

## MODELLING THROUGHPUT CROSSBAR SWITCH NODE WITH NONUNIFORM LOAD TRAFFIC

*Tasho Tashev*

Institute of Information and Communication Technologies, Bulgarian Academy of  
Sciences, Sofia, Bulgaria  
tashotr@bas.bg

**Abstract:** *The task which includes computing of non-conflict schedule for crossbar switch node of packet commutation is NP-hard. After checking their efficiency with throughput modelling by uniform load traffic, the check for nonuniform traffic is required. In this paper a family of pattern-matrices is proposed for the simulation of nonuniform traffic based on the hotspot traffic model. The applicability check is executed through computer simulated commutation using well known PIM-algorithm, spesified by Generalized nets.*

**Key words:** *Modelling, Communication Node, Crossbar Switch, Algorithms, Simlation.*

## МОДЕЛИРОВАНИЕ ПРОПУСКНОЙ СПОСОБНОСТИ КОММУТАЦИОННОГО УЗЛА ПРИ НЕСБАЛАНСИРОВАННОМ ВХОДЯЩЕМ ТРАФИКЕ

*T. Tashov*

Институт информационных и коммуникационных технологий, Болгарская  
Академия Наук, София, Болгария

Задача вычисления бесконфликтного расписания коммутации пакетов в коммуникационных узлах с матричным переключателем имеет неполиномиальной сложности. После проверки эффективности нового алгоритма вычисления моделированием пропускной способности коммутатора при равномерно распределенном сбалансированном входящем трафике, требуется проверка для неравномерного трафика. В этой работе предложено семейство шаблонов-матриц для симуляции несбалансированного трафика на базе модели “горячей точки”. Проверка их годности выполнена путем компьютерной симуляции коммутации с использованием известного PIM-алгоритма, специфицированного с применением аппарата Обобщенных сетей.

### 1. Введение

Для инженера пакетный коммуникационный узел с матричным переключателем (crossbar switch node - коммутатор) это устройство, основное

назначение которого – максимальное ускорение передачи данных за счет параллельно существующих потоков между узлами коммуникационной сети. В идеальном случае коммутатор передает пакеты между узлами, подключенными к его портам, с той скоростью, с которой узлы генерируют эти пакеты, не внося дополнительных задержек и не теряя ни одного пакета [1, 2]. Это должно обеспечить бесконфликтное расписание для коммутации, вычисляемое управляющим блоком коммутатора (Scheduler – Рис.1).

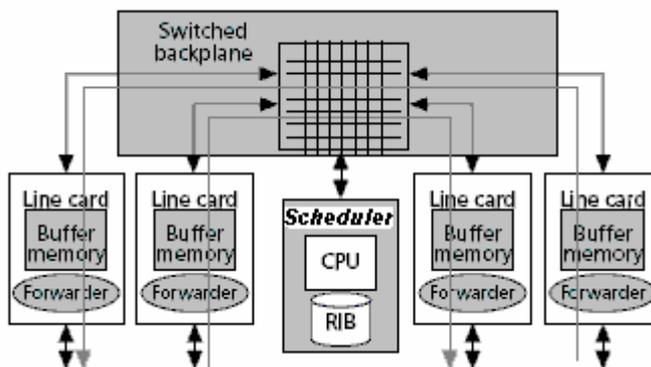


Рис 1. Общая архитектура коммутатора третьего поколения [1].

С математической точки зрения, эта задача является NP-полной [3] и для нее отсутствует единый эффективный метод решения. Существующие решения удовлетворяют часть целей [4], используя разные формализмы [5].

Повышение объемов коммуникационного трафика требует новых, более эффективных алгоритмов. Их надо проверять. Первый шаг при проверке эффективности нового алгоритма расчета бесконфликтного расписания - моделирование пропускной способности коммутатора при равномерно распределенном сбалансированном (uniform i.i.d. Bernoulli) входящем трафике. Следующий шаг требует проверки при несбалансированном трафике.

В настоящей работе предложено семейство шаблонов-матриц для симуляции несбалансированном (по выходам) трафика на базе модели "горячей точки" ([6]). Возможности его использования проверены посредством компьютерной симуляция коммутации пакетов с использованием известного РИМ-алгоритма [7], для которого с использованием аппарата Обобщенных сетей [8] построена модель [9]. В данном докладе представлены результаты вычислительных экспериментов для алгоритма, которые позволяют указать на заметное влияние размера входного буфера коммутатора на его пропускной способности. Они могут стать опорной базы для сравнения с результатами работы других алгоритмов.

