

# Характеристики QoS

## в пакетных сетях

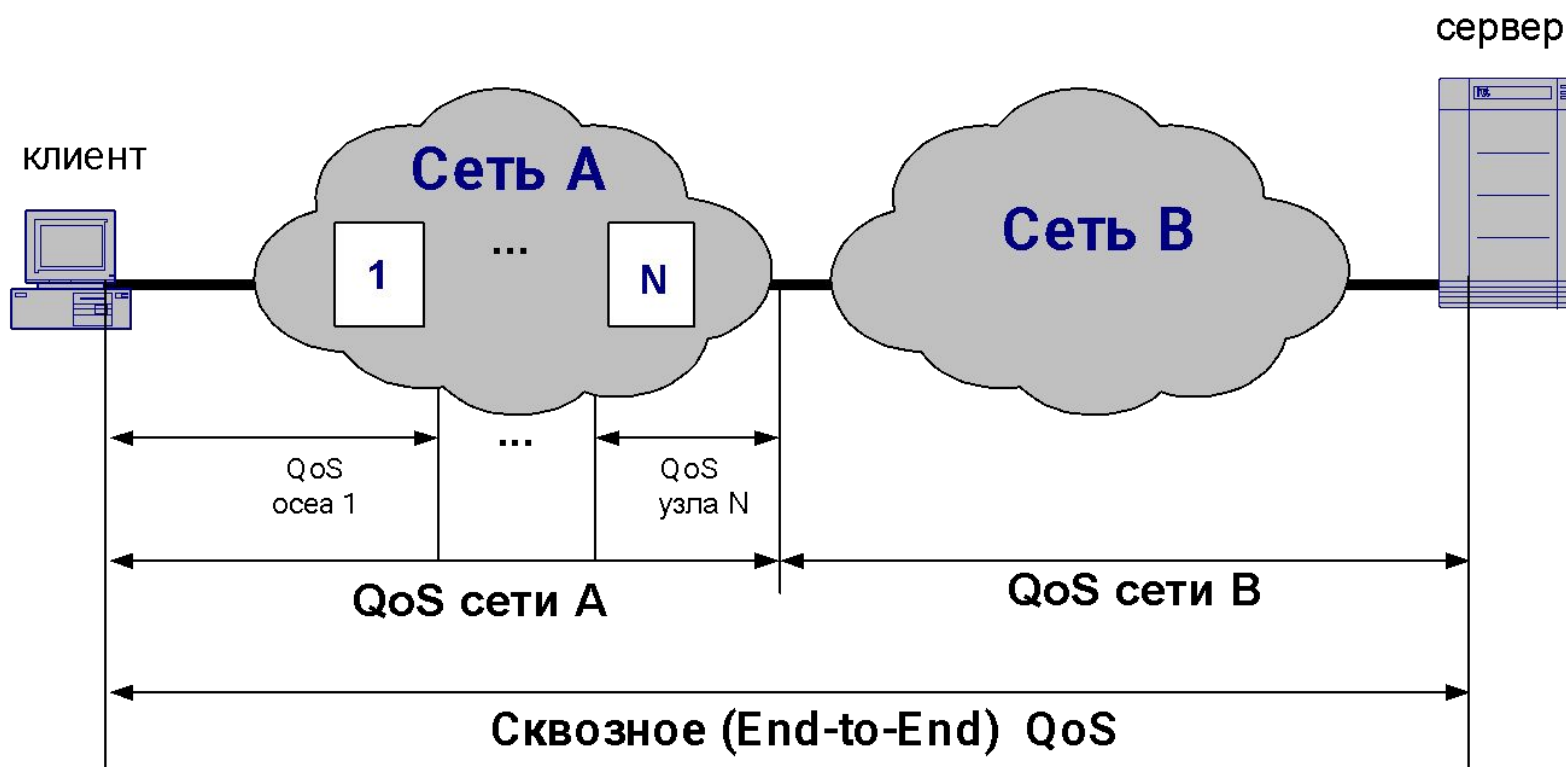
# КАЧЕСТВО ОБСЛУЖИВАНИЯ

**QoS (Quality of Servers)** рассматривается как «суммарный эффект *рабочих характеристик обслуживания*, который определяет степень удовлетворенности пользователя этой службой» (E.800)

