

Network Working Group
Request for Comments: 2816
Category: Informational

A. Ghanwani
Nortel Networks
W. Pace
IBM
V. Srinivasan
CoSine Communications
A. Smith
Extreme Networks
M. Seaman
Telseon
May 2000

A Framework for Integrated Services Over Shared and Switched IEEE 802 LAN Technologies

Интеграция служб в разделяемых и коммутируемых средах ЛВС IEEE 802

Статус документа

Этот документ содержит информацию для сообщества Internet. Документ не содержит каких-либо стандартов Internet. Разрешается свободное копирование данного документа.

Авторские права

Copyright (C) The Internet Society (2000). All Rights Reserved.

Аннотация

Этот документ посвящён вопросам организации интегрированных служб в разделяемых и коммутируемых инфраструктурах ЛВС. Документ включает базовое описание возможностей сетей типа IEEE 802, имеющих отношение к вопросам интегрированного сервиса (задержка доступа, вариации задержек, поддержка очередей в коммутаторах ЛВС). В документе обсуждаются вопросы интеграции служб (IETF Integrated Service), которые не удаётся простыми средствами адаптировать к различным средам ЛВС. Рассмотрена функциональная модель поддержки протокола резервирования ресурсов (RSVP) в таких средах ЛВС. Детальное рассмотрение расширений протокола RSVP для использования в ЛВС можно найти в документе [14]. Отображение различных интегрированных служб на локальные сети IEEE 802 рассмотрено в [13].

Оглавление

Интеграция служб в разделяемых и коммутируемых средах ЛВС IEEE 802.....	1
Статус документа.....	1
Авторские права.....	1
Аннотация.....	1
1. Введение.....	2
2. Структура документа.....	2
3. Определения.....	3
4. Рассылка кадров в сетях IEEE 802.....	3
4.1. Общая модель сервиса IEEE 802.....	3
4.2. Ethernet/IEEE 802.3.....	4
4.3. Token Ring/IEEE 802.5.....	4
4.4. FDDI.....	5
4.5. Demand Priority/IEEE 802.12.....	5
5. Требования и цели.....	5
5.1. Требования.....	5
5.2. Цели.....	6
5.3. Что не рассматривается в документе.....	6
5.4. Допущения.....	7
6. Базовая архитектура.....	7
6.1. Компоненты.....	7
6.1.1. Модуль запросов (RM).....	7
6.1.2. Модуль распределения полосы (BA).....	7
6.1.3. Коммуникационные протоколы.....	7
6.2. Централизованные и распределенные варианты.....	8
7. Модель VM в сети.....	8
7.1. Модель конечной станции.....	9
7.1.1. Модель клиента уровня 3.....	9
7.1.2. Запросы к L2 ISSLL.....	9
7.1.3. Отправитель L3.....	9
7.1.4. Получатель L3.....	10
7.2. Модель коммутатора.....	10

7.2.1. Централизованное выделение полосы.....	10
7.2.2. Распределенное выделение полосы.....	10
7.3. Управление доступом.....	11
7.4. Сигнализация QoS.....	11
7.4.1. Определения клиентских служб.....	12
7.4.2. Определения служб в коммутаторах.....	12
8. Вопросы реализации.....	13
8.1. Характеристики коммутаторов.....	13
8.2. Очереди.....	13
8.3. Отображение сервиса на приоритеты канального уровня.....	13
8.4. Повторное отображение некорректно агрегированных потоков.....	14
8.5. Изменение полученных значений user_priority.....	14
8.6. Различные стили резервирования.....	14
8.7. Неоднородность получателей.....	14
9. Варианты сетевой технологии.....	15
9.1. Полнодуплексные коммутируемые сети.....	16
9.2. Сети Ethernet с разделяемой средой.....	16
9.3. Полудуплексные коммутируемые сети Ethernet.....	16
9.4. Полудуплексные сети Token Ring с коммутируемой и общей средой.....	17
9.5. Сети Demand Priority с коммутируемой и разделяемой средой.....	17
10. Обоснование.....	18
11. Заключение.....	18
Литература.....	18
Вопросы безопасности.....	19
Благодарности.....	19
Адреса авторов.....	19
Полное заявление авторских прав.....	20
Подтверждение.....	20

1. Введение

В сети Internet для доставки трафика адресатам традиционно используется только метод “best effort” (лучшим из возможных способов). Однако последние достижения в области технологий канального уровня (link layer) и растущее число приложений, работающих в реальном масштабе времени (видеоконференции, Internet-телефония), вызвали повышенный интерес к разработке механизмов предоставления услуг в реальном масштабе времени через Internet. Основные требования к таким механизмам изложены в RFC 1633 [8], где содержатся также спецификации различных классов сетевого сервиса, разработанные группой Integrated Services (Интегрированные службы) в рамках IETF - к таким классам относятся Controlled Load (управляемая загрузка) и Guaranteed Service (управляемый сервис) [6,7]. Каждый из этих классов обслуживания предназначен для обеспечения некоторого уровня QoS¹ для трафика, соответствующего заданному набору параметров. Предполагается, что приложения будут выбирать один из таких классов в соответствии с требованиями QoS. Один из механизмов использования конечными станциями таких услуг в IP-сети обеспечивается за счёт сигнального протокола QoS - RSVP² [5], - разработанного группой RSVP в составе IETF. IEEE в своём проекте 802 определяет стандарты для множества различных технологий ЛВС. Все они обычно предлагают однотипные службы дейтаграмм уровня MAC [1] для обслуживания протоколов вышележащих уровней (таких, как IP), хотя динамические характеристики для разных технологий зачастую сильно отличаются (это важно принимать во внимание при рассмотрении вопросов поддержки сервиса в реальном масштабе времени). Ниже будут рассмотрены некоторые характеристики различных технологий ЛВС на MAC-уровне применительно к вопросам интеграции служб. Стандарт IEEE 802 определяет вопросы организации связи между различными сегментами ЛВС с помощью устройств, называемых мостами MAC-уровня (MAC Bridge) или коммутаторами (Switch) [2]. В недавних работах [3,4] были также определены классы трафика, групповая фильтрация (multicast filtering) и поддержка виртуальных ЛВС (VLAN) для таких устройств. Такие технологии ЛВС часто используются на последнем интервале (hop) между пользователями и сетью Internet, а также служат в качестве строительных блоков при создании кампусных сетей. Для использования таких технологий в системах со сквозной поддержкой сервиса в реальном масштабе времени требуется обеспечить механизмы стандартизации. Для решения задач стандартизации требуются механизмы управления ресурсами на уровне логических каналов данных. В контексте данного документа управление ресурсами трактуется как контроль доступа (admission control), планирование, управление трафиком (traffic policing) и т. п. Рабочая группа ISSLL³ в составе IETF была специально создана для разработки таких механизмов стандартизации применительно к различным технологиям канального уровня.

2. Структура документа

Этот документ посвящён вопросам интеграции сервиса в сетях на основе разделяемых и коммутируемых сред Ethernet/IEEE 802.3, Token Ring/IEEE 802.5, FDDI и т. п.. В параграфе 4 рассматриваются возможности различных технологий IEEE 802 на уровне MAC. В параграфе 5 рассмотрены цели, преследуемые при разработке систем с интегрированным сервисом, и предъявляемые к таким системам требования. Функции управления ресурсами, обсуждаемые в параграфе 5 обеспечиваются с помощью менеджера полосы BM (Bandwidth Manager). В параграфе 6 описана архитектурная модель BM, а компонентам этой модели посвящён параграф 7. Некоторые варианты реализации поддержки интегрированного сервиса на канальном уровне рассмотрены в параграфе 8. Вопросы топологии для различных технологий ЛВС с учётом возможностей поддержки интегрированного сервиса рассмотрены в параграфе 9. В данном документе не делается каких либо предположений о топологии сети на канальном уровне, поскольку ставилась задача подготовить максимально краткий документ общего характера - это означает, что часть обсуждаемых здесь функций может не поддерживаться при той или иной топологии. Однако это ограничение не относится к использованию обсуждаемой модели в целом.

¹Quality of Service - класс обслуживания.

²Resource Reservation Protocol - протокол резервирования ресурсов.

³Integrated Services over Specific Link Layers - интегрированные службы в конкретных канальных уровнях.

3. Определения

В настоящем документе и других документах ISSLL используются определённые ниже термины.

Link Layer, Layer 2, L2 (уровень 2, канальный уровень, уровень логического канала)

Технологии уровня логического канала данных (Data link layer) типа Ethernet/IEEE 802.3 и Token Ring/IEEE 802.5 обозначаются как Layer 2 или L2.

Link Layer Domain, Layer 2 Domain, L2 Domain (домен уровня 2, домен канального уровня)

Указывает на набор узлов и соединяющих эти узлы каналов без использования функций уровня L3. Домен L2 может содержать одну или несколько подсетей IP.

Layer 2, L2 Devices (устройства уровня 2, устройства канального уровня)

Устройства, реализующие только функции уровня 2, включая мосты IEEE 802.1D [2] и коммутаторы.

Internetwork Layer, Layer 3, L3 (уровень межсетевого взаимодействия, уровень 3, сетевой уровень)

Устройства, работающие на уровне 3 (сетевой) модели ISO/OSI. В данном документе рассмотрение сетевого уровня ограничивается сетями, использующими протокол IP.

Layer 3 Device, L3 Device, End Station (устройство уровня 3, устройство сетевого уровня, конечная станция)

Хосты и маршрутизаторы, использующие протоколы уровня L3 и выше, или прикладные программы, которым требуется резервирование ресурсов.

Segment (сегмент)

Физический сегмент уровня L2, содержащий одного или нескольких отправителей пакетов. Примерами сегментов могут служить сети Ethernet или Token Ring с разделением доступа к среде на базе CSMA или передачи маркера, а также полнодуплексные соединения между парами станций в коммутируемых средах.

Managed Segment (управляемый сегмент)

Управляемым сегментом будем называть сегмент с DSBM (designated subnet bandwidth manager - менеджер полосы указанной подсети) [14]), обеспечивающим управление допуском к ресурсам для каждого запроса. Управляемый сегмент включает соединённые между собой фрагменты ЛВС с разделяемой средой, которые не разделены с помощью DSBM.

Traffic Class (класс трафика)

Относится к потоку данных, для которых обеспечивается однотипный уровень сервиса в рамках коммутируемой сети.

Subnet (подсеть)

В контексте данного документа термин "подсеть" означает группу устройств сетевого уровня, имеющих одинаковый префикс адреса на сетевом уровне (адрес подсети) для всех сегментов и устройств канального уровня.

Bridge/Switch (мост/коммутатор)

Устройство рассылки пакетов на уровне 2, функционирующее в соответствии со стандартом IEEE 802.1D [2]. В контексте данного документа термины «мост» и «коммутатор» являются синонимами.

4. Рассылка кадров в сетях IEEE 802

4.1. Общая модель сервиса IEEE 802

Пользовательский приоритет (`user_priority`) представляет собой значение, связанное с передачей и приёмом всех кадров в модели сервиса IEEE 802¹. Это значение задаётся отправителем, использующим MAC-сервис, и передаётся вместе с данными приёмнику, также использующему сервис MAC. На самом деле это значение может и не передаваться через сеть в явном виде. Token Ring/IEEE 802.5 передаёт значение приоритета в октете FC, а Ethernet/IEEE 802.3 совсем не передаёт это значение. IEEE 802.12 может передавать или не передавать значение пользовательского приоритета в зависимости от принятого формата кадров. При использовании кадров формата IEEE 802.5 значение приоритета передаётся в явном виде, а при использовании кадров формата IEEE 802.3 поддерживаются только 2 уровня приоритета (высокий и низкий), служащие для определения приоритета доступа (для этого служит значение, помещаемое в стартовый ограничитель кадра IEEE 802.3).

IEEE 802.1D [3] определяет согласованный способ передачи значения `user_priority` через сеть с мостами, включающую Ethernet, Token Ring, Demand Priority, FDDI и другие среды MAC-уровня, путём использования расширенного формата кадров. Описание использования `user_priority` приведено ниже. Интересующимся читателям мы рекомендуем обратиться за дополнительной информацией к тексту стандарта IEEE 802.1D.

Если значение `user_priority` явно передаётся в пакетах, оно используется как простая метка, позволяющая сортировать пакеты в потоке данных по уровню приоритета без детального просмотра пакетов.

