

## АНАЛИЗ НАГРУЗКИ НА ВЫЧИСЛИТЕЛЬНЫЕ СЕТИ АВТОМАТИЗИРОВАННЫХ СИСТЕМ УПРАВЛЕНИЯ ВОЗДУШНЫМ ДВИЖЕНИЕМ

Институт компьютерных технологий Национального авиационного университета

*В работе рассмотрены характеристики трафика данных в вычислительных сетях современных автоматизированных и компьютеризованных систем УВД. Проанализированы статистические модели трафика данных, даны рекомендации по выбору характеристик и структуры сетей.*

### **Введение**

Как отмечается в [1], актуальность и перспективность объединения сетей различного назначения, предоставления через одну и ту же сеть услуг различного характера, гармонизации компьютерных и телекоммуникационных технологий уже не вызывают сомнений. В связи с постоянным изменением соотношения между объемами телефонного трафика и трафика передачи другой информации (в пользу второго) роль инфраструктуры передачи данных будет возрастать, а организация услуг (в частности, дополнительных видов обслуживания) будет все меньше связана собственно с транспортом информации. Как известно, такой подход является основополагающим при построении интеллектуальных сетей (ИС). Поэтому можно утверждать, что перечисленные тенденции развития – взаимосвязанные составляющие общего процесса модернизации существующих сетей и внедрения сетей новых поколений во всех областях человеческой деятельности, в том числе, конечно, и в сетях авиационной электросвязи, системах организации воздушного движения.

Основным принципом построения как существующих, так и перспективных (интегрированных) сетей объективно является модульный принцип. Это обусловлено множеством исторических, организационных, технических, производственных и других факторов. Характеристики каждого модуля, параметры систем управления и сигнализации согласовываются между собой, и модули объединя-

ются в корпоративные, региональные и общенациональные сети.

Прежде всего, следуя [1], дадим определение сетей новых поколений. Одна из основных целей, которые преследуются при их реализации – это передача и обработка разнородного трафика (речь, данные, видео) с качеством, которое обеспечивается в сетях с коммутацией каналов, в которых предоставляются услуги так называемого операторского класса (с коэффициентом готовности «пять девяток», т. е. не более одного отказа в обслуживании в течение года). В этом смысле можно говорить о пресловутой «конвергенции сетей», под которой подразумевается предоставление возможностей обмена информацией между самыми разными сетями [2,3]:

- телефонными сетями общего пользования (ТФОП);
- интеллектуальными сетями (ИС);
- сетями транкинговой и мобильной связи;
- сетями радиосвязи, в том числе космической связи;
- IP-сетями.

В соответствии с моделью инфокоммуникационной системы, предложенной международным союзом электросвязи (МСЭ), характерной ее особенностью является разделение уровней терминального оборудования клиента, сетей доступа и транспортных сетей, средств сигнализации и управления, средств создания услуг (см. рис. 1).

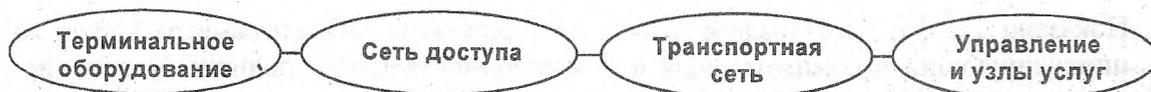


Рис. 1. Модель инфокоммуникационной системы

В перспективных корпоративных сетях систем УВД роль сетей доступа будут играть локальные сети, к которым подключены автоматизированные рабочие места (АРМ) операторов. В организационном плане в такой системе должна обеспечиваться поддержка всего спектра услуг, как на уровне транспортных сетей, так и на уровне сетей доступа для терминалов, находящихся в распоряжении оператора в настоящее время – обычных, мобильных или IP-телефонов, внутренних сетей громкоговорящей связи, персональных компьютеров. Кроме того, по заявкам клиентов необходимо быстро и без остановки работы системы заказчика (в нашем случае – системы УВД) организовывать и/или модифицировать индивидуальные наборы услуг, в том числе и дополнительные виды обслуживания. Другими словами, должна обеспечиваться так называемая «мультисервисность» [4].

Некоторые авторы считают, что конвергентные сети – промежуточный этап на пути к мультисервисным сетям [4]. Иногда и те, и другие сети называют «сетями следующего поколения» (*NGN – Next Generation Networks*) [5]. Однако как бы ни называли сети новых поколений, они характеризуются следующими принципиальными особенностями:

- многослойная инфраструктура с количеством независимых слоев от 4 до 6 (по оценкам разных специалистов), причем каждый из них может создаваться независимо от остальных по аналогии с эталонной моделью открытых систем;

- наличие открытых интерфейсов и стандартных протоколов обмена между аппаратурой доступа, коммутации, управления и сигнализации;

- поддержка старых и создание новых услуг с универсальным доступом из любой подсети того или иного вида – конвергенция услуг связи;

- независимость технологий создания аппаратуры и разработки программного обеспечения (ПО) от технологий передачи и обработки данных;

- решение проблем сигнализации и управления на качественно новом уровне;

- поддержка технологии коммутации пакетов с сохранением в течение некоторого, возможно, достаточно длительного периода, технологии коммутации каналов.