**Задача:** обеспечить заданное качество обслуживания в сквозном соединении (end-to-end) для различных видов трафика.

**Условие:** заданное качество обслуживания должны поддерживать все сетевые устройства на всем сквозном соединении.

# Эталонная модель сквозного QoS



# Организации, стандартизирующие модели обеспечения качества обслуживания

- **ITU-T:** International Telecommunication Union – Международный Союз Электросвязи
- **ETSI:** European Telecommunications Standardizations Institute - Европейский институт по стандартизации телекоммуникаций
- **IETF:** Internet Engineering Task Force – Инженерная группа по решению задач Internet
- **MMCF:** Multimedia Communications Forum - Форум по мультимедийным коммуникациям
- **EURESCOM:** European Institute for Research and Strategic Studies in Telecommunications - Европейский институт по исследованиям и стратегическому планированию в телекоммуникациях

# Анализ стандартизированных моделей QoS

<b>Разработчи к модели QoS</b>	<b>QoS из конца в конец</b>	<b>Классы QoS</b>	<b>Биллинг</b>	<b>Мониторинг и управление</b>	<b>Особенности предоставляе- мых услуг</b>
<b>ITU-T</b>	Да	Да	Нет	Да	Да
<b>ETSI</b>	Да	Да	Нет	Да	Да
<b>IETF</b>	Да	Нет	Нет	Да	Нет
<b>MMCF</b>	Да	Да	Нет	Да	Да
<b>EURESCOM</b>	Да	Нет	Нет	Да	Нет

# Характеристики QoS (Y.1540)

- Задержки и джиттер\* задержки
- Величина потерь
- Производительность сети
- Надежность сетевых элементов

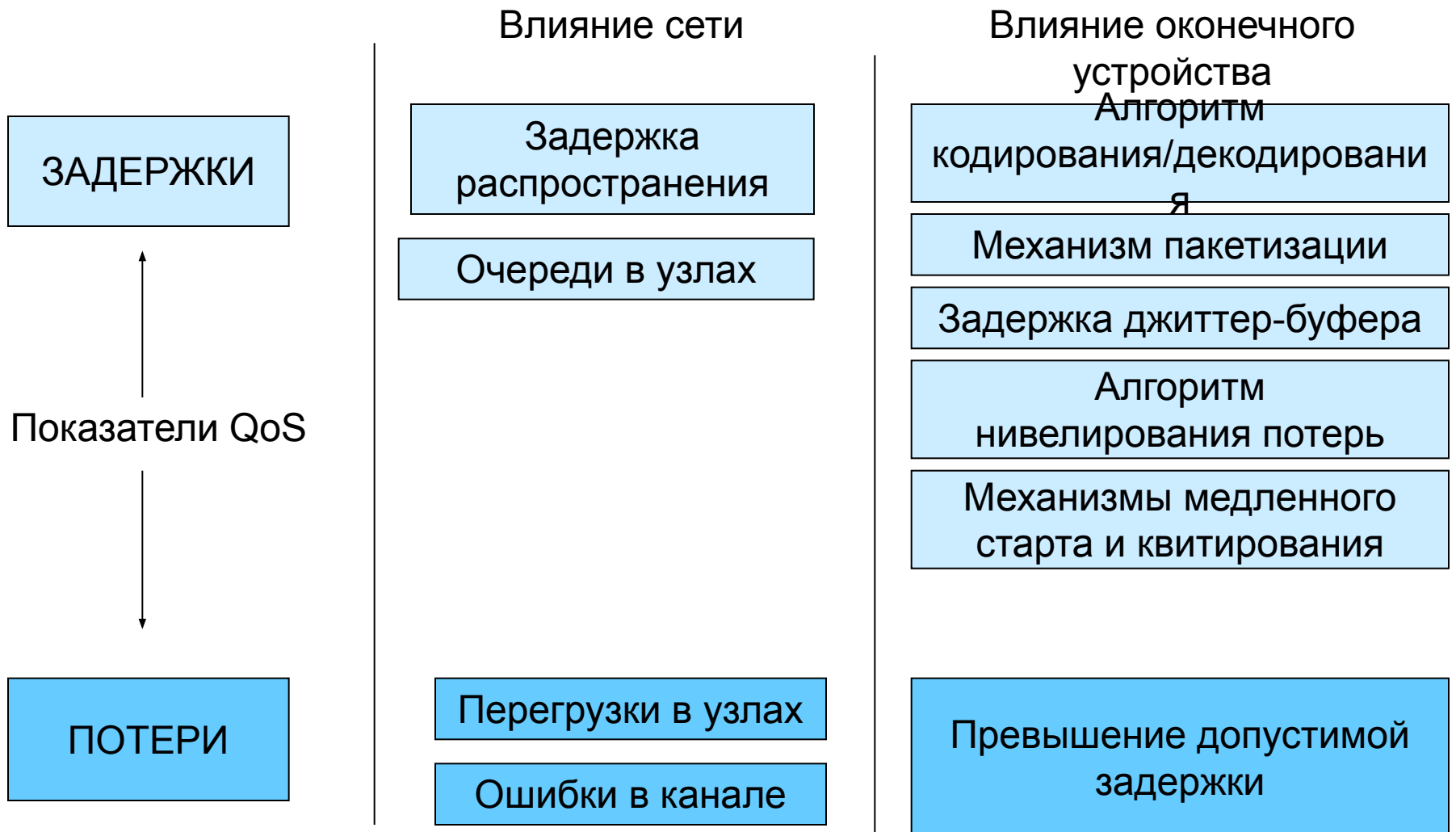
G.1000 – определяет структуру связей между рабочими характеристиками QoS.

