

Теоретическая модель Ethernet-коммутатора

© М.К. Бойченко, И.П. Иванов

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

Предложена теоретическая модель коммутатора, построенная на основе особенностей структуры кадров технологии Ethernet. Рассмотрены функционирование гипотетического коммутатора любого иерархического уровня топологической схемы построения транспортной системы компьютерной сети. Приведены расчетные зависимости для определения вероятности и математического ожидания времени взаимной блокировки кадров в коммутаторах, реализующих второй уровень эталонной модели ISO/OSI. Установлена связь между задержкой коммутатора и требованиями к объему потребной буферной памяти его входных и выходных интерфейсов для любых топологий пересечения информационными потоками в транспортных системах.

Ключевые слова: компьютерная сеть, транспортная система, транзитный узел, коммутатор, блокировка, задержка передачи информации, буферная память.

Известно, что большинство компьютерных сетей организованы по принципу Ethernet-коммутации. Каждый уровень транспортной системы корпоративной компьютерной сети, реализованной по иерархической схеме, строится с использованием различных Ethernet-коммутаторов, выбор которых зависит, прежде всего, от характеристик реализуемой транспортной системы, требуемой производительности, пропускной способности отдельных сегментов сети, времени реакции сети и т. д. Отдельные сегменты могут строиться на базе беспроводных технологий, например, Wi-Fi.

Среди указываемых фирмами-производителями характеристик нет информации о времени передачи кадров с порта на порт, что не позволяет разработчикам транспортной системы установить время передачи пакета через канал (Transmit Time), время реакции сети (Delay), колебание (вариация) задержки при передаче пакетов (Jitter), т. е. многие показатели качества обслуживания (QoS – Quality of Service).

Для повышения производительности транспортной системы компьютерной сети необходимо, прежде всего, предсказывать изменение показателей ее функционирования в случае изменения топологической схемы этой системы или изменения отдельных транзитных ее участков и элементов. Иными словами, отсутствие информации о временных характеристиках транзитных узлов не позволяет построить математическую модель как самих транзитных элементов, так и транспортной системы компьютерной сети в целом.

Для построения функциональной модели Ethernet-коммутатора прежде всего требуется определение размера передаваемых кадров L_K . Полезной нагрузкой для кадра, как правило, является IP-пакет [1]. В зависимости от версии IP-протокола переносимая в нем информация предваряется заголовком, размер которого для IPv4 в штатном режиме 20 байт.

Ethernet-кадр предваряется шестью байтами MAC-адреса назначения, шестью байтами MAC-адреса источника и двумя байтами размера кадра (EthernetDIX) [2], а заканчивается четырьмя байтами контрольной суммы (CRC — Cyclic Redundancy Check), что приводит к длине собственно кадра

$$L_K = L_P + 18 \text{ байт.} \quad (1)$$

где L_P — размер в байтах полезной нагрузки Ethernet-кадра.

В соответствии со стандартом для синхронизации источника и приемника перед каждым кадром передается семь байт преамбулы и один стартовый байт. Поэтому передача кадра длиной L_K требует $(L_K + 8)$ передаваемых байт, т. е.

$$L_t = L_K + 8 = L_P + 26. \quad (2)$$

Наконец, после каждого кадра (после передачи четырех байт CRC) выдерживается межкадровый интервал (IPG — Interpacket Gap) размером 12 байт. Иными словами, каждый из кадров размером L_K байт передается в технологическом окне (фрейме) размером L_f :

$$L_f = L_t + 12 = L_P + 38 \text{ байт.} \quad (3)$$

Формулы (1)–(3) позволяют исключить самоблокировку кадров при проведении экспериментальных исследований, определяя минимальный размер интервала времени для посылки служебных кадров. Так, для наименее быстрой пропускной способности технологии Ethernet $R = 10$ Мбит/с бит-таймом $\tau = 100$ нс для $L_P = 1500$ байт получим $L_f = 1538$ байт и $\Delta t_{\min} = 8 \cdot L_f \cdot \tau = 1,2304 \cdot 10^{-3} \text{ с} = 1,2304 \text{ мс}$,

что соответствует интенсивности $\lambda_{\max} = \frac{1}{\Delta t_{\min}} \cong 812$ кадр/с. Передача

кадров с более высокой интенсивностью от источника приводит к их самоблокировке, переполнению буферов в транзитных или оконечных узлах сети и возможной потере.

Рассмотрим более детально механизм передачи фрейма длиной L_f байт последовательно из выходного аппаратного буфера в кабельный сегмент. При пропускной способности интерфейсов $R = 10$

Мбит/с на физическом уровне эталонной модели ISO/OSI и технологии Ethernet применяется манчестерский код [1, 2], в соответствии с которым на каждый бит передаваемой полезной информации приходится один бит физического кода. Длительность одного бит-тайма для этой пропускной способности $\tau = 100$ нс, поэтому L_f байт кадра будут переданы за время

$$\tau_{10} = 8 \cdot L_f \cdot \tau = 800 \cdot L_f \text{ нс.} \quad (4)$$

Для Fast Ethernet пропускная способность $R = 100$ Мбит/с, а длительность одного бит-тайма $\tau = 10$ нс, но на физическом уровне используется код 4В/5В [2], для которого каждые четыре бита полезной информации требуют передачи в линию связи пяти бит, поэтому

$$\tau_{100} = \frac{5}{4} \cdot 8 \cdot L_f \cdot \tau = 100 \cdot L_f \text{ нс.} \quad (5)$$

Для пропускной способности Gigabyte Ethernet $R = 1000$ Мбит/с длительность бит-тайма $\tau = 1$ нс. При оптоволоконных кабелях связи на физическом уровне используется код 8В/10В, т. е. каждый байт передаваемой информации требует отправки в линию связи 10 бит. Для медных кабелей применяется код РАМ5 [3], в соответствии с которым восемь бит полезной информации на физическом уровне приводят к 10 отправляемым в линию битам, поэтому

$$\tau_{1000} = \frac{10}{8} \cdot 8 \cdot L_f \cdot \tau = 10 \cdot L_f \text{ нс.} \quad (6)$$

Формулы (4)–(6) необходимы для определения задержки коммутаторов при разработке математических моделей их функционирования. Рассмотрим более подробно алгоритм работы коммутатора в ненагруженном режиме.

