

М. К. Б о й ч е н к о, И. П. И в а н о в

**ВЕРОЯТНОСТНАЯ МОДЕЛЬ КОММУТАТОРОВ
ЛОКАЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СЕТЕЙ**

Рассмотрена методика расчета времени пребывания кадров в коммутаторах локальных вычислительных сетей с учетом доступности интерфейсов, зависящей от их взаимной блокировки для полnodуплексных режимов передачи информации.

E-mail: ivanov@bmstu.ru

Ключевые слова: интерфейс, транзитный узел, сеть, блокировка, ресурс, хост, транспортная подсистема, маршрутизатор.

В современных компьютерных сетях в качестве транзитных узлов используются коммутаторы и маршрутизаторы, при этом наблюдается тенденция сближения коммутации и маршрутизации, что подтверждается наличием коммутаторов 3-го и 4-го уровней эталонной модели ISO/OSI и маршрутизаторов, реализующих стандарт MPLS (MultiProtocol Label Switching) [1, 2]. Как правило [2, 3], локальные компьютерные сети (ЛВС) строятся на коммутаторах различных уровней, функционирующих в полnodуплексном режиме обмена информацией на всех уровнях сети, начиная от уровня доступа (микросегментирование) и заканчивая магистральным уровнем (ядро транспортной подсистемы). Особенностью реализации транспортных подсистем на коммутаторах является неизбежность буферизации данных сетью при востребованности общих ресурсов различными информационными процессами, протекающими в хостах компьютерной сети [3, 4]. Доступность ресурсов транспортных подсистем корпоративных сетей существенно влияет на их производительность. Построить полностью неблокирующую ЛВС нереально, так как взаимная блокировка информационных потоков неизбежна уже в самих транзитных узлах для определенных схем коммутации входных и выходных интерфейсов. В работе [4] рассмотрены вопросы влияния возможной блокировки интерфейсов коммутаторов на время передачи кадра с порта на порт, однако исследован частный случай распределения информационных потоков, при котором конкретный выходной интерфейс равновероятно должен быть скомутирован с одним из K входных интерфейсов при мультиплексировании потоков, или конкретный входной порт равновероятно коммутируется с одним из K выходных при демультиплексировании. Сам коммутатор рассматривается в этом случае как система массового обслуживания (СМО) и не учитывается полnodуплексность режимов передачи информации. Для более детального анализа функционирования коммутатора следует отказаться от его модели уровня системы и перейти к его модели уровня элементов, рассмотрев более

детально процесс передачи информации с порта на порт. Функционально любой коммутатор выполняет следующие операции.

1. Прием кадра из сегмента сети в аппаратный буфер порта, осуществляемый побитно. В технологии Ethernet, повсеместно применяемой в ЛВС, предварительно от источника к приемнику передается семь байтов преамбулы и один стартовый байт для синхронизации источника и приемника [3]. Время приема кадра зависит от его длины, измеряемой в байтах, и от пропускной способности входного интерфейса коммутатора, а точнее от пропускной способности сегмента между данным входным портом и взаимодействующим с ним интерфейсом (либо портом другого коммутатора, либо сетевым адаптером оконечного узла сети):

$$T_i^{in} = \frac{8L_i}{R_i}, \quad (1)$$

где T_i^{in} — время приема кадра (по биту) в аппаратный буфер i -го порта коммутатора, с; L_i — размер кадра (в байтах), поступающего в i -й порт коммутатора; R_i — пропускная способность i -го интерфейса коммутатора (бит/с).

2. После приема первых шести байтов, следующих за преамбулой и стартовым байтом, оказывается известным MAC-адрес узла назначения ЛВС, т.е. того узла сети, которому предназначается данный кадр. По таблице коммутации, хранящейся в памяти входного порта, устанавливается номер порта коммутатора, в который следует передать поступающий кадр. Если требуемый порт свободен в этот момент времени, то принимаемый кадр уже может быть передан в выходной интерфейс и начата его дальнейшая передача по сети. Необходимым условием для такой передачи является равенство пропускных способностей входного и выходного портов коммутатора ($R_i = R_j$). Если обозначить через T_{ij}^{com} время коммутации i -го входного и j -го выходного портов, включающее время определения по таблице коммутации потребного выходного интерфейса и собственно коммутации, то суммарное время задержки на коммутаторе будет равно

$$T_{ij}^{min} = T_{ij}^{com} + \frac{8 \cdot (8 + 6)}{R_i}. \quad (2)$$

Описанная процедура является так называемой коммутацией на лету, а T_{ij}^{min} — минимально возможным временем передачи кадра из сегмента в сегмент сети. Отметим, что все это время на входной порт продолжают поступать байты входного кадра (по биту). Если $R_i \neq R_j$, то коммутация “на лету” невозможна. Кроме того, ко времени определения потребного выходного интерфейса он может быть скомутирован с каким-либо иным портом коммутатора и требуется ожидание