\* джиттер задержки – отклонение значений задержки от заданной величины

# Классификация трафика мультисервисной IP-сети по приложениям

Тип трафика	Приложения	Требования	Протоколы транспортного уровня
Реального времени	IP-телефония и видеоконференцсвязь	<ul style="list-style-type: none"> <li>-Чувствительность к задержке</li> <li>-Чувствительность к джиттеру задержки</li> <li>-Малая чувствительность к потерям</li> </ul>	RSVP, RTP, RTCP, UDP
	Процессы управления, игры on-line	<ul style="list-style-type: none"> <li>-Чувствительность к задержке</li> <li>-Чувствительность к джиттеру задержки</li> <li>-Чувствительность к потерям</li> </ul>	UDP, TCP
Потоковый	<p>Аудио по требованию</p> <p>Видео по требованию</p> <p>Интернет-вещание</p>	<ul style="list-style-type: none"> <li>-Малая чувствительность к задержке</li> <li>-Чувствительность к джиттеру задержки</li> <li>-Чувствительность к потерям</li> </ul>	RSVP, SCTP, UDP, TCP
Эластичный	Конференция документов	<ul style="list-style-type: none"> <li>-Малая чувствительность к задержке</li> <li>-Малая чувствительность к джиттеру задержки</li> <li>-Высокая чувствительность к потерям</li> </ul>	TCP
	Анимация	<ul style="list-style-type: none"> <li>-Очень малая чувствительность к задержке</li> </ul>	

# Показатели качества обслуживания, учитываемые при передаче мультимедийного трафика, и механизмы их формирования





# Классы QoS и соответствующие им приложения (Y.1541)

- **Класс 0:** Приложения реального времени, чувствительные к джиттеру, характеризующиеся высоким уровнем интерактивности (VoIP, видеоконференции)
- **Класс 1:** Приложения реального времени, чувствительные к джиттеру, интерактивные (VoIP, видеоконференции)
- **Класс 2:** Транзакции данных, характеризующиеся высоким уровнем интерактивности (например, сигнализация)
- **Класс 3:** Транзакции данных, интерактивные приложения
- **Класс 4:** Приложения, допускающие низкий уровень потерь (короткие транзакции, массивы данных, потоковое видео)
- **Класс 5:** Традиционные применения сетей IP