1. Кадр побитно из кабельного сегмента поступает в буфер входного интерфейса.

2. После поступления первых 14 байт кадра в буфер (восемь байт преамбулы вместе со стартовым битом и шестью байтами MAC-адреса приемника) по таблице коммутации устанавливается номер выходного порта коммутатора.

3. В случае ненагруженного режима функционирования коммутатора для поступающего в буфер входного интерфейса кадра нет конкуренции за выходной порт с другими входными интерфейсами, поэтому сразу после установления номера выходного интерфейса может быть начата порционная передача информации из буфера входного в буфер выходного порта.

4. При коммутации «на лету» [2] после поступления первой порции информации из буфера входного интерфейса в буфер выход-

ного возможно ее дальнейшее побитное продвижение в сегмент транспортной системы, подключенный к выходному интерфейсу. Однако практически все современные коммутаторы используют тип коммутации с полной буферизацией поступающего кадра. Это объясняется, во-первых, возможным отличием пропускных способностей входного и выходного портов, что характерно при иерархическом построении транспортной системы компьютерных сетей и, во-вторых, желанием повысить производительность всей транспортной системы компьютерной сети, т. е. разрешением на эвакуацию кадра из выходного порта в подключенный к нему сегмент сети только после установления факта приема в буфер входного порта неискаженного кадра, т. е. после приема последних четырех байт коммутируемого кадра, в которых содержится CRC. Принятый CRC сравнивается с аппаратно-подсчитываемым во входном интерфейсе по мере поступления входного побитного потока. В случае расхождения значений CRC информация в частично заполненном буфере выходного интерфейса уничтожается (буфер очищается), предотвращая тем самым распространение по транспортной сети заведомо ложных кадров Ethernet.

5. При получении разрешения на эвакуацию в соответствии с реализуемой политикой приоритетов в организации очередей в аппаратном буфере выходного интерфейса начинается побитная передача кадра в кабельный сегмент транспортной системы компьютерной сети. Для ненагруженного режима функционирования коммутатора к этому моменту времени аппаратная очередь выходного порта чаще всего оказывается пустой (если только размер предшествующего передаваемого кадра несущественно превосходит размер коммутируемого). В любом случае за 96-битный межкадровый интервал должна закончиться передача оставшейся в буфере входного интерфейса части коммутируемого кадра и его CRC в буфер выходного порта, а сам входной интерфейс — подготовлен к приему следующего Ethernet-кадра из транспортной системы сети.

Таким образом, задержка кадра коммутатором при отсутствии самоблокировки оказывается равной времени приема фрейма, которое для различных пропускных способностей рассчитывается по формулам (4)–(6). Отметим, что время эвакуации кадра из буфера выходного порта может быть найдено по тем же самым зависимостям, что и время их приема во входной буфер. Это позволяет сделать заключение о том, что переполнение буфера выходного интерфейса возможно только при превышении пропускной способности входного интерфейса таковой для выходного порта. В этой ситуации «обрубание хвостов» приводит к потере кадров даже в ненагруженном коммутаторе.

Резюмируя вышесказанное, в качестве функциональной модели ненагруженных коммутаторов с N интерфейсами, каждый из которых работает в режиме full-duplex, допустимо использовать модель независимых параллельных одноканальных СМО (систем массового обслуживания) с временем обслуживания прямо пропорциональным размерам поступающих на коммутацию Ethernet-кадров. Число этих СМО равно количеству портов коммутатора.

Топология независимой параллельной передачи информационных потоков по N симплексным каналам для N -портового коммутатора весьма маловероятна в современных транспортных системах компьютерных сетей, построенных по иерархическому принципу. Значительно чаще встречаются топологические схемы передачи информационных потоков, для которых один и тот же выходной интерфейс востребован несколькими входными потоками, например, для передачи информации с уровня доступа на распределительный уровень сети (или с ее распределительного уровня на уровень ядра) при поступлении одновременных запросов от клиентов сети к какому-либо общему ее информационному ресурсу. Иными словами, конкуренция за какой-либо выходной интерфейс коммутатора между его входными информационными потоками является штатной ситуацией, следовательно, взаимная блокировка кадров практически неизбежна. По этой причине необходимо иметь методики расчета вероятностей блокировки кадров, а также времени возможной блокировки и математические модели функционирования Ethernet-коммутаторов, учитывающие мультиплексирование нескольких входных информационных потоков на одном выходном интерфейсе. В [4] предложена вероятностная модель Ethernet-коммутаторов локальных вычислительных сетей, справедливая для дуплексного режима их функционирования, при котором конкретный интерфейс коммутатора дуплексно связан с другим, т. е. двунаправленная передача информации возможна только между двумя скомутированными в данный момент времени портами и невозможна одновременная передача кадра из рассматриваемого порта в другой при одновременном приеме кадра в этот рассматриваемый порт из какого-либо третьего порта. Этой конструктивной особенностью обладают образцы коммутаторов, построенные на коммутационных матрицах. Вместе с тем результаты исследований свидетельствуют о том, что современные коммутаторы второго уровня реализуют режим full-duplex, при котором передача кадра из одного интерфейса в другой допускает одновременный прием в передающий порт кадра какого-либо иного порта коммутатора. Иными словами, передача кадра из буфера входного интерфейса в буфер выходного его порта невозможна только в случае, если в данный момент времени уже ведется передача в этот буфер кадра из какого-

либо другого интерфейса коммутатора, т. е. до окончания этой передачи, что соответствует принципу неразрывности процесса продвижения коммутатором передаваемых кадров [1, 2].