Кроме упрощения работы коммутаторов явное задание пользовательского приоритета позволяет, как правило, избежать замены оборудования и программ для реализации системы оценки и классификации пакетов (например, на основе новых протоколов или правил). Более сложные коммутаторы сетевого уровня, которые могут быть использованы в центре сети, способны обеспечить дополнительный сервис за счёт более тонкой классификации пакетов и, следовательно, более эффективного использования сетевых ресурсов и лучшей изоляции между различными потоками трафика.

В спецификациях IEEE 802 не рассматривается вопрос использования поля `user_priority` конечными станциями или сетью. Хотя IEEE 802.1D определяет статически очереди по уровню приоритета в качестве принятого по умолчанию режима работы коммутаторов, способных поддерживать несколько очередей, значение `user_priority` используется достаточно слабо, в силу зависимости от числа классов приоритета, реализованных в конкретном коммутаторе. Параметр `user_priority` определяется как 3-битовое значение, представляющее 8 уровней приоритета (7 для высшего и 0 для низшего приоритета). В общем случае коммутаторы используют следующий алгоритм - пакеты помещаются в очередь с соответствующим классом обслуживания на основе полученного значения `user_priority`, которое извлекается непосредственно из пакета (если используется заголовок IEEE 802.1Q или IEEE 802.5) или выделяется в соответствии с неким локальным набором правил. Очередь выбирается на основе отображения пользовательского приоритета `user_priority` (0 - 7) на номер одного из доступных классов трафика. Коммутатор может поддерживать несколько

¹Оригинальный стандарт IEEE 802.1D [2] содержит спецификации работы мостов уровня MAC. Эти спецификации были расширены с учётом поддержки классов трафика и динамической multicast-фильтрации [3]. В данном документе IEEE 802.1D означает ссылку на оригинальный стандарт [3], если иное не оговорено явно.

различных классов трафика. Для отображения user_priority на класс трафика в коммутаторе могут использоваться анонсируемые параметры IntServ и система управления доступом (admission control) в коммутаторе. Не предполагается реализации в коммутаторе других алгоритмов управления трафиком (типа взвешенных или циклических очередей).

IEEE 802.1D не содержит рекомендаций для отправителя по установке значения user_priority. Одной из основных задач данного документа является разработка таких правил и обсуждение семантики взаимодействия между коммутаторами и конечными станциями в части управления трафиком с различным уровнем приоритета. Далее по тексту документа термины «класс трафика» и user_priority трактуются как синонимы.

4.2. Ethernet/IEEE 802.3

В пакетах Ethernet не используется явная передача класса трафика или поля user_priority. Это означает, что поле user_priority должно быть регенерировано приёмником или коммутатором в соответствии с некоторыми принятыми по умолчанию правилами или на основании информации, содержащейся в полях более высоких уровней. Для явной передачи значения user_priority поверх базового формата кадров MAC может использоваться инкапсуляция IEEE 802.1Q [4].

Для различных вариантов инкапсуляции пакетов IP в сетях Ethernet/IEEE 802.3 необходимо согласовать расчёт управления доступом (admission control) в соответствии с требованиями кадрирования и заполнения, показанными в таблице 1. Значение ip_len в таблице указывает размер пакета IP с учётом его заголовков.

Таблица 1. Инкапсуляция Ethernet.

Инкапсуляция	Заголовок	IP MTU (байтов)
IP EtherType (ip_len<=46)	64-ip_len	1500
(1500>=ip_len>=46 байтов)	18	1500
IP EtherType over 802.1D/Q (ip_len<=42)	64-ip_len	1500 ¹
(1500>=ip_len>=42 байтов)	22	1500 ¹
IP EtherType over LLC/SNAP (ip_len<=40)	64-ip_len	1500
(1500>=ip_len>=40 байтов)	24	1500

4.3. Token Ring/IEEE 802.5

Стандарт Token Ring [6] обеспечивает механизм приоритизации, который можно использовать для управления как очередями на передачу, так и доступом к разделяемой среде. Этот механизм реализован на основе полей AC (Access Control - управление доступом) и FC (Frame Control - управление кадром) в кадре LLC. Первые три бита поля AC - биты приоритета маркера (Token Priority) вместе с тремя последними битами AC (Reservation - резервирование) определяют, какая станция получит доступ к кольцу. Последние три бита поля FC в кадре LLC (User Priority - пользовательский приоритет) принимаются от вышележащего уровня как параметр user_priority при запросе на передачу пакета. Этот параметр также определяет значение Access Priority (приоритет доступа), используемое MAC. Значение user_priority передаётся через сеть (end-to-end) с помощью битов User Priority в поле FC и обычно сохраняется при прохождении через все типы мостов Token Ring. Значение 0 во всех случаях соответствует низшему приоритету.

Token Ring использует также концепцию резервирования приоритета (Reserved Priority), связанную со значением приоритета, используемого станцией при резервировании маркера для следующей передачи в кольцо. Когда свободный маркер проходит по кольцу, только станция, имеющая значение Access Priority, которое больше или равно Reserved Priority в кольце, может забрать маркер для передачи кадра. Более подробную информацию по этим вопросам вы сможете найти в работе [14].

Таблица 2. Рекомендуемые значения пользовательского приоритета для Token Ring.

Приложение	Пользовательский приоритет
Некритично к задержкам	0
-	1
-	2
-	3
Управление ЛВС	4
Критичные к задержкам данные	5
Кадры приложений реального времени	6
MAC-кадры	7

Станция Token Ring теоретически способна поддерживать различные очереди для каждого из восьми уровней пользовательского приоритета user_priority и передавать кадры в соответствии с уровнем приоритета. Станция устанавливает биты Reservation в соответствии со значениями user_priority для кадров, помещённых в очередь на передачу с максимальным приоритетом. Это позволяет механизму доступа к среде обеспечить первоочередную передачу кадров с высоким приоритетом. Приложение Annex I к стандарту IEEE 802.5 Token Ring рекомендует станциям устанавливать уровни приоритета, как показано в таблице 2.

Во избежание проблем, связанных с передачей высокоприоритетного трафика, это приложение рекомендует передавать только один кадр на маркер и ограничивать размер информационного поля значением 4399 байтов при передаче через кольцо чувствительного к задержкам трафика. Большинство существующих реализаций мостов Token Ring пересылают все кадры LLC с принятым по умолчанию значением приоритета доступа (4). Приложение Annex I рекомендует таким мостам пересылать кадры LLC со значением user_priority > 4, сохраняя имеющееся значение user_priority (хотя стандарт IEEE 802.1D [3] позволяет системам сетевого управления менять значение этого поля). Возможности архитектуры Token Ring (такие, как User Priority и Reserved Priority) могут обеспечивать эффективную поддержку интегрированного сервиса для потоков, требующих гарантий QoS.

Для различных вариантов инкапсуляции пакетов IP в сетях Token Ring/IEEE 802.5 необходимо согласовывать расчёты управления доступом (admission control) с параметрами, приведёнными в таблице 3.

¹Отметим, что размер кадров Ethernet. Использующих спецификацию IEEE 802.1Q, превышает максимальный размер кадров 802.3 на 4 байта. Изменение размера MTU для кадров IEEE 802.1Q учтено в новом варианте стандарта IEEE 802.3ac [21].

Таблица 3. Инкапсуляция Token Ring.

Инкапсуляция	Размер заголовка (байт/пакет)	IP MTU (байт)
IP EtherType over 802.1D/Q	29	4370 ²
IP EtherType over LLC/SNAP	25	4370 ²

4.4. FDDI

Стандарт FDDI (Fiber Distributed Data Interface) [16] обеспечивает механизм установки приоритетов, который используется для управления очередями на передачу и доступом к разделяемой среде. Механизм организации приоритетного доступа реализован аналогично механизмам приоритетов Token Ring, описанным выше. Стандарт также задаёт условия для “синхронного” трафика с гарантированным доступом к среде и малыми задержками. Этот режим работы не обсуждается в данном документе, поскольку этим направлением занимается отдельная рабочая группа ISSLL. При обсуждении механизмов QoS далее в этом документе FDDI будет просто трактоваться как технология Token Ring 100 Мбит/с, использующая интерфейсы, которые совместимы с сетями IEEE 802.

4.5. Demand Priority/IEEE 802.12

Стандарт IEEE 802.12 [19] описывает ЛВС с разделяемой средой 100 Мбит/с. Пакеты данных передаются с использованием кадров формата IEEE 802.3 или IEEE 802.5. Протокол MAC носит название Demand Priority (приоритет запросов). Его основной характеристикой в контексте QoS является поддержка двух уровней приоритета на обслуживание - нормального и высокого (пакеты обслуживаются в зависимости от уровня приоритета). Пакеты данных от всех узлов сети (конечных станций, мостов и маршрутизаторов) обслуживаются с использованием циклического алгоритма (round robin).

Если для передачи данных используются кадры формата IEEE 802.3, значение `user_priority` помещается в стартовый ограничитель пакета данных IEEE 802.12. При использовании кадров формата IEEE 802.5 значение `user_priority` дополнительно помещается в биты YYY поля FC в заголовках кадров IEEE 802.5 (см. параграф 4.3). Более того, к сетям IEEE 802.12 применима также инкапсуляция IEEE 802.1Q с её собственным полем `user_priority`. Во всех случаях коммутаторы способны восстановить значение `user_priority`, установленное отправителем.

Такие же правила применимы к отображению IEEE 802.12 `user_priority` в мостах с различными типами сред. Единственным дополнением является то, что для кадров с нормальным приоритетом используются значения `user_priority` от 0 до 4, а для кадров с высоким приоритетом - от 5 до 7. Это обеспечивает отображение принятого по умолчанию в сетях Token Ring значения `user_priority` = 4 для мостов IEEE 802.5 в нормальный приоритет для сегментов IEEE 802.12.

Доступ к среде в сетях IEEE 802.12 является детерминированным. Механизм Demand Priority обеспечивает преимущественное право обработки для пакетов с высоким приоритетом. Если пакет с нормальным приоритетом находится в начале очереди на передачу в течение времени, превышающего PACKET_PROMOTION (200 - 300 мсек) [19], этот пакет автоматически обрабатывается как пакет с высоким приоритетом. Таким образом, пакеты с нормальным приоритетом также обрабатываются с наибольшей возможной скоростью даже при наличии в очереди пакетов с высоким приоритетом и гарантируется время доступа к среде для таких пакетов.

Интегрированный сервис может быть встроено поверх механизма доступа к среде IEEE 802.12. При объединении с механизмами управления доступом (admission control) и распределения полосы (bandwidth enforcement) гарантии задержки в соответствии с требованиями Guaranteed Service (гарантированное обслуживание) могут обеспечиваться без внесения каких-либо изменений в MAC-протокол IEEE 802.12.

Поскольку стандарт IEEE 802.12 поддерживает форматы кадров IEEE 802.3 и IEEE 802.5, для расчёта управления доступом на каналах IEEE 802.12 могут использоваться те же параметры, которые были приведены выше в параграфах 4.2 и 4.3.

5. Требования и цели

В этом разделе описываются требования и цели, которые должны приниматься во внимание при разработке архитектуры интегрированного сервиса в ЛВС. Требования относятся к функциям и возможностям, которые должны поддерживаться, а цели говорят о желаемых функциях и возможностях (они не являются абсолютно необходимыми, в отличие от требований). Многие из требований и целей обеспечиваются функциональностью, поддерживаемой с помощью Integrated Services и RSVP.

5.1. Требования

- **Резервирование ресурсов** - механизм должен обеспечивать резервирование ресурсов на одном или нескольких сегментах, к которым подключены мосты/коммутаторы. Резервирование ресурсов должно обеспечиваться как для индивидуальных (unicast), так и для групповых сеансов. Требуется также обеспечить возможность изменения уровня резервирования во время сеанса.
- **Управление доступом (Admission Control)** - механизм должен обеспечивать оценку уровня ресурсов, требуемого запрашиваемым уровнем QoS для сеанса, чтобы определить возможность допуска такого сеанса. Для решения задач управления полезно обеспечить возможность отклика на запросы о доступности ресурсов. Механизм должен обеспечивать принятие решений о допуске для различных типов сервиса (Guaranteed Service - гарантированное обслуживание, Controlled Load - управляемая загрузка и т.п.).
- **Разделение и планирование потоков** - необходимо обеспечить механизм разделения потоков трафика таким образом, чтобы потоки реального времени имели преимущество перед другими потоками трафика (потоки best effort). Пакеты потока реального времени могут изолироваться, а планирование этих потоков может осуществляться в соответствии с требованиями на обслуживание.

