

# СПЕЦИАЛИЗАЦИЯ КАК ПРОЦЕСС ДОСТИЖЕНИЯ РЕЗУЛЬТАТИВНОСТИ ИСПОЛЬЗОВАНИЯ СУПЕРКОМПЬЮТЕРОВ

В.С. Горбунов

ФГУП «НИИ «Квант»

## 1. Введение

Особенность суперкомпьютеров определяется их предназначением для решения таких задач, которые не могут быть решены на других вычислительных средствах по причине большого объема вычислений и/или большого объема обрабатываемых данных [1]. Такая исключительность предназначения связана с оптимизацией основной характеристики суперкомпьютеров – времени решения выбранной вычислительно сложной задачи или задач, что зависит от множества факторов.

Будущий владелец суперкомпьютера хотел бы чтобы его машина работала как можно быстрее и стоила меньше, требовала меньших эксплуатационных затрат (мощность, помещение, обслуживающий персонал и т.д.). В этой непростой ситуации будущему владельцу суперкомпьютера или разработчику такой машины приходится решать сложную оптимизационную задачу, пытаясь определить именно ту конфигурацию, которая наилучшим образом подходила бы для определенного набора задач и при этом помещалась в разумные ресурсные ограничения. Этот процесс назовем процессом специализации суперкомпьютера или суперкомпьютерного центра. Специализация суперкомпьютерного центра достигается объединением специализированных и проблемно-ориентированных суперкомпьютеров в гетерогенную вычислительную структуру. Эта структура состоит из различного типа вычислительных сегментов (суперкомпьютеров), ориентированных на заданные классы применений. При этом степень специализации таких фрагментов может быть разной, от «мягкой» специализации за счет подбора ключевых компонентов сегмента суперкомпьютера (память, коммуникации, процессоры, ускорители и пр.) до полной специализации, позволяющей, по сути, решать только одну или несколько задач [2].

Перспектива развития суперкомпьютерной отрасли достаточно уникальна и аспект специализации суперкомпьютеров, как и суперкомпьютерных центров на их основе, является важным для изучения и исследования. Действительно, в ближайшей перспективе качественно изменятся требования по функциональности и энергопотреблению будущих суперкомпьютеров, произойдет смена поколений элементно-конструкторской базы. На этом фоне значительно меняются архитектура и микроархитектура основных компонентов аппаратуры суперкомпьютеров, формируются новые сложные иерархические структуры их организации, значительно видоизменяются реализуемые ими модели вычислений и методы программирования, изменяются даже подходы к эксплуатации. Этот процесс значительно затруднен еще и тем, что особенностью современного этапа развития суперкомпьютеров является увеличение их сложности. Сложность суперкомпьютеров растет как в качественном, так и в количественном отношении. Например, в ближайшей перспективе суперкомпьютеры будут поддерживать выполнение порядка  $10^6$  –  $10^7$  параллельно функционирующих процессов, а для эксафлопсных суперкомпьютеров число процессов приблизится к  $10^9$ , для зеттафлопсных суперкомпьютеров количество параллельных процессоров может быть на уровне  $10^{12}$ . По всей вероятности изменятся и численные методы решения прикладных задач. На основе изложенного выше можно говорить о том, что в ближайшей перспективе, как никогда ранее, усиливается роль специализации суперкомпьютерных установок.

Обычно специализация суперкомпьютера – это сложный одновременно аналитический и творчески созидательный процесс “проб и ошибок”, связанный с глубоким проникновением в суть решаемых задач, а именно – изучением специфики выполняемых в них операций и организации управления выполнением этих операций, изучением возможностей элементно-конструктивной и микроэлектронной базы, выбором адекватных аппаратных решений и их оценкой. Уровень сложности современных и будущих суперкомпьютеров уже не позволяет прогнозировать их возможности и характеристики, путем простого экстраполирования, как нередко это делалось ранее, на базе полученных экспериментально данных на системах меньшей сложности.

