Лекция 7. Методы обеспечения качества обслуживания

Дисциплина: Сетевые технологии

Цель лекции

Изучить основные методы обеспечения качества обслуживания (Quality of Service, QoS) в компьютерных сетях. Рассмотреть требования приложений к скорости передачи данных, задержкам, потерям пакетов и искажению данных.

1. Понятие качества обслуживания в компьютерных сетях

Качество обслуживания (QoS) представляет собой совокупность характеристик сети, которые определяют условия передачи данных между устройствами и приложениями.

QoS позволяет управлять сетевыми ресурсами и обеспечивать необходимые параметры передачи данных для разных видов трафика.

Например, видеоконференция требует небольшой задержки и минимальных потерь пакетов. Передача большого файла, напротив, допускает некоторую задержку, но требует надёжной доставки всех данных.

К основным характеристикам качества обслуживания относятся:

- Пропускная способность — объём данных, который сеть может передать за единицу времени.
- Задержка — время, необходимое для передачи пакета от источника к получателю.
- Джиттер — изменение задержки пакетов при их передаче по сети.
- Потери пакетов — количество пакетов, которые не были доставлены получателю.
- Надёжность передачи — способность сети обеспечивать корректную и своевременную доставку данных.

Для оценки качества обслуживания могут использоваться следующие показатели:

Показатель	Единица измерения	Значение
Пропускная способность	бит/с, Мбит/с	Скорость передачи данных
Задержка	мс	Время доставки пакета
Джиттер	мс	Изменение задержки
Потери пакетов	%	Доля потерянных пакетов
Доступная полоса пропускания	Мбит/с	Ресурс сети, доступный приложению

Качество обслуживания особенно важно в сетях, где одновременно работают различные приложения. Без механизмов QoS загрузка сети может привести к задержкам, потере пакетов и ухудшению качества связи.

2. Приложения и качество обслуживания

Приложения предъявляют разные требования к сетевым ресурсам. Эти требования зависят от назначения приложения, объёма передаваемых данных и допустимого времени доставки.

Основные группы сетевых приложений:

1. Интерактивные приложения. К ним относятся видеоконференции, IP-телефония и онлайн-игры. Для них важны низкая задержка, небольшой джиттер и минимальные потери пакетов.

2. Потоковые приложения. Это интернет-телевидение, онлайн-видео и аудиотрансляции. Они требуют достаточной пропускной способности и стабильной передачи данных. Небольшая начальная задержка обычно допустима, поскольку перед воспроизведением может использоваться буферизация.

3. Приложения передачи файлов. К ним относятся загрузка документов, резервное копирование и передача архивов. Для них важны высокая скорость передачи и целостность данных. Задержка обычно не оказывает существенного влияния на результат.

4. Веб-приложения. К ним относятся сайты, интернет-магазины и облачные сервисы. Для них важны быстрое получение ответа от сервера и стабильная работа сети.

5. Критически важные приложения. Это системы промышленного управления, медицинского мониторинга и управления транспортом. Для них могут быть установлены строгие требования к задержке, потерям пакетов и доступности сети.

3. Требования к качеству обслуживания приложений разных типов

Для каждого типа приложения необходимо определить допустимые значения сетевых характеристик.

Тип приложения	Пропускная способность	Задержка	Потери пакетов
IP-телефония	Умеренная	Очень важна	Желательны минимальные
Видеоконференция	Высокая	Очень важна	Желательны минимальные
Онлайн-игры	Обычно умеренная	Критически важна	Нежелательны
Потоковое видео	Высокая	Важна стабильность	Допустимы в ограниченном объёме
Передача файлов	Зависит от размера файла	Менее критична	Важна корректная доставка
Электронная почта	Низкая или умеренная	Обычно не критична	Важна корректная доставка
Промышленное управление	Зависит от задачи	Может быть критически важна	Допустимый уровень определяется требованиями

Указанные характеристики являются общими. Конкретные требования зависят от протокола, настроек приложения, качества кодирования и условий эксплуатации сети.

Например, при видеоконференции высокая пропускная способность не всегда гарантирует хорошее качество. Если задержка слишком велика, участники будут слышать друг друга с опозданием. Если пакеты поступают с большими колебаниями задержки, звук и изображение могут воспроизводиться неравномерно.