## 2. Задача построения бесконфликтного расписания

Проблема конфликтов общепринятая : коммутатор имеет  $n$  входных каналов и  $n$  выходных каналов связи, то есть  $n$  источников пакетов через коммутатор могут посылать их к  $n$  получателям [3]. Заявки на передачу образуют матрицу трафика  $T$  с размерностью  $n \times n$ , причем  $T_{ij} = r$  ( $r = 1, 2, \dots$ ) если число заявок от  $i$ -того источника к  $j$ -тому приемнику равно  $r$ , и  $T_{ij} = 0$  если нет заявки. Для бесконфликтной коммутации надо вычислить последовательность бесконфликтных матриц  $Q_m$  (размерности  $n \times n$ ,  $m$  принадлежит интервалу  $[1, 2, \dots]$ ), сумма которых давала бы  $T$ . Каждая  $Q_m$  должна в любой строке или столбце иметь не более одной единицы. “Длина” последовательности является целое положительное число  $k$  – это число матриц в решении.

Эффективность работы коммутаторов на первом месте оценивают по использовании пропускной способности выходных каналов (throughput) – аналог Коэффициента полезного действия (Кпд). Для создания опорной базы для сравнения с результатами работы других алгоритмов, нам нужны такие матрицы трафика  $T$ , которые: (i) легко генерировались для любой размерности  $n \times n$  коммутатора; (ii) генерация не зависела бы от типа аппаратного и программного обеспечения и от операционной системы; (iii) их точное и оптимальное решение ( число  $k_{\text{опт}}$  ) для бесконфликтного расписания было известно.

Тогда возможно надежное сравнение и адекватность выводов.

Шаблоны матриц для равномерно распределенном (uniform i.i.d. Bernoulli) входящем трафике мы использовали в работе [10]. Продолжение исследований потребовало моделирование неравномерного входящего трафика.

Модель Чао для равномерно распределенного по входам (i.i.d.) и неравномерно распределенного по выходам трафика дается как [6] :  $\lambda_{ij} = \rho * 0,5$  для  $i = j$  и  $\lambda_{ij} = \rho * 0,5 / (n-1)$  в противном случае ( $i, j \in 1, \dots, n$ ), где  $\rho$  - интенсивность нагрузки на одной входящей линии (i.i.d. Bernoulli). По аналогии с матричным представлении для модели Чанга [10], построим базовую Трафичную матрицу для трафика типа “горячей точки”, которую будем обозначать  $Chao_1$ .

Матрица для шаблона  $Chao_1$  естественным образом развивается в виде  $Chao_i$ , где  $i=1, 2, \dots$ , как показано на Рис.2. Оптимальное решение для  $Chao_i$  равно  $k_{\text{опт}} = i * 2 * (n-1)$ , при размерности  $T$  равной  $n \times n$ . Это является эквивалентом решения модели Чао для дискретных значений трафика в виде матрицы  $T$ .

Развитие шаблона  $Chao_1$  в виде семейства шаблонов  $Chao_i$  позволяет, кроме получения значений для Кпд, исследовать влияние размера входного буфера на характеристики алгоритма для вычисления расписания.

$$T = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \dots \begin{bmatrix} k-1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & k-1 \end{bmatrix} \dots$$

$2 \times 2 \quad 3 \times 3 \quad k \times k$

$$T = \begin{bmatrix} i & i \\ i & i \end{bmatrix} \begin{bmatrix} i*2 & i & i \\ i & i*2 & i \\ i & i & i*2 \end{bmatrix} \dots \begin{bmatrix} i*(k-1) & \dots & i \\ \vdots & \ddots & \vdots \\ i & \dots & i*(k-1) \end{bmatrix} \dots$$

$2 \times 2 \quad 3 \times 3 \quad k \times k$

Рис.2 Матрицы трафика Chao<sub>1</sub> и Chao<sub>i</sub> для размерностей коммутатора до **k**

### 3. Построение Обобщенно-сетевой модели

РМ-алгоритм вычисляет ряд бесконфликтных матриц  $Q_k$ , для каждой проходя через три фазы [7]: Request, Grant и Асерт. Первое исполнение трех этапов формирует первую матрицу  $Q_1$  бесконфликтного расписания. Оставшиеся неудовлетворенные заявки подаются на второе исполнение. Трехэтапное исполнение продолжается, пока все заявки матрицы  $T$  не получат назначение на коммутацию. Тогда сформирован искомый ряд матриц  $Q_k$ , и  $k$  определено.

