

ПРОИЗВОДИТЕЛЬНОСТЬ БУФЕРНЫХ АСИНХРОННЫХ МНОГОСТУПЕНЧАТЫХ СЕТЕЙ С ПРОИЗВОЛЬНЫМ ЧИСЛОМ ПРИОРИТЕТНЫХ МОДУЛЕЙ ПАМЯТИ

Синтезируется математическая модель асинхронной многоступенчатой коммутирующей сети, которая применяется для оценки пропускной способности и времени отклика сети. Модель расширяется для случая произвольного числа приоритетных модулей памяти и случайно распределенного размера пакетов. Рассматривается влияние длины буфера на общую производительность системы.

Общая постановка проблемы

Многоступенчатые коммутирующие сети используются в качестве коммутирующей фабрики во многих коммерческих и экспериментальных многопроцессорных системах [1]. Производительность коммутирующей сети во многом определяет производительность системы в целом. Чтобы достичь оптимального соотношения скорости системы к ее стоимости, архитектору необходимо изучить влияние различных вариантов конфигурации сети на производительность многопроцессорной системы. Зная время отклика сети и ее пропускную способность, можно сделать вывод о производительности будущей многопроцессорной системы.

Анализ исследований и публикаций

Существующие результаты получены на основе математических моделей, построенных с использованием, преимущественно, вероятностных методов [2]. Такие модели могут быть применены только для многопроцессорных SIMD-систем с синхронно работающими процессорами [3].

Однако наиболее значительный интерес представляют MIMD-системы с работающими независимо друг от друга процессорами. Количество процессоров у таких систем нередко превышает десятки тысяч.

Для оценки производительности асинхронных коммутирующих сетей были построены математические модели с использованием теории очередей [3]. Недостаток предложенной в [3] модели состоит в ограниченности области ее применения. Модель применима только к системам с детерминированным временем коммутации пакетов и с не более чем одним приоритетным модулем памяти.

Цель исследования: разработать программный инструментарий для оценки производительности асинхронной буферной многоступенчатой сети.

Задача исследования: 1) синтезировать математическую модель многоступенчатой коммутирующей сети, которая позволяет оценить пропускную способность и время отклика асинхронной коммутирующей сети до начала имплементации. Модель должна быть применима к системам с недетерминированным временем коммутации пакетов и произвольным числом приоритетных модулей памяти; 2) реализовать модель в виде программного обеспечения.

Функционирование многоступенчатых коммутирующих сетей

Многоступенчатая коммутирующая сеть состоит из множества переключающих элементов, организованных в виде последовательных ступеней (рис. 1). Процессорные элементы (PE) генерируют пакеты независимо друг от друга. Данное предположение справедливо для MIMD-систем, где отдельный процессор управляется собственным потоком команд.

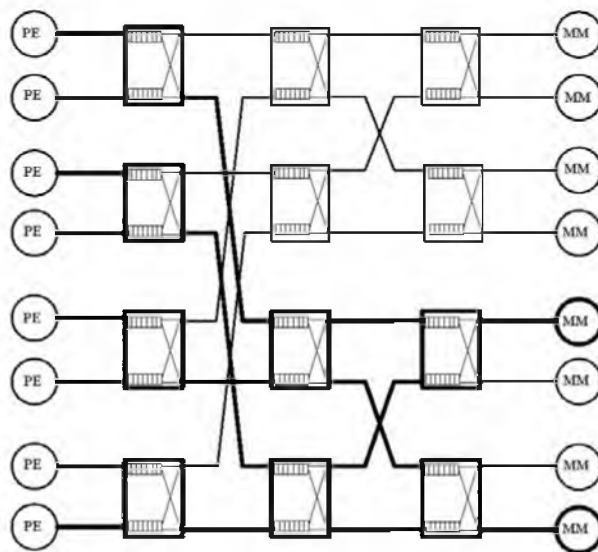


Рис. 1. Многоступенчатая коммутующая сеть 8×8

Чтобы не потерять входящие пакеты в случае высокой интенсивности трафика, каждый переключатель имеет буфер для временного хранения нескольких пакетов. Далее будет проанализировано влияние размера буфера на производительность многоступенчатой сети.

Математическая модель

Модель строится на основе следующих предположений:

- 1) процессорные элементы генерируют пакеты независимо друг от друга с интенсивностью λ согласно распределению Пуассона;
- 2) время, которое требуется переключателю для перенаправления пакета на один из своих выходных каналов, пропорционально размеру пакета;
- 3) размер пакетов распределен согласно закону Эрланга;
- 4) каждый пакет равновероятно перенаправляется к любому модулю памяти;
- 5) в случае необходимости перенаправить более одного пакета на один выходной канал один из пакетов блокируется.