Построение вероятностной модели коммутатора технологии Ethernet базируется на задании для каждого момента времени вероятности востребованности рассматриваемого j -го выходного порта всеми остальными его входными интерфейсами, т. е. знанием матрицы значений p_{ij} , $i=1, N$; $j=1, N$, где N — число портов коммутатора. Кроме того, должны быть известны интенсивности потоков информации λ_{ij} (кадр/с, 1/с) из i -го в j -й интерфейс и размеры передаваемых кадров L_{ij} байт. Диагональные элементы матриц p_{ii} , λ_{ii} и L_{ii} тождественно равны нулю для режима full-duplex, микросегментации и соединения «точка — точка» между портами каскада коммутаторов.

Введем понятие вероятности невостребованности j -го выходного интерфейса:

$$q_{ij} = 1 - p_{ij}, \quad i = 1, N, \quad j = 1, N. \quad (7)$$

Очевидно, что вероятность невостребованности j -го выходного порта ни одним из входных портов коммутатора равна произведению:

$$\Pr 0_j = \prod_i (1 - p_{ij}) = \prod_i q_{ij}, \quad i = 1, N, \quad j = 1, N. \quad (8)$$

Вероятность востребованности j -го выходного порта одним и только одним из его входных интерфейсов

$$\Pr 1_j = \sum_i p_{ij} \cdot \prod_{l \neq i} (1 - p_{lj}), \quad i = 1, N, \quad j = 1, N, \quad l = 1, N. \quad (9)$$

В формулах (8) и (9) $p_{ii} = 0$, $q_{ii} = 1$, $i = 1, N$. С вероятностью $(1 - \Pr 0_j - \Pr 1_j)$ j -й выходной интерфейс востребован более чем одним из входных портов коммутатора, поэтому вероятность взаимной блокировки кадров на j -м выходном интерфейсе

$$\Pr B_j = 1 - \Pr 0_j - \Pr 1_j. \quad (10)$$

При невостребованности j -го выходного порта коммутатора в рассматриваемый момент времени $\Pr 0_j = 1$ и $\Pr 1_j = 0$, что дает ожидаемый результат $\Pr B_j = 0$. Если j -й выходной порт востребован только одним из входных интерфейсов коммутатора, то $\Pr 0_j = 0$, $\Pr 1_j = 1$ и вероятность блокировки также равна нулю ($\Pr B_j = 0$). Отличие $\Pr B_j$ от нуля свидетельствует о возможной блокировке и дополнительной задержке входных информационных кадров в коммутаторе.

Рассмотрим ситуацию, при которой j -й выходной порт в рассматриваемый интервал времени с единичной вероятностью востребован одновременно K входными интерфейсами. Условие нормирования вероятностей в этом случае приводит к равенству

$$p_{ij} = \frac{1}{\sum_{l=1}^K P_{lj}} = \frac{1}{K}, \quad i = 1; K. \quad (11)$$

Следовательно, по формулам (8) и (9) находим $\text{Pr} 0_j = \left(1 - \frac{1}{K}\right)^K$ и

$$\text{Pr} 1_j = K \frac{1}{K} \left(1 - \frac{1}{K}\right)^{K-1}. \quad \text{При } K \rightarrow \infty \text{ имеем [1]}$$

$$\text{Pr} 0_\infty = \lim_{K \rightarrow \infty} \left(1 - \frac{1}{K}\right)^K = \frac{1}{e} \cong 0,367879\dots; \quad (12)$$

$$\text{Pr} 1_\infty = \lim_{K \rightarrow \infty} \left(1 - \frac{1}{K}\right)^{K-1} = \frac{1}{e} \cong 0,367879\dots \quad (13)$$

Наиболее неблагоприятный режим работы N -портового коммутатора с точки зрения возможной взаимной блокировки кадров реализуется, когда какой-либо j -й выходной интерфейс востребован потоками информации из всех остальных $(N - 1)$ входных его портов:

$$\text{Pr} 0_N = \left(\frac{N-2}{N-1}\right)^{N-1}; \quad (14)$$

$$\text{Pr} 1_N = \left(\frac{N-2}{N-1}\right)^{N-2}. \quad (15)$$

Зависимости $\text{Pr} 0_j$ и $\text{Pr} 1_j$ от числа входных процессов K , для которых одновременно востребован один и тот же j -й выходной порт приведены в [5], где установлена их быстрая сходимость к предельным значениям $\frac{1}{e}$. Так, для восьмипортового коммутатора имеем

$\text{Pr} 0_\infty \cong 0,33992$ и $\text{Pr} 1_\infty = 0,39657$, т. е. погрешности определения вероятностей не превосходят 7 % от предельных значений, определяемых зависимостями (12) и (13).