### **Параметры и статистические характеристики трафика конвергентных сетей**

В общем разнородном трафике (речь, данные, видео, другие типы) доля речевого трафика снижается, причем темп этих изменений нарастает их года в год. Например, в США и Западной Европе ежегодный темп роста трафика данных составляет до 30% в год, в то время как рост телефонного трафика – около 3% [2].

Перераспределение видов нагрузки на телекоммуникационные сети имеет такие последствия.

1. Появление второй моды в распределении длительности телефонного соединения: если для телефонных переговоров это 3...5 мин, то для пользователя Интернета – 20 мин (по другим оценкам – 40 мин). В корпоративных сетях длительность пользования Интернетом может быть еще больше.

2. Изменение статистических характеристик трафика. Для телефонных сетей широко используются марковские модели, потоки Пуассона и Эрланга [6,7]. В то же время результаты многочисленных экспериментальных исследования трафика данных [8-12] свидетельствуют о том, что он обладает весьма специфическими свойствами и не может быть удовлетворительно описан в рамках классической теории массового обслуживания.

Показано [13-16], что график данных, циркулирующий в компьютерных и телекоммуникационных цифровых сетях, и, в частности, в сетях с коммутацией пакетов, обладает самоподобными, или фрактальными, свойствами. «Самоподобие» представляет собой свойство процесса сохранять свое поведение и внешние признаки при рассмотрении в разном масштабе. Для временных последовательностей масштабируемой величиной является время. Исходя из определения самоподобия, можно утверждать, что временные и спектральные характеристики случайного процесса (в нашем случае – трафика) при изменении масштаба усреднения будут описываться одними и теми же уравнениями, функциями, но с соответствующими масштабными коэффициентами. Другими словами, самоподобие какого-либо процесса (явления) можно трактовать как инвариантность к изменениям масштаба или размера.

Реальные случайные процессы, конечно, сохраняют свойство самоподобия только до определенного предела. Этот предел или мера статистической устойчивости процесса при многократном масштабировании определяется так называемым параметром Херста или параметром самоподобия [13,15]. Случайный процесс  $x(t)$  является статистически самоподобным с параметром Херста  $H$  ( $0,5 \leq H \leq 1$ ), если для любого вещественного значения  $a > 0$  процесс  $x(at)/a^H$  обладает теми же статистическими характеристиками, что и сам процесс  $x(t)$ : математическое ожидание

$$M[x(t)] = \frac{M[x(at)]}{a^H}; \quad (1)$$

$$\text{дисперсия } D[x(t)] = \frac{D[x(at)]}{a^{2H}}; \quad (2)$$

корреляционная функция

$$R(t, \tau) = \frac{R(at, a\tau)}{a^{2H}}. \quad (3)$$

Чем больше  $H$ , тем дольше сохраняется свойство самоподобия при многократном масштабировании. При  $H = 0,5$  это свойство практически отсутствует.

Для оценивания параметра Херста по экспериментальным данным можно использовать довольно очевидные соотношения, вытекающие из выражений (1-2):

$$a^H = \frac{m(at)}{m(t)},$$

$$\text{откуда } H = \log_a \frac{m(at)}{m(t)}. \quad (4)$$

Корреляционные функции самоподобных процессов с большим параметром Херста затухают медленнее, чем у обычных случайных процессов, причем имеют, как правило, колебательный характер. Установлено, что убывание постоянной составляющей корреляционной функции происходит по закону  $c_1 t^{-c_2 a}$ , где  $c_1, c_2$  – константы,  $a$  – параметр масштаба. Соответственно и спектральная плотность процесса теоретически стремится к бесконечности при частоте, стремящейся к нулю [15].

Для описания плотностей вероятностей самоподобных потоков используют распределения с «тяжелыми хвостами»: логарифмически-нормальное, гамма-распределение, распределения Вейбулла, Парето. Последнее используется для описания самоподобного трафика наиболее часто [8]. Из всех медленно затухающих распределений оно описывается самыми простыми математическими формулами. (Конечно, простота не может служить основанием для использования той или иной модели процесса, поэтому в дальнейшем необходимо будет проверять принадлежность выборки самоподобного процесса к генеральной совокупности с тем или иным вероятностным распределением.)

Выражение для плотности вероятности распределения Парето имеет следующий вид:

$$f(x) = \frac{\alpha}{k} \left( \frac{k}{x} \right)^{\alpha+1}, \quad (4)$$

где  $k$  и  $\alpha$  ( $\alpha < 0$ ) – параметры распределения.

