

Текст к лекции 1 “Сети с малыми задержками и низким потреблением энергии. Качество обслуживания”.

Цель лекции – ознакомиться с новыми тенденциями в развитии представлений о качестве обслуживания и с влиянием характеристик качества обслуживания на построение сетей связи.

Исторически построение сетей связи общего пользования было ориентировано на обслуживание трафика передачи речи. Это было естественным для гомогенной Телефонной сети Общего Пользования (ТфОП), а для гетерогенной пакетной сети связи следующего поколения передача речи оказалась приложением, для которого требовались наиболее жесткие условия по задержкам (слайд 3). Так продолжалось до недавних пор, когда появились приложения, требующие меньших задержек, чем передача речи поверх IP. Последнее может оказать существенное влияние на характеристики сетей связи следующего поколения. Поэтому такие сети выделяются в новый класс сетей и называются сетями с малыми задержками.

Планирование аналоговых сетей ТфОП осуществлялось, в основном, на базе моделей с потерями с использованием формулы Эрланга для их расчета. При этом допустимые потери из конца в конец, например для городских сетей, составляли менее 5%, а потери на ступенях искания нормировались в пределах $(2-5) 10^{-3}$ (слайд 2). Системы с ожиданием также использовались на аналоговых ТфОП, но весьма редко, разве что на ступени абонентского искания для нормирования длительности ожидания сигнала “Ответ станции”. Создание цифровых ТфОП с точки зрения нормирования параметров качества обслуживания не привнесло новых принципиальных идей, поскольку сеть все равно осталась гомогенной.

Появление концепции сетей связи следующего поколения NGN привело не только к пакетизации сети связи общего пользования, но и к тому, что сеть стала гетерогенной со множеством приложений, потребовавших к тому же и дифференциации классов и параметров качества обслуживания QoS (Quality of Service) (слайды 3, 4). Заметим, что параметры качества обслуживания в соответствии с рекомендациями МСЭ-Т определяются для сети от интерфейса пользователь-сеть UNI (User-Network Interface) до другого интерфейса UNI (слайд 5). Кроме того, уже в последние годы в дополнение к имеющейся метрике QoS добавилась еще одна – качество восприятия QoE (Quality of Experience), определяющая субъективную оценку пользователем, например, качества видео (слайды 8, 9, 10).. Поскольку основное внимание в предлагаемой лекции уделяется новому классу сетей – сетям с малыми

задержками – сначала более подробно остановимся на характеристиках QoS, регламентируемых в рекомендации Y.1541 Сектора Стандартизации Телекоммуникаций Международного Союза Электросвязи (МСЭ-Т). На слайде 3 приведены наиболее жесткие требования к параметрам качества обслуживания для различных классов обслуживания, предусмотренных в NGN. IPTD (IP Time Delay) представляет собой задержку передачи пакетов, IPDV (IP Delay Variance) – вариацию задержки (джиттер), IPLR (IP Loss Ratio) – долю потерянных пакетов, IPER (IP Error Rate) – долю пакетов, переданных с ошибками. Таким образом, сети связи следующего поколения планируются и проектируются исходя из предположения, что задержка не должна превышать 100 мс, а джиттер - 50мс. Следует отметить, что, несмотря на требуемое малое значение джиттера, с указанным явлением достаточно успешно справляются за счет так называемых буферов деджиттеризации, созданных на базе одной из модификаций классического метода Token Bucket (ведро жетонов), позволяющего задавать любую функцию для времени обслуживания поступающих пакетов. Что касается требований по доле потерь пакетов и доле пакетов, переданных с ошибками, то на сегодняшний день выполнение приведенных в на слайде 3 норм не представляет больших сложностей.

Итак, 100 мс - наиболее критичный параметр качества обслуживания в рекомендациях МСЭ-Т и созданных на их основе к текущему моменту времени сетях NGN.