Предельное значение вероятности взаимной блокировки кадров в соответствии с формулой (10) будет

$$\Pr B_{\infty} = 1 - \frac{1}{e} - \frac{1}{e} \cong 0,264241... \quad (16)$$

Для сравнения, при восьмипортовом коммутаторе получаем $\Pr B_{\infty} = 0,26351$.

Для коммутаторов с большим количеством портов (16, 32, 48 и т. д.) расхождение в значениях вероятности взаимной блокировки кадров на выходном интерфейсе при предельной нагрузке этого порта со стороны остальных входных интерфейсов еще меньше отличается от $\Pr B_{\infty}$. Аналогично можно установить и вероятность блокировки впереди стоящим, так как она оказывается равной вероятности блокировки на предыдущем такте работы коммутатора при передачи кадра на любой из выходных его интерфейсов, отличный от рассматриваемого j -го выходного порта. Приведенные результаты для взаимной блокировки кадров аналогичны таковым, полученным Л. Робертсом еще в 1972 г. для метода «дискретная ALOHA» и метода CSMA/CD (Carrier Sense Multiple Access with Collision Detection), используемых в классическом стандарте технологии Ethernet для моноканала [1]. В оптимальной ситуации 36,8 % интервалов передачи оказываются пустыми (что соответствует $\Pr 0_{\infty}$), 36,8 % интервалов успешно передают кадры (что соответствует $\Pr 1_{\infty}$) и в оставшихся 26,4 % интервалов происходят коллизии (что соответствует $\Pr B_{\infty}$). Именно поэтому эффективная пропускная способность метода CSMA/CD составляла лишь треть заявленной пропускной способности в 10 Мбит/с. Такие же результаты в работе [5] распространены на любые системы, в основе функционирования которых используются общие ресурсы и названы «правилом доступности ресурсов».

В отличие от метода CSMA/CD, применявшегося в моноканальном Ethernet и локальных сетях Ethernet, построенных на концентраторах (hub-ax), в коммутируемых транспортных системах компьютерных сетей «столкнувшиеся» кадры не уничтожаются, а буферизуются во входных интерфейсах, задерживаясь до окончания взаимной блокировки. Время задержки зависит от дисциплины обслуживания очереди кадров, поступивших в буферы входных портов, при этом о готовности кадра к передаче в какой-либо выходной интерфейс свидетельствует выставление им «флага готовности» при совпадении передаваемой в последних четырех байтах кадра его контрольной суммы (CRC) со значением этой суммы, подсчитываемым аппаратно по мере поступления кадра во входной аппаратный буфер порта. Наиболее распространенным принципом обслуживания очереди является FIFO (акроним First In, First Out — «первым пришел — первым ушел»), при этом для равноприоритетных кадров применяется алгоритм

равновзвешенного циклического обслуживания (WRR – Weighted Round Robin), в соответствии с которым циклически опрашиваются входные интерфейсы коммутатора на предмет готовности к передаче в буфер рассматриваемого выходного интерфейса пришедших кадров информации и осуществляется их поочередная передача на основе принципа FIFO. Если в данный момент в буфере входного порта отсутствует готовый к передаче кадр, то он пропускается до момента следующего опроса. Реализация алгоритма WRR (или какого-либо иного, учитывающего приоритетность обслуживания) осуществляется процессорами портов коммутатора параллельно.

Пусть в рассматриваемый интервал времени с единичной вероятностью j -й выходной порт востребован K входными интерфейсами. Если в алгоритме WRR конкретный входной интерфейс опрашивается первым, то ожидающий в нем кадр передается без задержки в буфер выходного порта. Если в цикле опроса входной порт оказывается последним, то задержка в его передаче будет равна сумме времени обслуживания кадров остальных $(K - 1)$ портов. Логично допустить равновероятное нахождение в очереди на обслуживание i -го входного интерфейса. В предлагаемой модели коммутатора допускается, что обслуживание заканчивается с истечением межкадрового интервала (IPG), следуемого за выставлением флага готовности. В реальности подобным быстроедействием современные коммутаторы не обладают, однако использование конвейеризации при обработке передаваемых кадров допускает приемлемость данной модели. В любом случае соблюдение порядка в очереди буфера выходного интерфейса (или в нескольких очередях при учете приоритетности обрабатываемых кадров) позволяет найти максимальное значение времени задержки кадра как сумму длительностей их обработки на ненагруженном коммутаторе:

$$t_{ij}^{\max} = \sum_{l=1}^{K-1} t_{lj}, \quad l \neq i,$$

где t_{lj} — случайная величина времени обработки фрейма, передаваемого из l -го входного в j -й выходной порт, определяемая по формулам (4)–(6) в зависимости от пропускной способности интерфейсов и прямо пропорциональная длине L_{lj} , которая также является случайной величиной.

Математическое ожидание времени задержки будет

$$\overline{\Delta t_{ij}} = \frac{\overline{t_{ij}^{\max}}}{2} = \frac{1}{2} \sum_{l=1}^{K-1} \overline{t_{lj}}, \quad l \neq i \text{ нс}, \quad (17)$$

$$\overline{t_{lj}} = \alpha \cdot \overline{L_{fj}},$$

где $\overline{L_{fj}}$ — математическое ожидание размера фрейма при передаче кадра из l -го входного интерфейса в j -й выходной порт ($\alpha = \{800; 100; 10\}$ при $R = \{10, 100, 1000 \text{ Мбит/с}\}$).