Учитывая перечисленные проблемы, исследования и разработки по специализации суперкомпьютеров и суперкомпьютерных центров могли бы проводиться в соответствии с комплексным подходом, схема которого приведена на рис. 1. Он заключается в 4-х основных видах исследований, а именно: исследование свойств прикладных задач и их модификаций, тестирование и изучение особенности работы существующих вычислительных систем и их компонентов («Тестирование», рис.1), разработка новых архитектурных решений («Новые архитектура и структура», рис.1); создание и исследование аналитических моделей выполнения приложений («Задачи и модель производительности», рис.1); разработка и исследования имитационных и эмуляционных моделей («Имитационные и эмуляционные модели», рис.1). Для проведения этих исследований

необходимо создание экспериментальных исследовательских установок для адекватной оценки множества решений по разработке и применению будущих суперкомпьютеров.

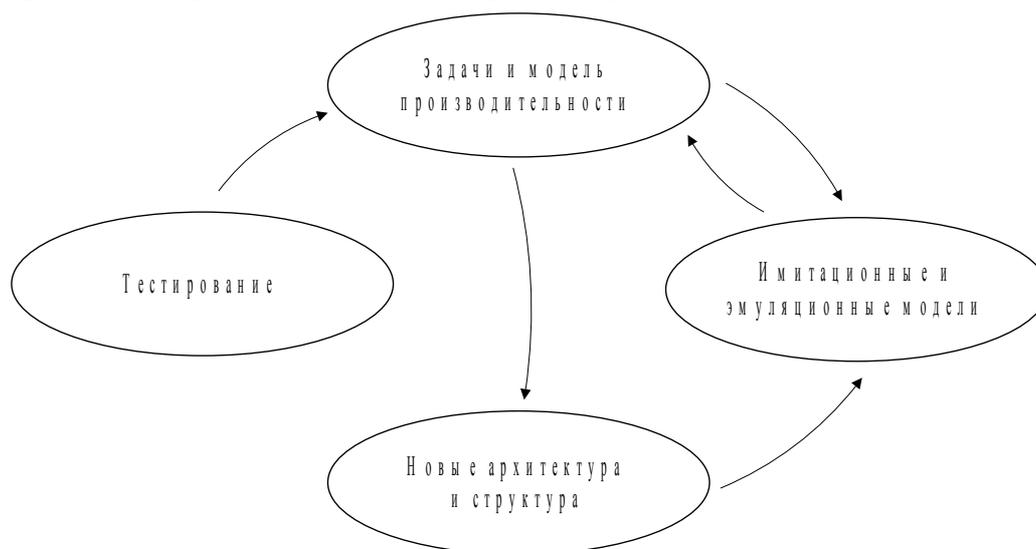


Рис. 1. Взаимосвязь основных направлений исследований

Рассмотрим далее особенности этих исследований и предложения по инструментальной установке, позволяющей проводить сложные эксперименты с применением имитационных и эмуляционных моделей суперкомпьютеров.

## 2. Исследование вычислительных задач и модель производительности суперкомпьютера

Для оценки времени решения суперкомпьютером задач часто используют аналитические модели, которые называют моделями производительности. Модель производительности строится на основе понимания того, как работает приложение на существующих суперкомпьютерных установках, изучения системных характеристик и особенностей этих установок. Процесс подготовки и построения такой модели, её верификации, был рассмотрен в журнале «Суперкомпьютеры» [3].

Прежде всего, определяются ключевые свойства приложения, от которых в наибольшей степени зависит время выполнения этого приложения на исследуемой установке. Важными свойствами обычно считают: распределение данных между параллельными процессами; объем памяти для каждого процесса; соотношение вычислительных операций и операций с памятью в каждом процессе; пространственно-временную локализацию работы с памятью; количество и тип обменов данными в расчете на один параллельный процесс; схема обменов данными между процессами и объемы пересылаемых данных, длины сообщений. Для рассматриваемого приложения учитываются только свойства первого порядка, те, которые самым непосредственным образом влияют на реальную производительность суперкомпьютера. Для определения свойств уже написанных приложений выполняется процедура профилирования, а для задач, приложения к которым еще не разработаны, можно построить модели их поведения на параллельной вычислительной системе.