²Предложенное в RFC 1042 [13] значение MTU составляет 4464 байта, но здесь есть вопросы, связанные с определением максимального значения MTU, поддерживаемого между любыми двумя точками внутри подсети Token Ring и между подсетями. Приведённые здесь значения MTU учитывают рекомендации IEEE 802.5 Annex I.

- **Правила и формирование трафика** - формирование и/или управление трафиком должно осуществляться с конечной станции (рабочая станция, маршрутизатор) для обеспечения соответствия согласованным параметрам трафика. Рекомендуется использовать формирование трафика (Shaping) для источника этого трафика. Маршрутизатор, иницирующий сеанс IntServ, должен поддерживать механизм управления трафиком в соответствии с требованиями IntServ, что позволяет обеспечить соответствие для всех потоков трафика, передаваемых маршрутизатором. Механизм IntServ на канальном уровне полагается на корректную реализацию механизмов управления (формирования) трафиком в устройствах вышележащих уровней, способных сделать это. Такое решение необходимо, поскольку мосты и коммутаторы обычно не способны поддерживать для каждого потока информацию о состоянии, требуемую для проверки соответствия потока заданным требованиям. Поддержка правил (Policing) в мостах и коммутаторах по прежнему остаётся лишь опцией - при реализации этой опции её можно использовать для повышения эффективности управления потоком трафика. Этот вопрос более подробно рассмотрен в параграфе 8.
- **Soft State** - этот механизм должен поддерживать информацию о состоянии резервирования. Информация о состоянии должна периодически обновляться для сохраняющихся резервов и удаляться по истечении заданного интервала времени.
- **Централизованная и распределенная реализация** - В случае централизованной реализации один элемент управляет ресурсами в масштабе всей подсети. Такое решение проще в реализации, поскольку для него не требуется модернизация мостов и коммутаторов с учётом поддержки дополнительных функций. Однако в этом случае возникают проблемы с масштабированием (географические размеры подсети и число станций в ней). В полностью распределённых системах каждый сегмент содержит локальный элемент управления ресурсами. Такое решение расширяет возможности масштабирования, однако требует модернизации всех мостов и коммутаторов в сети, поддерживающей данный механизм. Возможны и промежуточные варианты реализации, когда число управляющих ресурсами элементов больше одного, и каждый из таких элементов управляет ресурсами для нескольких сегментов и мостов/коммутаторов. В идеальном случае реализация должна быть гибкой, т. е. для небольших сетей может использоваться централизованное решение, а распределённые системы устанавливаются в больших подсетях. Примеры централизованных и распределённых систем рассмотрены в параграфе 6.
- **Масштабируемость** - механизм и протоколы не должны вносить большой объем служебной информации (overhead) и должны обеспечивать возможности масштабирования, достаточные для обслуживания больших групп приёмников информации так же, как это осуществляется для одного домена канального уровня.
- **Устойчивость к сбоям и восстановление** - этот механизм должен обеспечивать сохранение работоспособности при сбоях, т. е. в системе не должно быть критически важных для работы узлов без резервирования (single point of failure). Например, в централизованной системе должен обеспечиваться механизм резервирования и восстановления при сбоях.
- **Взаимодействие с существующими средствами управления ресурсами** - требуется обеспечить взаимодействие с существующими системами. Например, FDDI поддерживает механизм управления ресурсами "Synchronous Bandwidth Manager". Система управления ресурсами должна быть построена так, чтобы использовать преимущества этого механизма и взаимодействовать с его элементами управления.

5.2. Цели

- **Независимость от протоколов вышележащих уровней** - этот механизм должен (по мере возможностей) быть независимым от протоколов вышележащих уровней (типа RSVP и IP). Независимость от RSVP весьма желательна, поскольку это обеспечивает возможность взаимодействия с сетями, поддерживающими иные протоколы резервирования ресурсов (например, ST2 [10]). Независимость от протокола IP позволит взаимодействовать с сетями на основе других протоколов (IPX, NetBIOS и т. п.).
- **Поддержка разнородных приёмников** относится к групповому взаимодействию (multicast communication), при котором различные приёмники запрашивают разные уровни сервиса. Например, в multicast-группе со множеством приёмников разные приёмники могут запросить различные значения задержки. Снижение задержки может обеспечиваться за счёт расхода дополнительных ресурсов на пути к приёмнику, требующему минимальной задержки, без влияния на резервирование ресурсов для других приёмников. В более сложном случае поддержка неоднородных приёмников означает возможность одновременного предоставления различных уровней обслуживания по запросам разных приёмников. В простейшем варианте поддержка неоднородных приёмников будет позволять варианты, когда одним приёмникам предоставляется гарантированный сервис, а другим - только возможный (best effort service). Поддержка неоднородных приёмников весьма желательна и более подробно рассмотрена в параграфе 8.
- **Поддержка различных стилей фильтрации** - желательно обеспечить поддержку разных стилей фильтрации, определяемых протоколом RSVP (fixed filter - фиксированный фильтр, shared explicit - явное разделение и wildcard - шаблон). Некоторые вопросы поддержки различных стилей фильтрации в домене канального уровня более подробно рассмотрены в параграфе 8.
- **Выбор пути** - в локальных сетях с задаваемой отправителем маршрутизацией (source routed) типа Token Ring/IEEE 802.5 может оказаться полезным механизм выбора пути. Использование такого механизма позволяет более эффективно распределять ресурсы.

5.3. Что не рассматривается в документе

В данном документе описано отображение сервиса на существующие уровни MAC, определяемые стандартами IEEE и ANSI и использующие стандартные службы MAC-уровня (мосты IEEE 802.1). В документе не делается попыток использования или описания иных (фирменных - proprietary или стандартных) протоколов MAC-уровня, хотя следует отметить существование подобных работ (рассмотрение MAC-уровней, подходящих для отображения QoS). Эти вопросы выходят за пределы задач рабочей группы IntServ.

5.4. Допущения

В данном документе предполагается, что сеть, рассматриваемая в плане QoS, будет "богата коммутаторами", т. е. большинство станций, использующих интегрированный сервис, связаны по крайней мере через один коммутатор. Описанные в документе механизмы и протоколы легко распространить на коммуникационные системы с такой же разделяемой средой передачи, но очень важно не упустить рассмотрение вопросов, которые могут играть важную роль в практических приложениях на основе коммутируемых сетей. Принимаются во внимание и разработки в области расширения MAC для обеспечения детерминированной задержки при доступе в сеть (например, IEEE 802.12 [19] и некоторые фирменные реализации).

Хотя для примеров в основном используется протокол RSVP (как протокол вышележащего уровня для сигнализации QoS), на практике могут применяться и другие протоколы. RSVP можно заменить другим динамическим протоколом, а также делать запросы через систему управления или другие средства поддержки правил. Сигнальный протокол SBM [14], основанный на RSVP, разработан специально для архитектур, описываемых в данном документе.

В сети могут использоваться разнородные коммутаторы с разными наборами возможностей (все коммутаторы соответствуют стандарту IEEE 802.1D [2,3]) и отличающиеся механизмами организации и обслуживания очередей (от двух очередей на порт со статическим управлением приоритетами до более сложных систем со множеством очередей и поддержкой WFQ или других алгоритмов). Деление задачи на небольшие независимые фрагменты может обеспечить близкое к оптимальному использование сетевых ресурсов, но мы утверждаем, что такое решение чаще всего даёт лишь незначительное повышение эффективности работы сети. Следовательно, задача состоит в том, чтобы коммутаторы в сети работали с существенно меньшим объёмом информации, нежели машина RSVP в маршрутизаторах. В частности, предполагается, что в коммутаторах не реализованы очереди по потокам и правила (хотя их реализация никак не мешает).

Фундаментом модели IntServ является предположение об изоляции потоков трафика при их передаче через сеть. Предполагается также использование очередей на промежуточных узлах для формирования трафика и управления им с соответствием с заданными спецификациями трафика. В архитектуре, предлагаемой для отображения на канальный уровень, мы отошли от этого допущения в целях обеспечения простоты. Предполагается, что функции формирования/управления трафиком реализованы на конечных станциях. Для сред ЛВС разумно предположить, что конечные станции доверяют условиям согласования трафика на входе в сеть и мы будем просто выделять дополнительные ресурсы для компенсации выбросов трафика (связанных с самой природой трафика ЛВС) в процессе управления допуском к среде. Такое приближение имеет некоторое влияние на типы поддерживаемой однородности приёмников и использование статистического мультиплексирования (особенно для потоков Controlled Load). Данные вопросы рассматриваются в параграфе 8.7.

6. Базовая архитектура

Описанные в параграфе 5 функциональные требования реализуются с помощью объекта, названного менеджером полосы - BM (Bandwidth Manager). Менеджер BM отвечает за предоставление протоколам вышележащих уровней механизма запросов качества обслуживания (QoS) от сети. Компоненты менеджера полосы BM описаны ниже.

6.1. Компоненты

6.1.1. Модуль запросов (RM)

Модуль запросов (Requester Module - RM) устанавливается на каждой станции подсети. Одной из функций модуля является организация интерфейса между приложениями или протоколами вышележащих уровней (RSVP, ST2, SNMP и т. п.) и BM. Приложение может обращаться к различным функциям BM, используя примитивы для связи с RM и соответствующие параметры. Для инициирования резерва в домене канального уровня модулю запросов RM нужно передать ряд параметров: желаемый уровень сервиса (Guaranteed Service или Controlled Load), дескрипторы трафика, содержащиеся в Tspec и Rspec, которые задают количество резервируемых ресурсов [9]. Дополнительную информацию об используемых параметрах можно найти в работах [6,7,8,9]. Когда для сигнализации на сетевом уровне используется протокол RSVP, эта информация доступна и может быть получена из сообщений RSVP PATH и RSVP RESV¹. В дополнение к указанным параметрам должны быть заданы адреса сетевого уровня для конечных точек. Модуль RM должен преобразовать адреса сетевого уровня в адреса канального уровня и транслировать запрос в подходящий формат, который будет понятен другим компонентам, отвечающим за контроль доступа BM. Модуль RM также отвечает за возврат информации о состоянии запросов, обрабатываемых BM, приложениям или протоколу вышележащего уровня.

6.1.2. Модуль распределения полосы (BA)

Модуль распределения полосы (BA) отвечает за контроль над распределением полосы и поддержку информации о выделенных для подсети ресурсах. Конечная станция может запрашивать различный сервис (например, резервирование полосы, изменение заданного ранее резервирования, запрос о доступности ресурсов и т. п.), запросы обрабатываются модулем BA. Связь между конечными станциями и модулем BA осуществляется с использованием модуля RM. Местоположение модуля BA определяется прежде всего выбранным методом реализации. В централизованных моделях BA может работать на одной станции подсети, а в распределённых моделях функции модуля BA могут разделяться (при необходимости) между всеми конечными станциями и коммутаторами/мостами. Модуль BA также отвечает за принятие решения о способе маркировки протокола (например, на основе решения системы контроля доступа). Модуль BA может показывать модулю RM, что пакеты относящиеся к тому или иному потоку имеет соответствующий уровень приоритета, который отображается в определённый класс трафика.

6.1.3. Коммуникационные протоколы

Должны быть заданы коммуникационные протоколы для связи между различными компонентами BM, включая:

- **связь между протоколами вышележащего уровня и RM:** модуль BM должен определять для приложений примитивы, позволяющие инициировать резервирование, запрашивать BA о доступности ресурсов, менять или

¹ Этот вопрос подробно рассматривается в работе [5].

удалять сделанное ранее резервирование и т. п. Эти примитивы могут быть реализованы как интерфейс API для приложений. Вызывающих функции VM с помощью RM.

- **связь между RM и ВА:** должен быть определён механизм сигнализации для связи между RM и ВА. Протокол задаёт сообщения, которые передаются между RM и ВА для обслуживания различных запросов от объектов вышележащих уровней.
- **связь между равноправными ВА:** если в подсети существует более одного модуля ВА, требуется организация связи между разными модулями ВА. В частности, модули ВА должны иметь возможность распределения между собой ответственности за сегменты, мосты и коммутаторы подсети. При получении запроса на резервирование в домене с множеством модулей ВА эти модули должны иметь возможность корректной обработки таких запросов. Связь между ВА также обеспечивает резервирование и восстановление при сбоях.