Подход к решению проблем качества обслуживания, основанный на введении различных классов обслуживания и известный как дифференцированное обслуживание DiffServ, используется не только МСЭ-Т, но и другими стандартизирующими организациями. На слайдах 6 и 7 приведена информация о категориях доступа для сетей WiFi в соответствии со стандартом IEEE 801.11e. Внимательный взгляд на категории доступа показывает, что это тоже классы обслуживания в терминологии МСЭ-Т. С учетом особенностей радиоканалов в системе категорий и параметров качества обслуживания для WiFi выделяются следующие виды трафика (классы обслуживания): background (мусорный трафик), best effort (трафик наилучшей попытки), трафик видео и речи (слайд 7). Наличие мусорного трафика требуется в связи с тем, что радиоканалы допускают большие ошибки при передаче информации, являются нестабильными и мусорный трафик определяет класс трафика, который по классу обслуживания даже ниже, чем трафик наилучшей попытки. Два других вида трафика требуют обслуживания в реальном времени, что в принципе не соответствует первоначальной идеологии системы WiFi о конкурентном доступе к каналу.

Для разрешения этого противоречия вводятся два механизма. Первый – конкурентные преимущества для трафика реального времени в конкурентной борьбе за канал. В системах WiFi для конкурентной борьбы за канал выделяется так называемое окно конкуренции CW (Contention Window), которое по времени ограничивается, как правило, некими минимальными CW_{min} и максимальными CW_{max} значениями. С целью создания конкурентного преимущества для трафика реального времени его окно конкуренции имеет максимальное значение по времени меньшее или равное минимальному значению окна конкуренции для мусорного трафика и трафика наилучшей попытки. Точно также и среди видов трафика реального времени, как это видно из слайда 7, речь имеет преимущество перед видео. Второй механизм – период монопольного владения каналом TXOP (Transmission Opportunity), в течение которого может передаваться информация только конкретного приложения реального времени.

На слайде 11 представлены два новых вида трафика, которые окажут существенное влияние на развитие сетей связи в ближайшем будущем. Это – трафик игр реального времени и трафик системы e-health, подробно рассмотренной во второй лекции про Интернет Вещей.

В таблице на слайде 12 представлены требования 3GPP (Third Generation Partnership Project) к задержкам и потерям для различных видов трафика, передаваемых с различными приоритетами в системах современной мобильной связи при внедрении пакетной мультимедийной подсистемы IMS (IP Multimedia Subsystem). Наиболее приоритетным трафиком является трафик управления соединениями - трафик сигнализации, далее следует трафик передачи речи и третьим приоритетом обладает трафик игр в реальном времени. Представленная таблица весьма современна и рассчитана на использование в мобильных сетях связи систем длительной эволюции LTE (Long Term Evolution), в которых уже приняты меры по обеспечению длительности задержки, требуемой приложением игр в реальном времени, т.е. 50 мс. Как видим, это значение в два раза меньше, чем для передачи речи поверх IP, что требует и существенного увеличения скорости доступа пользователя в сеть. Следует заметить, что доля игр реального времени в цифровом контенте существенна уже сегодня (слайд 13).

Эволюция задержек в системах, стандартизованных 3GPP, происходила следующим образом (слайд 29) В системе третьего поколения 3G, реализованной по стандартам 3GPP 1999 года, задержка составляла 68 мс. Внедрение технологии высокоскоростного пакетного доступа HSPA (High Speed Packet Access) позволило снизить задержку в системах 3G до 51 мс, а при дальнейшей модификации этой технологии до HSPA+ задержка стала

менее 30мс. Принципиальные изменения в величине задержки в системах сотовой подвижной связи произошли при внедрении LTE. При использовании обслуживания трафика по расписанию удалось добиться значения задержки в 20 мс, а при резервировании с предварительным распределением ресурсов – менее 15 мс.

Однако только появлением услуги игр реального времени, изменения требований к задержкам для современных сетей связи на сегодняшний день не исчерпываются.

В течение последнего времени существенно возросло внимание к так называемым системам электронного здоровья (e-health). Система e-health представляет собой глобальный подход к информатизации медицины и медицинских услуг, предусматривающий использование современных сетей связи для обеспечения предоставления услуг e-health. При этом в качестве сетей связи достаточно высокого уровня развития для предоставления услуг системы e-health в настоящее время рассматриваются сети M2M (Machine-to-Machine). В рамках работ МСЭ-Т по классам и параметрам качества обслуживания для сетей M2M при предоставлении услуг системы e-health определены как новые для сетей связи классы обслуживания, так и их параметры (слайды 14 и 15). Как видим, среди параметров качества обслуживания для новых классов услуг системы e-health появляются требования к задержкам на уровне 10 мс. Следует отметить одну недоработку в таблице: очень сложно, а практически и невозможно, в современных условиях обеспечить задержку в 10 мс даже в мегабитных сетях.

