

## ПОВЫШЕНИЕ БЫСТРОДЕЙСТВИЯ СУПЕРКОМПЬЮТЕРА ЗА СЧЕТ ОПТИМИЗАЦИИ ИНФОРМАЦИОННОГО МЕЖПРОЦЕССОРНОГО ТРАФИКА

<sup>1</sup>Соломенцев Ю.М., <sup>1</sup>Шептунов С.А., <sup>2</sup>Кабак И.С., <sup>2</sup>Суханова Н.В.

<sup>1</sup>ИКТИ РАН, <sup>2</sup>МГТУ «СТАНКИН»

ship@ikti.ru

Прогнозы в области развития суперкомпьютерной техники показывают, что в течение ближайших 10 лет производительность суперкомпьютеров возрастает в среднем в 1000 раз. Решениями руководства страны планируются научно-технические работы, целью которых будет достижение производительности 1 эксафлопс ( $10^{18}$  оп/с) к 2018–2020 гг. Как ожидается, системы эксафлопсной производительности будут содержать до  $10^8$  процессорных ядер, что потребует решений, связанных с их энергопотреблением, надежностью и конструктивными размерами [1].

**Ключевые слова:** суперкомпьютеры, искусственные нейронные сети.

## INCREASE OF SPEED OF A SUPERCOMPUTER AT THE EXPENSE OF OPTIMIZATION OF THE INFORMATION INTERPROCESSOR TRAFFIC

<sup>1</sup>Solomentsev U.M., <sup>1</sup>Sheptunov S.A., <sup>2</sup>Kabak I.S., <sup>2</sup>Sukhanova N.V.

<sup>1</sup>IKTI RAS, <sup>2</sup>MSTU «STANKIN»

Forecasts in the field of development of supercomputer technics show that within the next 10 years productivity of supercomputers increases on the average in 1000 times. Country leaders decisions plan the scientific and technical works achievement of productivity of 1 Exaflops ( $10^{18}$  op/c) to 2018–2020 гг. will be which purpose. Predictably, systems exaflop productivity will contain to  $10^8$  processor kernels that will demand the decisions connected by their power consumption, reliability and the constructive sizes.

**Keywords:** supercomputer, artificial neurons networks.

Согласно оценкам ведущих зарубежных экспертов [2, 3], при использовании имеющихся в настоящее время технологий суперкомпьютер производительностью 1 эксафлоп/с (появление которых прогнозируется к 2020 г.) будет иметь до  $10^8$  процессорных ядер и потреблять до 100 МВт.

Использование такого количества процессоров в суперкомпьютерной системе ставит вопрос о передаче информации от одного процессора другим (при решении несколькими процессорами одной задачи). Современные технологии передачи информации позволяют передавать до 10–30 Гбит/сек.

Объем передаваемой по сети информации  $S$  бит в секунду можно оценить по формуле:  $S = S_c \times n \times P_c \times V$ , где  $S_c$  – среднее количество обменов информацией от одного процессорного блока другому за одну операцию;  $n$  – количество процессорных блоков в системе;  $P_c$  – производительность процессорного блока в операциях в секунду;  $V$  – размер передаваемого пакета в битах – средний объем информации, передаваемой от одного процессора другому.

Если для приблизительной оценки принять значения  $S_c = 0,01$ ;  $n = 10^8$ ;  $P_c = 10^{18}$  оп/с;  $V = 100$ , то объем передаваемой по сети информации  $S$  будет около  $10^{36}$  битов. Сравним с предполагаемой производительностью сети 10–30 Гбит в секунду и убедимся, что топологии сети, обеспечивающие последовательную передачу информации по кабелю (топологии типа «шина», «звезда» или «кольцо»), не подходят. Ячеистая топология также не подходит.

Рассмотрим суперкомпьютер с гиперкубической или аналогичной топологией межпроцессорной сети, способной обеспечить передачу информации между процессорными модулями. Количество линий связи между процессорными элементами возрастает, что позволяет разделить и распараллелить потоки межпроцессорной информации. Суммарная производительность межпроцессорной сети существенно возрастет.

Путь передачи информации от одного процессорного элемента другому может проходить через промежуточный процессорный элемент. Процессорный элемент включает не только собственно центральный процессор и локальную память, но и транспортный (сетевой) процессор, отвечающий за обмен информацией по сети (рис. 1).