Величина $\overline{\Delta t_{ij}}$ должна быть добавлена к времени задержки ненагруженного коммутатора с вероятностью $\text{Pr} B_j$, поэтому среднее время передачи кадра из i -го в j -й порт оказывается

$$\overline{t_{ij}}^* = \overline{t_{ij}} + \frac{1}{2} \cdot \text{Pr} B_j \cdot \sum_{l=1}^{K-1} \overline{t_{lj}}, \quad l \neq i \text{ нс}, \quad (18)$$

Введем понятие времени передачи в j -й порт всех K кадров входных интерфейсов, востребовавших данный выходной интерфейс

$$T_j = \sum_{i=1}^K \overline{t_{ij}}. \quad (19)$$

Фактически это время определяется длительностью цикла обслуживания коммутатором всех конфликтующих за j -й выходной порт входных интерфейсов в рассматриваемый интервал времени. Величина T_j является случайной, поэтому математическое ожидание времени передачи кадра из i -го интерфейса в j -й выходной порт можно найти по формуле

$$\overline{t_{ij}}^* = (1 - \text{Pr} B_j) \overline{t_{ij}} + \text{Pr} B_j \cdot \overline{T_j} = \overline{t_{ij}} + \frac{1}{2} \cdot \text{Pr} B_j \left(\sum_{i=1}^K \overline{t_{ij}} - \overline{t_{ij}} \right). \quad (20)$$

При ближайшем рассмотрении можно видеть совпадение полученного выражения с формулой (18), что лишний раз свидетельствует о правдоподобности полученных результатов. Однако с вычислительной точки зрения зависимость (20) оказывается предпочтительнее формулы (18). При $K = 1$, $\text{Pr} B_j = 0$ имеем ожидаемое тождество $\overline{t_{ij}}^* = \overline{t_{ij}}$. Выходные интерфейсы работают в параллельном режиме, поэтому сравнением можно установить максимальное значение времени передачи кадра из i -го входного порта

$$\overline{t_i}^* = \max_j (\overline{t_{ij}}^*). \quad (21)$$

Это время в итоге является определяющим для установления требуемых размеров буферов входных интерфейсов коммутатора.

Выше было получено выражение вероятности взаимной блокировки кадров для случая одновременной востребованности j -го выходного порта K входными информационными потоками с единичной вероятностью и установлено ее предельное значение. В действительности значения $\text{Pr}0_j$, $\text{Pr}1_j$ и $\text{Pr}B_j$ могут существенно меняться для небольшого количества конкурирующих за j -й выходной порт входных потоков информации, каждый из которых характеризуется своей пропускной способностью R_i , интенсивностью передачи кадров λ_{ij} и размерами передаваемых фреймов L_{fij} . В [5] предложена в качестве вероятности востребованности j -го выходного порта кадром i -го входного интерфейса использовать статистическое определение вероятности [4], в соответствии с которым за вероятность востребованности принимается доля пропускной способности, занимаемая данным кадром. Однако для пропускных способностей интерфейсов в 100 и 1000 Мбит/с расчетная формула должна быть изменена по причине особенностей используемых на физическом уровне эталонной модели ISO/OSI кадров, т. е.

$$\pi_{ij} = \begin{cases} \frac{8 \cdot \lambda_{ij} \cdot L_{fij}}{R_i} & \text{при } R_i = 10 \text{ Мбит/с,} \\ \frac{10 \cdot \lambda_{ij} \cdot L_{fij}}{R_i} & \text{при } R_i = 100 \text{ Мбит/с и } R_i = 1000 \text{ Мбит/с.} \end{cases} \quad (22)$$

Кроме того, необходимо учитывать условие нормировки [4], поэтому:

$$p_{ij} = \begin{cases} \pi_{ij} & \text{при } \sum_i \pi_{ij} \leq 1, \\ \frac{\pi_{ij}}{\sum_i \pi_{ij}} & \text{при } \sum_i \pi_{ij} > 1. \end{cases} \quad (23)$$

На уровне доступа к сети формулы (22) и (23) дают заниженные результаты, так как запросы пользователей отличаются крайне низкой интенсивностью (не более одного запроса в секунду), что объясняется человеческим фактором и малой длиной передаваемых фреймов из-за особенностей клиент-серверных технологии, применяемой повсеместно, в соответствие с которой нагрузка на транспортную систему компьютерной сети со стороны пользовательских хостов представляет собой в основном короткие запросы нужной информации, короткие кадры подтверждения (Acknowledge) для TCP-соединений и служебные кадры сети. Редкие пересылки графических, мультимедийных и прочих файлов можно считать нетипичными для штатного функционирования транспортных систем [5]. Более вероятным на

уровне доступа следует считать ситуацию, характеризующуюся одновременной востребованностью в каком-либо сетевом ресурсе для K пользователей с единичной вероятностью. В этом случае, как было показано выше, $p_{ij} = \frac{1}{K}$.

Очевидно, что j -й выходной порт не может функционировать со скоростью, превышающей его пропускную способность R_j , следовательно, нормальная его работоспособность гарантируется выполнением неравенства

$$R_j \geq \sum_i p_{ij} \cdot R_i, i \neq j. \quad (24)$$

Ясно, что случайный характер как интенсивности передачи кадров, так и размеров фреймов, поступающих на входные интерфейсы, не дает полной гарантии выполнения неравенства (24) (речь идет лишь о математических ожиданиях характеристик входных информационных потоков). Однако при выборе конкретной модели коммутатора необходимо оценить возможность нарушения ограничения (24) с учетом допустимой пульсации трафика в сегментах транспортной системы.