6.2. Централизованные и распределенные варианты

Приведены примеры, показывающие расположение компонент менеджера полосы в централизованной и распределенной среде. Отметим, что в любом случае требуется наличие RM на всех оконечных станциях, где нужно делать резервирование. По сути, централизованная или распределенная реализация относится к модулю распределения полосы ВА, отвечающему за резервирование ресурсов и управление доступом. Н приведенном ниже рисунке App обозначает приложение, использующее VM. Это может быть пользовательская программа или процесс вышележащего протокола (например, RSVP).

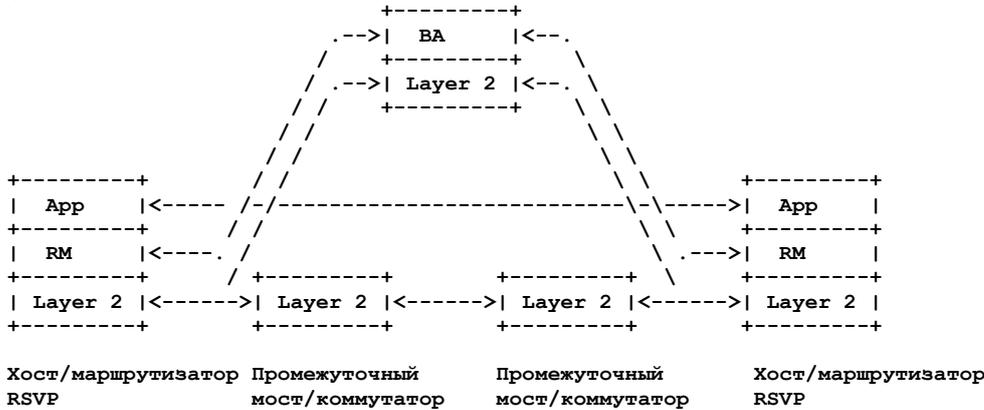


Рисунок 1. Менеджер полосы с централизованным модулем ВА.

На рисунке 1 показана централизованная реализация, где один модуль ВА отвечает за принятие решений по управлению доступом для всей подсети. Каждая оконечная станция включает модуль RM. Промежуточные мосты и коммутаторы в сети не выполняют каких-либо функций VM, поскольку они не принимают участия в управлении доступом. Модуль RM на оконечной станции, запрашивающей резервирование, инициирует взаимодействие с модулем ВА. В больших подсетях один модуль ВА может не справиться с обслуживанием резервов для всей подсети. В таких случаях потребуются развернуть множество модулей ВА, каждый из которых будет управлять ресурсами в одном из не перекрывающихся сегментов подсети. В централизованной реализации модуль ВА должен иметь информацию о топологии подсети на канальном уровне (Layer 2), например, данные об остоном дереве канального уровня (link layer spanning tree) для того, чтобы обеспечить резервирование ресурсов по сегментам подсети. Без такой информации модуль VM будет резервировать ресурсы во всех сегментах для каждого потока в коммутируемой сети, что на практике приведёт к неэффективному использованию ресурсов.

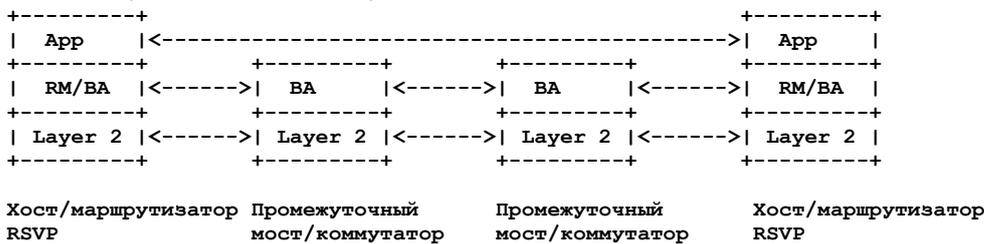


Рисунок 2. Менеджер полосы с распределенным модулем ВА.

На рисунке 2 показан вариант с полностью распределенным менеджером полосы. В этом случае все устройства подсети поддерживают функциональность VM. Для каждого оконечного по-прежнему требуется наличие RM. Кроме того, все станции активно участвуют в контроле доступа. В этой модели каждому модулю ВА достаточно иметь локальную топологическую информацию, поскольку он отвечает лишь за ресурсы того сегмента, к которому подключён напрямую. Эта топологическая информация (такая, как список активных портов на остоном дереве и доступных через них индивидуальных адресов) доступна в современных коммутаторах. Отметим, что на приведённых выше рисунках стрелки между однотипными уровнями служат для индикации логических соединений.

7. Модель VM в сети

В этом разделе описано, как рассмотренная выше модель вписывается в существующую схему интегрированных услуг IETF (Integrated Services) на хостах и маршрутизаторах IP. Сначала описываются реализации уровня 3 на хостах и маршрутизаторах. Затем рассматривается применение этой модели к коммутаторам уровня 2. Показаны различия между централизованной и распределенной реализацией. В тексте встречаются ссылки на термины, заимствованные из спецификации Subnet Bandwidth Manager [14].

7.1. Модель конечной станции

7.1.1. Модель клиента уровня 3

Предполагается использование такой же клиентской модели, как в IntServ и RSVP, где термин «клиент» обозначает элемент (сущность) обслуживающий QoS в устройстве уровня 3 на каждом конце домена уровня 2. В этой модели передающий клиент отвечает за локальное управление доступом и планирование передачи пакетов в канал в соответствии с согласованным уровнем обслуживания. Как и в модели IntServ это включает управление на уровне потока с возможностью применения политики/формирования трафика на каждом генерирующем трафик узле.

Предполагается, что клиент запускает процесс RSVP, который предоставляет приложениям интерфейс для организации сессий, обеспечивает сигнализацию через сеть, программирует планировщик и классификатор пакетов в драйвере, а также обеспечивает интерфейс с модулем управления политикой. В частности, RSVP поддерживает интерфейс с локальным модулем управления доступом, который и рассматривается в данном разделе.

На рисунке из спецификации RSVP показан процесс RSVP на передающих хостах.

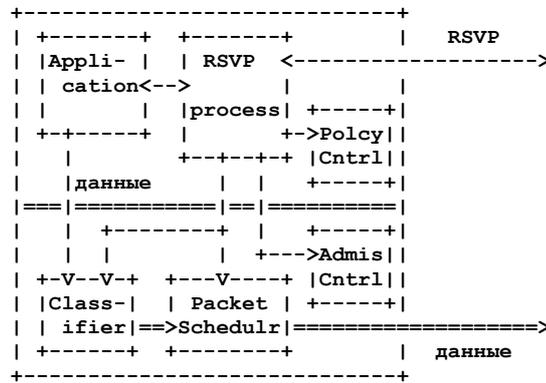


Рисунок 3. RSVP в передающих хостах.

7.1.2. Запросы к L2 ISSLL

Локальный элемент контроля доступа на клиенте отвечает за отображение запросов на организацию сессий уровня 3 на семантику уровня 2.

Вышележащий объект делает запрос в обобщённых терминах ISSLL вида:

"Могу я зарезервировать для трафика с <traffic characteristic> и <performance requirements> из <here> в <there> и как нужно помечать такой трафик?"

где:

- <traffic characteristic> = Sender Тspec (например, полоса, пиковый уровень, MTU)
- <performance requirements> = FlowSpec (например, задержка, границы её вариаций)
- <here> = IP-адрес (a)
- <there> = IP-адрес (a), который может быть групповым.

7.1.3. Отправитель L3

Функциональность ISSLL у отправителя показана на рисунке 4.

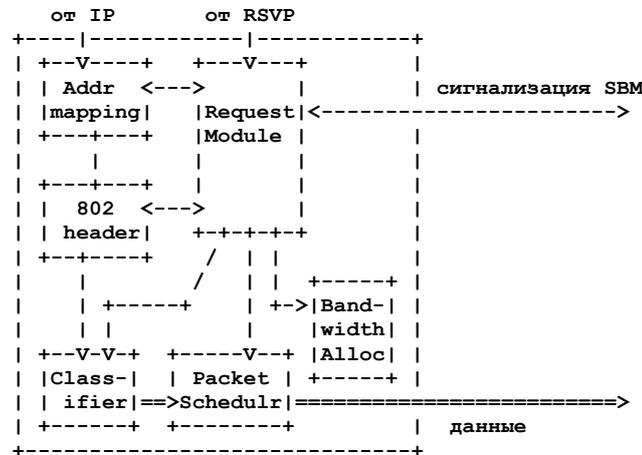


Рисунок 4. ISSLL на передающей конечной станции.

Функции модуля запросов RM перечислены ниже:

- Отображение конечных точек соединения на адреса уровня 2 в ЛВС, чтобы клиент мог определить, куда передавать трафик. Эта функция может ссылаться на кэш ARP для индивидуальных адресов или выполнять алгоритмическое отображение для групповых адресов.
- Взаимодействие с локальным модулем ВА для принятия решений в части контроля доступа.
- Форматирование запросов SBM в сеть с отображёнными адресами и спецификациями фильтров и потоков.
- Получение откликов из сети и передавать решения по управлению доступом вышележащим сетевым элементам вместе с согласованными изменениями параметров сессий.

- Сохранение возвращённых значений `user_priority` для их связывания с данной сессией в таблице заголовков 802. Это будет применяться при создании заголовков уровня 2 для последующих пакетов данных этой сессии. Таблица может индексироваться, например, по идентификаторам потока RSVP.

Модуль распределения полосы (BA) присутствует только в реализациях с распределенной моделью BA. При наличии такого модуля его функция заключается в основном в применении локального контроля доступа к полосе исходящих каналов и очередям драйверов.

7.1.4. Получатель L3

Функциональность ISSLL на приёмной стороне проще (рисунок 5):

- Обработка всех полученных индикаций протокола SBM.
- Связь со всеми локальными BA в части принятия решений по контролю доступа.
- Передача индикаций в RSVP для нормальных ситуаций.
- Восприятие подтверждений от RSVP и передача их через сигнальные механизмы SBM в сторону запрашивающего.

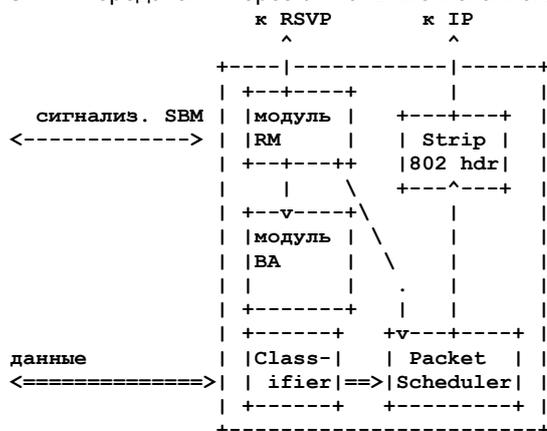


Рисунок 5. ISSLL на приёмной конечной станции.

- Возможность программирования приёмного классификатора и планировщика (если он применяется) для идентификации и соответствующей трактовки классов трафика (например, резервирование буферов для определённого класса).
- Программирование получателя на «вырезание» информации из заголовков канального уровня в принятых пакетах.

Модуль BA, присутствующий только в распределенных реализациях, применяет локальный контроль доступа, чтобы видеть, может ли запрос быть поддержан с имеющимися на приёмной стороне ресурсами.

7.2. Модель коммутатора

7.2.1. Централизованное выделение полосы

В централизованных реализациях BA коммутаторы не участвуют в процессе управления доступом. Этот контроль реализуется центральным модулем BA (например, SBM¹), как описано в [14]. В реальности такой модуль BA может быть совмещён с коммутатором, но его функции не обязательно будут близки к функциям пересылки коммутатора как в описанной ниже модели с распределенным BA.

