ОПРЕДЕЛЕНИЕ ИНТЕРВАЛА ЭФФЕКТИВНОСТИ ПАРАМЕТРОВ ПРОЦЕДУРЫ ВЫЧИСЛЕНИЯ ГРАНИЦЫ ПРОПУСКНОЙ СПОСОБНОСТИ ПАКЕТНОГО КОММУТАТОРА

Ташо Ташев, Арсений Баканов, Радостина Ташева

Аннотация: Задача вычисления бесконфликтного расписания для коммутатора пакетов с матричным переключателем является NP-полной с математической точки зрения. Поэтому постоянно предлагаются новые алгоритмы вычисления расписания. Их надо проверять. Пропускная способность коммутатора при работе "хорошего" алгоритма монотонно стремится к некоторому пределу. Нами была предложена процедура для определения количественных значений сходимости к пределу, использующая данные компьютерного моделирования. В этой работе представлены результаты моделировании одного алгоритма для конкретной модели входящего потока пакетов, выполненных на грид-структуре CERN-а (Ixplus.cem.ch). Проверялась гипотеза, что вычисление сходимости зависит от выбранного параметра т процедуры. Проверка проводилась для значений т = 4 и т = 3. Полученные данные подтверждают валидность гипотезы для выбранных значений т. В результате мы высказываем новое утверждение: полученные значения ошибки для вытотенней переменной δ дают (независимую) оценку ошибок моделирования. На этой базе мы можем планировать такие эксперименты, которые дадут нам требуемую точность для определения границы при заданном ограничении доступной компьютерной мощности.

Ключевые слова: Моделирование, Коммуникационный узел, Матричный переключатель, Алгоритмы.

Классификация ключевых слов ACM: B.4.4 Performance Analysis and Design Aids, C.2.1 Network Architecture and Design, C.4 Performance of Systems

Введение

Коммуникационный узел для коммутации пакетов с матричным переключателем (crossbar switch node - коммутатор) имеет основную задачу - максимальное ускорение передачи пакетов данных, за счет параллельно распределенных потоков между узлами коммуникационной сети [Deb et al, 2006]. В идеальном случае узел передает пакеты между другими узлами, подключенными к его портам, с той скоростью с которой эти узлы генерируют пакеты, не внося дополнительных задержек и не теряя ни одного пакета [Elhanany and Hamdi, 2007]. Для достижения этой цели необходимо обеспечить бесконфликтное расписание для коммутации, вычисляемое управляющим блоком (Scheduler – Рис.1).

Формально, математическим путем доказано, что задача вычисления бесконфликтного расписания является NP-полной [Chen et al, 1990]. Существующие решения – алгоритмы PIM [Anderson et al, 1993], iSLIP [Gupta and McKeown, 1999], DISQUO [Shunyuan et al., 2010], CTC(N) [Chang et al, 2012] и др. дают частичное достижение результата. При этом используются формализмы такие как теория массового обслуживания, генетические алгоритмы, матрицы-маски и др. [Elhanany, Hamdi, 2007; Kolchakov, 2011].

Достижения научно-технического прогресса ведут к постоянному повышению скоростей и объемов коммуникационного трафика [Атанасова и др., 2010], что требует новых алгоритмов. Их надо проверять на эффективность.



Рис 1. Логическая структура коммутатора с входящими буферами [Deb et al, 2006].

Первый шаг при проверке эффективности нового алгоритма для вычисления бесконфликтного расписания – это моделирование пропускной способности (throughput) коммутатора при равномерно распределенном (uniform i.i.d. Bernoulli) входящем трафике. Следующий шаг требует проверки при неравномерно распределенном трафике [Chang et al, 2012].

