

УДК 681.3

**АНАЛИЗ ВЛИЯНИЯ БЫСТРОДЕЙСТВИЯ ПАМЯТИ НА
ЭФФЕКТИВНОСТЬ МНОГОПОТОЧНЫХ СЕТЕВЫХ
МУЛЬТИПРОЦЕССОРОВ С ОДНОКРИСТАЛЬНОЙ
КОНВЕЙЕРНОЙ АРХИТЕКТУРОЙ**

Д.Д. Моргайлов, Моатаз Юнис

Донецкий национальный технический университет
Кафедра прикладной математики и информатики

Описана марковская модель для оценки эффективности однокристалльного многопоточного сетевого мультипроцессора. Получены зависимости, характеризующие влияние быстродействия подсистемы памяти на эффективность сетевого процессора.

Постановка задачи исследования

Главное преимущество однокристалльных многопроцессорных архитектур сетевых мультипроцессоров (МСП)[1] заключается в значительном приросте производительности за счет размещения на кристалле нескольких пакетных процессоров и блоков памяти и формирования таким способом распределенной общей внутренней памяти.

Однако, несмотря на потенциальную масштабируемость и высокую степень параллелизма однокристалльной архитектуры, ее узким местом остается подсистема памяти [2]. Поэтому анализ влияния быстродействия памяти на эффективность сетевого процессора является актуальной научной задачей.

Аналитическая модель однокристалльного МСП

Модель однокристалльного МСП [2, 3] состоит из I процессоров, каждый из которых выполняет J потоков (рис. 1).

Входными данными модели являются следующие параметры:

- интенсивность входного потока процессора (λ , млн. пакет/с);
- интенсивность обслуживания j -го потока (μ_j , млн. пакет/с);

– интенсивность обращения j -го потока с запросами к подсистеме памяти (r_j , млн. запрос/с);

– интенсивность обслуживания запросов подсистемой памяти (m , млн. запрос/с).

Интенсивности обслуживания и обращения к памяти у потоков с одинаковым индексом имеют равные значения для всех процессоров структуры в силу их гомогенного характера.

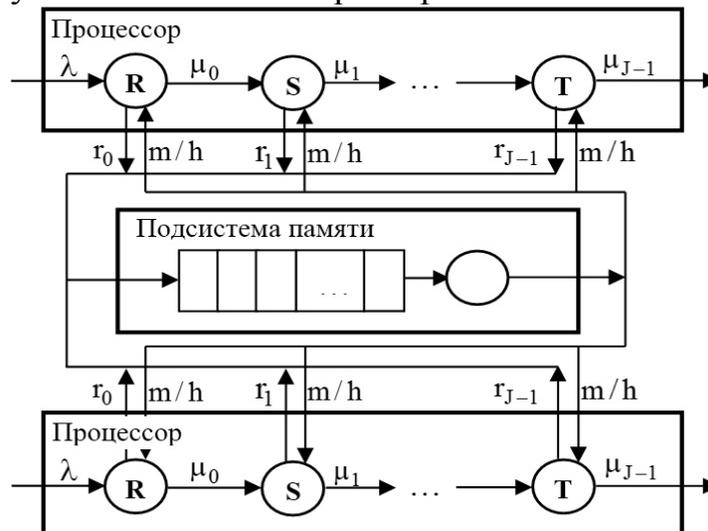


Рисунок 1 – Структура модели на двух процессорах

Пространство состояний модели есть множество:

$$S = (S_{00}, \dots, S_{0j}, \dots, S_{ij}), 0 \leq i < I, 0 \leq j < J. \quad (1)$$

Элемент S_{ij} представляет собой состояние j -го потока на i -м процессоре и может принимать следующие значения (в скобках указаны их коды):

$$S_{ij} \in \{\text{свободен (0), активен (1), запрос-к-памяти (2), готов-к-выполнению (3), завершен (4)}\}. \quad (2)$$

Пусть $S(k)$ – множество элементов, которые в текущий момент времени находятся в состоянии k :

$$S(k) = \{S_{ij} \mid S_{ij} = k, 0 \leq i < I, 0 \leq j < J\}. \quad (3)$$

Представим очередь запросов i -го процессора к подсистеме памяти как множество $h(i)$:

$$h(i) = \{S_{ij} \mid S_{ij} = 2, 0 \leq j < J\}. \quad (4)$$

Тогда длина очереди запросов i -го процессора к памяти в текущем состоянии будет равна мощности этого множества ($|h(i)|$).

Строится и рассчитывается непрерывная марковская модель [2]. Вычисляются показатели эффективности: загрузка пакетных

процессоров и подсистемы памяти, среднее число обращений к памяти, среднее число готовых к выполнению потоков на пакетном процессоре и доля отказов в обслуживании пакетов.