Соответственно функция вероятности  $F(x) = 1 - \left( \frac{k}{x} \right)^\alpha$  ( $x > k; \alpha > 0$ ), (5)

среднее значение  $E[X] = \frac{\alpha}{\alpha - 1} k$  ( $\alpha > 1$ ).

Такие специфические характеристики присущи не только трафику данных (протоколы *TCP*, *FTP*), но и сигнальному трафику (протокол *SS7*), *VBR*-видео, *Ethernet/ISDN* и некоторых других [13,16]. Физически они объясняются высокой степенью группирования пакетов на клиентских участках, в маршрутизаторах и узлах коммутации компьютерных и телекоммуникационных сетей. Даже если источник порождает регулярный поток пакетов, данные до потребителя доставляются сериями, которые перемежаются интервалами простоя. Причинами этого являются ограниченная скорость работы сетевых устройств, недостаточный объем буферов и др. В частности, для сетей *Ethernet* трафик приобретает самоподобные свойства при обработке коллизий и формировании *jam*-последовательностей в условиях большой загрузки сети.

Кроме того, самоподобный трафик имеет особую структуру, сохраняющуюся при многократном масштабировании – в реализации, как правило, присутствует некоторое количество выбросов при относительно небольшом среднем уровне трафика (см. рис. 2). Из-за таких всплесков нагрузки характеристики сети также ухудшаются: увеличиваются потери, задержки, джиттер («дрожание» или нарушение синхронизации) пакетов при прохождении через узлы сети.

Методы расчета требований к сетям новых поколений (пропускной способности каналов, емкости буферов, коэффициента использования и пр.), основанные на марковских моделях и формулах Эрланга, которые с успехом использовались при проектировании телефонных сетей, могут давать неоправданно оптимистические решения и приводить к недооценке нагрузки.

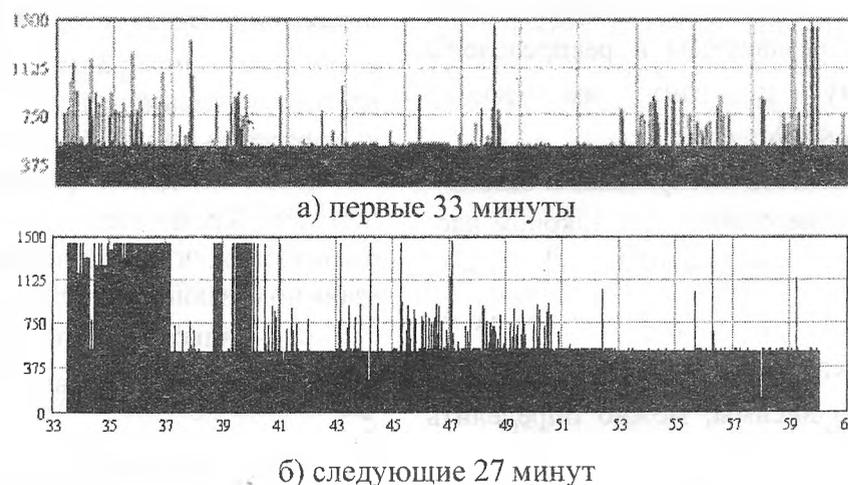


Рис. 2. Пример реализации высокоскоростного трафика данных в локальной сети *Ethernet*. Зависимость длин кадров от времени [9]

Для самоподобного трафика результаты классической теории массового обслуживания [6, 7] нужно применять с некоторыми оговорками. Учитывая пульсирующий характер самоподобного трафика, в общем нельзя считать поток заявок простейшим, поскольку на интервале наблюдения не выполняется условие стационарности.

Однако, исходя из логики предоставления услуг с гарантированным сквоз-

ным качеством обслуживания *QoS*, нужно требовать обеспечения качества обслуживания на интервале любой длительности, случайно выбранном из общего сеанса передачи данных. Как на интервалах с низкой интенсивностью трафика, так и на интервалах, где наблюдаются всплески нагрузки, трафик можно с достаточной для практики точностью считать локально-стационарным. Вся проблема заключается в том, чтобы определить моменты

перехода от одного интервала к другому. В принципе, конечно, можно предложить алгоритмы адаптации к изменениям нагрузки, например, с оцениванием корреляционных свойств потока данных, однако вряд ли стоит ожидать приемлемой точности, а, следовательно, и высокой эффективности таких алгоритмов.

На наш взгляд, больший интерес представляет получение асимптотических сравнительных оценок для классического пуассоновского и самоподобного потоков.