В настоящей работы использовано семейство шаблонов-матриц для моделирования неравномерно распределенного трафика на базе модели "горячей точки" (hotspot - Чао-модель [Chao-Lin et al., 2007]). Представлены результаты вычислительных экспериментов для PIM-алгоритма, модель которого построена нами с использованием аппарата Обобщенных сетей [Atanassov, 1997]. Пропускная способность коммутатора (Кпд) при работе PIM-алгоритма стремится к определенному пределу. Нами была предложена процедура для определения количественных значений сходимости к пределу, позволяющая вычислить этот предел однозначно [Ташев и др., 2013]. Там была высказана гипотеза, что вычисление сходимости зависит от выбираемого параметра *m* процедуры. Здесь представлены результаты проверки гипотезы. С помощьи процедуры были получены численные значения для предела Кпд для двух значений параметра *m* процедуры. Они были сравнены и гипотеза получила подтверждение. На этой базе сделан вывод как планировать такие эксперименты, которые дадут требуемую точность для определения границы.

Задача построения бесконфликтного расписания

Проблемма конфликтов определяется так: коммутатор имеет *n* входных каналов и *n* выходных каналов связи, то есть *n* источников пакетов через коммутатор могут посылать их к *n* получателям [Gupta and McKeown, 1999], Заявки на передачу пакетов формируют матрицу трафика *T* с размерностью *n* x *n*, причем $T_{ij} = r$ (r =1,2,..) когда число заявок от i-того источника к j-тому приемнику равно r, и $T_{ij} = 0$ если нет заявки. Для бесконфликной комутации вычисляеться последовательность бесконфликтных матриц Q_1 , $Q_2,..., Q_i,..., Q_q$ (матрицы размерности *n* x *n*, *q* – неизвестное целое), сумма которых давала бы T. Каждая Q_i ($1 \le i \le q$) такая, что в любой строке или столбце имеет не более одной единицы. "Длина" последовательности является целое положительное число k_{pew} – это число матриц в решении (последовательности $Q_1 \div Q_q$). Отметим, что матрица *T* представляет однонаправленный поток заявок – от входам к выходам коммутатора, представляющих пакетов с одной и той же длины.

Модель Чао (hotspot traffic) для равномерно распределенного по входам (i.i.d. Bernoulli) и неравномерно распределенного по выходам трафика дается как [Chao-Lin et al., 2007]: $\lambda_{ij} = \rho.0,5$ для i = j и $\lambda_{ij} = \rho.0,5/(n-1)$ в противном случае ($i, j \in 1, ..., n$), где ρ - интенсивность нагрузки одной входящей линии. Для модели Чао строится базовая Трафичная матрица (для трафика типа "горячей точки"), которую будем обозначать Chao₁. Матрица для шаблона Chao₁ развертывается в виде семейства шаблонов Chao_i, где i=1, 2,..., как показано на Рис. 2. Оптимальное решение для Chao_i равно k_{ont} =2.i.(n-1) при размерности T равной $n \ge n$ [Ташев, 2013], и вычисление Клд производится как Клд = k_{ont}/k_{pew} .

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

Рис. 2 Матрицы трафика Chao₁ и семейства Chao_i при размерности коммутатора до k x k

Компьютерное моделирование РІМ-алгоритма

Формально PIM-алгоритм нами специфицирован с помощью аппарата Обобщенных сетей (ОС) [Atanassov, 1997]. Основными элементыми ОС являются переходы и позиции. В позициях находятся ядра, которые проходят через переходы соответствено с логических условий, имея необходимые характери-стики Любая ОС-модель имеет графическое представление (кроме формальной спецификации). Для модели PIM-алгоритма они представлены в [Tashev and Monov, 2012]. Переход от ОС-модели к вычислительным экспериментам выполнен сообразно [Tashev and Vorobiov, 2007]. Для программирования использован программный пакет Vfort, предоставленный для свободного пользования Московским Институтом математического моделирования РАН. Исходный код был компилирован средствами грид-структуры CERN-а (Ixplus.cern.ch) и результирующий код был выполнен там же. Основным ограничением при компьютерном моделировании являлось время выполнения кода программы (до 50 часов). Входными данными для алгоритма являются шаблоны Chao_i, матриц входящих заявок на коммутирование.