Очевидно, что с развитием информационно-коммуникационных технологий и их внедрением во все сферы деятельности, меняется и характер распределения потоков в компьютерных сетях, порождаемых как информационным обменом между клиентами и внутренними и внешними серверами, так и межсерверным обменом многих сетевых служб. В коммутируемых транспортных системах принято выделять типовые участки, к которым, прежде всего, относятся ядро иерархически построенной системы (магистраль), пулы серверов (централизованные или сосредоточенные на более низких иерархических уровнях), уровень распределения транспортной системы и ее уровень доступа. Для каждого типового участка существует свое распределение трафика для входных и выходных интерфейсов коммутатора. Коэффициент перегрузки показывает, во сколько раз интенсивность входного потока превышает предельную интенсивность выходного. В нашем случае коэффициент перегрузки для j -го выходного порта может быть вычислен по формуле

$$\rho_j = \frac{\sum_i p_{ij} R_i}{R_j}, i \neq j. \quad (25)$$

Превышение коэффициента перегрузки единичного значения указывает точки транспортной системы, являющиеся потенциально узким местом для переносимого трафика. Различные участки иерар-

хически построенных транспортных систем на коммутаторах предъявляют разные требования к значению коэффициента перегрузки. Магистральные участки сети и серверные пулы наиболее чувствительны к перегрузкам, так как различные варианты распределения трафика здесь практически равновероятны. Компания ExtremeNetworks рекомендует придерживаться значения $\rho_j \leq 1$ даже для кратковременных пульсаций трафика [7]. На распределительном уровне для информационных всплесков допустимо $\rho_j \leq 2$, а для уровня доступа — $\rho_j \leq 3$. Это объясняется тем, что среднестатистический пользовательский компьютер генерирует трафик, не превышающий 30 % от пропускной способности его сетевого адаптера, что в свою очередь позволяет установить потребные пропускные способности интерфейсов уровня распределения. Ясно, что рекомендации компании ExtremeNetworks носят статистический характер и не гарантируют отсутствия заторов в сегментах транспортной системы. В [4] предпринята попытка математически обосновать допустимые кратковременные превышения коэффициента перегрузки единичного значения, однако в выводе допущена неточность. Логика математического построения базируется на приведенных выше значениях вероятности невостробованности j -го выходного порта ни одним из входных интерфейсов $\text{Pr}0_j$ (8), одним и только одним входным интерфейсом $\text{Pr}1_j$ (9) и вероятностью блокировки $\text{Pr}B_j$ (10). При этом утверждается, что эффективная пропускная способность j -го интерфейса должна быть принята $R_{\text{эф}} = R(\text{Pr}0 + \text{Pr}1)$. Скорее всего, в тексте допущена опечатка, так как j -й выходной порт осуществляет передачу кадров в сегмент транспортной системы и при блокировке, но с иного входного интерфейса коммутатора. Поэтому значение вероятности $\text{Pr}0$ должно быть заменено значением $\text{Pr}1$. Лишь во время невостробованности j -го выходного порта ни одним из входных интерфейсов не осуществляется передача информации в буфер j -го выходного порта, следовательно, доля времени отсутствия эвакуации кадров из выходной очереди и доля времени их эвакуации со скоростью, равной пропускной способности интерфейса пропорциональны соответственно значениям $\text{Pr}0_j$ и $(1 - \text{Pr}0_j) = \text{Pr}1_j + \text{Pr}B_j$. Следовательно, эффективная пропускная способность может быть найдена как математическое ожидание скорости эвакуации (схема Бернулли [6])

$$R_{\text{эф}j} = \overline{R}_j = 0 \cdot \text{Pr}0_j + R_j(1 - \text{Pr}0_j) = R_j(1 - \text{Pr}0_j). \quad (26)$$

С учетом изложенных соображений неравенство (24) приобретает вид:

$$\overline{R}_j \geq \sum_i p_{ij} R_i, i \neq j. \quad (27)$$

Если принять во внимание, что сумма правой части полученного ограничения, по сути, является математическим ожиданием скорости суммарного входного потока информации на входные интерфейсы, востребовавшие j -й выходной порт, то выражение (27) вполне логично — скорость эвакуации кадров из выходного буфера не должна быть ниже скорости их поступления из буферов входных портов, поэтому для штатного режима следует использовать топологию подключения сегментов транспортной системы компьютерной сети к надлежащим портам коммутатора, строго соблюдая выполнение ограничения (27).

На уровне доступа типовым считается подключение конечных пользователей, при котором каждому клиенту выделяется отдельный порт коммутатора (микросегментация), один из его интерфейсов предназначается для взаимодействия с уровнем распределения и еще один (возможно) — для подключения сервера рабочей группы. Если допустить, что потребность во «внешней» информации и в информации «внутренней» соотносятся как 50 % на 50 % (при правильной организации информационного обеспечения рабочей группы соотношение выглядит как 20 % к 80 % [2]), то с учетом приведенного фирмой ExtremeNetworks факта о непревышении интенсивности трафика пользователей 30 % уровня пропускных способностей сетевых адаптеров клиентских хостов имеем

$$(1 - \text{Pr } 0_j) R_j \geq \sum_i \frac{1}{K} \cdot \frac{1}{2} \cdot 0,3 R_i, i \neq j, i = 1, 2, \dots, K.$$

Или при одинаковых пропускных способностях входных и выходных интерфейсов $R_i = R_j = R$ получим для $K \rightarrow \infty$ ($\text{Pr } 0_j = 0,368$) $0,632 > 0,15$. Иными словами, в отсутствие пульсации трафика вполне допустимо и клиентские хосты, и сервер рабочей группы, и сегмент уровня распределения подключать к интерфейсам коммутатора с одинаковой пропускной способностью (например, 100 Мбит/с). Это положение подтверждается практикой эксплуатации коммутаторов в транспортной системе корпоративной сети МГТУ им. Н.Э. Баумана [5].

