым поддерживаемым типом PSN. Эта модель способствует максимальному использованию уже существующей функциональности. Например, модули MIB уровня эмулируемого сервиса не переопределяют заново существующие модули сервисного уровня для эмулируемого сервиса. Эти модули просто связываются с псевдопроводами, используемыми для доставки эмулируемого сервиса через заданный тип PSN. В результате архитектура PWE3 MIB соответствует общей архитектуре PWE3.

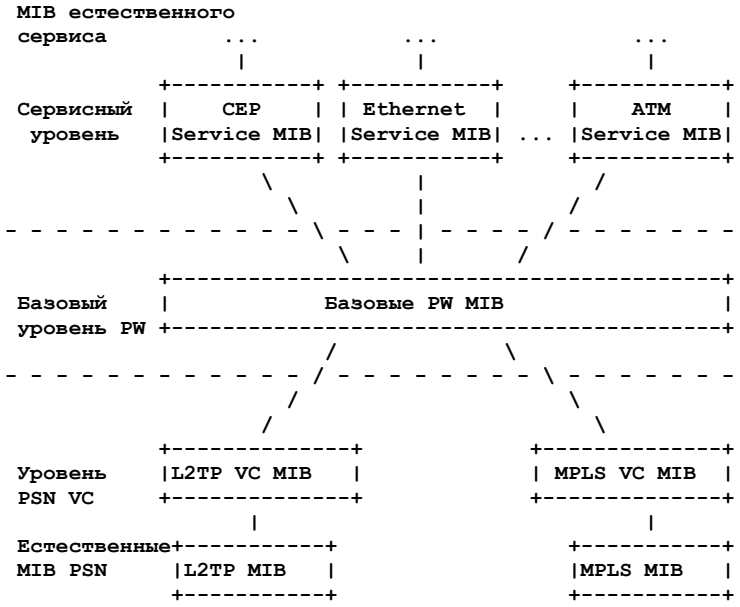

Рисунок 12. Уровни модулей MIB.

Архитектура позволяет добавлять не поддерживаемые типы сервиса или PSN путём простого определения модулей MIB для связи новых типов с существующими. Эти новые модули могут быть впоследствии стандартизованы. Отметим, что существует отдельный модуль MIB для каждого эмулируемого сервиса, а также отдельные модули для каждого типа PSN. Эти модули MIB могут использоваться в различных комбинациях в соответствии с потребностями.

На рисунке 13 показан пример для SONET PW при передаче через MPLS Traffic Engineering Tunnel и LSP с сигнализацией LDP.

8.2.2. Модули сервисного уровня

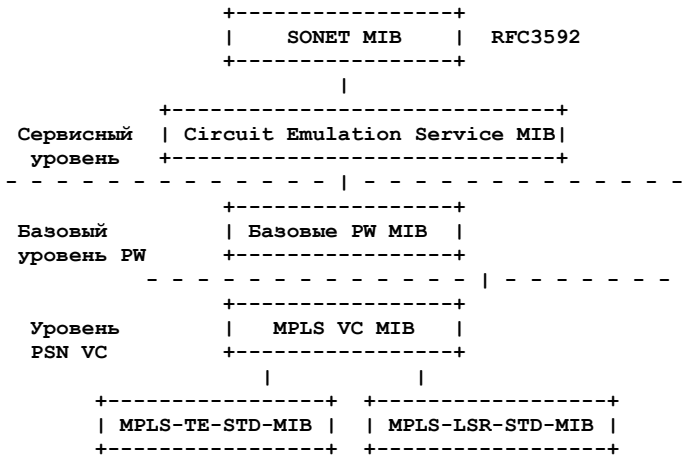

Рисунок 13. Пример для SONET PW через MPLS PSN.

Этот концептуальный уровень модели содержит модули MIB, используемые для представления отношений между эмулируемыми PWE3 службами (такими, как Ethernet, ATM или Frame Relay) и псевдопроводом, используемым для доставки сервиса через PSN. Данный уровень содержит соответствующие модули MIB, используемые для адаптации эмулируемых служб к базовому представлению псевдопровода, которое показано блоком "Базовые PW MIB" на рисунке 13. Рабочей группе PWE3 не следует создавать какие-либо модули MIB для управления базовым сервисом, а следует создать модули, которые обеспечат интерфейс или адаптацию в систему управления PWE3, как показано выше. Например, стандартная база SONET-MIB [RFC3592] разработана и поддерживается другой группой. База SONET-MIB создана для управления естественным сервисом без эмуляции PW. Задачей рабочей группы PWE3

является подготовка стандартов, которые показывают, как эмулировать существующие технологии (такие, как SONET/SDH) через псевдопроводные соединения, а не повторная разработка модулей.

8.2.3. Базовые модули PW MIB

Средний уровень архитектуры носит название “Базовый уровень PW” (Generic PW Layer). MIB этого уровня отвечают за представление связанных с псевдопроводом счётчиков и сервисные модели, используемые для мониторинга и настройки сервиса PWE3 через любую поддерживаемую сеть PSN. Т. е., данный уровень обеспечивает общую модель абстракции PWE3 для решения задач управления. Эта база MIB используется для соединения модулей MIB сервисного уровня с базами PSN VC Layer MIB (см. параграф 8.2.4).