Анализ эффективности однокристалльного МСП

Будем рассматривать модель однокристалльного сетевого мультипроцессора с параметрами, отражающими характер использования его аппаратных ресурсов приложением SimpleForwarding [3].

Исходя из этого, соотношение процессорной нагрузки (P) и нагрузки на подсистему памяти (M) будет представлено следующими значениями: P = 235 такт/пакет, M = 12 запрос/пакет.

Тактовую частоту системы примем за C = 600 МГц, интенсивность поступления пакетов на сетевые интерфейсы считаем равной E = 7300 такт/пакет.

Рассчитаем параметры исследуемой структуры сетевого процессора с двумя пакетными процессорами и двумя потоками (I = 2, J = 2) в соответствии с формулами, предложенными в работе [3]:

$$\lambda = \frac{1}{E} * C = 0,08 \text{ млн. пакет/с}; \quad (5)$$

$$\mu_i : \mu_0 = \mu_1 = C * \frac{J}{P} = 5,11 \text{ млн. пакет/с}; \quad (6)$$

$$r_i : r_0 = r_1 = \mu_i * \frac{M}{J} = 30,66 \text{ млн. запрос/с}. \quad (7)$$

Интенсивность обслуживания запросов подсистемой памяти зависит от типа применяемого запоминающего устройства: со статической (SRAM) либо динамической (SDRAM) памятью (табл. 1).

Таблица 1 – Показатели быстродействия SRAM и SDRAM для сетевого процессора Intel IXP 1200

	SRAM	SDRAM
Скорость обслуживания (L, такт/запрос)	20	40
Интенсивность обслуживания (m = C * 1/L, млн. запрос/с)	30	15

Выполним анализ эффективности сетевого процессора, приняв в качестве критериев степень загрузки пакетных процессоров и

подсистемы памяти, а также долю отказов в обслуживании пакетов. Считаем, что для достижения высокой эффективности необходимо минимизировать вероятность появления отказов в обслуживании и добиться максимальной загрузки процессоров и памяти.

Графики зависимости эффективности сетевого процессора при использовании статической и динамической памяти приведены на рис. 2 и 3.

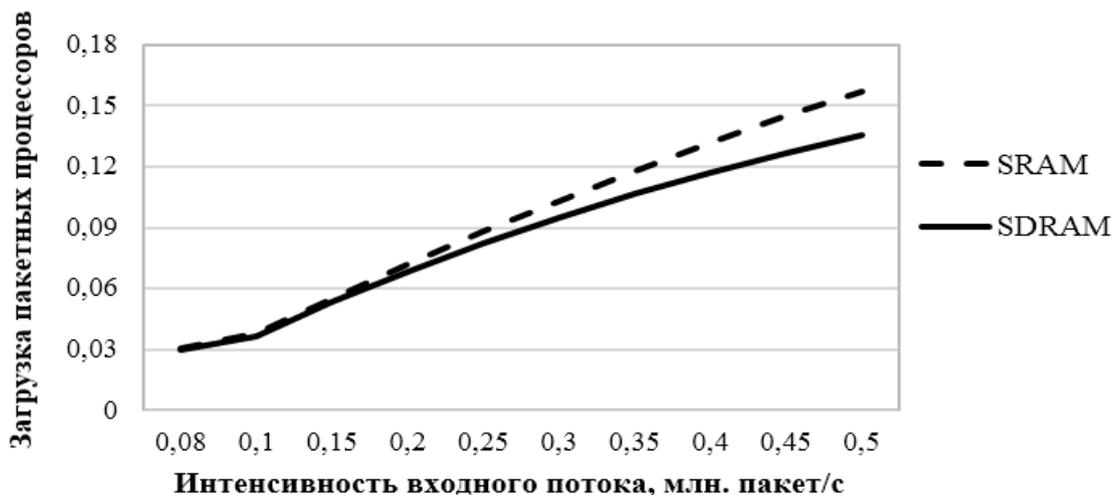


Рисунок 2 – Зависимость загрузки пакетных процессоров от интенсивности входного потока