# Нормы на параметры доставки пакетов IP с разделением по классам обслуживания, модель ITU-T

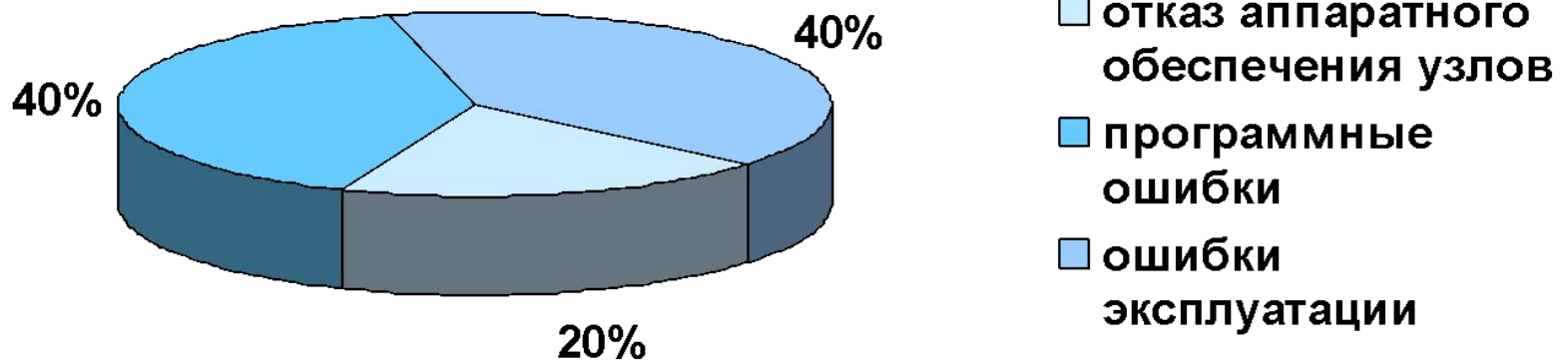
Сетевые характеристики	Классы QoS					
	0	1	2	3	4	5
<b>Задержка доставки пакета IP, IPTD</b>	100 мс	400 мс	100 мс	400 мс	1 с	Н
<b>Вариация задержки пакета IP, IPDV</b>	50 мс	50 мс	Н	Н	Н	Н
<b>Коэффициент потери пакетов IP, IPLR</b>	$1 \times 10^{-3}$	$1 \times 10^{-3}$	$1 \times 10^{-3}$	$1 \times 10^{-3}$	$1 \times 10^{-3}$	Н
<b>Коэффициент ошибок пакетов IP, IPER</b>	$1 \times 10^{-4}$	$1 \times 10^{-4}$	$1 \times 10^{-4}$	$1 \times 10^{-4}$	$1 \times 10^{-4}$	Н

**Примечание.** Н - не нормировано. Значения параметров представляют собой верхние границы для средних задержек, джиттера, потерь и ошибок пакетов.

# Коэффициенты готовности и значения времени простоя оборудования

Коэффициент готовности		Время простоя
0,99	“две девятки”	3,7 дней в год
0,999	“три девятки”	9 часов в год
0,9999	“четыре девятки”	53 минуты в год
0,99999	“пять девяток”	5,5 минут в год
0,999999	“шесть девяток”	30 секунд в год

# Причины системной ненадежности



Источник: Gartner Group

# Причины отказов в IP-сетях



# Службы QoS

- **Best effort** – обработка информации как можно быстрее, но без дополнительных усилий (FIFO, drop tail)
- **Мягкий QoS** – сервис с предпочтениями. Приоритетное обслуживание, значения параметров QoS зависят от характеристик трафика.
- **Жесткий QoS** – гарантированный сервис. Основан на предварительном резервировании ресурсов для каждого потока.