Следуя [7], рассмотрим одноканальную систему массового обслуживания (СМО) с ожиданием класса  $GI/G/1$ . Поскольку, как отмечено выше, корреляционная функция самоподобного трафика не является экспоненциальной, входящий поток заявок следует считать потоком с ограниченным последствием. Заявки поступают в последовательные дискретные моменты  $t_i, t_{i+1}, \dots, t_n, \dots, t_j \leq t_{j+1}$  для любого  $j$ , интервалы между ними  $\tau_n = t_n - t_{n-1}$  независимы и распределены по одному и тому же закону  $F_n(\tau) = P\{\tau_n < \tau\}, n \geq 2$ .

Длительности обслуживания заявок – независимые величины  $\zeta_n$  с законом распределения  $\Psi_n(\zeta) = P\{\zeta_n < \zeta\}, n \geq 1$ . Обозначим  $\xi_n = \zeta_{n-1} - \tau_n$ . Тогда при условии что последовательности  $\{\tau_n\}$  и  $\{\zeta_n\}$  взаимно независимы, можно определить вероятность

$$\Theta(\tau) = P\{\xi_n < \tau\} = \int_0^{\infty} \overline{F}_n(\eta - \tau) d\Psi_n(\eta), \quad (6)$$

где  $\overline{F}_n(\eta - \tau) = 1 - F_n(\eta - \tau)$ .

Обозначим длительность ожидания  $n$ -й заявки через  $\omega_n$ . Если  $n$ -я заявка поступит сразу вслед за  $(n-1)$ -й, ей, с учетом величины интервала  $\tau_n$ , придется ждать обслуживания  $\omega_{n-1} + \zeta_{n-1} - \tau_n = \omega_{n-1} + \xi_n$  единиц времени. Однако при достаточно больших  $\tau_n$  величина  $\omega_{n-1} + \xi_n$  может формально стать отрицательной. Ясно, что в этом случае действительное время

ожидания  $n$ -й заявки будет равно нулю – очереди нет, и заявка поступает на обслуживание сразу по приходе. Следовательно, выполняется рекуррентное соотношение  $\omega_n = \max_n \{\omega_{n-1} + \xi_n, 0\}$ . (7)

Обозначим  $G_n(x) = P\{\omega_n < x\}$ . Тогда соотношение (7) можно выразить через функции распределения следующим образом:

$$G_{n+1}(x) = \begin{cases} \int_0^x G_n(x-y) d\Theta(y), & x > 0, n \geq 2; \\ 0, & x \leq 0, n \geq 1. \end{cases} \quad (8)$$

Дополним выражения (8) очевидным соотношением для функции распределения времени ожидания первой заявки:

$$G_{n+1}(x) = \begin{cases} 1, & x > 0; \\ 0, & x \leq 0. \end{cases} \quad (9)$$

Выражения (6) и (8) представляют собой интегралы Стильеса, которые в случае непрерывных почти всюду распределений  $\Psi_n(t)$  и  $\Theta_n(t)$  превращаются в обычные интегралы. Таким образом, используя выражения (8 – 9), можно рекуррентно вычислять распределения длительности ожидания для заявки с любым номером. Кроме того, как отмечено в [7], оказывается, что они применимы и при взаимной зависимости последовательностей случайных величин  $\{\tau_n\}$  и  $\{\zeta_n\}$ . Имеет значение лишь независимость величин  $\xi_n$ .

Логично предположить, что время обслуживания заявки, например, время обработки пакета в программном коммутаторе, связано функциональной зависимостью с длиной пакета. Тогда, зная характеристики длительности пакетов на входе коммутатора как СМО, можно на участках локальной стационарности входящего трафика конкретизировать параметры распределения времени обслуживания. Например, при группировании однородных пакетов (что характерно для самоподобного трафика), можно сделать допущение о детерминированном времени обслуживания (модель  $GI/D/1$ ).

Выше рассматривалась СМО с ожиданием, ограниченным только длиной очереди. В такой СМО заявка, поступившая в очередь, уже не покидает ее и «терпеливо» дожидается обслуживания [6]. Однако в сетях систем УВД весь трафик сети с разнородными потоками уже не может считаться «эластичным», т.е. не имеющим ограничений на время ожидания. Для таких изохронных приложений, как передача голоса, при превышении времени задержки более чем на 100... 150 мс резко снижается качество воспроизведения [17]. Ухудшается разборчивость речи, что в некоторых ситуациях может привести к недопустимым последствиям. Теоретически при превышении порога ожидания заявка может уйти из очереди (так называемые «нетерпеливые» заявки). Практически, конечно, в системах УВД такая ситуация, как правило, недопустима. Для СМО с «нетерпеливыми» заявками понятие «вероятность отказа» не имеет смысла – каждая заявка становится в очередь, но может и не дожидаться обслуживания, уйдя раньше времени. Для предотвращения потери таких «нетерпеливых» заявок и, соответственно, ухудшения качества обслуживания целесообразно применять алгоритмы обслуживания с приоритетами типа маркировки потоков [7] или ранжирования пакетов по длительности [13].