Следовательно, при настройке QoS необходимо учитывать не только скорость передачи данных, но и остальные характеристики сети.

4. Предсказуемость скорости передачи данных

Предсказуемость скорости передачи данных означает возможность обеспечить приложению ожидаемый уровень пропускной способности в течение определённого периода времени.

В компьютерных сетях фактическая скорость передачи может изменяться под воздействием различных факторов:

- увеличения количества пользователей;
- перегрузки каналов связи;
- одновременной передачи большого объёма данных;
- ограниченной пропускной способности сетевого оборудования;
- конкуренции между приложениями за сетевые ресурсы.

Например, канал связи имеет пропускную способность 100 Мбит/с. Если несколько пользователей одновременно загружают большие файлы, доступная скорость для видеоконференции может снизиться. В результате качество изображения ухудшится.

Для повышения предсказуемости скорости применяются следующие методы.

Резервирование полосы пропускания. Для определённого вида трафика выделяется необходимая часть сетевых ресурсов.

Управление очередями. Пакеты размещаются в очередях и передаются в соответствии с установленными правилами.

Формирование трафика (Traffic Shaping). Скорость передачи пакетов регулируется для предотвращения резких всплесков нагрузки.

Ограничение скорости (Policing). Трафик контролируется на соответствие установленным ограничениям. Пакеты сверх допустимого уровня могут быть отброшены или перемаркированы.

Контроль перегрузки. При высокой нагрузке сеть регулирует объём передаваемых данных и распределение ресурсов.

Предсказуемость скорости особенно важна для видеоконференций, потокового видео и приложений реального времени.

5. Чувствительность трафика к задержкам пакетов

Задержка пакета представляет собой время, которое проходит с момента отправки пакета источником до его получения адресатом.

Полная задержка может включать несколько составляющих:

$$D_{\text{общ}} = D_{\text{обраб}} + D_{\text{очер}} + D_{\text{перед}} + D_{\text{расп}}$$

где:

- $D_{\text{обраб}}$— задержка обработки пакета;
- $D_{\text{очер}}$— задержка ожидания в очереди;
- $D_{\text{перед}}$— время передачи пакета в канал;
- $D_{\text{расп}}$ — время распространения сигнала по среде передачи.

Задержка может увеличиваться при перегрузке сети, поскольку пакеты дольше ожидают передачи в очередях маршрутизаторов и коммутаторов.

Чувствительность к задержкам зависит от типа приложения.

- Для IP-телефонии высокая задержка затрудняет естественный разговор.
- Для видеоконференций задержка вызывает рассинхронизацию звука и изображения.
- Для онлайн-игр задержка приводит к запаздыванию реакции на действия пользователя.
- Для электронной почты задержка в несколько секунд обычно не имеет существенного значения.
- Для передачи файлов задержка влияет на время завершения загрузки, но обычно не нарушает целостность данных при корректной работе протокола.

Для уменьшения задержки применяются приоритетная обработка пакетов, управление очередями, выбор эффективных маршрутов и снижение перегрузки каналов связи.

Приоритетная обработка пакетов

Приоритетная обработка позволяет передавать более важный трафик раньше менее чувствительного к задержкам трафика.

Например, маршрутизатор может предоставить высокий приоритет пакетам IP-телефонии, а загрузку больших файлов обслуживать с более низким приоритетом.

Однако чрезмерный приоритет одного вида трафика может привести к длительному ожиданию остальных пакетов. Поэтому при настройке QoS необходимо соблюдать баланс между разными классами трафика.

6. Чувствительность трафика к потерям и искажению данных

Потери пакетов возникают, когда пакеты не достигают получателя. Основными причинами являются перегрузка сети, переполнение очередей, ошибки передачи и неисправности оборудования.

Последствия потерь зависят от типа приложения и используемого протокола.

Для голосовой связи потеря пакетов может привести к пропуску звуков, появлению шумов и ухудшению разборчивости речи.

Для видеоконференций потери могут вызвать искажение изображения, пропуск отдельных кадров и кратковременные остановки видео.

Для потокового видео потери могут вызвать снижение качества изображения или прерывание воспроизведения.