7.2.2. Распределенное выделение полосы

Описанная здесь модель поведения коммутатора уровня 2 использует терминологию протокола SBM, как примера протокола управления доступом. Эта модель применима и с другими механизмами (например, статической конфигурацией или сетевым управлением) управления доступом. В коммутаторе определяются следующие объекты:

- Локальный модуль управления доступом. Один экземпляр на каждом порту учитывает доступную полосу подключённого к порту канала. Для полудуплексных соединений учитываются ресурсы, выделяемые для обоих направлений. Для полнодуплексных каналов задача учёта на входе становится тривиальной.
- Входной модуль SBM. Один экземпляр на каждом порту реализует «сетевую» часть сигнального протокола для партнерских отношений с клиентами или другими коммутаторами. Он также поддерживает информацию об отображении классов IntServ на `user_priority`.
- Модуль распространения SBM. Передаёт запросы, полученные контролем доступа на входных портах, соответствующим модулям SBM на выходных портах. Это требует доступа к таблице пересылки коммутатора и данным о состоянии остовного дерева для портов.
- Выходной модуль SBM. Пересылает запросы на следующий интервал уровня 2 или 3.
- Модуль классификации, очередей и планировщика. Функции этого модуля совпадают с описанными для процесса пересылки (Forwarding Process) стандарта IEEE 802.1D (параграф 3.7 в [3]). Модуль классификатора идентифицирует имеющую отношение к делу информацию QoS во входящих пакетах и использует её вместе с обычной базой пересылки моста для выбора выходного порта и класса трафика при постановке пакета в очередь. Различные типы коммутаторов будут использовать разные методы идентификации потоков (см. параграф 8.1). В коммутаторах IEEE 802.1D эта информация представляет собой регенерированный параметр `user_priority`, который

¹Subnet Bandwidth Manager - менеджер полосы для подсети.

уже был декодирован принимающим сервисом MAC и мог быть отображён процессом пересылки (см. параграф 3.7.3 в [3]). Это не исключает более изощрённой классификации (например, классификации отдельных потоков IntServ). Модули Queue и Scheduler реализуют выходные очереди для портов и обеспечивают алгоритмы обслуживания очередей для обеспечения сервиса IntServ. Коммутаторы поддерживают одну или несколько очередей для каждого порта и, по крайней мере, базовый механизм приоритизации в соответствии с IEEE 802.1D.

- **Модуль отображения классов и политики для входного трафика.** Функции этого модуля описаны в параграфе 3.7 стандарта IEEE 802.1D. Этот необязательный модуль может распределять входящий трафик по классам в соответствии с согласованными параметрами, а также может отбрасывать пакеты или менять отображение user_priority. По умолчанию модуль пропускает трафик без изменений.
- **Модуль отображения классов для выходного трафика.** Функции этого модуля описаны в параграфе 3.7 стандарта IEEE 802.1D. Этот необязательный модуль может заново отображать классы трафика для каждого выходного порта. По умолчанию модуль пропускает трафик без изменений.

На рисунке 6 показаны все модули поддерживающего ISSLL коммутатора. Модель ISSLL является надмножеством модели моста IEEE 802.1D.

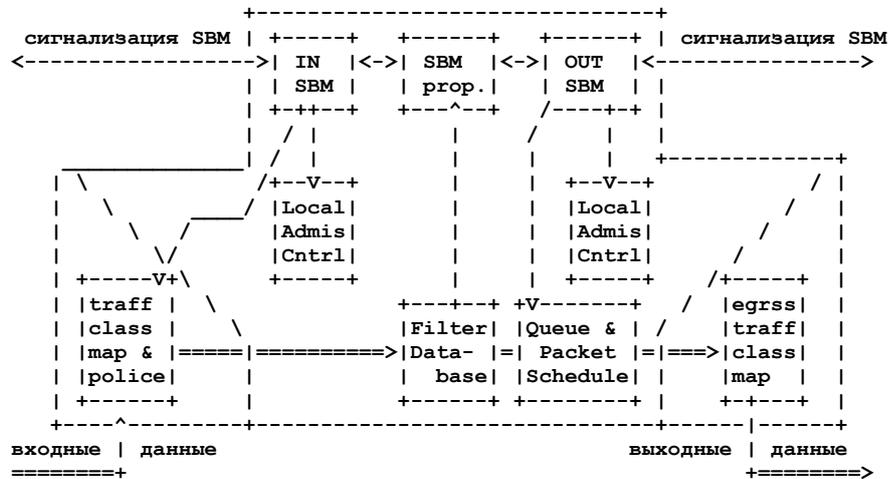


Рисунок 6. ISSLL в коммутаторе.

7.3. Управление доступом

При получении запроса на управление доступом, коммутатор выполняет перечисленные ниже действия (в качестве примера используется протокол SBM). Поведение будет меняться в зависимости от наличия для данного сегмента означенного (Designated) SBM. Более подробное описание действий DSBM/SBM приведено в [14].

- Если входной SBM является Designated SBM для этого канала, он транслирует все принятые значения user_priority или выбирает класс трафика уровня 2, совместимый с запросом, использование которого не нарушает действующих административных правил. По сути, это будет выбором «лучшего» класса из числа доступных и соответствующих запросу. При успешном резервировании это гарантирует передачу клиенту значения user_priority, соответствующего классу трафика.
- Входной DSBM наблюдает состояние распределения ресурсов на входном порту/канале и определяет возможность выделения новых ресурсов для отображённого класса трафика. Если запрос принимается, он будет передан распространителю резервирования (reservation propagator).
- Если входной SBM не является Designated SBM для канала, тогда он напрямую передаёт запрос распространителю резервирования.
- Распространитель резервирования передаёт запрос системе учёта полосы на каждом выходном канале коммутатора, к которому это резервирование может относиться. Это предполагает интерфейс с таблицей маршрутизации/пересылки.
- Система учёта исходящей полосы наблюдает текущее состояние распределения ресурсов очередей на выходном порту и полосы канала для определения возможности нового выделения ресурсов. Отметим, что это решение является единственным локальным решением для данного коммутатора, при дальнейшем пути через сеть L2 другие коммутаторы могут отвергнуть этот запрос.
- Будучи принятым данным коммутатором, запрос распространяется на все выбранные исходящие каналы. Любое значение user_priority, указанное в пересылаемом запросе, должно транслироваться в соответствии с таблицей исходящих отображений.
- Если запрос принят, коммутатор должен сообщить клиенту значение user_priority для использования с пакетами, относящимися к данному потоку. Следует отметить, что и в этом случае нисходящие коммутаторы могут отвергнуть запрос.
- Если коммутатор хочет отвергнуть запрос, он может сделать это, уведомив запросившего клиента по адресу L2.

7.4. Сигнализация QoS

Описанные в этом документе механизмы используют сигнальный протокол на устройствах для передачи их запросов на управление доступом через сеть. Ниже приведены определения услуг, обеспечиваемых таким протоколом (например, [14]). Показаны примитивы и информация, которой нужно обмениваться с помощью сигнального протокола. Во всех примерах также указываются механизмы удаления/сброса для разрыва организованных сессий.

7.4.1. Определения клиентских служб

Перечисленные ниже интерфейсы можно видеть на рисунках 4 и 5.

- **SBM <-> Отображение адресов**

Это простая функция просмотра, которой может потребоваться взаимодействие с протоколом ARP или алгоритмическим отображением. Адреса уровня 2 нужны SBM для включения в сигнальные сообщения, чтобы избежать необходимости использования коммутаторами данных уровня 3 для отображения.

```
l2_addr = map_address( ip_addr )
```

- **SBM <-> Заголовок сессии/канального уровня**

Эта функция служит для уведомления пути передачи о способе добавления данных из заголовков L2 (например, значений user_priority) к трафику каждого исходящего потока. Путь передачи будет предоставлять значение user_priority при запросе для каждого пакета передачи на уровне MAC. Параметр user_priority является одним из передаваемых в пакете примитивов, определённых сервисной моделью IEEE 802.

```
bind_l2_header( flow_id, user_priority )
```

- **SBM <-> Классификатор/планировщик**

Эта функция служит для уведомления передающего классификатора/планировщика о любой дополнительной информации L2, связанной с планированием передачи потока пакетов. Данный примитив может не использоваться в некоторых реализациях или может служить, например, для предоставления планировщику пакетов информации о необходимости выполнения операций на уровне класса трафика в дополнение к операциям на уровне потока пакетов, обусловленной требованиями IntServ; заголовок L2 может служить образцом (в дополнение к FilterSpec) для идентификации потока трафика.

```
bind_l2schedulerinfo( flow_id, , l2_header, traffic_class )
```

- **SBM <-> Локальное управление доступом**

Эта функция применяется локальной системой контроля доступа для сессии (например, для решения вопроса о наличии достаточной для новой сессии полосы, свободных буферов и т. п.). При положительном ответе следует выделять требуемые ресурсы; при отказе на последующих узлах все выделенные ресурсы должны быть освобождены. Вызывать эту функцию может, например, Designated SBM для сегмента.

```
status = admit_l2session( flow_id, Tspec, FlowSpec )
```

- **SBM <-> RSVP**

Эта функция рассмотрена в параграфе 7.1.2 и полностью описана в [14].

- **Интерфейсы управления**

Некоторые или все описанные в этой модели модули будут требовать настройки конфигурации. Предполагается, что детали объектов управления будут разработаны в будущем группой ISSLL WG.

7.4.2. Определения служб в коммутаторах

Перечисленные ниже интерфейсы показаны на рисунке 6.

- **SBM <-> Классификатор**

Служит для уведомления приёмного классификатора о способе связывания входной информации L2 с классом трафика. В некоторых случаях может состоять из группы неизменных отображений, используемых по умолчанию.

```
bind_l2classifierinfo( flow_id, l2_header, traffic_class )
```

- **SBM <-> Планировщик очередей и пакетов**

Служит для уведомления передающего планировщика о дополнительной информации L2, связанной с данным классом трафика. В некоторых случаях может не использоваться (см. предыдущий параграф).

```
bind_l2schedulerinfo( flow_id, l2_header, traffic_class )
```

- **SBM <-> Локальное управление доступом**

Идентично описанной выше функции для хоста.

- **SBM <-> Отображение класса трафика и правила**

Необязательная настройка любого повторного отображения user_priority, которое может быть выполнено на входе и выходе портов коммутатора. Для коммутаторов IEEE 802.1D такие отображения будут согласованными на всех портах.

```
bind_l2ingressprimap( inport, in_user_pri, internal_priority )
bind_l2egressprimap( outport, internal_priority, out_user_pri )
```

Необязательная любой функции исполнения правил L2 будет применяться на уровне класса трафика, для пакетов, соответствующих заголовку L2. Если коммутатор может управлять правилами на уровне потока, существующие модели IntServ/RSVP будут обеспечивать определение сервиса для такой настройки.

```
bind_l2policing( flow_id, l2_header, Tspec, FlowSpec )
```

- **SBM <-> База данных фильтрации**

Правила распространения SBM требуется доступ к базе данных пересылки L2 для решения вопросов о пересылке сообщений SBM. Это похоже на интерфейс RSRR в L3 RSVP.

```
output_portlist = lookup_l2dest( l2_addr )
```

- **Интерфейсы управления**

Некоторые или все описанные в этой модели модули будут требовать настройки конфигурации. Предполагается, что детали объектов управления будут разработаны в будущем группой ISSLL WG.

8. Вопросы реализации

Как было отмечено выше, рабочая группа Integrated Services определила различные классы сервиса, предлагающие разные уровни гарантий QoS. На начальном этапе работа концентрировалась вокруг классов Controlled Load¹ [6] и Guaranteed Service² [7]. Сервис Controlled Load обеспечивает слабые гарантии и неформально его можно выразить словами «как в модели best effort для незагруженной сети». Сервис Guaranteed Service ограничивает максимальную задержку пакетов в сети. Расширения, которые эти службы могут поддерживать на канальном уровне, зависят от множества факторов, включая топологию сети и используемые технологии. Некоторые вопросы отображений рассматриваются ниже с учётом развития стандартов канального уровня и функций, поддерживаемых протоколами вышележащих уровней. С учётом присущих некоторым топологиям ограничений реализация всех требований Integrated Services для заданной топологии может оказаться невозможной. В таких случаях целесообразно рассмотреть приближение, которое может быть реализовано на практике. Например, управление трафиком на всех элементах сети в соответствии с требованиями Controlled Load может оказаться невозможным. Но эту задачу можно оставить окончательным станциям, что обеспечит достаточно хорошее приближение к желаемому сервису.

8.1. Характеристики коммутаторов

Имеется множество мостов/коммутаторов ЛВС с существенно различающимися возможностями поддержки QoS. Ниже рассмотрены различные типы устройств, которые можно встретить в средах ЛВС.