В качестве еще одного примера можно рассмотреть функционирование коммутатора уровня распределения, типовая схема подключения которого отличается выделением двух высокоскоростных интерфейсов: одного для связи с магистралью транспортной системы компьютерной сети, а другого — либо для подключения сервера уровня нескольких подразделений (например, уровня факультета в университете), либо для каскадного подключения какого-либо ком-

мутатора такого же уровня распределения. Более низкоскоростные порты выделяются для связи с коммутаторами уровня доступа. В этой ситуации даже для коэффициентов перегрузки входных интерфейсов, равных единице, неравенство (24) выполняется минимум для десяти входных интерфейсов, так как $R_j = 10 \cdot R_i$ (например, $R_j = 1000$ Мбит/с, $R_i = 100$ Мбит/с или $R_j = 100$ Мбит/с, $R_i = 10$ Мбит/с). С учетом очевидного соотношения $p_{ij} < 1$ вполне допустимо на уровне распределения число низкоскоростных интерфейсов увеличить до 20. Что касается выполнения неравенства (27), то следует учитывать тот факт, что входные интерфейсы уровня распределения являются выходными для уровня доступа, и их пропускные способности в свою очередь должны быть выбраны с учетом возможной блокировки кадров информационных потоков пользователей (клиентов).

Выполнение ограничений (24) и (27) предусматривается для штатных режимов работы коммутаторов и может отсутствовать при пульсациях трафика, возникающих, к сожалению, с ненулевой вероятностью в любых компьютерных сетях.

Потери кадров в коммутаторах второго уровня возможны в том случае, если интенсивность входящих потоков информации превышает соответствующую интенсивность исходящих потоков и объем превышения за определенный промежуток времени превысит объем буферной памяти, предназначенной, в первую очередь, для сглаживания пульсации трафика. Применительно для любого входящего интерфейса с пропускной способностью R_i предельное число байт, поступивших в его входной буфер за время перегрузки (congestions), может быть найдено по формуле

$$B_i = \begin{cases} \frac{R_i \cdot \tau_c}{8} & \text{при } R_i = 10 \text{ Мбит/с,} \\ \frac{R_i \cdot \tau_c}{10} & \text{при } R_i = 100 \text{ Мбит/с и } R_i = 1000 \text{ Мбит/с,} \end{cases} \quad (28)$$

где τ_c — время перегрузки.

Очевидно, что при задержке кадров t_i^* (см. (21)) именно это значение является предельным, т. е. $\tau_c = t_i^*$.

Если B_i оказывается больше размера буфера входного порта, то в соответствии с алгоритмом «обрубания хвостов» очередной кадр, поступающий на входной интерфейс, отбрасывается. В случае использования при перегрузках алгоритмов RED (Random Early Detection), ECN (Explicit Congestion Notification) [2] или более гибких алгоритмов с установлением пороговых значений (threshold) [1] по-прежнему возможны переполнения буферной памяти входных интерфейсов.

Знание потребных объемов буферов входных интерфейсов влияет на выбор той или иной модели коммутатора. К сожалению, на стадии проектирования транспортных систем компьютерных сетей эти характеристики могут быть определены весьма приблизительно. При этом возможные ошибки приведут в дальнейшем в лучшем случае к необходимости перетрассировки информационных потоков в каждом транзитном узле транспортной системы, а в худшем — к замене отдельных транзитных узлов.

Для выходных интерфейсов коммутатора определение потребных объемов буферной памяти не столь очевидно. При одновременной востребованности j -го выходного порта K входными интерфейсами скорость поступления информации в буфер этого порта равна $\sum_i p_{ij} R_i$, а скорость его опорожнения (эвакуации из порта) равна его пропускной способности, поэтому

$$B_j = \left(\sum_i p_{ij} \cdot \frac{R_i}{\alpha_i} - \frac{R_j}{\alpha_j} \right) \tau_c, \quad (29)$$

где $\alpha_i = 8$ при $R_i = 10$ Мбит/с и 10 при $R_i = 100$ Мбит/с и $R_i = 1000$ Мбит/с.

Аналогично определяется значение α_j .

Время перегрузки можно оценить исходя из следующих логических рассуждений. При задержке кадра из i -го входного интерфейса во j -й выходной порт осуществляется передача информации в этот порт из других входных интерфейсов с вероятностью p_{lj} , $l \neq i$. Весьма правдоподобно в качестве времени перегрузки принять значение

$$\tau_c = \sum_i t_{ij}^*. \quad (30)$$

Однако следует иметь ввиду, что перегрузка может прекратиться вместе с передачей не всех кадров из входных интерфейсов, востребовавших j -й выходной порт в рассматриваемый интервал времени, а лишь какого-либо одного из них. В этом случае в качестве оценки времени перегрузки следует использовать максимальное из t_{ij}^* значение, т. е.:

$$\tau_c = t_j^* = \max_i(t_{ij}^*). \quad (31)$$

Длительность задержки кадров влияет на общую производительность компьютерной сети, и ее минимизация путем перетрассировки информационных потоков в транзитных узлах, а также с помощью механизма виртуальных сетей на втором уровне эталонной модели не

только повышает эффективность всей транспортной системы, но и обеспечивает более высокий уровень надежности и безопасности функционирования всей информационно-коммуникационной системы предприятия, снижая уровень потерь информационных кадров в буферах коммутаторов.