# Логические плоскости механизмов QoS

## Контрольная плоскость

- Управление допустимостью соединения
- QoS-маршрутизация
- Резервирование ресурсов

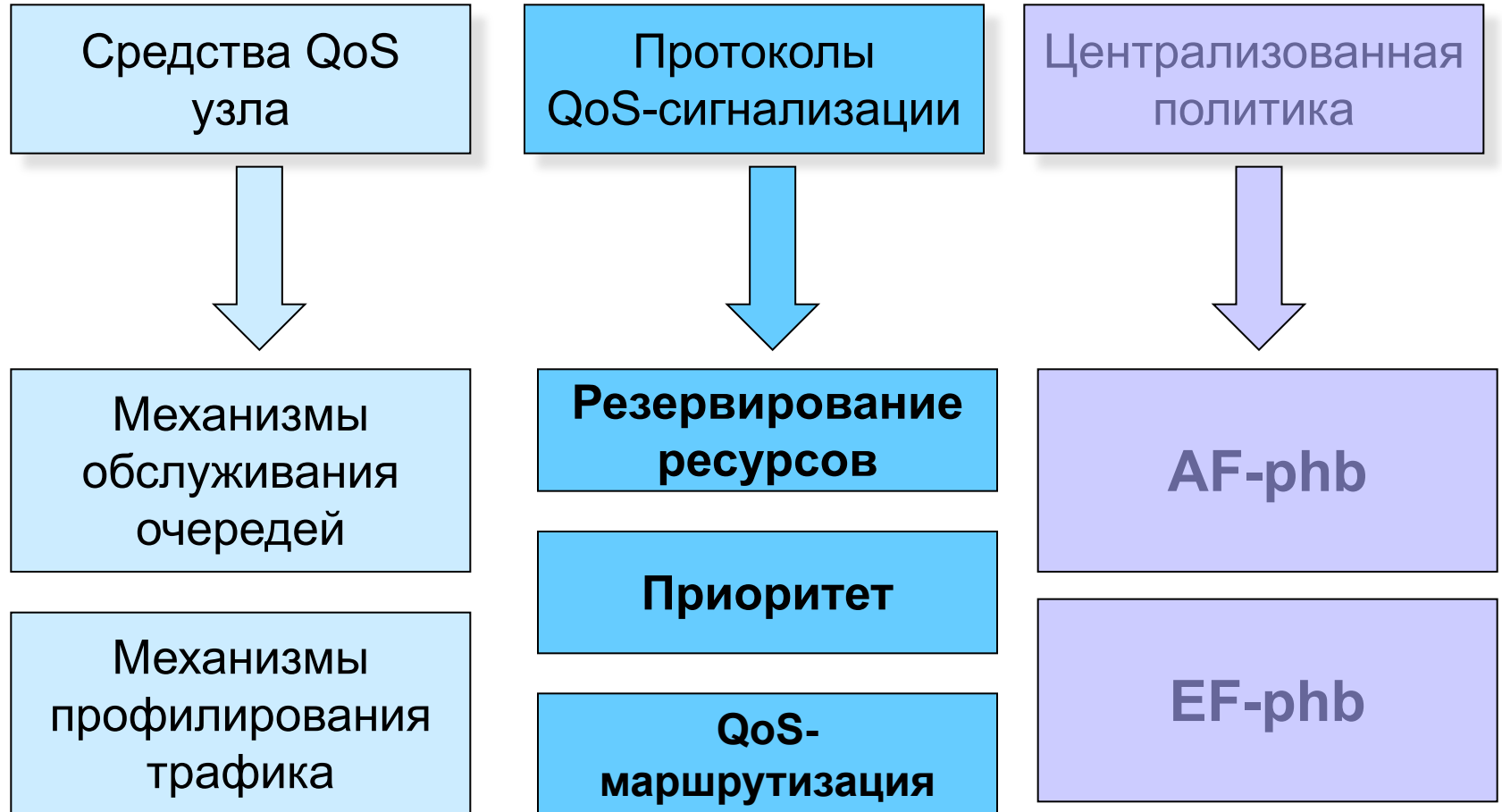
## Плоскость данных

- Предотвращение перегрузок
- Управление буфером
- Классификация трафика
- Маркировка пакетов
- Управление характеристиками трафика
- Организация и планирование очередей