Входы выполняют параллельно первую фазу. Выходы выполняют параллельно вторую фазу. Входы работают параллельно и в третьей фазе. Этот параллелизм есть наше основание для применения Обобщенных сетей (ОС).

Формально РМ-алгоритм специфицирован с помощью аппарата ОС [8]. Основные элементы ОС – это переходы – моделирующие действия, и позиции – моделирующие состояния. В позициях находятся ядра, которые проходят через переходы соответственно с логических условий. Ядра имеют необходимые характеристики, которые могут изменяться при прохождении через переходов. Любая Обобщенно-сетевая модель имеет графичное представление.

Используемая нами ОС-модель РМ-алгоритма представлена в [9]. Основное предимство этой модели – явное представление Виртуальных выходных очередей (VOQ). VOQ есть решение, вызванное необходимостью преодоления HOL-блокировки в коммутаторе [3]. Они формируются по входам коммутатора и их содержимое будет входными данными для алгоритма вычисления бесконфликтного расписания.

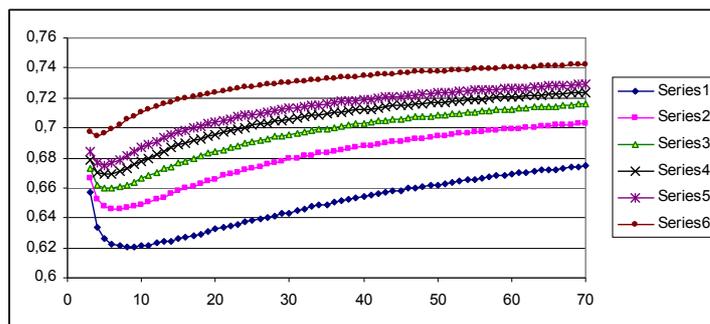
При представлении входящего трафика матрицей  $T$ , каждая виртуальная очередь является ее строкой, причем число  $r$  определяется количеством заявок в виртуальной очереди, а число ячеек в строке отвечает числу Виртуальных очередей (числу входов равно числу выходов коммутатора –  $n$ ).

### 4. Результаты вычислительных экспериментов

Последовательность вычисления матриц  $Q_k$  построена по методике [9]. Для программирования был использован программный пакет Vfort,

предоставленный московским Институтом математического моделирования РАН для свободного пользования [10]. При вычислительных экспериментах использовалась IBM совместимая вычислительная машина типа Pentium IV с тактовой частотой 3000 МГц и 2 Гб оперативной памяти. Ограничением для экспериментов явилось время выполнения кода программы. Входными данными для алгоритма является матрица входящих заявок на коммутирование - матрица  $T$ . При размерности  $n \times n$  ( $n = 2, 3 \dots \in N$ ) каждый ее элемент принимает значение сообразно выбранного  $Chao_i$ -шаблона.

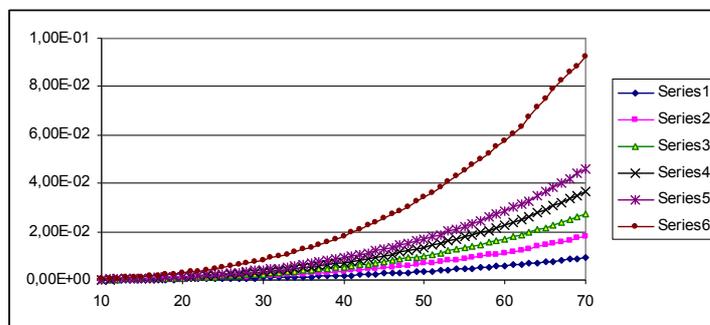
На Рис. 3 и Рис.4 показаны результаты работы алгоритма. По горизонтальной оси координат показана размерность  $n$  входной матрицы  $T$  (от 3 до 70). По левой вертикальной оси Рис.3 показано число  $k$  выходных матриц  $Q_m$ , приведенное к пропускной способности коммутатора- принятой за 1. По левой вертикальной оси Рис.4 - время для вычисления расписания (в секундах). Кривые соответствуют результатам, полученными с разными шаблонами  $Chao_i$  трафика : Series 1 – это трафик вида  $Chao_1$ , Series 2 - трафик вида  $Chao_2$ , и т.д. Series6 – трафик вида  $Chao_{10}$ .