На Рис. З показаны результаты работы алгоритма. По горизонтальной оси координат показана размерность *п* входной матрицы *T* (от 3 до 60). Использованы обозначения: C₁ – это Кпд входящего трафика вида Chao₁, C₅ - трафика вида Chao₅, и т.д.



Рис. 3 Пропускная способност для шаблонов вида Chao1-5 и Chao100.

По вертикальной оси показано число **k**_{реш} вычисленной последовательности выходных матриц **Q**_q, приведенное к пропускной способности коммутатора (т.е. Кпд = **k**_{опт} / **k**_{реш}). Кривые соответствуют результатам, полученными с шаблонами Chao_i трафика. Каждое значение для Chao₁ до Chao₅ – среднее от 10 000 "проходов" шаблона, а значения для Chao₁₀₀ – среднее от 10000 "проходов" шаблона, а значения для Chao₁₀₀ – среднее от 1000 "проходов" шаблона. Результаты компьютерного моделирования указывают как на наличие предела Кпд для каждой исследованной размерности комутационного поля **n** при *i*→∞, так и при **n** клонящему к бесконечности. Для точного вычисления первого предела предложена соответствующая процедура.

Определение сходимости к верхней границе пропускной способности

Предложенная процедура вычисляет сходимость пропуской способности коммутатора к определенной верхней границы по результатам вычислительных экспериментов (при моделировании конкретного алгоритма с заданной семьей шаблонов для входящего трафика). Далее мы воспользуемся ею без формального описания. Для нас важно проверить гипотезу δ=*m*^½ (при m=10, δ≈3,15 [Ташев и др., 2013]).

Шаг 1. На основании данных Рис. 3 решаем что эсть основания существовании верхней границы Кпд.

Шаг 2. Выбираем "шаг" для индекса і шаблона равен 4 (*m* = 4) в том смысле, что будем исследовать Клд для ряда шаблонов Chao₁, Chao₄, Chao₆, Chao₂₅₆, Резултаты моделирования для первых четырех шаблонов даны на Рис. 4 Значения для С₁ и С₄ – среднее от 10 000 "проходов", для остальных - 1000 "проходов".

Шаг 3. Вычисляем разницу между Кпд для соседних шаблонов. Строим график как показано на Рис. 5. Имеем Δ₁=Кпд(Chao₄) - Кпд(Chao₁), Δ₂=Кпд(Chao₁₆) - Кпд(Chao₄), Δ₃=Кпд(Chao₆₄) - Кпд(Chao₁₆)



Рис.4 Пропускная способность для Chao1 до Chao64

Рис.5 Разности пропускной способности (m = 4)

Шаг 4. Вычисляем отношение разностей $\delta_1 = \Delta_1 / \Delta_2$, $\delta_2 = \Delta_2 / \Delta_3$. Графики показаны на Рис. 6 (влево).

Оба отношения сходятся к значению 2, но у обоих видно влияние смещения минимума Кпд С₁ и С₄ влево (при увеличении индекса шаблона). Второе отношение устремляется быстрее к 2, но ошибка большая (так как число "проходов" при С₁₆ и С₆₄ меньшее). Поэтому вычислим $\delta_3 = \Delta_3/\Delta_4$, Δ_4 =Кпд(Chao₂₅₆) - Кпд(Chao₆₄) используя шаблон Chao₂₅₆. Результат показан на Рис. 6 (вправо): δ_3 уже колеблется около значения 2.



Рис.6 Отношения разностей пропускной способности (*m* = 4)

Шаг 5. Результаты моделирования показывают, что исключая зоны смещения минимума Кпд (2 < n < 15), разности Кпд между "шагами" шаблонов уменьшаются в геометрической прогрессии. Из теории известно, что бесконечный числовой ряд вида $1/a + 1/a^2 + 1/a^3 + ... + 1/a^i + ...,$ где a>1, сходится при i→∞ к значению 1/(a-1). В нашем случае разности Δ_2 , Δ_3 ,... образуют такой ряд с a = 2. Следовательно, коэффициент сходимости для наших результатов Sum(2) = 1/(2-1) =1. А гипотеза $\delta = m^{\frac{1}{2}}$ подтверждается: 2 = 4^{1/2}.