его освобождения или в буферной памяти входного интерфейса могут присутствовать ранее поступившие на него и еще не переданные кадры. Не менее важной причиной непопулярности коммутации “на лету” является необходимость проверки контрольной суммы кадра, передаваемой дополнительно вместе с ним в концевики размером 4 байта (CRC — Cyclic Redundancy Check) [3]. В случае несовпадения передаваемого вместе с кадром концевика с аппаратно-вычисляемым по мере приема кадра в аппаратный буфер входного порта, принятый кадр считается искаженным и подлежит уничтожению в целях повышения общей производительности сети, снижающейся при необходимости повторной передачи искаженного кадра. Проверка совпадений CRC может быть выполнена лишь после приема в аппаратный буфер порта всего кадра, включая концевик. Отмеченные особенности коммутации “на лету” существенно ограничивают ее применимость [1].

3. По мере побитного приема кадра из сегмента сети в аппаратный буфер входного интерфейса могут выполняться различные операции, а именно:

- порционное переписывание кадра из аппаратного буфера в очередь, организуемую в оперативной памяти входного порта, с проверкой возможности размещения кадра в этой очереди и принятием решения об уничтожении этого кадра в соответствии с алгоритмами борьбы с перегрузками (TD – Tail Drop или WRED — Weighted Random Early Detect [2, 5]);

- выполнение запроса на коммутацию с требуемым выходным портом в соответствии с используемым алгоритмом обслуживания интерфейсов (например, WRR — Weighted Round Robin [2, 5]);

- продвижение кадра к началу (голове) очереди, организуемой в оперативной памяти входного порта, по мере эвакуации из нее информации в выходные интерфейсы и т.д.

После проверки целостности кадра сравнением концевика с аппаратно-вычисленным CRC устанавливается флаг готовности кадра к передаче в выходной порт и завершается его запись в оперативную память входного интерфейса. Обозначим через T_j^w время ожидания кадра с момента выставления флага готовности до достижения им головы очереди в оперативной памяти входного порта.

4. Осуществляется коммутация i -го входного порта с j -м выходным за время T_{ij}^{com} при условии отсутствия коммутации j -го порта с каким-либо иным интерфейсом коммутатора. В случае невозможности коммутации имеем режим блокировки, которая может длиться в течение времени T_{ij}^B .

5. После прекращения возможной блокировки и выполнения коммутации кадр порционно из головы очереди входного порта передается в хвост очереди выходного порта с последующей его эвакуацией.

Время передачи зависит от конструктивного выполнения коммутатора [1]. В случае коммутаторов с общей разделяемой памятью это время не зависит от размера передаваемого кадра. Для коммутаторов с общей шиной из-за ограничения разрядности последней время передачи оказывается пропорциональным размеру кадра. Аналогичная ситуация имеет место для коммутационных матриц. В работе [6] излагается методика определения времени передачи кадров коммутаторами различных производителей. Однако при этом рассмотрение коммутаторов как СМО уровня системы дает возможность рассчитать не собственно время передачи кадра из головы очереди входного порта в хвост очереди выходного интерфейса, а суммарное время пребывания кадра в ненагруженном коммутаторе. Если из полученных значений вычесть T_i^{in} и T_{ij}^{com} , то могут быть установлены собственно значения времени передачи кадров с i -го входного порта на j -й интерфейс как функции размера передаваемого кадра $T_{ij}^S = T_{ij}^S(L_i)$.

Общее время пребывания передаваемого кадра в коммутаторе можно найти суммированием соответствующих составляющих:

$$T_{ij} = T_i^{in} + T_{ij}^{com} + T_i^w + T_{ij}^S + T_j^w + T_{ij}^B, \quad (3)$$

где T_i^w — время пребывания передаваемого кадра в очереди i -го входного порта; T_j^w — время пребывания переданного кадра в очереди j -го выходного интерфейса.