## Плоскость менеджмента

- Измерения
- Восстановление трафика
- Соглашение об уровне обслуживания

# Базовая архитектура службы QoS





# Механизмы обслуживания очередей

- **FIFO** (First In First Out) – без использования дополнительных возможностей, используется в best effort
- **PQ** (Priority Queuing) – приоритетные очереди, вводится приоритет трафика (1-8)
- **CQ** (Custom Queuing) – настраиваемые очереди, используется при резервировании ресурсов
- **WFQ** (Weighting Fair Queuing) – взвешенное справедливое обслуживание, позволяет динамически управлять ресурсами

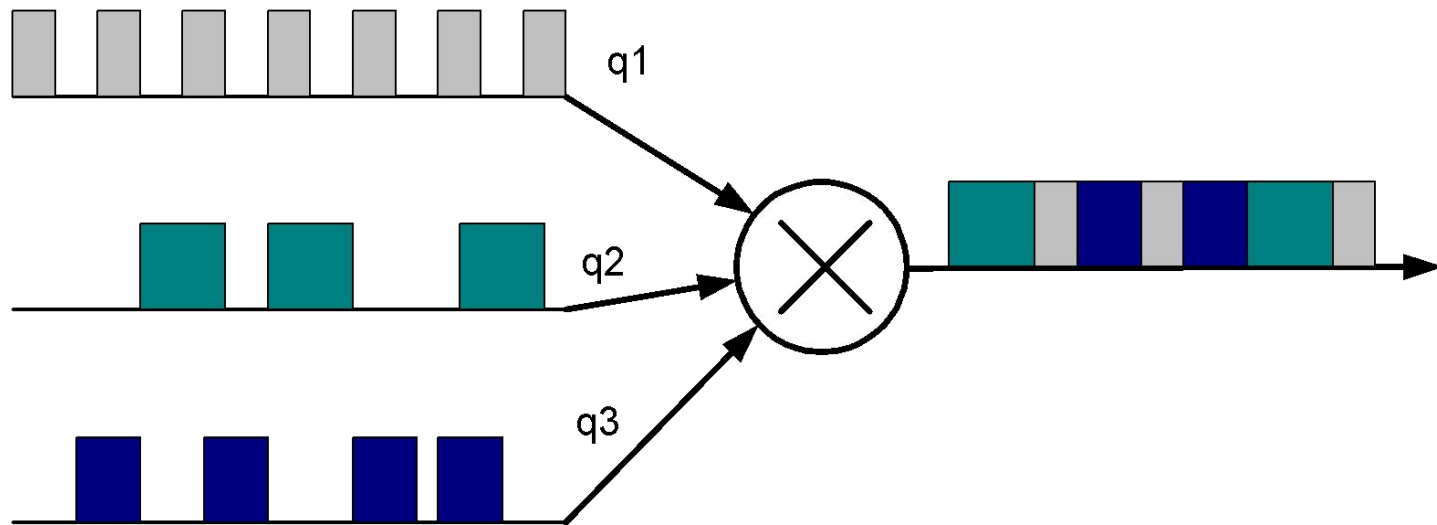
# Организация очередей WFQ

## Приоритет:

7-8 сигнализация, транзакции

5-6 трафик реального времени

1-4 эластичный трафик



# Механизмы профилирования трафика

- **Drop tail** – отбрасывание хвоста: отбрасываются все пакеты, заставшие буфер полным. Используется в best effort.
- **RED** – случайное раннее обнаружение: при угрозе перегрузки пакеты из буфера отбрасываются с ненулевой вероятностью.
- **Дырявое ведро** – отбрасываются пакеты, не обслужившиеся за установленный период.
- **Корзина маркеров** (токенов) – дозирование трафика с целью уменьшения неравномерности продвижения пакетов

# Управление потоками

- **Прерывание передачи:** при перегрузке передача пакетов источниками трафика прерывается на случайный интервал времени, затем возобновляется с той же интенсивностью.
- **Использование динамического окна:** размер окна (количество пакетов, посылаемых источником за период) изменяется в зависимости от загрузки буфера.
- **Медленный старт:** в случае перегрузки источники трафика прекращают передачу, затем посылают пакеты, постепенно увеличивая размер окна.