Поэтому ввычисляем значения верхней границы для конкретных *n*, используя шаблон с самым большим номером - Chao₂₅₆, таким образом:

ГраницаКпд(*n*) = Кпд(Chao₂₅₆ (*n*)) + Sum (2). Δ₄ (*n*)= Кпд(Chao₂₅₆ (*n*)) + 1. Δ₄ (*n*)

Результат показан на Рис. 7: с (=δ граница) обозначено вычисление использующее δ = const = 2. Самое большое значение Кпд(Chao₂₅₆(70))=0,76767, Δ₄(70) = 0,00618, **Граница** Кпд(70) = 0,77385.



Рис.7 Граница пропускной способности (*m* = 4)

Шаг 6. Рассматриваем на Рис. 7 тенденции нарастания значения границы Кпд. Имея в виду самое большое значение 0,7738 для Кпд, полученное на Шаге 5, оцениваем значение при *п* →∞ на 0,775.

Шаг 7. Первичный анализ ошибок. На Рис. 7 справа показаны значения границы, вычисленные при Sum(δ =const=2) и Sum(δ_3 =variable) (обозначено ~ δ граница). Ошибка Граница Кпд(n) = ±0,0003 - уменьшается больше чем в два раза при вычислении использующее δ = const = 2. Ошибка больше зависит – при таком выборе – от точности предьидущего шаблона (Chao₆₄ (n)).

Для проверки гипотезы δ = *m*^½ при *m* = 3 повторим Шаги 2 - 7.

Шаг 2. Будем исследовать Кпд для ряда шаблонов Chao₁, Chao₃, Chao₉, Chao₂₇, Chao₈₁, Резултаты моделирования для первых четырех шаблонов даны на Рис.8. Повысим точность – все значения для C1 до C₈₁ – среднее от 10 000 "проходов".

Шаг 3. Вычисляем разницу между Кпд для соседних шаблонов. Строим график как показано на Рис. 9.

Шаг 4. Вычисляем отношение разностей δ₁ = Δ₁/ Δ₂ , δ₂ = Δ₂/ Δ₃. Графики показаны на Рис. 10 (влево). Сходимост – медленне, ошибки – меньше, по отношении (*m* = 4). Поэтому опять вычислим δ₃: уже колеблется около 1,73 (Рис. 10 влраво).

Шаг 5. В случае m = 10 получили, что δ_3 колебалось около 3,15. Выше за случая $m = 4 \delta_3$ было около 2. Сейчас за случая m = 3 получили значение для δ_3 около 1,73. Имеем еще одно подтверждение гипотезы: 1,732 = 3^{1/2}. Тогда колебания значения δ_3 – это следствие ошибок компьютерного моделирования (в данном исследовании – из-за случайного выбора, используемого в PIM-алгоритме). В результате мы приходим к **Утверждению:** полученные отклонения значений внутренней переменной δ от точного значения $m^{1/2}$ дают (независимую) оценку численных ошибок компьютерного моделирования.

Закончим шаг 5. Принимаем, что разности Δ_3 , Δ_4 ,... образуют ряд с а = 3^{1/2}. Следовательно, коэффициент сходимости для *m* = 3 есть Sum(3^{1/2}) = 1/(1,732-1) = 1,366.









Рис. 10 Отношения разностей пропускной способности (*m* = 3)

Вычисляем значения верхней границы, используя шаблон Chao₈₁. Результат показан на Рис. 11: с (= δ граница) обозначено вычисление использующее δ = const = 1,732. Самое большое значение Кпд(Chao₈₁(70)) = 0,76281, Δ₄(70) = 0,00819, 1,366. 0,00819 = 0,1119 → **Граница** Кпд(70)=0,77400.

Шаг 6. Рассматриваем на Рис. 11 тенденции нарастания значения границы Кпд. Имея в виду самое большое значение 0,7740 для Кпд, полученное на Шаге 5, подтверждается значение при *n* →∞ как 0,775.