Отметим, что время T_i^w явно зависит от T_{ij}^B — от числа очередей во входном интерфейсе и установленного на этом интерфейсе алгоритма обслуживания очереди. Время T_j^w также определяется числом очередей в выходном интерфейсе и приоритетностью их обслуживания при передаче кадров в следующий сегмент транспортной подсистемы ЛВС.

Рассмотрим функционирование коммутатора с N портами, реализующего полнодуплексные режимы передачи информации на всех своих интерфейсах. Для определенности будем считать, что в стационарном режиме на i -й входной порт поступает поток информации с интенсивностью λ_i (кадр/с), складывающийся из кадров, предназначенных для разных выходных интерфейсов:

$$\lambda_i = \sum_j \lambda_{ij}, \quad i \in 1, N, \quad j \in 1, N, \quad (4)$$

где λ_{ij} — интенсивность потока кадров из i -го входного в j -й выходной интерфейс коммутатора.

Каждый поток моделируем по Бернулли, т.е. с вероятностью p_{ij} кадр имеет размер $L_{\min} = 84$ байт и с вероятностью $q_{ij} = 1 - p_{ij}$ — размер $L_{\max} = 1538$ байт для технологии Ethernet [2, 3]. Средняя длина кадров, передаваемых с i -го входного на j -й выходной порт, может

быть найдена как математическое ожидание размера передаваемых кадров $L_{ij} = p_{ij}L_{\min} + (1 - p_{ij})L_{\max}$.

Введем понятие востребованности портов заданием вероятностей π_{ij} . За некоторый промежуток времени τ на вход i -го входного порта для передачи на j -й выходной интерфейс поступает объем информации, бит:

$$S_{ij} = 8\lambda_{ij}L_{ij}\tau. \quad (5)$$

За этот же промежуток времени в i -й порт может поступить (в битах)

$$B_i = R_i\tau. \quad (6)$$

Тогда вероятность востребованности j -го порта i -м может быть найдена как отношение S_{ij} к B_i , т.е.

$$\pi_{ij} = \frac{8\lambda_{ij}L_{ij}}{R_i}. \quad (7)$$

Таким образом, взаимная востребованность портов может быть описана квадратной матрицей $[\pi_{ij}]$ размера $N \times N$. Диагональные элементы этой матрицы для полнодуплексного режима передачи информации тождественно равны нулю, т.е. $\pi_{ii} \equiv 0, i = 1, N$. Также равны нулю элементы π_{ij} , для которых $\lambda_{ij} = 0$, т.е. отсутствуют потоки информации из i -го в j -й порт.

Вероятность невостребованности j -го выходного порта ни одним из входных интерфейсов составляет

$$\text{Pr}0_j^{\text{in}} = \prod_i (1 - \pi_{ij}), \quad i = 1, N; j = 1, N \dots \quad (8)$$

Вероятность того, что этому j -му выходному порту за рассматриваемый интервал времени не требуется ни один из входных портов для полнодуплексного режима, равна

$$\text{Pr}0_j^{\text{out}} = \prod_i (1 - \pi_{ji}), \quad i = 1, N; j = 1, N \dots \quad (9)$$

Вероятность востребованности j -го выходного интерфейса только одним из входных портов

$$\text{Pr}1_j^{\text{in}} = \sum_i \pi_{ij} \cdot \prod_i (1 - \pi_{ij}), \quad i = 1, N; j = 1, N \dots \quad (10)$$

Вместе с тем вероятность связи j -го интерфейса для обратного потока информации в полнодуплексном режиме хотя бы с одним из входных интерфейсов составляет

$$\text{Pr}1_j^{\text{out}} = 1 - \text{Pr}0_j^{\text{out}}, \quad j = 1, N \dots \quad (11)$$

С вероятностью $(1 - \text{Pr}0_j^{\text{in}} - \text{Pr}1_j^{\text{in}})$ j -й выходной интерфейс требуется одновременно более чем одному из входных портов, а с ве-

роятностью $\Pr 1_j^{out}$ этот порт уже связан с каким-либо из входных интерфейсов для передачи обратного потока информации при полнодуплексном режиме. Следовательно, вероятность блокировки на j -м выходном интерфейсе

$$\Pr B_j = 1 - \Pr 0_j^{in} - \Pr 1_j^{in} + \Pr 1_j^{out}, \quad j = 1, N \dots \quad (12)$$