# Модели обеспечения качества обслуживания в сетях IP

- **Модель предоставления интегрированных услуг (IntServ)**  
**RFC-2205, 1994-1997 г.**
- **Модель предоставления дифференцированных услуг (DiffServ)**  
**RFC 2475, 1998 г.**
- **MPLS (Multi-Protocol Label Switching)**

# Интегрированные услуги IntServ

Разработана IETF, 1994-1997 г.

RFC 2205, RFC 2210, RFC 2211, RFC 2212

**Цель:** предоставление приложениям возможности запрашивать сквозные требования по ресурсам.

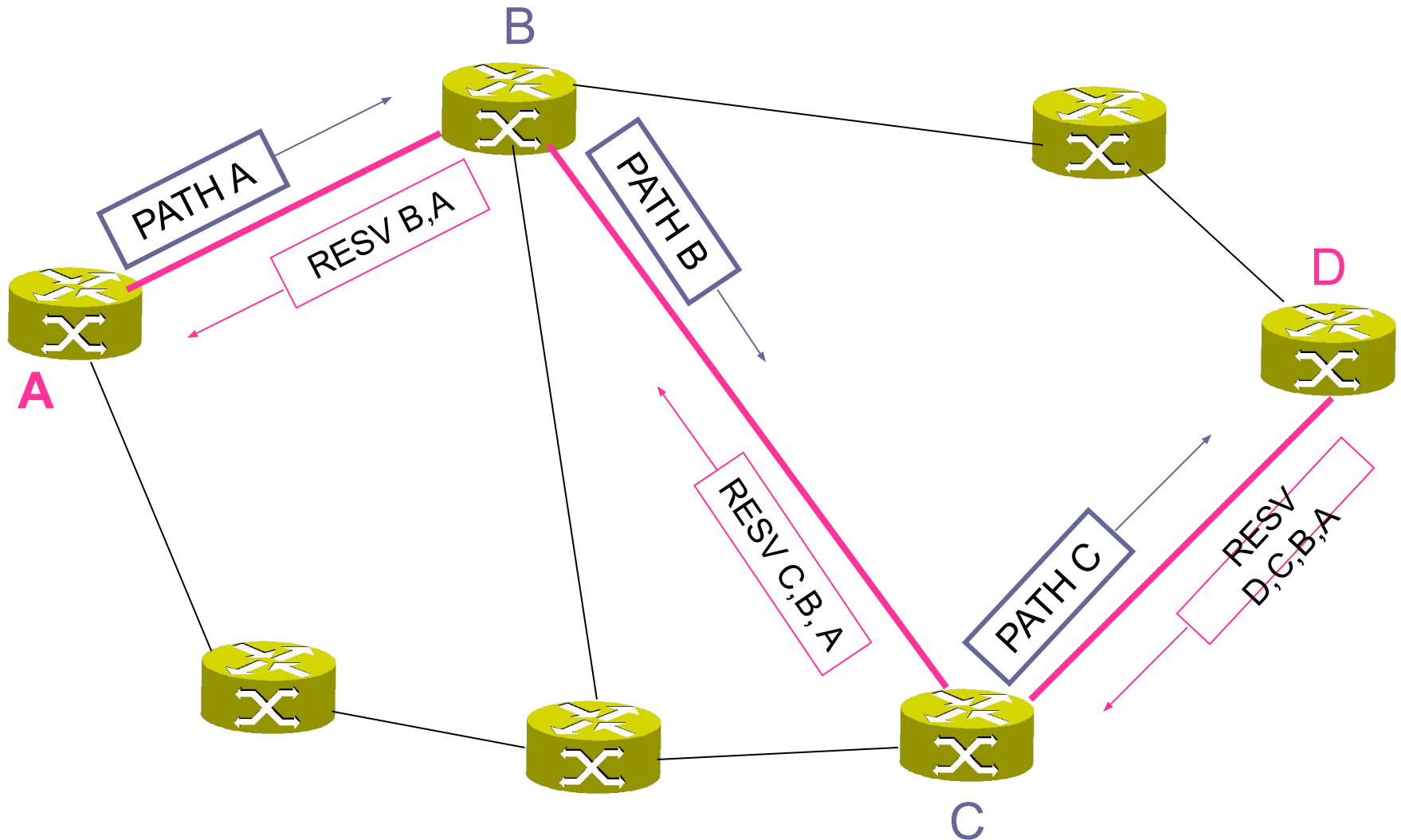
**Недостатки:** проблемы масштабирования.