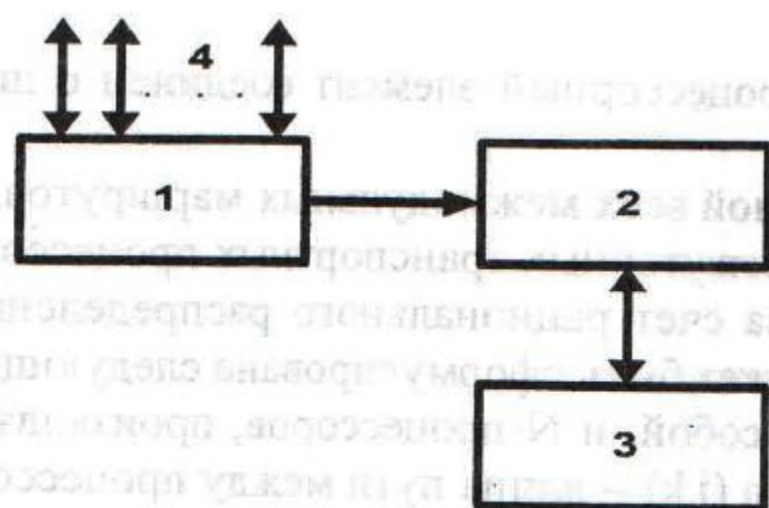


Рис. 1. Архитектура процессорного элемента суперкомпьютера:

- 1 – транспортный процессор;
- 2 – центральный процессор;
- 3 – локальная память; 4 – линии связи (сеть)

Транспортный процессор задействован как в передачах собственной информации (отправителем или получателем является данный элемент), так и в передаче промежуточной информации (ни отправителем, ни получателем не является). Чем меньше количество промежуточных процессорных элементов, тем выше эффективность информационной межпроцессорной сети.

При работе процессорных элементов с общей для них памятью (рис. 2) блок оперативной памяти можно рассматривать с точки зрения межпроцессорной сети как некоторый процессорный элемент, который только хранит, но не обрабатывает данные.

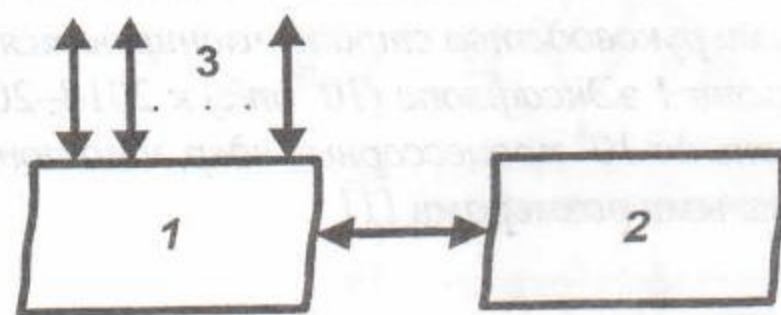


Рис. 2. Архитектура элемента оперативной памяти суперкомпьютера:  
1 – транспортный процессор; 2 – устройство (модуль) оперативной памяти; 3 – сеть

Длина маршрута межпроцессорного информационного потока зависит от количества промежуточных процессорных элементов. Загрузка каждого сетевого процессора процессорного элемента зависит от количества проходящих через него маршрутов.

Будем считать длиной маршрута количество промежуточных процессорных элементов, через которые он проходит. Можно поставить несколько оптимизационных задач, позволяющих снизить нагрузку на процессорные элементы и повысить быстродействие суперкомпьютера.

Первая оптимизационная задача – снизить нагрузку на транспортный процессор каждого процессорного элемента и обеспечить уровень нагрузки, не превышающий заданный.

Вторая оптимизационная задача – минимизировать длину пути каждого маршрута.

Величина транспортного информационного потока определяется решаемой на суперкомпьютере задачей, степенью распараллеливания вычислений и рядом других, менее значимых параметров. Для примера рассмотрим кубическую (трехмерную) транспортную систему, состоящую из  $N = i \times j \times k$  процессорных элементов.

Задача, решаемая на компьютере, состоит из ряда параллельно выполняемых нитей. Нить загружается на отдельный процессор и им выполняется. Для совместной работы нити должны обмениваться информацией между собой, а также с оперативной памятью. Обмен происходит периодически. В терминах процессорных элементов и транспортной сети это означает, что процессорные элементы обмениваются информацией друг с другом. Каждый такой обмен представляет собой маршрут в системе и проходит через ряд промежуточных процессорных элементов.

Для определения таких маршрутов в суперкомпьютерной системе имеется специальное устройство. Это устройство называется маршрутизатором и может быть выполнено как отдельное аппаратное устройство или как исполняемая одним или несколькими процессорными элементами программа.