Если все каналы обслуживания заняты и имеется очередь заявок, то, как известно [7], поток обслуженных заявок можно считать простейшим. Сделаем также допущение о простейшем характере потока «нетерпеливых» заявок в общем потоке. Относительная пропускная способность системы  $q$  вычисляется из предположения, что будут обслужены все заявки, кроме тех, которые уйдут из очереди досрочно. Поэтому для нахождения среднего числа заявок, покидающих очередь досрочно, вычислим среднее число заявок в очереди:

$$\bar{r} = 1 \cdot p_{n+1} + 2 \cdot p_{n+2} + \dots + r \cdot p_{n+r} + \dots$$

На каждую из них действует «поток уходов» с интенсивностью  $v$ . Значит, из среднего числа  $\bar{r}$  заявок в очереди в среднем будет уходить, не дождавшись обслуживания,  $v\bar{r}$  заявок в единицу времени; всего в единицу времени в среднем будет обслужено  $A = \lambda - v\bar{r}$  заявок.

Относительная пропускная способность СМО будет  $q_p = 1 - \frac{v\bar{r}}{\lambda}$ , среднее число занятых каналов (из общего числа  $n$ )  $\bar{z} = \rho - \beta\bar{r}$ , среднее число заявок в очереди  $\bar{r} = \frac{\rho}{\beta} - \frac{\bar{z}}{\beta}$ . Здесь  $\lambda, \mu$  – интенсивности потока заявок и обслуживания соответственно;  $\rho = \frac{\lambda}{\mu}$ ,  $\beta = \frac{v}{\mu}$ .

### Оценка требований к параметрам сетевого оборудования

С учетом полученных выше результатов проанализируем требования к терминальному и коммутационному оборудованию сетей для простейшего и самоподобного входного потоков.

При самоподобной природе трафика зависимость средней длительности очереди (соответственно, необходимого размера буфера)  $q$  от среднего коэффициента использования имеет следующий вид [13]:

$$q = \frac{\rho^{1/2(1-H)}}{(1-\rho)^{H/(1-H)}} \quad (10)$$

При  $H=0,5$  эта формула упрощается:

$$q = \rho / (1 - \rho), \quad (11)$$

что представляет собой классический результат СМО с простейшим входным потоком и экспоненциально распределенным временем обслуживания (М/М/1). Для системы с детерминированным временем обслуживания (М/D/1) классический результат выглядит следующим образом:

$$q = \frac{\rho}{1-\rho} - \frac{\rho^2}{2(1-\rho)}. \quad (12)$$

По формулам (10-12) проведены расчеты [1], результаты которых изображены на рис. 3.

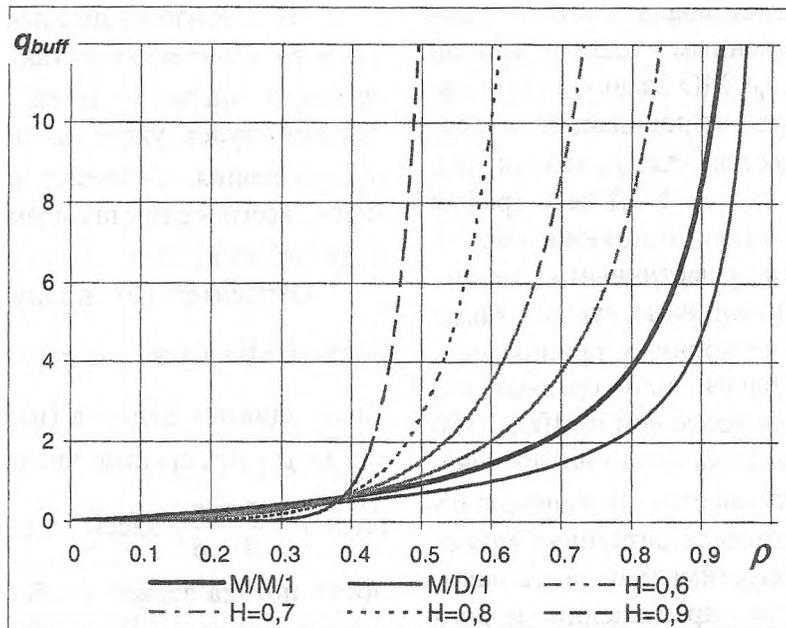


Рис. 3. Зависимости длины очереди заявок (потребной памяти буфера  $q_{buff}$ ) от коэффициента использования  $\rho$  для различных моделей входного трафика

На графиках хорошо видно, что для самоподобного трафика уже при  $\rho = 0,4$  требуется больший ресурс памяти буферных устройств, чем для классической модели  $M/M/1$ , которая считается наименее благоприятной по сравнению с другими (например, с постоянным или нормально распределенным временем обслуживания). Скорость роста требуемого объема памяти растет при увеличении параметра Херста, который обусловлен, в основном, степенью группирования однородных пакетов и всплесками нагрузки на сеть.