Рис. 11 Граница пропускной способности (*m* = 3)

Шаг 7. Первичный анализ ошибок. Более точны данные (10 000 сичуляций) ведут к меньшей ошибке границы при δ_3 = variable (обозначено ~ δ граница). Ошибка ГраницаКпд(n) = ±0,00015 при δ = const = 3^{1/2} тоже меньше. Сравним оба графика Граница Кпд(n) для δ =const как показано на Рис. 12.

Вычисленная **Граница** Кпд(*n*) при *m* = 3 более "гладкая". Затраченное компьютерное время для моделирования: 81+27=108 условных единиц времени (10 000 "проходов"). Затраченное компьютерное время для моделирования при *m* = 4: (265+64)/10 = 32 условных единиц (1000 "проходов"), а потеря точности ~2 раза. Имея в виду, что нас ограничивает время моделирования, то пока не достигнут этот потолок, при стремлении достигнуть большей точности надо планировать эксперименты с *m*=3, при увеличенном числе "проходов" для каждого шаблона.



Рис. 12 Граница пропускной способности при m = 4 и m = 3

Заключение

Исследование подтверждает зависимость коэффициента сходимости от параметра *m* для исследованных значений *m* = 4 и *m* = 3. В результате полученные значения ошибки для внутренней переменной δ можно использовать как (независимую) оценку ошибок моделирования. На этой базе при планировании

дальнейших экспериментов для получения требуемой точности для определения границы Кпд, при заданном ограничении доступной компьютерной мощности, предлагается использовать значение *m* = 3. Описанную процедуру следует апробировать при исследовании других алгоритмов и моделей трафика.

Благодарности

Авторы благодарны доценту к.ф.н. Леандру Литову, Заведующий секцией Атомная Физика, Физический Факультет, Софийский Университет "Св. Климент Охридски", Болгария, за предоставленный доступ к компьютерным ресурсам грид-структуры CERN.

Библиография

- [Anderson et al, 1993] 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, 1993, pp.319-352.
- [Atanassov, 1997] Atanassov K. Generalized Nets and System Theory. Akad. Press "Prof.M. Drinov", Sofia, Bulgaria, 1997.
- [Chang et al, 2012] H.J.Chang, G.Qu, S.Zheng. Performance of CTC(N) Switch under Various Traffic Models. Springer, Lecture Notes in Electrical Engineering, Volume 126, 2012, pp. 785-793.
- [Chao-Lin et al., 2007] 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, October 2007. pp.1659–1671.
- [Chen et al, 1990] T.Chen, J.Mavor, Ph.Denyer, D.Renshaw. Traffic routing algorithm for serial superchip system customisation. IEE Proc 137:[E]1, 1990.
- [Deb et al, 2006] S.Deb, D.Shah, S.Shakkottai. Fast Matching Algorithms for Repetitive Optimization: An Application to Switch Scheduling // Information Sciences and Systems, 2006 40th Annual Conference on, 22-24 March 2006, P. 1266-1271.
- [Elhanany and Hamdi, 2007] I.Elhanany, M.Hamdi. High-performance packet switching architectures. Springer-Verlag London Limited, 2007.
- [Gupta and McKeown, 1999] P.Gupta, N.McKeown. Designing and Implementing a Fast Crossbar Scheduler. IEEE Micro, Jan-Feb 1999, pp. 20-28.
- [Kolchakov, 2011] Kolchakov K. An Algorithm Synthesis of Non-Conflict Schedule by Diagonal Connectivity Matrix Activation. Proceedings of the International Conference AUTOMATICS AND INFORMATICS `11, John Atanasoff Society of Automatics and Informatics, Bulgaria, Sofia 03.10-07.10.2011., pp. B-247 – B251.
- [Shunyuan et al., 2010] Shunyuan Ye, Yanming Shen, Shivendra Panwar, DISQUO: A Distributed 100% Throughput Algorithm for a Buffered Crossbar Switch, Proceedings of IEEE Conference on High-Performance Switching and Routing (HPSR 2010), Richardson, TX, USA, 2010. pp. 63 – 68.
- [Tashev and Monov, 2012] Tashev T., Monov V. Modeling of the Hotspot Load Traffic for Crossbar Switch Node by Means of Generalized Nets. Proc. of Intelligent Systems (IS), 2012 6th IEEE International Conference, 6-8 Sept. 2012, Sofia, Bulgaria. pp.187-191.
- [Tashev and Vorobiov, 2007] Tashev T.D., Vorobiov V. M. Generalized Net Model for Non-Conflict Switch in Communication Node. Proc. Of Int. Workshop "DCCN'2007" September 10-12, 2007, Moscow, Russia. IPPI Publ., Moskow, 2007, pp.158-163.
- [Атанасова и др., 2010] Т. Атанасова, Т. Н. Савченко, Г.М. Головина, А С. Баканов. Интеллектуальная информационная среда обитания и субъективное восприятие качества жизни. Методы исследования психологических структур и их динамики: Вып. 5: / Под редакцией Т. Н. Савченко, Г.М. Головиной. М.: Издательство «Институт психологии РАН», 2010. стр. 164-170. (Труды Института психологии РАН)