Наиболее широко распространены коммутаторы, соответствующие спецификации IEEE 802.1D 1993 г. [2]. Такие устройства поддерживают по одной очереди на порт и используют алгоритм spanning tree для предотвращения петель. В построенных на таких устройствах сетях не следует ожидать каких-либо гарантий сервиса по причине полного отсутствия возможностей изоляции трафика.

Следующий уровень образуют мосты, коммутаторы, соответствующие пересмотренному стандарту IEEE 802.1D [3]. Они поддерживают до восьми очередей для разных классов трафика. Уровень изоляции трафика достаточно груб, поскольку все потоки одного класса объединяются. Кроме того, возможно отображение нескольких уровней приоритета на один класс трафика в зависимости от числа реализованных в коммутаторе очередей. Такому устройству будет сложно обеспечить защиту от потоков с некорректным поведением. Область группового трафика может быть ограничена с помощью GMRP лишь теми сегментами, которые входят в путь к интересующим получателям.

Следующий уровень образуют коммутаторы, мосты, реализующие необязательные функции IEEE 802.1D типа отображения принятых значений user_priority на некий внутренний набор канонических значений по входным портам. Коммутаторы могут также поддерживать отображение этих внутренних канонических значений на передаваемые значения user_priority по выходным портам. С помощью таких дополнительных средств администраторы могут организовать отображения классов трафика между парами конкретных портов, что обеспечивает более гибкий контроль доступа в защищённые классы трафика.

К другим опциональным возможностям мостов и коммутаторов относится поддержка классификации потоков IntServ с использованием полей заголовков сетевого уровня, применение правил к потокам и формирование трафика для поддержки Guaranteed Service, а также более изощрённые алгоритмы планировщиков типа взвешенных очередей для ограничения полосы конкретного класса трафика. Отметим, что эти возможности обеспечивают преимущества в плане изоляции трафика и управления каждым потоком для поддержки сервиса Controlled Load и Guaranteed Service.

8.2. Очереди

Пакетные сети без явной организации соединений (Connectionless), в целом, и ЛВС, в частности, работают благодаря масштабированию сетевого сервиса. Обычно в сети предоставляется избыточная полоса и буферы для потребления трафика, генерируемого протоколами вышележащих уровней, зачастую достаточные для того, чтобы перегрузки в сети были редкими и кратковременными, а средняя нагрузка была достаточно мала.

С появлением критичного к задержкам трафика предложить клиента избыточные ресурсы стало сложнее. Критичные к задержкам пакеты могут занимать очередь продолжительное время (не только в период выбросов), создавая проблемы для остальных пакетов (особенно в узких местах сети типа переходов 100 или 10 Мбит/с между коммутатором в стойке и настольным коммутатором, к которому подключены пользователи). В таких случаях, если заранее известно (особенности приложений, статистика или административный контроль), что доля критичного к задержкам трафика достаточно мала, целесообразно обеспечить ему приоритет перед остальным трафиком. При этом в самом неблагоприятном случае задержка критичного трафика составит не более времени передачи одного кадра (меньше 1 мсек для Ethernet 10 Мбит/с) и будет существенно меньше задержки, воспринимаемой человеком.

Когда сетевой элемент предлагает более одного приоритетного сервиса (например, поддержка Controlled Load и Guaranteed Service), требования к дисциплине планирования усложняются. Для обеспечения требуемой изоляции между классами обслуживания может потребоваться организация отдельных очередей. Возникает проблема обслуживания очередей, для которых требуется комбинация контроля доступа с более интеллектуальными дисциплинами очередей. Как спецификации сервиса, так и спецификации алгоритмов управления очередями выходят за пределы данного документа.

8.3. Отображение сервиса на приоритеты канального уровня

Число поддерживаемых классов трафика и методов доступа рассматриваемой технологии будет определять число и тип поддерживаемых услуг. Например, в сетях Token Ring/IEEE 802.5 поддерживается 8 уровней приоритета, которые могут отображаться на один или множество классов трафика. Сети Ethernet/IEEE 802.3 не поддерживают в кадрах сигнализации уровней приоритета. Однако комитет по стандартизации IEEE 802 недавно разработал новый стандарт для мостов и коммутаторов, поддерживающий передачу multimedia-трафика и динамическую фильтрацию при групповой передаче (multicast) [3]. Формат передачи полей user_priority во всех средах ЛВС IEEE 802 LAN в настоящее время определён в документе [4]. Эти стандарты позволяют во всех средах поддерживать до 8 классов трафика. Передаваемые в кадрах биты user_priority отображаются на конкретный класс трафика в мостах и коммутаторах. Поле user_priority обеспечивает сквозную сигнализацию, если коммутаторы/мосты на пути передачи не меняют его значение.

¹Управляемая нагрузка.

²Гарантированное обслуживание.

Класс трафика, используемый для потока, зависит от запрошенного качества обслуживания и результатов резервирования. Следовательно, отправителю нужно использовать значение `user_priority`, которое отображается на класс `best effort` (по возможности), если иное не задано VM. Модуль VM при успешном резервировании задаёт значение `user_priority`, которое отправитель будет применять для данных сессии. В документе [13] решается задача отображения разных интегрированных служб (Integrated Service) на подходящие классы трафика.

8.4. Повторное отображение некорректно агрегированных потоков

Другой темой обсуждения в контексте IntServ является обслуживание трафика для потоков данных от источников, генерирующих избыточный по сравнению с контрактом трафик. Решение, имеющие некоторые перспективы, заключается в трактовке такого трафика, как имеющего сниженный приоритет¹ для защиты трафика, которому обычно предоставляется сервис `Best effort`. Трафик `Best effort` часто является адаптивным и использует те или иные механизмы контроля перегрузок, и будет некорректно «наказывать» такие потоки в результате «плохого поведения» зарезервированных потоков, которые зачастую организуются неадаптивными приложениями.

Возможным решением является присвоение обычного класса `Best effort` для одного из `user_priority` и отнесение избыточного не соответствующего требованиям трафика к более низкому уровню `user_priority`, хотя возможное нарушение порядка в результате такого подхода может оказаться нежелательным, особенно для потоков TCP. По этой причине для сервиса `controlled load` рекомендуется отбрасывать избыточный трафик вместо снижения приоритета для него. Этот вариант будет рассмотрен ниже.

8.5. Изменение полученных значений `user_priority`

В некоторых случаях сетевой администратор может не доверять содержащимся в пакетах значениям `user_priority` и отображать их на более подходящий набор значений. Кроме того, отображение значений `user_priority` может потребоваться в результате присущих оборудованию ограничений или перехода на другой диапазон и будет зависеть от областей сети, с которыми происходит обмен пакетами.

Некоторые коммутаторы могут реализовать такую функцию на входе, отображая полученные `user_priority` на некий внутренний набор значений. Эта функция реализуется с помощью таблицы, которая в стандарте IEEE 802.1D называется `User Priority Regeneration Table`² (таблица 3-1 в [3]). Полученные таким образом значения могут отображаться с использованием описанной выше выходной таблицы с исходящие значения `user_priority`. Те же самые отображения могут использоваться для управления доступом применительно к запросам, использующим `user_priority` (см., например, [14]). Возможны и более изощрённые решения, когда устройство контролирует потоки трафика и устанавливает значения `user_priority` с соответствиями со своей политикой и спецификациями трафика.

8.6. Различные стили резервирования

На рисунке 7 SW показывает мост/коммутатор в домене канального уровня. S1, S2, S3, R1 и R2 - оконечные станции, являющиеся членами группы, которая связана с одним потоком RSVP. S1, S2 и S3 - станции восходящего направления, а R1 - R2 - станции нисходящего, которые получают трафик от всех отправителей. RSVP позволяет получателям R1 и R2 задать резервирование, которое может быть применено: (а) к одному заданному отправителю (фиксированный фильтр), (б) любому из явно заданного множества (два или более) отправителей (разделяемый явный фильтр) и (с) любому отправителю в группе (разделяемый шаблонный фильтр). Поддержка фиксированных фильтров проста - для трафика каждого отправителя организуется своё резервирование. Однако поддержка двух других стилей фильтрации подразумевает правила, т. е. для слитых потоков трафика от разных отправителей должны применяться правила, которые соответствуют параметрам трафика, заданным в Rspec фильтра. Дополнительные сложности возникают в тех случаях, когда запросы на обслуживание от R1 и R2 различаются. Следовательно, при отсутствии поддержки правил в мостах/коммутаторах на канальном уровне возможно только резервирование с фиксированными фильтрами.

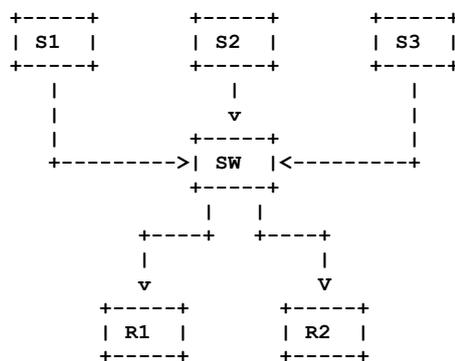


Рисунок 7. Стили фильтрации

8.7. Неоднородность получателей

На уровне 3 модель IntServ поддерживает разнородных получателей для групповых потоков, где разные ветви дерева могут использовать различные типы резервирования для данного группового получателя. В этой модели также поддерживается сосуществование ветвей с резервированием потоков и ветвей с услугами `best effort`. Если мы будем трактовать подсеть уровня 2, как один элемент сети в соответствии с определением [8], все ветви распределительного дерева, лежащие в подсети, могут рассматриваться, как требующие одинаковой трактовки QoS и трактоваться как неделимый элемент (атом) в плане управления доступом и т. п. С учётом такого допущения модель и протоколы, определённые в IntServ и RSVP, уже обеспечивают достаточную поддержку разнородности групп. Отметим, однако, что запросы управления доступом могут отвергаться, поскольку перегрузка одного из каналов в подсети ведёт к отказу при резервировании для всей подсети.

¹В оригинале - *somewhat less than best effort* - несколько меньше, нежели *best effort*. Прим. перев.

²Таблица регенерации пользовательских приоритетов.

В качестве примера рассмотрим рисунок 8, где SW представляет собой устройство уровня 2 (мост или коммутатор), участвующее в резервировании ресурсов, S является станцией-источником восходящего направления, а R1 и R2 - конечными получателями нисходящего направления. R1 хочет организовать резервирование для потока, а R2 устраивает сервис best effort. Станция S передаёт сообщения RSVP PATH, которые являются групповыми для R1 и R2. R1 передаёт сообщение RSVP RESV станции S с запросом резервирования ресурсов.

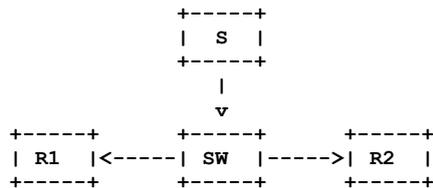


Рисунок 8. Пример разнородности получателей.

При успешном резервировании на уровне 2 кадры, адресованные группе, будут отнесены к классу трафика для сервиса, запрошенного R1. На устройстве SW должны присутствовать некие механизмы, обеспечивающие пересылку трафика в соответствии с зарезервированным классом на интерфейсе в сторону R1, при этом для интерфейса в сторону R2, обеспечивающие класс best effort. Это может реализоваться путём изменения содержимого кадров или игнорирования уровней приоритета на интерфейсе R2.

Другим способом поддержки разнородных получателей может служить создание отдельных групп MAC-адресов для каждого класса обслуживания. По умолчанию получатель включается в группу best effort, для трафика которой предоставляется одноимённый тип обслуживания. Если получатель зарезервировал ресурсы, он его адрес переносится в группу для соответствующего класса обслуживания. Средства динамической групповой фильтрации в коммутаторах и мостах, поддерживающих стандарт IEEE 802.1D, будут очень полезны для такой модели. Данный поток будет передаваться только в те сегменты, которые включаются в путь между отправителем и получателями данного потока. Очевидным недостатком такого подхода является необходимость передачи отправителем множества копий одного пакета для разных классов обслуживания, что может приводить к дублированию трафика на некоторых ветвях дерева распространения.

Описанная выше модель будет обеспечивать близкое к оптимальному использование ресурсов с учётом указанного ожидаемого размера и сложности подсетей уровня 2. Следовательно, желательно будет разрешать коммутаторам использовать разные QoS для различных исходящих ветвей дерева, которые разделяет данный коммутатор.