Фигура 3. Пропускная способность для шаблонов вида  $Chao_{1-5}$  и  $Chao_{10}$ .

Время вычисления имеет линейную зависимость от нагрузки и пропорционально меньше кубу  $n$  (и нарастает - в степень от 2,60 для  $Chao_1$  до 2,67 - для  $Chao_{10}$ ). Время – на однопроцессорном вычислителе, поэтому временная сложность надо понизить на одну степен  $n$ . Тоест для  $Chao_i$  время выполнения РИМ-алгоритма меньше второй степени от размерности коммутатора, что указывает на сходимость – наличие теоретического предела (меньшего 100%) пропускной способности для исследованного трафика типа “горячей точки”.

Для РИМ-алгоритма известен теоретический предел для uniform трафика ~ 63,2% (точно =  $1-1/e$ ). Полученные результаты – смотри Рис.3 – позволяют заключить, что для исследованного шаблона он выше 74 %.



Фигура 4. Время вычисления расписания.

Выскажем предположение – из простых соображений – что он меньше 81,6 %  $\sim (100+63,2)/2$ . Это надо исследовать более точно, а для проверки утверждений провести симуляцию для более высоких размерностей коммутатора на грид-структуре.

## 5. Заключение

Компьютерное моделирование подтверждает применимость шаблонов, продемонстрировав линейную зависимость времени вычисления решения от степени загрузки входов. Результаты указывают на заметное влияние размера входного буфера коммутатора на его пропускной способности. Вопрос наличия границы пропускной способности для предложенных шаблонов - открыт.

## Литература REFERENCES

1. Csaszar A., Enyedi G., Retvari G., Hidell M., Sjodin P. Converging the Evolution of Router Architecture and IP Networks // IEEE Network, 2007, V.4, P. 8-14.
2. Ермаков А.С. Модель балансировки коммутационных процессов // Proc. of Int. Workshop "DCCN-2007", Sofia, Bulgaria, 2007. Moscow, Technosfera, 2007, P. 171-176.
3. Chen T, Mavor J, Denyer Ph, Renshaw D. Traffic routing algorithm for serial superchip system customisation // IEE Proc.- E,part, Jan 1990, Vol. 137, No.1. pp.65-73.

4. Qu G., Chang H.G., Wang J., Fang Z., Zheng S.Q. Designing Fully Distributed Scheduling Algorithms for Contention-Tolerant Crossbar Switches // Proc. of IEEE Conf. HPSR' 2010, Richardson, TX, USA, 2010. P. 69 - 74.
5. Kolchakov K.H. An Approach for Non-Conflict Schedule Synthesis - Modeling and Optimization // Научная сессия "МИФИ-2008" 21-25 Января 2008 года, Москва, Россия. Том 10. Интеллектуальные системы и технологии, С.100-101.
6. Chao-Lin Yu; Cheng-Shang Chang; Duan-Shin Lee; CR Switch: A Load-Balanced Switch With Contention and Reservation // IEEE/ACM Transactions on Networking, vol. 17 no.5, P. 1659–1671, October 2007.
7. T.Anderson, S.Owicki, J.Saxe, and C.Thacker. High speed switch scheduling for local area networks // ACM Trans. Comput. Syst., vol. 11, no.4, P.319-352, Nov. 1993.
8. Atanassov K. Generalized Nets and System Theory // Akad. Press "Prof.M. Drinov", Sofia, Bulgaria, 1997.
9. Ташев Т., Железов Р. Моделирование виртуальных выходных очередей для коммутационного узла Обобщенными сетями // Proc. Of Int. Workshop "DCCN '2010" October 26-28 2010, Moscow, Russia. RDC "INT", Moscow, 2010. P. 198-203.
10. <http://www.imamod.ru/~vab/vfort/download.html>