Большинство существующих моделей предполагают наличие трафика, равномерно распределенного между всеми модулями памяти. Данное предположение выражено в пункте 4. После построения упрощенной модели данное ограничение будет снято для получения более общей математической модели.

Введем следующие обозначения:

- 1) λ – интенсивность генерации пакетов;
- 2) μ – интенсивность перенаправления пакетов на выходные каналы;
- 3) $P_k^{(L)}$ – вероятность того, что в буфере размера L находится k пакетов;
- 4) $P_{k,s}^{(L)}$ – вероятность того, что в буфере размера L находится k пакетов и пакет, который перенаправляется на один из выходных каналов, находится в фазе s . $1 \leq s \leq r$.

$$\text{Очевидно, что } P_k^{(L)} = \sum_{s=1}^r P_{k,s}^{(L)}.$$

Путь, который проходит пакет, удобно представить в виде последовательно соединенных друг с другом очередей. Уровень выходного потока предыдущей очереди равен уровню входного потока последующей очереди.

После прохождения первой ступени поток более не подчиняется условиям Пуассона по причине влияния ненулевого времени переключения пакета. Однако имитационные эксперименты показали, что погрешность, вызванная аппроксимацией потока к пуассоновскому, пренебрежительно мала.

Введем дополнительные обозначения:

- 1) λ_i – интенсивность входного потока ступени i ;
- 2) $P_k^{(L)}(i)$ – вероятность нахождения k пакетов в буфере на ступени i ;
- 3) x_i – вероятность блокировки пакета на ступени i .

Пусть τ_i – случайная величина, которая представляет собой длину интервала времени прошедшего между отправками двух последовательных пакетов. Пусть ϕ – событие, обозначающее опустошение буфера после отправки последнего пакета. Тогда

$$f_{\tau_i}(t) = P_0^{(L)}(i) \cdot f_{\tau_i|\phi}(t) + (1 - P_0^{(L)}(i)) \cdot f_{\tau_i|\bar{\phi}}(t),$$

где $f_{\tau_i}(t)$ – функция распределения случайной величины τ_i . Используя свойства линейных операторов, имеем: $E[\tau_i] = P_0^{(L)}(i) \cdot E[\tau_i|\phi] + (1 - P_0^{(L)}(i)) \cdot E[\tau_i|\bar{\phi}]$. Далее находим, что

$$E[\tau_i] = P_0^{(L)}(i) \cdot \left(\frac{1}{\lambda_i} + \frac{1}{\mu} \right) + (1 - P_0^{(L)}(i)) \frac{1}{\mu}.$$

В случае, когда последний отправленный пакет оставляет за собой пустой буфер, время следующей отправки пакета будет состоять из двух составляющих: времени, необходимого для поступления следующего пакета, и времени, необходимого для перенаправления пакета.

Так как интенсивность входного потока равна интенсивности выходного потока предыдущей ступени, имеем: $\lambda_{i+1} = \frac{1}{E[\tau_i]} = \frac{(1 - x_i)\lambda_i\mu}{P_0^{(L)}(i)\mu + (1 - x_i)\lambda_i}$.

Используя теорему Литла, получаем: $E[T_i] = \frac{\sum_{k=1}^{L+1} k \cdot P_k^{(L)}(i)}{\lambda_i}$. Среднее время отклика

сети рассчитывается как время пребывания пакета на ступени i , просуммированное по всем i . Пропускная способность сети рассчитывается как λ_i последней ступени.

Найдем выражение для $P_k^{(L)}(i)$. Воспользуемся результатами, приведенными в [4]:

$$P_{1,s} = \frac{\lambda}{\Gamma\mu} \left(1 + \frac{\lambda}{\Gamma\mu}\right)^{s-1} \cdot P_0, \quad P_{n,s} = \frac{\lambda}{\Gamma\mu} \left(1 + \frac{\lambda}{\Gamma\mu}\right)^{s-1} \left(\sum_{i=1}^r P_{n-1,i} - \sum_{i=1}^{s-1} \left(\frac{1}{1 + \frac{\lambda}{\Gamma\mu}}\right)^i \cdot P_{n-1,i} \right),$$

$$P_{N,s} = \frac{\lambda}{\Gamma\mu} \cdot \sum_{i=s}^r P_{N-1,i},$$

где $P_{n,s}$ – вероятность того, что в буфере находятся n пакетов и пакет в процессе переключения находится в фазе s . $1 \leq n \leq N$, $1 \leq s \leq r$.