8.2.4. MIB-модули уровня PSN VC

Третий уровень архитектуры управления PWE3 называется уровнем PSN VC. Этот уровень включает базы MIB, которые специально разработаны для связывания псевдопроводов с транспортными технологиями нижележащей PSN, которые передают данные из псевдопровода через сеть PSN. В общем случае это означает, что модуль MIB обеспечивает отображение эмулируемого сервиса, который связан с псевдопроводом через сервисный уровень и базовый уровень PW MIB, на естественный сервис PSN. Например, для случая MPLS требуется, чтобы общий сервис VC отображался на MPLS LSP через MPLS-LSR-STD-MIB [RFC3813] или на туннели TE[1] через MPLS-TE-STD-MIB [RFC3812]. В дополнение к этому может использоваться MPLS-LDP-STD-MIB [RFC3815] для выявления меток MPLS, которые распространяются через MPLS PSN для поддержки сервиса PW. Как сказано выше модули MIB естественного сервиса, используемые для управления естественным сервисом PSN, разрабатываются другими группами, которые разрабатывают естественный сервис PSN. В эти модули MIB следует включать соответствующие механизмы для мониторинга и настройки сервиса PSN, чтобы эмулируемый сервис PWE3 работал корректно.

8.3. Проверка и трассировка соединений

В PW следует поддерживать механизм проверки соединений. Такая проверка, наряду с другими механизмами сигнализации может информировать оператора о том, что в PW прервалась связь с удалённой стороной. Закрытая природа PW означает невозможность в общем случае задать механизм проверки или трассировки соединений, который будет передавать информацию о состоянии устройств CE через PW. Если статус проверки соединения PW нужен устройству CE, он должен отображаться на естественный метод проверки состояния соединения.

В целях поиска неисправностей зачастую желательно иметь точную информацию о функциональном пути PW между устройствами PE. Эта информация обеспечивается средствами трассировки (traceroute) нижележащей сети PSN. Закрытая природа PW означает, что трассировочная информация доступна только в сети провайдера (например, на устройствах PE).

9. Взаимодействие с IANA

Согласование с IANA потребуется для документов PWE3, которые определяют протоколы инкапсуляции, управления и контроля PWE3.

10. Вопросы безопасности

PWE3 не обеспечивает средств защиты целостности и конфиденциальности, а также не гарантирует доставку блоков данных естественного сервиса. Использование PWE3 может, следовательно, подвергать конкретную среду риску угроз безопасности. Допущения, сделанные для случаев, когда все взаимодействующие системы соединены каналами «точка-точка» или через сеть с коммутацией каналов, не могут применяться при соединении устройств эмулируемыми псевдопроводами через некоторые типы PSN. Полный анализ и обзор рисков, связанных с использованием PWE3, выходит за рамки этого документа, особенно в тех аспектах, которые зависят от PSN. Для большей ясности приведём пример. Многие стандарты IETF обеспечивают сравнительно слабые механизмы защиты в предположении, что взаимодействующие узлы соединены между собой через одну локальную сеть. Одним из примеров может служить протокол VRRP[2] [RFC3768]. Сравнительно слабые механизмы защиты представляют более серьезные уязвимости в эмулируемой среде Ethernet, использующей PW-соединения.

Использование уязвимостей со стороны PSN может быть направлено против конечных точек туннеля PW с целью нарушения работы демультиплексора PW и туннеля PSN. Контроль доступа из PSN к конечной точке туннеля PW является одним из способов защиты. Предоставляя доступ к конечной точке туннеля PW лишь легитимным удаленным PE-источникам трафика, устройство PE может отвергать трафик, который будет оказывать вредное воздействие на работу демультиплексора PW и туннеля PSN.

Следует также обеспечить механизм защиты от подмены туннелируемых данных PW. Проверка трафика, адресованного конечной точке демультиплексора PW, является основой защиты целостности инкапсуляции PW. На уровне демультиплексора PW могут использоваться защищённые протоколы (например, IPSec [RFC2401]) для обеспечения аутентификации и целостности данных между конечными точками PW Demultiplexer.

IPSec может обеспечить аутентификацию, а также защиту целостности и конфиденциальности данных, передаваемых между двумя PE. Однако этот протокол не может обеспечить эквивалентные функции для естественного сервиса.

Базируясь на типе передаваемых данных, PW может указывать уровню демультиплексирования PW требуемые функции защиты. Уровень демультиплексирования PW может определить множество профилей защиты на основе требования эмулируемого сервиса. Сигнализация между устройствами CE и управляющие события, эмулируемые PW, а также некоторые типы данных могут потребовать дополнительной защиты. В дополнение к сказанному уровень демультиплексирования PW может использовать аутентификацию партнёра для каждого пакета PSN, чтобы предотвратить подмену естественных блоков данных, передаваемых CE-адресату.