Очевидно, простое наращивание буферной памяти (аппаратным или программным способом) является малоэффективным. При ожидаемом увеличении доли трафика данных в общем объеме степень самоподобия будет увеличиваться, и зависимость  $\rho(q_{buff})$  будет все более резкой. Кроме того, ожидание в очереди на обработку больших массивов информации для таких систем реального времени, как АС УВД, вообще недопустимо.

Повышение производительности новых образцов программных коммутаторов типа *Softswitch* внушает некоторый оптимизм, однако нельзя забывать, что

любые сетевые ресурсы неожиданно быстро истощаются при непрерывном появлении новых услуг и приложений. Хорошие возможности снижения коэффициента использования возникают при увеличении числа независимых параллельных каналов программного коммутатора с организацией общей очереди к нескольким входным портам. При этом упрощаются алгоритмы обработки приоритетных потоков трафика данных программными методами. Поэтому важной задачей является постоянное совершенствование программного обеспечения узлов коммутации, в частности, интерфейсов прикладного программирования *API*.

Поскольку локальные сети систем УВД строятся, как правило, по технологии *Ethernet*, проблема ограничения коэффициента использования сети является еще более острой по причинам, связанным именно с принципами работы, заложенными в этой технологии. Как отмечалось выше, при увеличении коэффициента использования сети  $k_{ucn}$  — отношения пропускной способности к интенсивности трафика — все большая часть ресурса тратится на обработку коллизий. При стремлении  $k_{ucn}$  к единице будет стремиться к

единице вероятность коллизий и, соответственно, появление все новых и новых jam-последовательностей. Пытаясь обрабатывать коллизии, сеть перестанет пропускать полезную информацию и будет работать «на себя». При обосновании предельно допустимого коэффициента использования сети необходимо учитывать риск возникновения такой ситуации.

Рассмотрим задачу оценивания вероятности возникновения коллизии в сети Ethernet. Пусть к сети подключено  $N$  компьютеров, которые выдают потоки пакетов с интенсивностями  $\lambda_n(t)$ ,  $n = \overline{1, N}$ . Будем считать, что пакеты, выдаваемые  $n$ -м компьютером, имеют длительность  $\tau_n$ . Чтобы не связывать себя необходимостью точного учета момента появления каждого пакета, предположим, что они происходят в случайные моменты времени с одинаковым вероят-

ностным распределением, и на конечном отрезке времени  $T \gg \tau_n$  образуют поток Эрланга  $k$ -го порядка. Как известно [6], задаваясь порядком потока Эрланга, можно получить любую степень последствия: от полной взаимной независимости между моментами появления событий при  $k=0$  до детерминированной функциональной связи при  $k \rightarrow \infty$ .

Пусть в точке  $t_j$ , находящейся внутри интервала анализа  $T$ , появился пакет  $f_j$  с длительностью  $\tau_j$ . Вероятность появления этого пакета обозначим  $P(t_j)$ . Выведем выражение для условной плотности вероятности частичного перекрытия пакета  $f_j$  другим пакетом  $f_i$  с длительностью  $\tau_i$ . Возможные положения пакетов изображены на рис. 4.

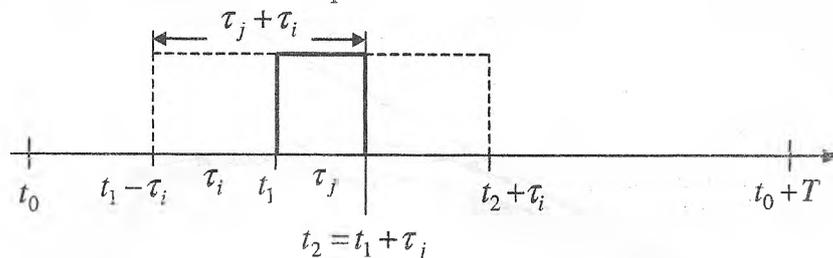


Рис. 4. Возможные положения пакетов

Будем считать, что перекрытие происходит уже при соприкосновении пакетов. Тогда необходимо вычислить совместную вероятность  $P(t_i, t_j)$  того, что пакет  $f_i$  попадет в интервал  $[t_1 - \tau_i, t_1 + \tau_j]$  при условии, что пакет  $f_j$  начинается в точке  $t_1$ . С учетом временных соотношений (см. рис. 4) эта вероятность определяется как

$$P(t_i, t_j) = P(t_1 - \tau_i \leq t \leq t_1 + \tau_j / t_1 = t_j) \quad (13)$$

Логично предположить, что процессы появления в сети пакетов, выдаваемых разными источниками, взаимно независимы, а моменты появления пакета  $f_j$  на интервале  $[t_0, t_0 + T]$  равновероятны. Тогда

$$P(t_i, t_j) = P(t_1 - \tau_i \leq t \leq t_1 + \tau_j / t_1 = t_j) = P(t_1 - \tau_i \leq t \leq t_1 + \tau_j) P(t_1) \quad (14)$$

где  $P(t_1) = P(0 \leq t_j < t_j + \tau_j \leq T) = \tau_j / T$  (15) – вероятность того, что пакет  $f_j$  «накроет» интервал длительностью  $\tau_j$  в какой-либо точке  $t_j$ .