Используя эти результаты, имеем: $P_k^{(L)} = \sum_{s=1}^r P_{k,s}$.

Приоритетные модули памяти

В многопроцессорных системах поток пакетов распределяется между приоритетными модулями памяти. Это происходит вследствие того, что наиболее важные данные находятся на одном или нескольких модулях памяти. В результате происходит насыщение фрагмента многоступенчатой сети и общее падение производительности системы [5]. Для анализа производительности сети при работе с неравномерным трафиком расширим описанную выше модель.

Рассмотрим случай единственного приоритетного модуля памяти MM_h . Методом половинного деления разобьем группу модулей памяти на подгруппы, пока в некоторой подгруппе не останется один модуль памяти MM_h . Процессоры генерируют поток пакетов независимо друг от друга. Некоторая часть h_0 сгенерированных пакетов направляется к приоритетному модулю памяти. Остальные пакеты, равномерно распределены между всеми модулями памяти.

Пусть процессоры генерируют пакеты с интенсивностью λ_0 .

Этап 1. Обозначим через λ_1 – интенсивность потока после прохождения первой ступени. Поток разделяется на два потока: один поток содержит пакеты как равномерно распределенные между всеми модулями памяти, так и предназначенные для приоритетного модуля памяти; второй поток содержит исключительно пакеты равномерно распределенные между всеми модулями памяти. Пусть λ_1^h – интенсивность первого потока, λ_1^r – интенсивность второго потока.

Этап 2. Находим λ_1^h и λ_1^r . $\lambda_1^h = \lambda_1(1 - h_0) + 2h_0\lambda_1$, $\lambda_1^r = \lambda_1(1 - h_0)$. Перепишем выражение для λ_1^h . $\lambda_1^h = \lambda_1^h(1 - h_1) + \lambda_1^h h_1$. Отсюда $h_1 = \frac{\lambda_1^h - \lambda_1(1 - h_0)}{\lambda_1^h}$.

Этап 3. После прохождения второй ступени поток с пакетами для приоритетного модуля памяти расщепляется на два потока: λ_2^h и λ_2^r . Находим λ_2^h и λ_2^r , $\lambda_2^h = \lambda_2(1 - h_1) + 2h_1\lambda_2$, $\lambda_2^r = \lambda_2(1 - h_1)$. Получаем $\lambda_2^h = \lambda_2^h(1 - h_2) + \lambda_2^h h_2$, где $h_2 = \frac{\lambda_2^h - \lambda_2(1 - h_1)}{\lambda_2^h}$.

Повторяем этап 3 до тех пор, пока получим интенсивность потока пакетов на последней ступени сети λ_N^h .

Процедуру выполняем для каждой подгруппы модулей памяти. Пропускная способность сети равняется арифметическому среднему величин λ_N^h , найденных для каждой подгруппы модулей памяти.

Применим данный метод для нахождения пропускной способности сети 8×8 с двумя приоритетными модулями памяти (см. рис. 1).

Разбиваем модули памяти на подгруппы: $M_0 : \{MM_0, MM_1, MM_2, MM_3\}$, $M_1 : \{MM_4\}$, $M_2 : \{MM_5\}$, $M_3 : \{MM_6\}$, $M_4 : \{MM_7\}$. Модули памяти MM_4 и MM_7 из групп M_1 и M_4 являются приоритетными с долей пакетов h_0^4 и h_0^7 соответственно. Оставшаяся часть пакетов $(1 - h_0^4 - h_0^7)$ равномерно распределена между всеми модулями памяти. Необходимо найти λ_3 для каждой подгруппы:

$$1) M_0: \quad \lambda_1^{M_0} = \lambda_1(1 - h_0^4 - h_0^7), \quad \lambda_2^{M_0} = \lambda_2, \quad \lambda_3^{M_0} = \lambda_3.$$

$$2) M_1: \quad \lambda_1^{M_1} = \lambda_1(1 - h_0^4 - h_0^7) + 2h_0^4\lambda_1 + 2h_0^7\lambda_1;$$

$$\lambda_1^{M_1} = \lambda_1^{M_1}(1 - h_1^4 - h_1^7) + \lambda_1^{M_1}h_1^4 + \lambda_1^{M_1}h_1^7,$$