Функция маршрутизатора – определение траектории передачи информации от одного блока другому. Поставленные оптимизационные задачи должны решаться именно в этом блоке.

Исходя из предположения, что «узким местом» в суперкомпьютере является транспортная система, снизим нагрузку транспортных процессоров процессорных блоков. Для этого используем решения, реализованные в протоколе «АТМ-статические маршруты передачи информации».

В процессе маршрутизации определяется маршрут (последовательность) процессорных блоков, через которые будет происходить передача информации. Чем меньше количество промежуточных элементов, тем быстрее будет передаваться информация и меньше использоваться совокупная транспортная система. В этом и заключается вторая оптимизационная задача.

Транспортный процессор получает от маршрутизатора таблицу коммутации, где для каждого маршрута указан следующий процессорный блок. Функции транспортного процессора сводятся к проверке номера маршрута и определению следующего промежуточного транспортного процессора. При гиперкубической топологии соединения процессорный элемент соединен со всеми соседними процессорными элементами. Количество соседних процессорных элементов  $n_c$  (и, соответственно, количество сетевых входов/выходов модуля) можно определить по формуле:

$$n_c = 2xp,$$

где  $p$  – размерность системы.

Для кубического случая, когда  $p$  равно трем, каждый процессорный элемент соединен с шестью другими элементами.

Эффективность транспортной системы определяется суммой всех межмодульных маршрутов, в то время как длина каждого маршрута напрямую зависит от количества промежуточных транспортных процессоров. Эффективность транспортной системы можно существенно увеличить за счет рационального распределения процессоров между процессорными модулями. Математически эта задача может быть сформулирована следующим образом:

Имеется  $J$  процессов, информационно связанных между собой, и  $N$  процессоров, производящих обработку этих процессов. Имеется матрица  $M$ , каждый элемент которой  $m(i,k)$  – длина пути между процессорными элементами в гиперкубической транспортной системе – определяется формулой:

$$\sum_{i=1}^K (l_j^i + l_k^i)$$

где  $K$  – количество координат гиперкуба транспортной системы;  $l_j^i$  и  $l_k^i$  определяют номера начального ( $j$ ) и конечного ( $k$ ) процессорных блоков маршрута по  $i$ -ой координате гиперкуба.

Решение этой задачи возможно и использованием оптимизации на графах или другим способом, как, например, в [4]. Но заметим, что решение задачи с использованием оптимизации на графах требует существенного объема вычислений, как, впрочем, и решение, предложенное в [4], хотя это решение требует значительно меньше ресурсов.

В суперкомпьютерной системе процессы создаются и уничтожаются динамически. Распределение процессорных элементов осуществляется программным обеспечением, операционной системой, и длительный расчет распределения процессоров может крайне негативно сказаться на производительности суперкомпьютерной системы в целом.

Решением данной задачи могло бы стать использование специального высокоскоростного аппаратного блока (маршрутизатора) в составе суперкомпьютера, построенного по технологии, описанной в работах [5–7].

#### Выводы

1. Показано, что быстродействие суперкомпьютерной системы можно увеличить за счет оптимизации трафика информации между процессорами этой системы.
2. Решение задачи оптимизации трафика информации возможно на базе теории графов или эвристическими методами.
3. Существенно более эффективным решением является применение специального аппаратного блока.

#### Библиография

1. Протокол заседания комиссии при Президенте Российской Федерации по модернизации и технологическому развитию экономики России от 18 июня 2009 г. № 1 [Электронный ресурс]. – Режим доступа: <http://www.i-russia.ru/sessions/decisions/24.html>.
2. Dongarra J. Impact of architecture and technology for extreme scale on software and algorithm design. Cross-cutting technologies for computing at the exascale. February 2–5, 2010. – Washington, D.C.
3. Benton V. Exascale Evolution. IBM. March 15, 2010.
4. Степанов С.Ю., Кабак И.С. Алгоритм фрагментации больших нейронных сетей и исследование его сходимости // Информационные технологии. – 2012. – № 7. – С. 73–78.
5. Кабак И.С., Суханова Н.В. Моделирование надежности программного обеспечения систем управления автоматизированными технологическими комплексами на базе искусственного интеллекта // Вестник МГТУ «Станкин». – 2012. – № 1 (19). – С. 95–99.
6. Кабак И.С., Суханова Н.В. Технология реализации автоматизированных систем управления на базе больших искусственных нейронных сетей МОДУС-НС // Межотраслевая информационная служба. – 2012.
7. Кабак И.С. Создание больших аппаратно-программных нейронных сетей для систем управления // Авиационная промышленность. – 2012. – № 4.