Полученные теоретические зависимости дают возможность оценить варианты построения топологии транспортной системы, однако их использование возможно только при знании вероятностей востребованности тех или иных информационных ресурсов предприятия со стороны конечных пользователей различных категорий и трафика информационного обмена между подразделениями. Кроме того, рекомендации по применению полученных моделей могут быть даны только после экспериментальной проверки приведенных результатов.

ЛИТЕРАТУРА

- [1] Таненбаум Э., Уэзеролл Д. *Компьютерные сети*. 5-е изд. Санкт-Петербург: Питер, 2012, 960 с.
- [2] Олифер В.Г., Олифер Н.А. *Компьютерные сети*. 4-е изд. Санкт-Петербург: Питер, 2011, 944 с.
- [3] Олифер В.Г., Олифер Н.А. *Основы компьютерных сетей*. Санкт-Петербург: Питер, 2009, 352 с.
- [4] Бойченко М.К., Иванов И.П. Вероятностная модель коммутаторов локальных вычислительных сетей. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*, 2009, № 2, с. 84–92.
- [5] Иванов И.П. *Математические модели, методы анализа и управления в корпоративных сетях*. Автореф. дис. ... д-ра техн. наук. Москва: МГТУ им. Н.Э. Баумана, 2010, 34 с.
- [6] Вентцель Е.С. *Теория вероятностей*: учебник для вузов. 11-е изд. Москва: КноРус, 2010, 664 с.
- [7] Олифер В.Г., Олифер Н.А. *Новые технологии и оборудование IP-сетей*. Санкт-Петербург: БХВ-Петербург, 2000, 512 с.

Статья поступила в редакцию 03.10. 2014

Ссылку на эту статью просим оформлять следующим образом:

Бойченко М.К., Иванов И.П. Теоретическая модель Ethernet-коммутатора. *Инженерный журнал: наука и инновации*, 2014, вып. 10.

URL: <http://engjournal.ru/catalog/it/network/1311.html>

Иванов Игорь Потапович родился в 1955 г., окончил в 1979 г. МВТУ им. Н.Э. Баумана. С 1996 г. по настоящее время — проректор по информатизации и модернизации МГТУ им. Н.Э. Баумана, с 2000 г. заведующий кафедрой «Теоретическая информатика и компьютерные технологии», д-р техн. наук, доцент. Автор 40 научных работ в области информационно-коммуникационных технологий. e-mail: ivanov@bmstu.ru

Бойченко Максим Константинович родился в 1978 г., окончил МГТУ им. Н. Э. Баумана в 2001 г. Начальник ИЦ УИ-ВИЦ. Специалист в области информационно-коммуникационных технологий. e-mail: noc@bmstu.ru

Theoretical model of Ethernet-switch

© М.К. Boychenko, I.P. Ivanov

Bauman Moscow State Technical University, Moscow, 105005, Russia

The article describes a proposed theoretical model of the switch. The built model is based on the structural features of the frame of Ethernet technology. We considered the operation of a hypothetical switch of any hierarchical level of a topological scheme of constructing transport system of network. The calculated dependences are aimed at determination of probability and mathematical expectation of the time interlocking frames in switches that implement the second level of the reference model ISO/OSI. A connection is established between the magnitude of the switch delay and requirements of buffer memory required of its input and output ports for all topologies of intersection information flows in the transport systems.

Keywords: computer network, transport system, transit node, switch, blocking, delay, buffer memory.

REFERENCES

- [1] Tanenbaum A., Wetherall D. Computer Networks. 5th ed. Prentice Hall, Inc., Upper Saddle River, New Jersey, 2011.
- [2] Olifer V.G., Olifer N.A. Kompyuternye seti [Computer Networks]. 4th ed. St.-Petersburg, Piter Publ., 2011, 944 p. [in Russian].
- [3] Olifer V.G., Olifer N.A. Osnovy kompyuternykh setei [Basics of Computer Networks]. St.-Petersburg, Piter Publ., 2009, 352 p. [in Russian].
- [4] Boychenko M.K., Ivanov I.P. Vestnik MGTU im. N.E. Baumana. Priborostroenie – Herald of the Bauman Moscow State Technical University. Instrument Engineering, 2009, no. 2, pp. 84–92.
- [5] Ivanov I.P. Matematicheskie modeli, metody analiza i upravleniya v korporativnykh setiakh. Avtoreferat dis. dokt. tehn. nauk. Moscow, Bauman Moscow State Technical University, 2010, 34 p.
- [6] Ventcel' E.S. Teoriya veroyatnostei. 11th ed. Moscow, KnoRus Publ., 2010, 664 p.
- [7] Olifer V.G., Olifer N.A. *Novye tehnologii i oborudovanie IP-setei*. St.-Petersburg, BHV Publ., 2000, 512 p. [in Russian].

Ivanov I.P., Dr. Sci. (Eng.), vice-rector in informatization and modernization of the Bauman Moscow State Technical University, head of the Theoretical Informatics and Computer Technologies Department at the Bauman Moscow State Technical University. Author of more than 40 publications in the field of data-communication technologies. e-mail: ivanov@bmstu.ru

Boychenko M.K., leading programmer of the IT laboratory of the Administration on Informatization – Computing Center of the Bauman Moscow State Technical University. Specializes in the field of data-communication technologies. e-mail: noc@bmstu.ru