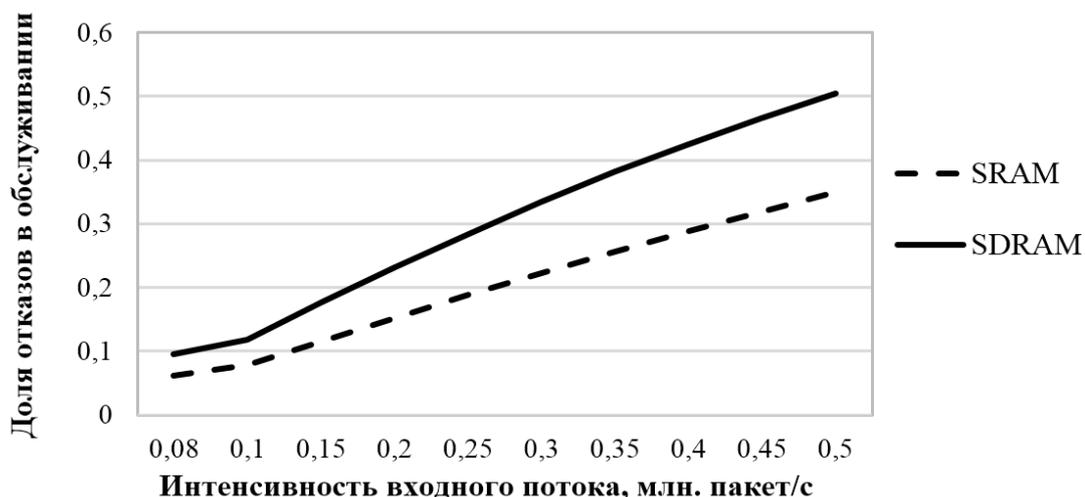


Рисунок 3 – Зависимость доли отказов в обслуживании от интенсивности входного потока

Анализ графиков показывает, что применение статической памяти приводит к заметному снижению доли отказов в обслуживании пакетов и замедлению темпов ее роста при повышении

интенсивности входного потока. Использование быстрой статической памяти, способной обрабатывать 5 запросов за такт ($L = 5$ такт/запрос), позволяет добиться приемлемой (менее 0,1) доли отказов уже при интенсивности входного потока в 0,2 млн. пакет/с.

Кроме того, исследования показали, что увеличение числа потоков в пакетных процессорах обуславливает сокращение накладных расходов на доступ к подсистеме памяти (рис. 4) и, как следствие, снижает долю отказов в обслуживании пакетов. Данный подход демонстрирует возможность повышения эффективности МСП за счет усложнения архитектуры пакетных процессоров.

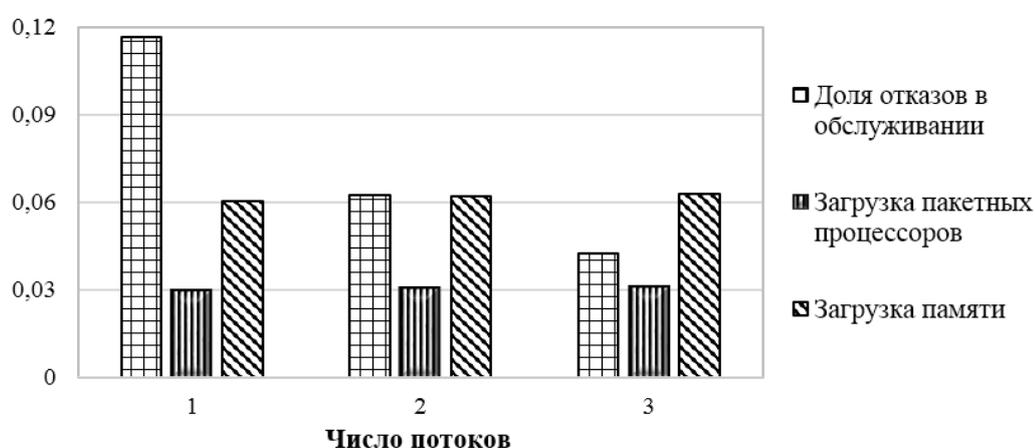


Рисунок 4 – Зависимость показателей эффективности от числа потоков в конвейере

Выводы

Использование статической памяти приводит к повышению пропускной способности однокристалльного МСП. Как следствие, растет порог интенсивности входного потока, при котором доля отказов в обслуживании сохраняется в допустимых пределах.

Увеличение длины конвейера на пакетных процессорах позволит снизить вероятность блокировки первых потоков процессоров. Это приведет к сокращению накладных расходов на доступ к памяти и, таким образом, к уменьшению доли отказов в обслуживании пакетов.

Список литературы

1. Crowley P., Fiuczynski M.E., Baer J.-L., Bershad B.N. Characterizing Processor Architectures for Programmable Network Interfaces / P. Crowley, M.E. Fiuczynski, J.-L. Baer, B.N. Bershad // Proceedings of the 14th international conference on Supercomputing. – 2000. – pp. 54-65.
2. Ладыженский Ю.В. Марковская модель для анализа производительности многопоточных архитектур сетевого мультипроцессора / Ю.В. Ладыженский, Д.Д. Моргайлов // Научные труды Донецкого национального технического университета. Серия: «Информатика, кибернетика и вычислительная техника». – 2013. – № 2. – С. 35-40.
3. Lin Y.-N., Lin Y.-D., Lai Y.-C. Thread Allocation in CMP-based Multithreaded Network Processors / Y.-N. Lin, Y.-D. Lin, Y.-C. Lai // Parallel Computing. – 2010. – vol. 36 (2-3). – pp. 104-116.