Отличие значения $\Pr B_j$ от нуля свидетельствует о дополнительной задержке входных информационных кадров в коммутаторе. Если недоступность j -го выходного порта вызвана уже существующей его коммутацией с каким-либо из входных интерфейсов, то математическое ожидание длительности его освобождения может быть найдено как

$$T_{ji}^{out} = \sum_i \pi_{ji} T_{ji}^S, \quad i = 1, N; \quad j = 1, N \dots \quad (13)$$

Если дополнительная задержка вызвана конкуренцией входных портов, то ее значение зависит от дисциплины обслуживания интерфейсов коммутатора. Так, например, для равновзвешенного циклического обслуживания (WRR) время равновероятно меняется от нуля (если первым обслуживается данный входной порт) до суммы времен передачи кадров всех остальных потоков (если данный входной порт оказывается последним в очереди на обслуживание), т.е.

$$T_{ij}^{in} = \frac{1}{2} \sum_i \pi_{ij} T_{ij}^S, \quad i = 1, N; \quad j = 1, N \dots \quad (14)$$

Длительность блокировки можно найти как математическое ожидание рассчитанных задержек

$$\begin{aligned} T_{ij}^B &= \Pr 1_j^{out} T_{ji}^{out} + (1 - \Pr 1_j^{out}) T_{ij}^{in} = \\ &= \Pr 1_j^{out} T_{ji}^{out} + \Pr 0_j^{out} T_{ij}^{in}, \quad i = 1, N; \quad j = 1, N \dots \quad (15) \end{aligned}$$

Аналогичные формулы расчета математического ожидания времени блокировки могут быть выведены и для иных дисциплин обслуживания интерфейсов коммутатора.

Для экспериментальной проверки правильности полученных зависимостей необходимо создание сети, включающей в себя не менее трех хостов, подключенных к исследуемому коммутатору, и проведение большого объема экспериментов в реальной ЛВС, что планируется в ходе дальнейших исследований, проводимых в корпоративной сети МГТУ им. Н.Э. Баумана.

СПИСОК ЛИТЕРАТУРЫ

- О л и ф е р В.Г., О л и ф е р Н. А. Новые технологии и оборудование IP-сетей. – СПб.: БХВ, 2001. – 512 с.

2. Т а н е н б а у м Э. Компьютерные сети. – СПб.: Питер, 2009. – 992 с.
3. О л и ф е р В. Г., О л и ф е р Н. А. Компьютерные сети. – СПб.: Питер, 2011. – 944 с.
4. Б о й ч е н к о М. К., И в а н о в И. П., К о н д р а т ь е в А. Ю. Доступность ресурсов транспортных подсистем корпоративных сетей // Вестник МГТУ им. Н.Э.Баумана. Сер. Приборостроение. – 2010. – № 3. – С. 103–118.
5. Ф и л и м о н о в А. Ю. Построение мультисервисных сетей Ethernet. – СПб.: БХВ, 2007. – 592 с.
6. И в а н о в И. П. Математические модели коммутаторов локальных вычислительных сетей // Вестник МГТУ им. Н.Э.Баумана. Сер. Приборостроение. – 2009. – № 2. – С. 84–92.

Статья поступила в редакцию 25.10.2011

Максим Константинович Бойченко родился в 1978 г., окончил МГТУ им. Н.Э.Баумана в 2001 г. Ведущий программист лаборатории ИТ Управления информатизации МГТУ им. Н.Э.Баумана. Специализируется в области информационных и телекоммуникационных технологий.

M.K. Boychenko (b. 1978) graduated from the Bauman Moscow State Technical University in 2001. Leading programmer of “Information Technologies” laboratory of Administration on Informatization of the Bauman Moscow State Technical University. Specializes in the field of information and telecommunication technologies.

Игорь Потапович Иванов родился в 1955 г., окончил в 1979 г. МВТУ им. Н.Э. Баумана. Проректор по информатизации, заведующий кафедрой “Теоретическая информатика и компьютерные технологии” МГТУ им. Н.Э. Баумана. Д-р техн. наук, доцент. Автор более 30 научных работ в области информационно-коммуникационных технологий.

I.P. Ivanov (b. 1955) graduated from the Bauman Moscow Higher Technical School in 1979. Ph. D. (Eng.), vice-rector on informatization, head of “Theoretical Information Technology and Computer Technologies” department of the Bauman Moscow State Technical University. Author of more than 30 publications in the field of data and communication technologies.