Исследование работы существующих суперкомпьютеров важно так же и при разработке новых или адаптации существующих численных методов решения прикладных задач. Это позволяет выполнить адаптацию алгоритмов, или найти новые, иногда неожиданные, методы и решения.

Часто считается, что для построения модели производительности суперкомпьютера, который появится в ближайшие 10-15 лет, нужно соответственным образом увеличить системные характеристики существующих его частей и получить рассмотренную выше «упрощенную» модель производительности. Так, например, в Лос-Аламосе это выполняют для архитектуры базового суперкомпьютера BlueGene. [4]. В частности, ими рассматриваются варианты изменения трех основных характеристик: увеличение скорости обменов на всех уровнях иерархии коммуникаций, уменьшение задержки межузловой коммуникационной сети и увеличения количества ядер в процессорах.

Однако при оценке перспектив на 10-15 лет этот способ может быть весьма неточен. Причиной тому, является ранее отмеченное ожидаемое значительное изменение облика суперкомпьютера будущего. Для оценок производительности суперкомпьютеров с прицелом на 10-15 лет вперед необходимы дополнительные усилия и необычные вычислительные ресурсы. Эти вычислительные ресурсы должны позволять эмулировать те функции будущих машин, которые в настоящее время еще не реализованы. Таким инструментом может быть специально организованный массив программируемых логических интегральных микросхем (ПЛИС), которые позволят на аппаратном уровне создавать элементы будущих функциональных устройств. Предложения по этому инструменту будут изложены в разделе 4 настоящей статьи.

### 3. Комплексная методика тестирования суперкомпьютера

Для определения возможностей реальных суперкомпьютерных установок, что, например, в контексте данной работы необходимо для получения исходных данных для построения моделей задач и производительности, исследуются их свойства путем выполнения серии тестов. Определяются скорости выполнения операций в различных режимах, скорости работы подсистемы памяти на разных уровнях иерархии памяти. Определяются характеристики обмена данными между вычислительными узлами. Для примера, в Лос-Аламосской национальной лаборатории США, в качестве тестов используются несколько модельных задач разных предметных областей. Представляется, что такой подход недостаточен. Нами предложен подход, основанный на многоуровневой методике, базирующейся на углубленном исследовании памяти, функциональных устройств и коммуникационной подсистемы суперкомпьютерных установок [5].

В основу этой методики положен принцип оценки «снизу вверх», от основных элементов до суперкомпьютера в целом. Методика отображает и степень детализации, с которой требуется получить знания о тестируемой системе – на нижних уровнях находятся тесты компонентов системы, а на верхних – пользовательские программы. Причем начиная с четвертого уровня для тестирования части реальных задач, что достаточно важно. На этих уровнях происходит поэтапное приближение к реальным приложениям от базовых алгоритмов (матричные операции, операции редукции и пр.) до реальных расчетных программ, оформленных в виде тестов. Уровни тестирования можно выбирать в зависимости от целей тестирования. Детализация тестовой нагрузки в методике делается целенаправленно, со знанием ожидаемых особенностей работы оборудования.

*Первый уровень* – оценка подсистемы памяти, в различных режимах её функционирования.

*Второй уровень* – граничные тесты, соответствующие четырем предельным значениям пространственно-временной локализации обращений к памяти: Linpack (хорошая пространственная и временная локализация), Random Access (одновременно плохая пространственная и временная), PTRANS и TRIAD (плохая временная, хорошая пространственная), FFT (хорошая временная, плохая пространственная локализация).

*Третий уровень* – специально подобранные тесты процессорных функциональных устройств выполнения арифметико-логических операций и операций с памятью, внутренней и внешней сети вычислительных узлов, системы ввода-вывода.

*Четвертый уровень* – общие и специальные базовые алгоритмы прикладных программ.