Положим, что в момент времени  $t_1 - \tau_i$  пакет  $f_i$  не наблюдается. Вычислим вероятность  $P(t_1 - \tau_i \leq t_i \leq t_1 + \tau_j)$  появления  $f_i$ -го пакета за отрезок времени  $[t_1 - \tau_i, t_1 + \tau_j]$ , используя выражения для закона Эрланга  $k$ -го порядка [6], в которых для учета нестационарности введем переменную интенсивность потока  $\lambda_i(t)$ :

$$P(t_1 - \tau_i \leq t \leq t_1 + \tau_j) = \frac{\lambda_i(t_1 - \tau_i)(\lambda_i(t_1 - \tau_i + t))^{k-1}}{(k-1)!} \exp\left(-\int_{t_1 - \tau_i}^{t_1 + \tau_j} \lambda_i(t) dt\right). \quad (16)$$

Если считать интенсивность потока  $\lambda_i(t)$  медленно меняющейся величиной, то можно с приемлемой точностью усреднить ее на интервале интегрирования  $[t_1 - \tau_i, t_1 + \tau_j]$ :  $\lambda_i(t) \approx \lambda_{icp}$ . Тогда выражение (16) упрощается:

$$P(t_1 - \tau_i \leq t \leq t_1 + \tau_j) = \frac{\lambda_{icp} (\lambda_{icp} t)^{k-1}}{(k-1)!} \exp(-\lambda_{icp} t). \quad (17)$$

После несложных, но громоздких преобразований выражение (17) преобра-

зуется для случая самоподобного трафика по аналогии с выражением (10):

$$P(t_1 - \tau_i \leq t \leq t_1 + \tau_j) = \frac{\lambda_{icp}^{H/(1-H)} (\lambda_{icp}^{H/(1-H)} t)^{k-1}}{(k-1)!} \exp\left[-\lambda_{icp}^{H/(1-H)} [H/(1-H)] t\right], \quad (18)$$

где  $H$  – параметр Херста. По формулам (16 – 18) рассчитаны зависимости полезной пропускной способности сети от коэффициента использования (рис. 5) для Пуассоновского трафика и самоподобного трафика с разными параметрами Херста. Под идеальным трафиком мы понимаем отсутствие коллизий. Видно, что производительность сети заметно убывает уже при коэффициенте использования, превышающем значение 0,3...0,4 (особенно для самоподобного трафика).

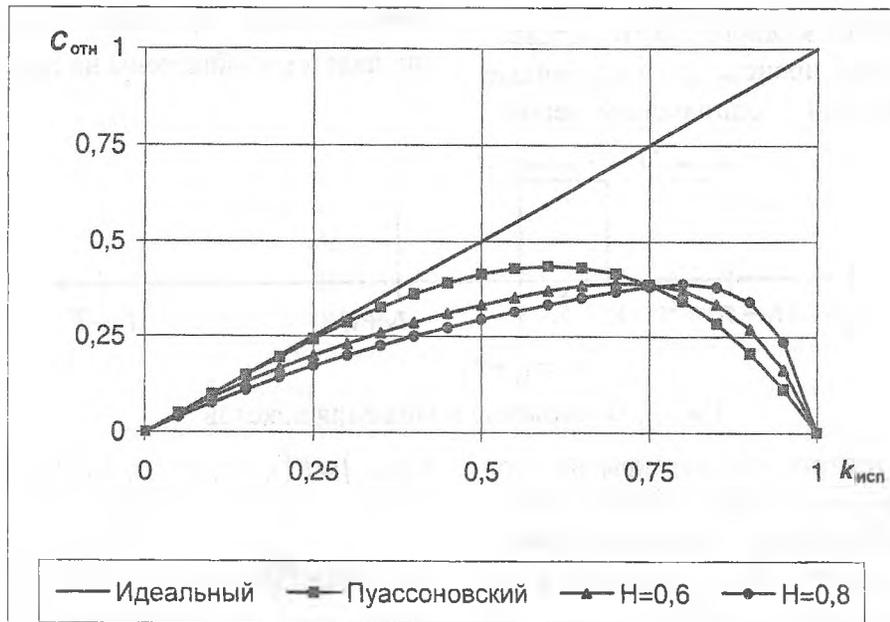


Рис. 5. Зависимость относительной пропускной способности сети *Ethernet* от коэффициента использования при разных видах трафика данных ( $H$  – параметр Херста самоподобного трафика)

### Вывод

Сети новых поколений – ковергентные, мультисервисные, интеллектуальные – необычайно расширяют возможности информационного обмена, однако порождают и множество новых проблем. Необходимость сочетания в одних сетевых устройствах и узлах новых технологий передачи данных с существующими требует совершенных и весьма сложных технологических и алгоритмических решений.