Для передачи файлов потери могут потребовать повторной передачи данных. Протокол TCP обычно обеспечивает обнаружение потерь и повторную передачу, если это необходимо.

Для систем управления потеря даже небольшого количества пакетов может быть недопустимой, если она нарушает установленный цикл управления.

Искажение данных означает изменение передаваемой информации в результате ошибок. Такие ошибки могут возникать на физическом уровне сети. Для обнаружения повреждённых данных применяются контрольные суммы и механизмы обнаружения ошибок.

Необходимо различать потерю пакета и искажение данных. При потере пакет не доставляется получателю. При искажении пакет может быть доставлен, но его содержимое окажется повреждено. На практике повреждённые кадры часто отбрасываются на канальном уровне.

Для уменьшения потерь применяются:

1. Увеличение доступной пропускной способности.
2. Управление очередями и предотвращение их переполнения.
3. Приоритизация критически важного трафика.
4. Повторная передача потерянных данных, если она допустима.
5. Использование резервных каналов и альтернативных маршрутов.
6. Контроль ошибок и состояния сетевого оборудования.

Повторная передача повышает надёжность доставки, но может увеличить задержку. Поэтому для приложений реального времени она не всегда является лучшим решением.

7. Основные методы обеспечения качества обслуживания

Для обеспечения QoS используются специальные механизмы, которые позволяют распределять сетевые ресурсы между приложениями.

1. Классификация трафика

Пакеты распределяются по категориям в зависимости от протокола, адреса, порта или типа приложения. Это позволяет применять разные правила обработки к разным видам трафика.

2. Приоритизация

Для важных пакетов устанавливается более высокий приоритет. Этот метод особенно полезен для голосовой связи и видеоконференций.

3. Управление пропускной способностью

Сетевые ресурсы распределяются между приложениями. При необходимости для отдельных потоков резервируется определённая полоса пропускания.

4. Управление очередями

Очереди определяют порядок передачи пакетов. Для разных классов трафика могут применяться отдельные очереди и алгоритмы обслуживания.

5. Мониторинг сетевых характеристик

Контролируются задержка, джиттер, потери пакетов и пропускная способность. Полученные данные используются для выявления проблем и изменения настроек сети.

Основные модели QoS

Best Effort (обслуживание по возможности). Сеть передаёт пакеты без специальных гарантий по задержке, потерям и скорости. Такой подход часто применяется для обычного интернет-трафика.

IntServ (Integrated Services). Модель предусматривает резервирование ресурсов для отдельных потоков. Она позволяет обеспечить определённые параметры обслуживания, но требует дополнительных ресурсов и контроля состояния потоков.

DiffServ (Differentiated Services). Модель предусматривает разделение трафика на классы. Каждый класс получает определённые правила обработки. Такой подход широко применяется в корпоративных и операторских сетях.

DiffServ удобна для крупных сетей, поскольку маршрутизаторы могут обрабатывать пакеты по установленным классам без отдельного резервирования ресурсов для каждого потока.

Заключение

Качество обслуживания определяет, насколько хорошо компьютерная сеть удовлетворяет требованиям приложений. Для голосовой связи и видеоконференций важны низкая задержка, небольшой джиттер и минимальные потери пакетов. Для передачи файлов важны пропускная способность и корректная доставка данных.

Для обеспечения QoS применяются классификация трафика, приоритизация, управление очередями, регулирование скорости передачи и резервирование ресурсов. Выбор методов зависит от назначения сети и требований приложений.

Таким образом, правильная настройка QoS позволяет рационально использовать пропускную способность сети, уменьшить влияние перегрузок и повысить качество работы сетевых приложений.

Контрольные вопросы

1. Что такое качество обслуживания QoS?
2. Какие основные характеристики определяют качество обслуживания сети?
3. Почему приложения предъявляют разные требования к сетевым ресурсам?
4. Что означает предсказуемость скорости передачи данных?
5. Как задержка влияет на работу IP-телефонии и видеоконференций?
6. Что такое джиттер и как он влияет на качество связи?
7. По каким причинам возникают потери пакетов?
8. Чем отличаются потеря пакета и искажение данных?
9. Какие методы применяются для уменьшения задержки и потерь?
10. Чем отличаются модели Best Effort, IntServ и DiffServ?