*Пятый уровень* – ядра разных приложений, а именно: научные расчеты (линейная алгебра и аэрогидродинамика); сжатие текстов и изображений; верификационный анализ программного обеспечения; обработка текстов и изображений; задачи оптимизации и поиска; задачи на высокорегулярных структурах (сеточные методы и клеточные алгоритмы); задачи на деревьях и графах; тесты разных моделей организации параллельных программ.

*Шестой уровень* – тесты модельных приложений, а также операционной системы и системы планирования прохождения заданий.

*Седьмой уровень* – тесты в виде прикладных программ, немного упрощенных для проведения тестирования.

Таким образом, оценочное тестирование позволяет получить разноплановое знание о возможностях улучшения исследуемого оборудования, которые далее обобщаются в моделях производительности, что позволяет определиться с множеством исследуемых далее вариантов новых архитектур.

### 4. Моделирование новых архитектур и инструментальная система для проведения экспериментов

Рассмотрим два важных направления исследований для будущих суперкомпьютеров, выделенных по архитектурно-программным принципам. Эти два направления представляются наиболее характерными и значимыми для развития суперкомпьютеров в ближайшем будущем. Каждый из этих подходов имеет свой уровень специализации, поэтому они взаимно дополняют друг друга, не являясь альтернативными.

Первое направление эксафлопсного суперкомпьютера, «В-1», ориентировано на использование однородных “легких” ядер и глобально-адресуемой памяти, причем реализуются специальные теги управления для доступа к ячейкам этой памяти. Примером такой архитектуры являются суперкомпьютеры BlueGene (ф. IBM). Эта модель пригодна также для моделирования многоядерных массово-мультитредовых систем типа Cray XMT (ф. Cray Inc.). В этой модели существенна реализация быстрого переключения программных потоков (тредов) при выполнении операций с локальной и глобально-адресуемой памятью, мелкозернистая синхронизация параллельных процессов на тегированных ячейках памяти.

Второе направление «В-2», реализует эксафлопсный суперкомпьютер с применением идеологии гибридных мультитредовых процессоров и асинхронных схем. Такие процессоры включают “легкие” мультитредовые ядра для доступа к данным, а также “тяжелые” ядра. “Тяжелые” ядра предназначены для выполнения интенсивных вычислений в виде мультитредовых ядер с большим набором функциональных устройств. В ядрах параллельно запускается на выполнение, причем это делается за такт работы процессора, множество операций. В качестве “тяжелых” ядер могут быть настраиваемые решающие поля, подобия клеточных автоматов, специализированные блоки, непосредственно реализующие тот или иной алгоритм. В качестве примера такого суперкомпьютера можно считать современные суперкомпьютеры с мощными

гибридными вычислительными узлами, где в качестве “легких” ядер можно использовать суперскалярные процессоры общего назначения (CPU), а в качестве “тяжелых” графические массово-мультитредовые ускорители (GPGPU) и/или ускорители на ПЛИС (FPGA).

Для проведения масштабных экспериментов по моделированию экзафлопсных суперкомпьютерных установок будущего необходим соответствующий инструментарий. С целью создания такого инструмента Российской академией наук развернут проект построения вычислительной системы на ПЛИС (моделирующей гетерогенной вычислительной системы – МГВС) [6].

Основная идея МГВС заключается в следующем [7]. Поскольку совершенно точно известно, что на выпускаемой в настоящее время электронно-компонентной базе построить суперкомпьютер будущего с требуемыми показателями удельной эффективности (50 Гфлопс/Ватт) будет невозможно, то придется проводить глубокие исследования по аппаратно-программным принципам создания такого суперкомпьютера. Одна из главных целей – поддержание в таком суперкомпьютере одновременного выполнения уровня  $10^9$  параллельных процессов. Инструментальные установки для экспериментов с моделями такого уровня параллелизма должны быть сами очень мощными. Для этого можно использовать большой массив ПЛИС (программируемых логических микросхем, выполняющих аппаратную эмуляцию элементов вычислителя). Эти ПЛИС можно использовать при создании действующей модели суперкомпьютера тремя основными способами:

- для реализации легких вычислительных ядер, то есть «правильных», с точки зрения использования в составе суперкомпьютера, процессоров, логически похожих на процессоры общего назначения;
- для реализации тяжелых программируемых вычислителей (“тяжелых” ядер) в том числе с нетрадиционной архитектурой, логически похожих на сегодняшние GPGPU, но только со сложной иерархической структурой;
- для прямой схемной реализации вычислений, то есть в качестве специализированных ускорителей, реконфигурируемых непосредственно под набор методов решения прикладных задач.