Перераспределение соотношения различных видов трафика приводит к изменениям его статистических характеристик, которые необходимо учитывать при обосновании требований к аппаратуре коммутации и управления. Гипотеза самоподобия пакетного трафика подтверждена теоретически и экспериментально, однако, исследования в этом направлении далеко не завершены. Например, остается открытым вопрос: справедлива ли эта гипотеза на интервалах, длительность кото-

рых сравнима с длительностью интервалов оценивания мгновенной пропускной способности программного коммутатора.

Простое наращивание (буферной) памяти и производительности программных коммутаторов не может дать долговременного эффекта без разработки новых алгоритмов оптимальной обработки и передачи трафика: перераспределения, сегментации, адаптации к изменению характеристик. Для обеспечения требуемой эффективности функционирования сети и качества обслуживания необходим постоянный контроль пропускной способности и предотвращение перегрузок на отдельных фрагментах и во всей сети.

Эффективным методом уменьшения бесполезной нагрузки на сети Ethernet является логическая структуризация с помощью коммутаторов или разделение на изолированные фрагменты с помощью мостов. При этом в каждом фрагменте уменьшается количество компьютеров и, соответственно, уменьшаются размеры доменов коллизий. Окончательный выбор метода уменьшения размеров доменов зависит от конкретных технических, экономических и других факторов.

#### Список литературы

1. Виноградов Н.А. Анализ потенциальных характеристик устройств коммутации и управления сетями новых поколений // – К.: Зв'язок, 2004. – №4. – С. 10-17.
2. Гольдштейн Б.С. Интеллектуальная сеть плюс IP – шаг к конвергенции сетей // Компьютерная телефония, № 6, 2000. – С. 11-17.
3. Гольдштейн Б.С., Орлов О.П., Ошев А.Т., Соколов Н.А. Эволюция услуг в сетях следующего поколения // Вестник связи № 7, 2003. – С. 23-28.
4. Гольдштейн А.Б. Проблемы перехода к мультисервисным сетям // Вестник связи № 12, 2002. – С. 14-18.
5. Хиленко В.В., Михайлов В.Ф. Сетевой интеллект и сети нового поколения // «Зв'язок», № 1, 2003. – С. 2-5.
6. Венцель Е. С. Исследование операций. – М.: Советское радио, 1972. – 552 с.
7. Гнеденко Б.В., Коваленко И.Н. Введение в теорию массового обслуживания, 2-е изд. – М.: Наука, 1987. – 336 с.
8. Paxson V. and Floyd S. Wide Area traffic: The Failure of Poisson Modeling. // IEEE/ACM Transactions on Networking, v.3. – N3, 1995. – PP. 321-327.
9. H. J. Fowler, W. E. Leland. Local Area Network Traffic Characteristics, with Implications for Broadband Network Congestion Management, IEEE JSAC, 9(7), September 1991. – PP. 1139-1149.
10. Theiler J. Some comments on the correlation dimension of  $1/f_a$  noise // Phys. Lett. A. 155 (1991). – PP. 480-493.
11. A. Erramilli, J. Gordon, W. Willinger. Applications of Fractals in Engineering for Realistic Traffic Processes / Proceedings, International Telecommunications Conference (ITC-14), Amsterdam: Elsevier, 1994. – PP. 56-63.
12. A. Erramilli, O. Narayan, W. Willinger. Experimental Queuing Analysis with Long-Range Dependent Packet Traffic / IEEE/ACM Transactions on Networking, April 1996. – PP. 85-93.
13. Столлинс В. Современные компьютерные сети. 2-е издание. – СПб.: Питер, 2003. – 783 с.
14. M. Garrett and W. Willinger. Analysis, Modeling, and Generation of Self-Similar VBR Video Traffic / Proceedings, SIGCOMM'94, August 1994. – PP. 181-190.
15. Городецкий А.Я., Заборовский В.С. Информатика. Фрактальные процессы в компьютерных сетях / – С.Пб.: Изд-во СПбГТУ, 2000. – 102 с.
16. Криштофович А.Ю. Построение прогноза нагрузки сети ОКС №7 // Доклады 5-й международной конференции DSPA-2003 – М., 2003. – С. 3-5.
17. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. 2-е изд. – С.Пб.: Питер, 2003. – 884 с.