**Основной механизм:** протокол резервирования ресурсов **RSVP**, в узлах используется WFQ.

# RSVP – Resource Reservation Protocol

- Протокол резервирования ресурсов. Позволяет посылать в сеть информацию о требованиях QoS для каждого потока. Работает совместно с IP.
- Резервирование проводится по адресу получателя. В случае отказа маршрута резервирование происходит заново.
- Работает с двумя видами сообщений:
  - PATH: запрос на резервирование. Содержит:
    - скорость передачи данных;
    - максимально допустимый размер пульсации трафика.
  - RESV: запрос резервирования. Содержит:
    - скорость передачи данных;
    - максимально допустимый размер пульсации трафика.
    - QoS

# Организация RSVP-пути





# Процесс резервирования пути

- Узел-отправитель посылает запрос PATH как обычный пакет.
- Каждый маршрутизатор прописывает в своей памяти адрес предыдущего и посылает свой адрес в PATH-запросе.
- Получатель в ответ на PATH генерирует RESV и отправляет по прописанному в PATH пути. Т.о. резервирование происходит в обратном порядке, от получателя к отправителю.
- Маршрутизаторы обрабатывают RESV-запросы, пытаясь предоставить требуемые ресурсы. В случае невозможности предоставления ресурсов резервирование начинается сначала.
- Путь считается установленным, когда отправитель получает RESV. После этого начинается сеанс.

# Дифференцированные услуги DiffServ

Разработана IETF, 1998 г.

RFC 1349, RFC 2475, RFC 2597, RFC 2598

**Цель:** поддержка легко масштабируемых дифференцируемых услуг в Internet

**Недостатки:** отсутствие гарантированного QoS

**Основной механизм:** маркировка трафика с использованием бита ToS (Type of Service).  
Поддерживает политики поведения сетевого узла: AF-phb и EF-phb (Per-Hop Behavior)

# Политики поведения сетевого узла - phb

- **AF-phb** (Assured Forwarding): политика гарантированной доставки – средство, позволяющее обеспечить несколько различных уровней надежности доставки IP-пакетов.

Механизмы: эффективное управление полосой пропускания за счет организации собственной очереди для каждого типа трафика; 3 уровня приоритетов пакетов; RED.

- **EF-phb** (Expedited Forwarding): политика немедленной доставки – обеспечение сквозного QoS для приложений реального времени.

Механизмы: приоритезация трафика; WFQ; распределение ресурсов; RED.

# MPLS

## (Multi-Protocol Label Switching)

Разрабатывается IETF

RFC 2702, RFC 2283, RFC 2547

**Цель:** отделение процесса маршрутизации пакета от необходимости анализа IP-адресов в его заголовке, что существенно уменьшает время пребывания пакетов в маршрутизаторе и обеспечивает требуемые показатели QoS для трафика реального времени.

**Недостатки:** ориентирован на топологию

**Основной механизм:** коммутация по меткам, туннелирование