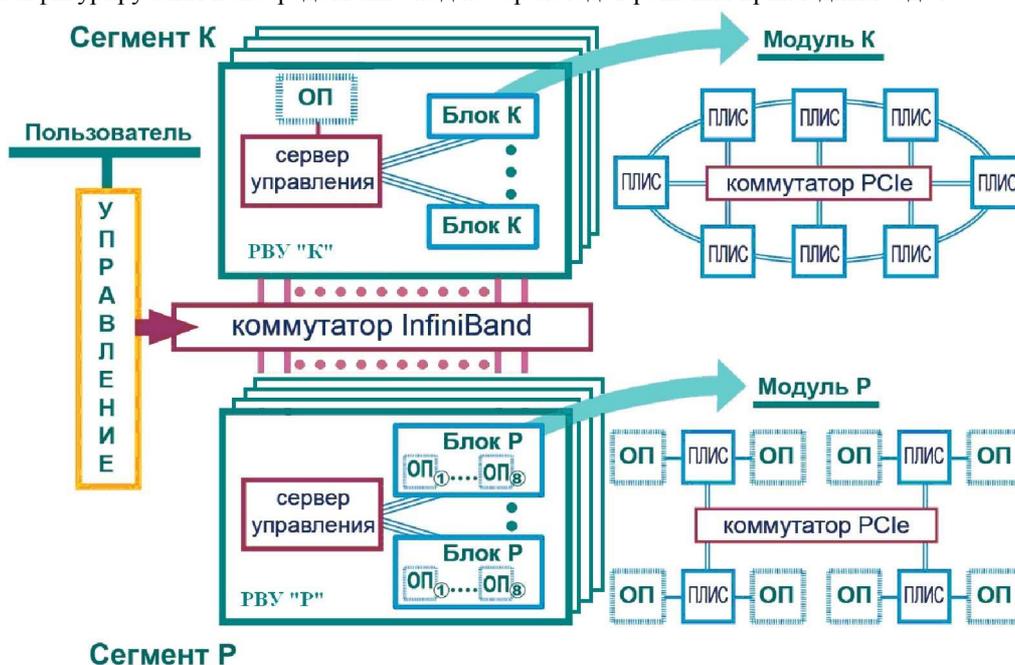


Рис. 2 Структура МГВС

МГВС представляет собой гибридно-гетерогенную вычислительную систему, в которую входят сегменты двух типов с множеством реконфигурируемых вычислительных узлов (РВУ, см. рис. 2). Отличие этих сегментов заключается в особенностях подключения основной памяти и особенностях используемых ПЛИС. В сегменте первого типа (сегмент Р) память реализуется на микросхемах, подключенных непосредственно к ПЛИС. За счет этого можно имитировать традиционную память процессоров общего назначения. В сегменте второго типа (сегмент К) память подключается к процессорам общего назначения серверов управления РВУ. Всего в системе могут использоваться сотни РВУ и десятки тысяч ПЛИС разного типа. Из зарубежных установок на МГВС больше похожи проекты DEEP [8] и RAMP Blue [9], но МГВС отличается большим размером и специальным образом подобранной неоднородностью сегментов. Такая неоднородность позволяет моделировать на МГВС существенно большее количество различных вариантов суперкомпьютеров.

Первый сегмент (сегмент Р) в большей степени ориентирован на построение эмулятора суперкомпьютера, создаваемого в рамках идеологии использования “легких” однородных ядер - «В-1».