Стандарт IEEE 802.1D задаёт базовую модель для групповой пересылки, когда коммутатор принимает решение на базе адреса получателя. В этом случае создаётся список выходных портов, в которые пакет следует пересылать. В принятом по умолчанию режиме коммутатор будет использовать значение user_priority из принятых пакетов или значения, восстановленные с учётом входных портов в отсутствие заданных явно значений, для размещения пакетов в выходных очередях каждого порта. Все коммутаторы IEEE 802.1D, поддерживающие классы группового трафика, могут поддерживать такую операцию.

Если коммутатор выбирает очереди для выходных портов только на основе входящих значений user_priority, как описано в IEEE 802.1D, он должен для всех ветвей всех групповых сессий в данном классе user_priority использовать один механизм очередей. Поддержка неоднородных получателей в таком случае не возможна и могут возникать отказы для запросов управления доступом в масштабе групповой сессии в целом по причине перегрузки на одном из каналов. Отметим, что для уровня 2, в отличие от уровня 3 с RSVP/IntServ, вариант предоставления некоторым получателям сервиса QoS в то время, как другим предоставляется только best effort не возможен, поскольку базовые коммутаторы IEEE 802.1 не способны отображать значения user_priority на отдельные каналы. Это может вызывать проблемы перегрузок в случаях использования динамических групповых сессий. Если коммутатор поддерживает отдельные отображения user_priority для каждого выходного порта, тогда (в некоторых случаях) в резервировании могут применяться разные классы трафика на различных ветвях, разделяемых таким коммутатором для обеспечения разным получателям разных параметров QoS. Это будет возможно, если все потоки данного класса (на входе в коммутатор) будут выходить с таким же классом трафика. Например, трафик может пересылаться с использованием user_priority 4 на одной ветви, где получатели выполняют контроль доступа, и с user_priority 0 там, где такого контроля нет. Предполагается, что размещение в очереди по значениям user_priority обеспечивается минимальной функциональностью коммутаторов ЛВС в соответствии со стандартом IEEE 802.1D, а коммутаторы уровня 2 или даже 3 (т. е., маршрутизаторы) могут поддерживать более сложные варианты неоднородностей, обеспечивая более эффективное использование ресурсов. Поведение коммутаторов уровня 3 в таком контексте уже стандартизовано IETF.

9. Варианты сетевой технологии

Уровень обеспечения гарантий сервиса в сети в значительной степени зависит от поддержки ключевых функций идентификации и планирования потоков в дополнение к управлению доступом и политикой. В этом разделе обсуждаются некоторые возможности рассматриваемых технологий ЛВС и приводится классификация вариантов совершенствования с целью расширения возможностей поддержки этих функций. Для рассматриваемых здесь ситуаций базовой технологией¹ ЛВС может быть разделяемая, коммутируемая полудуплексная и коммутируемая полнодуплексная сетевая среда. В разделяемой среде множество получателей использует один сегмент сетевой среды. Доступ к среде контролируется специальным протоколом (CSMA/CD в Ethernet, передача маркера в Token Ring и FDDI). Полудуплексная коммутируемая среда является частным случаем разделяемой с совместным использованием общей среды двумя устройствами. В полнодуплексной коммутируемой среде контроля доступа не требуется и канал передачи с полной полосой предоставляется обоим передатчикам, подключённым к разным концам соединения. Следовательно, в таких средах не применяются протоколы управления доступом к среде типа CSMA/CD или передачи маркеров и нет борьбы передатчиков за доступ к среде. Обычно этот вариант обеспечивает наилучший вариант поддержки QoS. Другим важным элементом при обсуждении сетевых технологий является возможность поддержки множества классов трафика. Этот вопрос был рассмотрен выше в параграфе 4.1. В зависимости от базовой технологии и возможности поддержки множества классов трафика возможны шесть вариантов:

¹В оригинале вместо термина «технология» не совсем корректно используется термин «топология». Прим. перев.

1. разделяемая среда без поддержки классов трафика;
2. разделяемая среда с поддержкой классов трафика;
3. коммутируемая полудуплексная среда без поддержки классов трафика;
4. коммутируемая полудуплексная среда с поддержкой классов трафика;
5. коммутируемая полнодуплексная среда без поддержки классов трафика;
6. коммутируемая полнодуплексная среда с поддержкой классов трафик.

Возможно также одновременное использование нескольких технологий. Например, в одной подсети могут применяться коммутаторы, часть которых поддерживает классы трафика, а остальные не поддерживают. Если поток проходит через коммутаторы обоих типов, преобладать будет «наименьший общий знаменатель». Иными словами, будут использоваться свойства технологии с меньшими из возможностей, поддерживаемых коммутаторами на пути прохождения потока. В последующих параграфах эти варианты рассматриваются более подробно для разных типов сетей IEEE 802 с учётом их возможностей поддержки услуг IntServ.

9.1. Полнодуплексные коммутируемые сети

В полнодуплексных коммутируемых ЛВС протокол MAC не имеет значения в плане управления доступом, но должен приниматься во внимание при учёте анонсируемых устройством параметров, поскольку время доступа определяется продолжительностью передачи самого длинного пакета. Приблизительные параметры разных сред приведены в представленных ниже таблицах. Задержки следует также учитывать ограниченность скорости света, вызывающую задержку приблизительно в 400 нсек для типовых линий UTP длиной 100 м и 7 мсек для многомодового оптического кабеля длиной 2 км.

Полнодуплексные коммутируемые технологии обеспечивают лучшие параметры QoS как для сервиса Controlled Load, так и для Guaranteed при поддержке коммутаторами подходящих стратегий управления очередями.

Таблица 4. Задержка при доступе к полнодуплексным коммутируемым средам.

Тип	Скорость	Макс. длительность пакетов	Максимальная задержка доступа
Ethernet	10 Мбит/с	1,2 мсек	1,2 мсек
	100 Мбит/с	120 мсек	120 мсек
	1000 Мбит/с	12 мсек	12 мсек
Token Ring	4 Мбит/с	9 мсек	9 мсек
	16 Мбит/с	9 мсек	9 мсек
FDDI	100 Мбит/с	360 мсек	8.4 мсек
Demand Priority	100 Мбит/с	120 мсек	120 мсек

9.2. Сети Ethernet с разделяемой средой

До этого момента мы не рассматривали сложности, связанных с разделяемой средой CSMA/CD. Как только начинается использование любого алгоритма CSMA/CD возможности обеспечения любой формы гарантированного обслуживания существенно снижаются, если многочисленные отправители на канале не связаны жёстко между собой. Есть целый ряд причин, в силу которых здесь не предлагается лучшего решения этой задачи.

Во-первых, мы не считаем эту проблему корректно разрешимой без внесения изменений в протокол MAC. Комитет IEEE 802.1 провёл моделирование гарантий производительности для разделяемой среды CSMA/CD Ethernet без изменения протокола MAC и получил неутешительные результаты. Были внесены предложения по усовершенствованию протоколов уровня MAC (например, BLAM) и улучшению управления потоками данных в IEEE 802.3. Однако любое решение, включающее более эффективные программы MAC, работающие на основе традиционного IEEE 802.3 MAC или фирменных протоколов MAC, выходит за рамки деятельности группы ISSLL и данного документа. Во-вторых, мы не уверены в том, что эта проблема действительно представляет интерес. Хотя станции, подключённые к разделяемым средам, ещё будут использоваться некоторое время, число станций, подключённых к коммутируемым средам, растёт гораздо быстрее по сравнению с числом станций в разделяемых сегментах. Это ведёт к тому, что все сетевые коммуникации, требующие резервирования ресурсов будут проходить по крайней мере через один коммутатор или маршрутизатор. Иными словами, проще обновить для поддержки QoS существующую инфраструктуру уровня 2 путём установки коммутаторов. И лишь после такого обновления следует заниматься более сложными решениями, включающими управление доступом. В-третьих, ядро кампусных сетей обычно строится на основе коммутаторов, а не повторителей. Возможно в будущем возникнут иные обстоятельства (например, появятся буферизуемые гигабитные повторители), однако характеристики таких устройств в любом случае будут отличаться от характеристик существующих повторителей CSMA/CD.

Таблица 5. Задержка доступа в разделяемой среде Ethernet.

Тип	Скорость	Макс. длительность пакета	Макс. задержка доступа
Ethernet	10 Мбит/с	1,2 мсек	Не ограничена
	100 Мбит/с	120 мсек	Не ограничена
	1 Гбит/с	12 мсек	Не ограничена

9.3. Полудуплексные коммутируемые сети Ethernet

Многие из аргументов для субоптимальной поддержки сервиса Guaranteed в разделяемых средах Ethernet применимы и к полудуплексным коммутируемым средам Ethernet. Существенно то, что в данном случае разделяемая среда совместно используется по меньшей мере двумя отправителями, конкурирующими за передачу пакетов. Пока действия отправителей не скоординированы, существует вероятность того, что трафик best effort одного отправителя будет конфликтовать с зарезервированным трафиком другого. Координация действий отправителей требует внесения изменений в протокол MAC.

Не противореча сказанному выше, можно считать, что полудуплексные коммутируемые технологии дают шанс обеспечить сервис Controlled Load. С учётом наличия в точности двух потенциальных передатчиков и использования обоими приоритизации для своего трафика Controlled Load в потоках best effort, а также управления доступом на

основе этой информации характеристики доступа к среде являются вполне предсказуемыми, хотя и не детерминированными. Это обеспечивает достаточно близкое приближение для сервиса Controlled Load.

Таблица 6. Задержка доступа в коммутируемой среде Ethernet

Тип	Скорость	Макс. длительность пакета	Макс. задержка доступа
Ethernet	10 Мбит/с	1,2 мсек	Не ограничена
	100 Мбит/с	120 мсек	Не ограничена
	1 Гбит/с	12 мсек	Не ограничена

9.4. Полудуплексные сети Token Ring с коммутируемой и общей средой

В сетях Token Ring с разделяемой средой время доступа для высокоприоритетного трафика от любой станции ограничено и определяется значением $(N+1) \cdot THT_{max}$, где N - число станций, передающих трафик с высоким приоритетом, а THT_{max} - максимальное время удержания маркера [14]. Это предполагает, сетевые адаптеры имеют приоритетные очереди и резервирование маркера осуществляется для высокоприоритетного трафика, находящегося в очереди адаптера. Легко видеть, что время доступа к среде можно уменьшить, снижая значения N или THT_{max}. Рекомендуемое значение THT_{max} составляет 10 мсек [6]. N может принимать значения от 2 до 256 для разделяемого кольца и равно 2 для коммутируемой полудуплексной среды. Аналогичный анализ применим для сетей FDDI.

Таблица 7. Задержка доступа в полудуплексной коммутируемой и разделяемой среде Token Ring и FDDI.

Тип	Скорость	Макс. длительность пакета	Макс. задержка доступа
Token Ring	4/16 Мбит/с, разделяемая	9 мсек	2570 мсек
	4/16 Мбит/с, коммутируемая	9 мсек	30 мсек
FDDI	100 Мбит/с	360 мсек	8 мсек

С учётом ограниченности времени доступа можно обеспечить верхний предел для сквозной задержки в соответствии с требованиями Guaranteed Service, предполагая, что для этого класса трафика используется высший приоритет, дозволённый пользовательскому трафику. Реальное число станций, передающих трафик, который отображается в тот же класс, который применяется для Guaranteed Service, может существенно меняться с течением времени, но с точки зрения контроля доступа это число нужно знать заранее. Следовательно, элементы контроля доступа должны использовать фиксированное значение N, которое может совпадать с общим числом станций в кольце или быть несколько меньше, если хочется снизить величину задержки. Если принятое значение N меньше числа станций в кольце, контроль доступа должен гарантировать, что число станций, передающих высокоприоритетный трафик, не будет превышать этого значения. Такое решение позволяет системе контроля доступа оценить максимальную задержку в предположении, что все N станций передают пакеты с высоким приоритетом. В реальности для большинства случаев задержка может оказаться существенно меньше.

В предположении использования для трафика Controlled Load меньшего уровня приоритета по сравнению с Guaranteed Service, для потоков Controlled Load верхняя граница задержки обеспечиваться не может. Однако для потоков Controlled Load будет обеспечиваться лучшее по сравнению с best effort обслуживание.

Отметим, что во многих существующих сетях Token Ring с разделяемой средой мосты пересылают кадры с использованием для Access Priority (см. параграф 4.3) значения 4, безотносительно к значению user_priority в управляющем поле кадра. Следовательно, такие мосты нужно переконфигурировать или изменить для того, чтобы стало возможным описанное выше ограничение времени доступа.