$$\text{где } h_1^4 = \frac{\lambda_1^{M_1} - \lambda_1(1 - h_0^4 - h_0^7) - 2h_0^7\lambda_1}{\lambda_1^{M_1}}, \quad h_1^7 = \frac{\lambda_1^{M_1} - \lambda_1(1 - h_0^4 - h_0^7) - 2h_0^4\lambda_1}{\lambda_1^{M_1}};$$

$$\lambda_2^{M_1} = \lambda_2(1 - h_1^4 - h_1^7) + 2h_1^4\lambda_2; \quad \lambda_2^{M_1} = \lambda_2^{M_1}c_2^{M_1} + \lambda_2^{M_1}h_2^4,$$

$$\text{где } h_2^4 = \frac{\lambda_2^{M_1} - \lambda_2(1 - h_1^4 - h_1^7)}{\lambda_2^{M_1}}, c_2^{M_1} = \frac{\lambda_2^{M_1} - \lambda_2^{M_1} h_2^4}{\lambda_2^{M_1}}; \lambda_3^{M_1} = 2h_2^4 \lambda_3.$$

$$3) M_2: \quad \lambda_3^{M_2} = 2c_2^{M_1} \lambda_3.$$

$$4) M_3: \quad \lambda_2^{M_3} = \lambda_2(1 - h_1^4 - h_1^7) + 2h_1^7 \lambda_2;$$

$$\lambda_2^{M_3} = \lambda_2^{M_3} c_2^{M_3} + \lambda_2^{M_3} h_2^7,$$

$$\text{где } h_2^7 = \frac{\lambda_2^{M_3} - \lambda_2(1 - h_1^4 - h_1^7)}{\lambda_2^{M_3}}, c_2^{M_3} = \frac{\lambda_2^{M_3} - \lambda_2^{M_3} h_2^4}{\lambda_2^{M_3}}; \lambda_3^{M_3} = 2c_2^{M_3} \lambda_3.$$

$$5) M_4: \quad \lambda_3^{M_4} = 2h_2^7 \lambda_3.$$

$$6) \lambda = \frac{4\lambda_3^{M_0} + \lambda_3^{M_1} + \lambda_3^{M_2} + \lambda_3^{M_4}}{8}.$$

Анализ производительности

Используем математическую модель для получения некоторых показателей производительности многоступенчатой коммутирующей сети, а именно: *пропускной способности, времени отклика и системной производительности.*

Влияние размера буфера на пропускную способность сети показано на рис. 2. При размере буфера в две ячейки насыщение сети происходит при входной интенсивности в 0.6. Дальнейшее увеличение интенсивности входного потока не приводит к увеличению интенсивности потока на выходе по причине блокировки пакетов.

При размере буфера в пять ячеек насыщение сети происходит медленнее. Если величина интенсивности входного потока пакетов для проектируемой системы не планируется быть больше 0.5, дальнейшее увеличение размера буфера не приведет к росту пропускной способности сети.

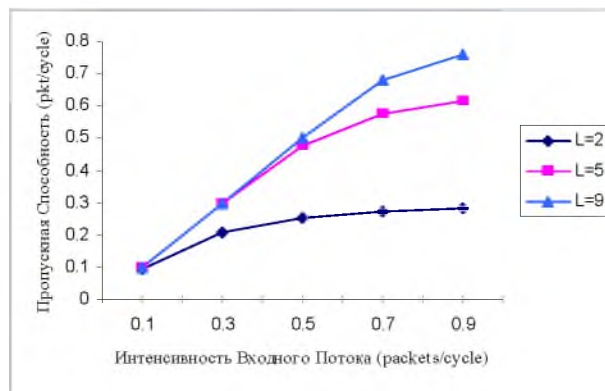


Рис. 2. Пропускная способность многоступенчатой сети 1024×1024

Для системы с размером буфера в девять ячеек увеличение интенсивности входного потока после 0.5 дает лишь незначительный прирост пропускной способности. Вместе с тем, время отклика сети увеличивается значительно (рис. 3).