Второй сегмент (сегмент К) ориентируется, в основном, на реализацию суперкомпьютера, создаваемого по архитектурной идеологии гибридных многопоточных процессоров «В-2». Ядра для “тяжелых” вычислений

будут эмулироваться посредством мощных конвейерных устройств и решающих полей, реализуемых в ПЛИС этого сегмента.

Предусматривается и совместное использование сегментов МГВС при моделировании, например, гетерогенных суперкомпьютеров. По направлению "В-2" РВУ сегмента Р могли бы эмулировать работу модулей интеллектуальной памяти со встроенными в нее процессорами непосредственной "сверхлегкой" обработки данных. МГВС планируется запустить в 2014 году.

## 5. Заключение

Ближайшая задача, которую нужно решить в направлении подготовки к использованию, выбору и созданию ориентированных на конкретные научно-технические расчеты суперкомпьютеров, - это активизация исследований в области совершенствования алгоритмов, системного программного обеспечения, освоения возможностей новых технологий и выработка на основе этого эффективных архитектурных и технических решений. Эти исследования невозможны без исследований существующих приложений (профилирование), глубокого тестирования существующих систем (многоуровневое тестирование), разработки новых архитектур и структур суперкомпьютеров (направления «В-1» и «В-2»), построения масштабных имитационных и эмуляционных моделей. Таким образом, в процессе специализации суперкомпьютеров, под которой понимается настройка их архитектуры на конкретные научно-технические расчеты, способствуют работы по тестированию, исследованию вычислительных свойств задач и построению моделей суперкомпьютеров с использованием моделирующей вычислительной системы МГВС.

## ЛИТЕРАТУРА:

1. В.К. Левин, В.С. Горбунов, С.В. Яблонский "Суперкомпьютеры сегодня и завтра" // Материалы Международной научно-технической конференции «Суперкомпьютерные технологии: разработка, программирование, применение (СКТ-2010)», с.Дивноморское, Геленджик, 27 сентября - 2 октября 2010 г., Изд-во ТТИ ЮФУ, Том 1, с. 57-62
2. Интернет ресурс: [techno.co.jp/english/NEW\\_EN/20100708/18477](http://techno.co.jp/english/NEW_EN/20100708/18477)
3. В.С. Горбунов "НПС Оценить, измерить, оптимизировать" // Суперкомпьютеры. №3 (15), ООО «Издательство СКР-Медиа», осень-2013, ст. 46-51
4. K.J. Barker, K. Davis, A. Hoisie, D.J. Kerbyson, M. Lang, S. Pakin, J. Sancho "Using performance modeling to design large-scale systems" // Computer, vol 42, no. 11, Published by the IEEE Computer Society, 2009, pp. 42-46
5. В.С. Горбунов, Л.К. Эйсымонт "Комплексная методика тестирования производительности суперкомпьютеров" // Вычислительные методы и программирование, Том 14, Изд-во Московского Университета, 2013, ст. 115- 121
6. Интернет ресурс: [keldysh.ru/exaflops.pdf](http://keldysh.ru/exaflops.pdf)
7. В. Горбунов, С. Елизаров, В. Корнеев, А. Лацис "Футурология суперкомпьютеров" // Суперкомпьютеры №1 (17), ООО «Издательство СКР-Медиа», весна-2014., ст. 46-51
8. J. Ributzka, Y. Hayashi, F. Chen, G.R. Gao "DEEP: An Interactive FPGA-based Many-Core Emulation System for Chip Verification and Architecture Research" // In Proceeding 19<sup>th</sup> ACM/SIGMA International Symposium on Field-Programmable Gate Arrays (FPGA'11), Monterey, CA, USA, February 27- march 1, 2011
9. D. Burke, J. Wawrzynek, K. Asanovic, A. Krasnov, A. Schultz, G. Gibeling, P.Y. Droz "RAMP Blue: Implementation of a Manycore 1008 Processor FPGA System" // Proceeding. Reconfigurable Systems Summer Institute 2008 (RSSI 08), July 2008, pp. 54-61