- [Ташев и др., 2013] Ташев Т., Баканова Н., Ташева Р. Исследование верхней границы пропускной способности коммутаци-онного узла при входящем трафике типа "горячей точки". International Journal "Information Technologies & Knowledge", Vol.7, No. 2, 2013, pp. 182-189.
- [Ташев, 2013] Ташев Т.Д. Моделирование пропускной способности MiMa-алгоритма для пакетного коммутатора при входящем трафике типа "горячей точки". Proc. of International Conference "DCCN'2013", October 7-10 2013, Moscow, Russia. JSC "TECHNOSPHERA", Moscow, 2013. pp.257-264.

Информациа об авторах



Ташев Ташо Димитров – инж.-математик, главный ассистент, Институт информационных и коммуникационных технологий Болгарской академии наук, ул. "Акад.Г.Бончев" бл.2, София 1113, Болгария; e-mail: <u>ttashev@iit.bas.bg</u>

Основные области научных исследований: Проектирование расспределенных информационных систем, Методы и средства исследования сетевых моделей



Баканов Арсений Сергеевич – к.т.н., Институт психологии Российской академии наук, Москва, Россия; e-mail: <u>arsb2000@pochta.ru</u>

Основные области научных исследований: Интелигентные системы, Системы принятия решений, Софтуерные технологии



Ташева Радостина Пенева – к.ф.н., Технический университет-София, бул. "Климент Охридски" 8, бл.10, София 1000, Болгария; e-mail: <u>rpt@tu-sofia.bg</u>

Основные области научных исследований: Процессы звездообразования, Галактики с активными ядрами, Компьютерное моделирование

Efficiency Interval Determination of a Calculation Procedure for Throughput Boundary of Crossbar Switch

Tasho Tashev, Arsenij Bakanov, Radostina Tasheva

Abstract: The problem connected to computation of non-conflict schedule for crossbar switch node is NPcomplete from mathematical point of view. It is the reason for constantly suggested new algorithms for schedule calculations. The problem is that their efficiency has to be checked. The throughput of a "good" algorithm tends monotonically to some boundary. We suggest a procedure for quantitative determination of the convergence toward the boundary, utilizing data from computer simulation. In this paper are presented results of computer simulation for PIM algorithm with a certain model of incoming traffic (Chao model), made by means of CERN grid structure (lxplus.cern.ch). We check the hypothesis that convergence computation depends on the chosen parameter m of the procedure. Checking is made for m=4 and m=3. The obtained data confirm validity of the hypothesis for the chosen values m. As a result we reach a new conclusion: the derived values of the error for the internal variable δ give (an independent) estimation of the simulation errors. Having this in mind we can plan experiments that will give us the required accuracy for the determination of the boundary under the limited computer power available.

Keywords: Modeling, Communication node, Crossbar switch, Algorithms.