Как видим, внедрение новых услуг, таких как игры в реальном времени и услуги системы e-health, требует существенно меньших задержек, чем при передаче речи поверх IP. Рассмотрим далее, какие новые требования могут быть предъявлены к сетям связи с малыми задержками и как эти требования поддерживаются технологически уже в настоящее время.

На слайде 16 изображена гипотетическая сеть, состоящая из восьми узлов, что соответствует рекомендации Y.1541, причем узлы со второго по седьмой соединены линиями связи с пропускной способностью 10 Гб/с, а первый со вторым и седьмой с восьмым – линиями связи с пропускной способностью 4 Мб/с (сети доступа).

В качестве модели узла используем Систему Массового Обслуживания (СМО) с ожиданием и постоянным временем обслуживания. На узлы подается фоновый трафик, определяющий использование (загрузку) соответствующего узла. Этот трафик обслуживается соответствующим узлом и выводится из системы (завершается). Наряду с фоновым трафиком, на вход первого узла поступает испытательный трафик, который после обслуживания

передается на вход последующего узла. Интенсивность испытательного трафика много меньше интенсивности фонового трафика. Значения параметров функционирования оцениваются для испытательного трафика.

В качестве фонового трафика были выбраны два типа потоков: простейший поток и самоподобный поток с коэффициентом Хёрста $H=0,8$. В результате эксперимента получены зависимости средней величины задержки от нагрузки в Эрлангах. При моделировании загрузки всех узлов принимались равными.

На слайде 18 приведены результаты моделирования для сети доступа со скоростью 4 Мб/с. Как видим при достаточно больших нагрузках значения задержек измеряются в десятках миллисекунд как для простейшего, так и для самоподобного потоков. Последнее свидетельствует о том, что сети доступа с такими параметрами вряд ли могут обеспечить требования по качеству обслуживания не только для услуг системы e—health, но и для услуги игр реального времени.

На слайде 17 представлены результаты моделирования для участка гипотетической сети со скоростью 10 Гб/с. В этом случае задержки измеряются в долях миллисекунд, что позволяет говорить о готовности такой сети к предоставлению услуг e-health, а тем более и услуги игр реального времени. Таким образом, сети доступа в условиях поступления нового трафика должны быть гигабитными (слайд 19).

Остается только проанализировать технологические достижения в области сетей доступа, что сделаем применительно к беспроводным сетям доступа. В рамках работ по 12 версии систем длительной эволюции предусмотрено достижение пиковых скоростей в 3 Гб/с (слайд 27). В новых стандартах рабочей группы IEEE 802.11, которые именуются как IEEE 802.11ac и IEEE 802.11 ad, предусматриваются скорости доступа в 3,6 Гб/с и 7 Гб/с соответственно (слайд 20). Таким образом, технологически внедрение сетей с малыми задержками поддержано уже сегодня.

Далее более подробно остановимся на технологических решениях для построения сотовых сетей подвижной связи с целью определения их роли в процессе развития сетей с малыми задержками и иных перспективных сетей связи. В течение первого десятилетия 21-ого века широкое распространение получили сети третьего поколения 3G (Third Generation). Их численные характеристики к концу десятилетия приведены на слайде 21, включая системы высокоскоростного пакетного доступа HSPA (High Speed Packet Access). Однако, уже в это время были начаты работы по сетям четвертого