9.5. Сети Demand Priority с коммутируемой и разделяемой средой

В сетях IEEE 802.12 коммуникации между конечными узлами и концентратором, а также между концентраторами основываются на обмене сигналами управления каналом. Эти каналы служат для контроля доступа к среде передачи. Если концентратор, например, получает запрос с высоким приоритетом в то время, как другой концентратор обрабатывает запросы с обычным приоритетом, доступ к среде может быть перенят первым концентратором для того, чтобы сначала обслужить высокоприоритетный запрос. По завершении обработки всех запросов с высоким приоритетом обслуживание обычных запросов продолжается с той точки, где оно было прервано.

Время доступа к среде для пакетов с высоким приоритетом определяется временем, которое нужно для того, чтобы перенять обычное обслуживание. Этот интервал ограничен и зависит от физического уровня и технологии разделяемой среды. Физический уровень оказывает существенное влияние при работе в полудуплексном режиме (например, по кабелям UTP¹, поскольку в процессе передачи пакета через канал обмен сигналами управления не возможен). Следовательно, для рассматриваемой здесь сетевой технологии в больших сетях сигнал может пройти через несколько соединений и концентраторов прежде, чем достигнет концентратора с функциями управления сетью. Это может задерживать перехват обслуживания и, следовательно, увеличивать верхнюю границу времени доступа.

Ниже приведены максимальные значения времени доступа для приоритетных пакетов в физической среде UTP с длинами кабелей между конечными узлами и концентраторами по 100 м и максимальной задержке распространения 570 нсек, как определено в [19]. Эти значения представляют наихудший вариант перекрытия сигналов в предположении передачи с обычным приоритетом пакетов максимального размера.

Таблица 8. Задержка доступа к полудуплексным средам Demand Priority UTP.

Тип	Скорость	Максимальная длительность пакетов	Максимальная задержка доступа
Demand Priority	100 Мбит/с	Пакеты 802.3, UTP	120 мсек
		Пакеты 802.5, UTP	360 мсек
			254 мсек
			733 мсек

Разделяемые среды IEEE 802.12 можно классифицировать по «глубине» каскадирования концентраторов (N). Простейшим вариантом будет сеть на основе одного концентратора (N = 1). Для физической среды UTP стандарт разрешает максимальную глубину каскадирования N = 5. Большие сети с сотнями узлов можно построить, используя уровень каскадирования 2. Средства управления полосой пропускания должны получать информацию о глубине каскадирования через систему сетевого управления и могут использовать эти сведения в своих алгоритмах управления доступом.

¹Unshielded twisted pair - скрученные пары без экрана.

В отличие от UTP, оптические системы работают в симплексном режиме до двум волокнам. Значения верхней границы времени доступа для высокоприоритетного трафика приведены ниже для многомодовых кабелей длиной 2 км при задержке на распространение 10 мксек.

Для разделяемой среды с кабелями длиной до 2 км между всеми конечными узлами и концентраторами стандарт IEEE 802.12 разрешает максимальный уровень каскадирования 2. Увеличение глубины каскадирования требует сокращения максимальной длины кабелей [15].

Ограниченная задержка доступа и детерминированный доступ к сети позволяют поддерживать услуги Guaranteed и Controlled Load даже в разделяемых средах. Однако поддержка стандартом 802.12 лишь двух уровней приоритета ограничивает число разных услуг, которые могут быть одновременно реализованы в сети.

Таблица 9. Задержка доступа к разделяемой среде Demand Priority UTP.

Тип	Скорость	Пакеты	Максимальная длительн. пакетов	Максимальная задержка доступа	Каскад
Demand Priority	100 Мбит/с	802.3	120 мксек	262 мксек	1
				554 мксек	2
				878 мксек	3
				1,24 мксек	4
				1,64 мксек	5
		802.5	360 мксек	722 мксек	1
				1,41 мксек	2
				2,32 мксек	3
				3,16 мксек	4
				4,03 мксек	5

Таблица 10. Задержка доступа к полудуплексной коммутируемой оптической среде Demand Priority.

Тип	Скорость	Пакеты	Максимальная длительн. пакетов	Максимальная задержка доступа
Demand Priority	100 Мбит/с	802.3	120 мксек	139 мксек
		802.5	360 мксек	379 мксек

Таблица 11. Задержка доступа к разделяемой оптической среде Demand Priority.

Тип	Скорость	Пакеты	Максимальная длительн. пакетов	Максимальная задержка доступа	Каскад
Demand Priority	100 Мбит/с	802.3	120 мксек	160 мксек	1
				202 мксек	2
		802.5	360 мксек	400 мксек	1
				682 мксек	2

10. Обоснование

Очевидной проблемой является сложность предложенной модели. Почему мы считаем, что на уровне 2 мы можем достичь лучших результатов, чем при выполнении тех же действий на уровне 3 с помощью протокола RSVP?

Основная причина заключается в наличии множества случаев, когда на уровне 2 обеспечивается простое решение значительной части реальных проблем QoS. Решение большинства проблем обеспечивается существенно более дешёвыми способами. Для полного решения всех проблем может потребоваться поддержка RSVP/IntServ на уровне очередей для потоков в коммутаторах и маршрутизаторах высокого уровня, однако устройства на основе описанной здесь архитектуры позволят существенно упростить сети.

11. Заключение

Этот документ определяет модель предоставления интегрированных услуг (Integrated Services) в ЛВС с разделяемой и коммутируемой средой. Возможность обеспечения гарантий QoS требует той или иной формы контроля доступа и управления ресурсами. Идентифицированы и рассмотрены требования и цели управления ресурсами на уровне подсетей. Отмечена схема управления ресурсами в целом, названная менеджером пропускной способности (Bandwidth Manager). Рассмотрены вопросы архитектуры этой схемы и примеры её возможных реализаций. Рассмотрены также некоторые вопросы отображения служб вышележащих уровней на канальный уровень. Сопутствующие документы рабочей группы ISSLL решают вопросы отображения служб [13] и спецификации протокола для Bandwidth Manager [14] на базе рассмотренных в этом документе требований и целей.

Литература

- [1] IEEE Standards for Local and Metropolitan Area Networks: Overview and Architecture, ANSI/IEEE Std 802, 1990.
- [2] ISO/IEC 10038 Information technology - Telecommunications and information exchange between systems - Local area networks -Media Access Control (MAC) Bridges, (also ANSI/IEEE Std 802.1D-1993), 1993.
- [3] ISO/IEC 15802-3 Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 3: Media Access Control (MAC) bridges (also ANSI/IEEE Std 802.1D-1998), 1998.
- [4] IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks, IEEE Std 802.1Q-1998, 1998.
- [5] Braden, B., Zhang, L., Berson, S., Herzog, S. and S. Jamin, "Resource Reservation Protocol (RSVP) - Version 1 Functional Specification", [RFC 2205](#), September 1997.
- [6] Wroclawski, J., "Specification of the Controlled Load Network Element Service", RFC 2211, September 1997.
- [7] Shenker, S., Partridge, C. and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, September 1997.
- [8] Braden, R., Clark, D. and S. Shenker, "Integrated Services in the Internet Architecture: An Overview", RFC 1633, June 1994.
- [9] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", [RFC 2210](#), September 1997.

- [10] Shenker, S. and J. Wroclawski, "Network Element Service Specification Template", RFC 2216, September 1997.
- [11] Shenker, S. and J. Wroclawski, "General Characterization Parameters for Integrated Service Network Elements", RFC 2215, September 1997.
- [12] Delgrossi, L. and L. Berger (Editors), "Internet Stream Protocol Version 2 (ST2) Protocol Specification - Version ST2+", RFC 1819, August 1995.
- [13] Seaman, M., Smith, A. and E. Crawley, "Integrated Service Mappings on IEEE 802 Networks", RFC 2815, May 2000.
- [14] Yavatkar, R., Hoffman, D., Bernet, Y. and F. Baker, "SBM Subnet Bandwidth Manager): Protocol for RSVP-based Admission Control Over IEEE 802-style Networks", RFC 2814, May 2000.
- [15] ISO/IEC 8802-3 Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications, (also ANSI/IEEE Std 802.3-1996), 1996.
- [15] ISO/IEC 8802-5 Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 5: Token Ring Access Method and Physical Layer Specifications, (also ANSI/IEEE Std 802.5-1995), 1995.
- [17] Postel, J. and J. Reynolds, "A Standard for the Transmission of IP Datagrams over IEEE 802 Networks", STD 43, [RFC 1042](#), February 1988.
- [18] C. Bisdikian, B. V. Patel, F. Schaffa, and M. Willebeek-LeMair, The Use of Priorities on Token Ring Networks for Multimedia Traffic, IEEE Network, Nov/Dec 1995.
- [19] IEEE Standards for Local and Metropolitan Area Networks: Demand Priority Access Method, Physical Layer and Repeater Specification for 100 Mb/s Operation, IEEE Std 802.12-1995.
- [20] Fiber Distributed Data Interface MAC, ANSI Std. X3.139-1987.
- [21] ISO/IEC 15802-3 Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements - Supplement to Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications - Frame Extensions for Virtual Bridged Local Area Network (VLAN) Tagging on 802.3 Networks, IEEE Std 802.3ac-1998 (Supplement to IEEE 802.3 1998 Edition), 1998.

Вопросы безопасности

Реализация описанной здесь модели не открывает новых возможностей для атак на сетевую инфраструктуру. Однако читателям рекомендуется обратиться к параграфу 2.8 спецификации RSVP [5], где рассмотрено влияние использования сигнальных протоколов управления доступом на безопасность сети.

Благодарности

Значительная часть работы, представленной в этом документе, явилась результатом плодотворных дискуссий в рамках рабочей группы ISSLL¹. Авторы хотели бы отметить вклад многих участников этих обсуждений как на встречах группы, так и в почтовой переписке. Особая благодарность Eric Crawley, Don Hoffman и Raj Yavatkar за их вклад в подготовку предварительных версий (Internet draft) и Peter Kim за текст о сетях Demand Priority.

Адреса авторов

Anoop Ghanwani

Nortel Networks
600 Technology Park Dr
Billerica, MA 01821, USA
Phone: +1-978-288-4514
E-Mail: aghanwan@nortelnetworks.com

Wayne Pace

IBM Corporation
P. O. Box 12195
Research Triangle Park, NC 27709, USA
Phone: +1-919-254-4930
E-Mail: pacew@us.ibm.com

Vijay Srinivasan

CoSine Communications
1200 Bridge Parkway
Redwood City, CA 94065, USA
Phone: +1-650-628-4892
E-Mail: vijay@cosinecom.com

Andrew Smith

Extreme Networks
3585 Monroe St
Santa Clara, CA 95051, USA
Phone: +1-408-579-2821
E-Mail: andrew@extremenetworks.com

Mick Seaman

Telseon

¹Integrated Services over Specific Link Layers - интегрированные услуги для конкретных канальных уровней.

480 S. California Ave
Palo Alto, CA 94306, USA
Email: mick@telseon.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

Copyright (C) The Internet Society (2000). Все права защищены.

Этот документ и его переводы могут копироваться и предоставляться другим лицам, а производные работы, комментирующие или иначе разъясняющие документ или помогающие в его реализации, могут подготавливаться, копироваться, публиковаться и распространяться целиком или частично без каких-либо ограничений при условии сохранения указанного выше уведомления об авторских правах и этого параграфа в копии или производной работе. Однако сам документ не может быть изменён каким-либо способом, таким как удаление уведомления об авторских правах или ссылок на Internet Society или иные организации Internet, за исключением случаев, когда это необходимо для разработки стандартов Internet (в этом случае нужно следовать процедурам для авторских прав, заданных процессом Internet Standards), а также при переводе документа на другие языки.

Предоставленные выше ограниченные права являются бессрочными и не могут быть отозваны Internet Society или правопреемниками.

Этот документ и содержащаяся в нем информация представлены "как есть" и автор, организация, которую он/она представляет или которая выступает спонсором (если таковой имеется), Internet Society и IETF отказываются от каких-либо гарантий (явных или подразумеваемых), включая (но не ограничиваясь) любые гарантии того, что использование представленной здесь информации не будет нарушать чьих-либо прав, и любые предполагаемые гарантии коммерческого использования или применимости для тех или иных задач.

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.