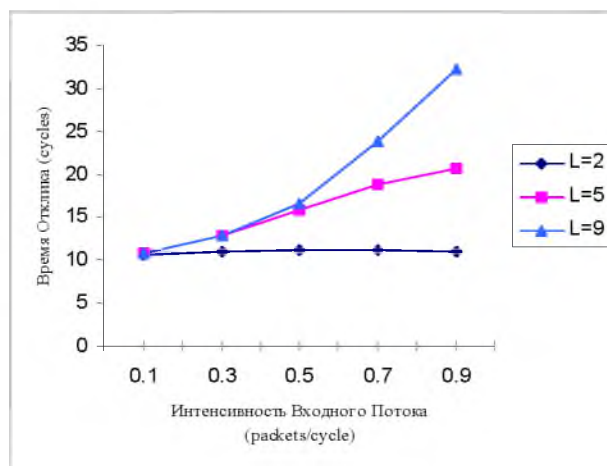


Рис. 3. Время отклика многоступенчатой сети 1024×1024

Отдельно взятые пропускная способность и время отклика не являются достаточными показателями производительности системы. При возрастании пропускной способности и насыщении сети также увеличивается время отклика сети. Определим системную производительность как отношение пропускной способности к времени отклика сети.

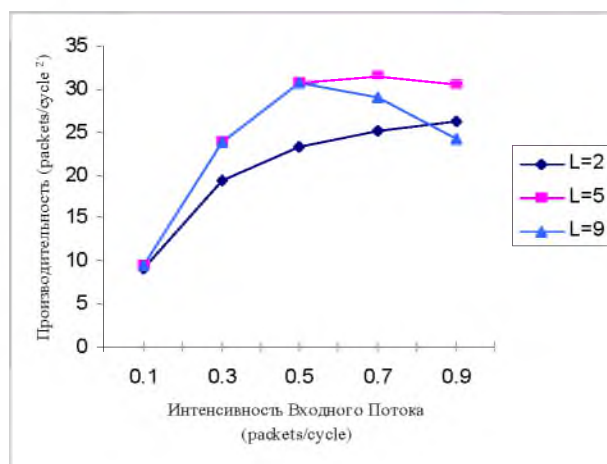


Рис. 4. Производительность многоступенчатой сети 1024×1024

Из графика для системы с буфером на девять ячеек (рис. 4) делаем вывод, что размер буфера в девять ячеек оправдан для систем, ориентированных исключительно на пропускную способность, для которых время отклика не критично.

Выводы

Основные результаты. Разработан программный пакет, который позволяет оценить производительность асинхронной многопроцессорной системы и оптимизировать ее путем подбора оптимальных размеров буферов многоступенчатой коммутационной сети.

Сравнение с лучшими аналогами. Применение реализованной в виде программного пакета математической модели не ограничено случаем единственного приоритетного модуля памяти и детерминированным временем коммутации пакета. Данный программный пакет также применим для случая произвольного числа приоритетных модулей памяти и недетерминированным временем коммутации.

Научная новизна. В данной работе обобщена модель, ранее предложенная в [3]. Впервые получена модель, которая применима к асинхронным многоступенчатым сетям с случайным временем коммутации, распределенным по закону Эрланга, и с произвольным числом приоритетных модулей памяти.

Практическая значимость. Программный комплекс позволяет ускорить разработку MIMD-системы благодаря возможностям расчета ключевых показателей производительности многоступенчатой сети на ранних этапах проектирования. Благодаря предварительной оценке производительности будущей коммутирующей сети, сокращаются затраты на имплементацию пробной версии сети, что уменьшает общую стоимость многопроцессорной системы.

Список литературы: 1. *J. Konicek, T. Tilton, et al, The organization of the Cedar system // Proc. Int' Conf. Parallel Processing, Aug. 1991. P. 49-56.* 2. *Y. Mun and H.Y. Youn. Performance analysis of finite buffered multistage interconnection networks // IEEE Trans. Computers, Feb. 1994. P. 153-162.* 3. *Prasant Mohapatra, Chita R. Das. Performance Analysis of Finite-Buffered Asynchronous Multistage Interconnection Networks, IEEE Transactions on Parallel and Distributed Systems, Vol. 7, No. 1. 1996. P. 18-25.* 4. *El-Paoumy, On Truncated Erlangian Queuing Systems with State-Dependent Service Rate, Balking and Reneging, Applied Mathematical Sciences, Vol. 2. 2008. P. 1161-1167.* 5. *Dickey S.R. and Percus O.E. Performance differences among combining switch architectures, Proc. I&Z, Conf. on Parallel Processing, Aug. 1992. Vol. 1. P. 110-117.*

Поступила в редколлегию 15.12.2008

Евграфов Вячеслав Николаевич, аспирант кафедры ПМ ХНУРЭ. Научные интересы: многопроцессорные системы, многоступенчатые сети. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 8(063)95-82-985.