поколения, построенных на базе пакетной коммутации и известных в настоящее время как системы длительной эволюции LTE (Long Term Evolution). Необходимость такого перехода обосновывается на слайдах 22, 23 и 24. Помимо очевидной необходимости в увеличении скорости доступа, тем более в условиях появления сетей с малыми задержками, существенную роль сыграла и стоимость передачи речи поверх IP (слайд 24). Системы 3G, несмотря на все их уникальные для сетей связи показатели, были построены на принципах коммутации каналов, что и неудивительно, если внимательно посмотреть на слайд 23. Действительно, в конце 90-х годов (время разработки систем 3G) доля трафика речи поверх IP составляла только единицы процентов от общего речевого трафика, и строить систему на коммутации пакетов было нерационально. Существенно проще было дополнить 3G модулями высокоскоростного пакетного доступа HSPA, а затем и HSPA+, чем делать полностью пакетную систему. Через десять лет ситуация принципиально изменилась: доля трафика VoIP превысила 40% и продолжала увеличиваться, что и сделало необходимостью разработку систем 4G. К технологиям 4G относятся и системы WiMAX, разработанные в соответствии со стандартом IEEE 802.16m. Ничем не уступающие по техническим характеристикам системам LTE, системы WiMAX не могут найти себе широкий рынок сбыта в связи с тем, что операторы сотовой подвижной связи обязаны до 2018 года обслуживать не только новые терминалы, но и терминалы 2-2,5 G (слайд 25).

К настоящему времени практически завершены работы по 11-ой версии LTE и ведется разработка решений и спецификаций для 12-ой версии LTE (слайд 26). Заметим, что в обеих версиях предусматривается тесная интеграция с сетями USN (всепроникающими сенсорными сетями). На слайде 27 можно увидеть, что уже в 11-ой версии LTE достигается пиковая скорость 3Гб/с. Существенно улучшается и эффективность передачи информации в бит/герц (слайд 28). На слайде 28 DL (Down Load) означает передачу информации в нисходящем направлении, а UL (Upper Load) – в восходящем. Аббревиатура MIMO (Multiple Input Multiple Output) означает множественные антенны. Слайд 29 об эволюции задержек в системах 3G и 4G был рассмотрен ранее.

Не меньший интерес вызывает и обслуживание системами LTE приложений Интернета Вещей, в частности и в первую очередь в их физической реализации – сетях M2M (Machine-to-Machine). Подобные сети создают трафик полярный по отношению к сетям с амалой задержкой. Действительно, в сетях M2M существует очень большое число устройств,

которые потенциально могут создавать трафик для загрузки сети. Но величина трафика для каждого конкретного устройства, как правило, невелика – речь идет о передаче целого числа или дроби.

На слайде 30 приведена информация о результатах моделирования для обслуживания трафика M2M системами LTE. Как видим, в зоне действия одной базовой станции системы LTE может быть размещено до 30 тысяч устройств M2M без ущерба для качества предоставления иных услуг связи. Много это или мало? На слайде 31 приведена информация о плотности жителей на кв. км для различных районов С.Петербурга. Будем считать, что на каждого жителя приходится до 10 устройств M2M (смотрите лекцию 3). Тогда можно признать, что сеть из множества базовых станций системы LTE, размещенных с традиционной для современных сетей связи плотностью, сможет обслужить эту нагрузку без ущерба для качества предоставления иных услуг связи. Проблемой использования системы LTE для выполнения этих задач является стоимость точки доступа для устройства M2M, которая составляет примерно 10 \$. Для сетей же WiFi и USN (на протоколе ZigBee) оценка стоимости дает 5\$. Поэтому, одной из основных задач на ближайшее будущее в рамках развития системы LTE является создание экономичной (low cost) системы (слайд 32).

Последнее положение перекликается с задачами, которые стоят сегодня по созданию низкоскоростных мало потребляющих сетей, чему посвящен слайд 33. Все вышесказанное приводит к разделению сетей на гигабитные и низкоскоростные, и одной из актуальных задач в ближайшее время будет разработка систем, предлагающих комплексное решение для обслуживания высокоскоростного и низкоскоростного трафика (слайд 34).

В заключении остановимся на еще одном новом применении системы LTE и виде сетей, которые называются кооперативными. На слайде 35 изображена гетерогенная зона базовой станции LTE, в зоне действия которой в соответствии с современными воззрениями могут находиться и шлюзы сетей USN, и узлы сети VANET, и узлы медицинских сетей и т.д. Естественным образом возникает возможность использования каждого из них в интересах любых членов любых сетей, входящих в зону действия конкретной базовой станции LTE. Способы кооперативной передачи приведены на слайдах 36-38, а целью организации кооперативных сетей является улучшение качества обслуживания пользователей, находящихся у границы зоны базовой станции или в местах с ухудшенными возможностями по приему сигналов.