Неограниченная возможность преобразований в NSP может трактоваться как дополнительный риск. На практике тип операций, которые может выполнять NSP, будет ограничен тем набором, который реализован на пути передачи

[1]Traffic-Engineered.
[2]Virtual Router Redundancy Protocol - протокол резервирования виртуальных маршрутизаторов.

данных. Устройства PE, разработанные и управляемые с учётом набранного опыта, будут обеспечивать средства защиты и проверки своей конфигурации и этого будет достаточно для обеспечения подобающего функционирования NSP.

11. Благодарности

Мы благодарим Sasha Vainshtein за работу над NSP и советы в части передачи битовых потоков через PW и Thomas K. Johnson за работу по основам и мотивации PW.

Мы также благодарим Ron Bonica, Stephen Casner, Durai Chinnaiah, Jayakumar Jayakumar, Ghassem Koleyni, Danny McPherson, Eric Rosen, John Rutemiller, Scott Wainner и David Zelig за их комментарии и вклад в работу.

12. Литература

12.1. Нормативные документы

[RFC3931] Lau, J., Townsley, M., and I. Goyret, "Layer Two Tunneling Protocol - Version 3 (L2TPv3), RFC 3931[1], March 2005.

[RFC768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.

[RFC2401] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474[2], December 1998.

[RFC3592] Tesink, K., "Definitions of Managed Objects for the Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) Interface Type", RFC 3592, September 2003.

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

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE", RFC 2890, September 2000.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550[3], July 2003.

12.2. Дополнительная литература

[DVB] EN 300 744 Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television (DVB-T), European Telecommunications Standards Institute (ETSI).

[RFC3815] Cucchiara, J., Sjostrand, H., and J. Luciani, "Definitions of Managed Objects for the Multiprotocol Label Switching (MPLS), Label Distribution Protocol (LDP)", RFC 3815, June 2004.

[RFC3813] Srinivasan, C., Viswanathan, A., and T. Nadeau, "Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base (MIB)", RFC 3813, June 2004.

[MPLSIP] Rosen et al, "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", Work in Progress[4], March 2004.

[RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.

[RFC1958] Carpenter, B., "Architectural Principles of the Internet", RFC 1958, June 1996.

[RFC1981] McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

[RFC2022] Armitage, G., "Support for Multicast over UNI 3.0/3.1 based ATM Networks", RFC 2022, November 1996.

[RFC3768] Hinden, R., "Virtual Router Redundancy Protocol (VRRP)", RFC 3768[5], April 2004.

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, January 2001.

[RFC3448] Handley, M., Floyd, S., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification", RFC 3448[6], January 2003.

[RFC3812] Srinivasan, C., Viswanathan, A., and T. Nadeau, "Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)", RFC 3812, June 2004.

[RFC3916] Xiao, X., McPherson, D., and P. Pate, Eds, "Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)", RFC 3916, September 2004.

13. Соавторы

Ниже перечислены соавторы этого документа.

[1]Этот документ частично обновлён в RFC 5641. *Прим. перев.*
[2]Этот документ частично обновлён в RFC 3260. *Прим. перев.*
[3]Этот документ частично обновлён в RFC 5506. *Прим. перев.*
[4]Работа опубликована в RFC 4023, который обновлён в RFC 5332. *Прим. перев.*
[5]Этот документ заменён RFC 5798. *Прим. перев.*
[6]Этот документ заменён 5348. *Прим. перев.*

Thomas K. Johnson
Litchfield Communications

Kireeti Kompella
Juniper Networks, Inc.

Andrew G. Malis
Tellabs

Thomas D. Nadeau
Cisco Systems

Tricci So
Caspian Networks

W. Mark Townsley
Cisco Systems

Craig White
Level 3 Communications, LLC.

Lloyd Wood
Cisco Systems

14. Адреса редакторов

Stewart Bryant
Cisco Systems
250, Longwater
Green Park
Reading, RG2 6GB,
United Kingdom
EMail: stbryant@cisco.com

Prayson Pate
Overture Networks, Inc.
507 Airport Boulevard
Morrisville, NC, USA 27560
EMail: prayson.pate@overturenetworks.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

Интеллектуальная собственность

IETF не принимает какой-либо позиции в отношении действительности или объема каких-либо прав интеллектуальной собственности (Intellectual Property Rights или IPR) или иных прав, которые, как может быть заявлено, относятся к реализации или использованию описанной в этом документе технологии, или степени, в которой любая лицензия, по которой права могут или не могут быть доступны, не заявляется также применение каких-либо усилий для определения таких прав. Сведения о процедурах IETF в отношении прав в документах RFC можно найти в BCP 78 и BCP 79.

Копии раскрытия IPR, предоставленные секретариату IETF, и любые гарантии доступности лицензий, а также результаты попыток получить общую лицензию или право на использование таких прав собственности разработчиками или пользователями этой спецификации, можно получить из сетевого репозитория IETF IPR по ссылке http://www.ietf.org/ipr.

IETF предлагает любой заинтересованной стороне обратить внимание на авторские права, патенты или использование патентов, а также иные права собственности, которые могут потребоваться для реализации этого стандарта. Информацию следует направлять в IETF по адресу ietf-ipr@ietf.org.

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.