

СПЕЦИАЛИЗИРОВАННЫЕ КОМПАКТНЫЕ ВЫСОКОПРОИЗВОДИТЕЛЬНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ

*А. С. Рыбкин, Т. А. Агапова, И. А. Крючков, Ю. В. Логвин, А. Г. Ломтев,
Р. М. Шагалиев, В. В. Южаков*

ФГУП «РФЯЦ-ВНИИЭФ», г. Саров Нижегородской обл.

Введение

На текущий момент в мире уже созданы и эксплуатируются высокопроизводительные гибридные вычислительные системы с арифметическими ускорителями на базе графических процессоров. В РФЯЦ-ВНИИЭФ ведутся интенсивные исследования по применению арифметических ускорителей для научных расчетов. Работы выполняются с использованием:

– NVIDIA CUDA на графических процессорах G80, GT200, GF100/GF110, специализированных серверных модулях Tesla S1070/S2050, выполненных на основе процессоров T10 и T20;

– OpenCL на графических процессорах AMD RV870, Cypress, а также универсальных процессорах фирм Intel и AMD.

Возможность использования графических арифметических ускорителей для выполнения вычислений общего назначения совместно с универсальными процессорами открыла огромные возможности для прикладного программного обеспечения.

Графические арифметические ускорители обладают некоторыми архитектурными особенностями, которые определяют их высокие характеристики. Для сравнения приведем в табл. 1 основные параметры, характеризующие современные универсальные процессоры и арифметические ускорители.

Высокая производительность, наряду с возможностью одновременно выполнять огромное количество потоков позволяет эффективно использовать графические арифметические ускорители как в гибридных вычислительных системах петафлопного класса, так и в специализированных компактных вычислительных системах.

Таблица 1

Характерные параметры ускорителя и универсального процессора

	Универсальный процессор	Арифметический ускоритель	Соотношение
Наименование процессора	AMD Opteron 6168	NVIDIA Tesla C2050	
Тактовая частота процессора, ГГц	1,9	1,15	1,7
Вычислительные ядра, шт.	12	14	1,2
Потребляемая мощность АСР, Вт	80	238	3,0
Теоретическая пиковая производительность (64-бит), ГФлоп/с	91,2	515,2	5,7
Удельная производительность на Вт	1,14	2,16	1,9
Максимальное количество обрабатываемых потоков, шт.	12	21504	1792
Одновременное количество исполняемых потоков, шт.	12	448	37,3
Пропускная способность памяти, ГБ/с	42,7	144	3,4

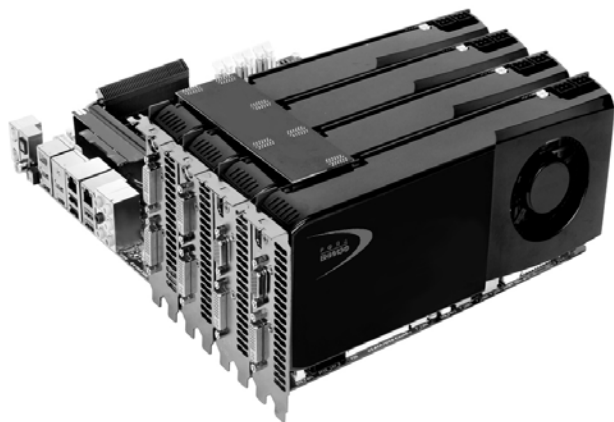


Рис. 1. Типичная компоновка арифметических ускорителей

Ряд специализированных компактных суперЭВМ

С 2007 года в РФЯЦ-ВНИИЭФ ведутся интенсивные исследования возможности применения графических арифметических ускорителей для научных расчетов.

В 2009 году в РФЯЦ-ВНИИЭФ под руководством Ю. Г. Бартенева и С. А. Степаненко был разработан первый опытный образец компактной гибридной вычислительной системы ГВС-10 «Кубань» (рис. 2,а). Он построен на базе универсального процессора Intel и восьми графических процессорах GT200 от NVIDIA. Для снижения уровня шума было реализовано жидкостно-воздушное охлаждение. Данная архитектура позволила получить рекордную на момент выпуска производительность до 700 ГФлоп/с в физическом объеме обычного персонального ком-

пьютера. Для сравнения, универсальная компактная суперЭВМ АПК-1 [1] обладает производительностью 1044 ГФлоп/с.

В 2010 году создан серийный образец ГВС-18А [2] (рис. 2,б), обладающий уникальными характеристиками, такими как пиковая производительность более 3,5 ТФлоп/с, уровень акустического шума не превышает 39 дБА, потребляемая мощность менее 1,5 кВт и низкая стоимость по сравнению с аналогичными разработками.

Также с 2010 года выпускаются модели специализированной компактной суперЭВМ ГВС-14 и ГВС-14Т. Они обладают производительностью до 2 ТФлоп/с, построены на современных графических процессорах NVIDIA GF100/GF110. Система охлаждения применяется как воздушная, так и жидкостная. Большое количество модификаций данной машины позволяет варьировать параметрами производительности, уровнем акустического шума, потребляемой мощностью и стоимостью.

В табл. 2 приведено сравнение специализированных компактных суперЭВМ разработки РФЯЦ-ВНИИЭФ различных типов и вычислительной системы Supermicro [3] (типичной для данного класса).

Приведенные выше системы используются как в РФЯЦ-ВНИИЭФ, так и поставлены в целый ряд научных и промышленных предприятий.

При проектировании вычислительных систем используются современные системы построения 3D моделей (рис. 3). Это позволяет уже на стадии разработки выявить достоинства и недостатки будущей вычислительной системы, определить, какие компоненты можно использовать стандартные, а какие необходимо заказывать или реализовывать на собственном производстве.



а



б

Рис. 2. Внешний вид: а – ГВС-10 «Кубань»; б – ГВС-18А

Характеристики специализированных компактных суперЭВМ

	ГВС-10 «Кубань», 2009 г.	ГВС-18А, 2010 г.	ГВС-14, 2010 г.	ГВС-14Т, 2010 г.	Supermicro США
Теоретическая пиковая производительность (двойная точность), ТФлоп/с	0,7	3,7	0,7	2	2
Теоретическая пиковая производительность (одинарная точность), ТФлоп/с	7	18,5	5,4	4	4
Арифметический ускоритель	NVIDIA GeForce GTX 295	AMD RADEON 5970	NVIDIA GeForce GTX 480	NVIDIA Tesla C2050	NVIDIA Tesla C2050
Количество вычислительных ядер, шт.	244	164	66	66	66
Объем ОЗУ вычислительного модуля, ГБ	12	до 48			
Объем ОЗУ арифметического ускорителя, ГБ	7,2	до 16	до 8	до 24	
Операционная система	Linux / MS Windows				
Акустический уровень шума, дБА	48	до 45			–
Габариты (В x Ш x Г), мм	660 × 250 × 605	495 × 210 × 490			–
Система охлаждения	воздушная/жидкостная			воздушная	
Подключение у локальной сети предприятия	Gigabit Ethernet				
Электропитание	220 В, 50 Гц				
Потребляемая мощность, кВт	до 1,5				
Цена, тыс. руб.	–	до 500	до 400	до 900	1000

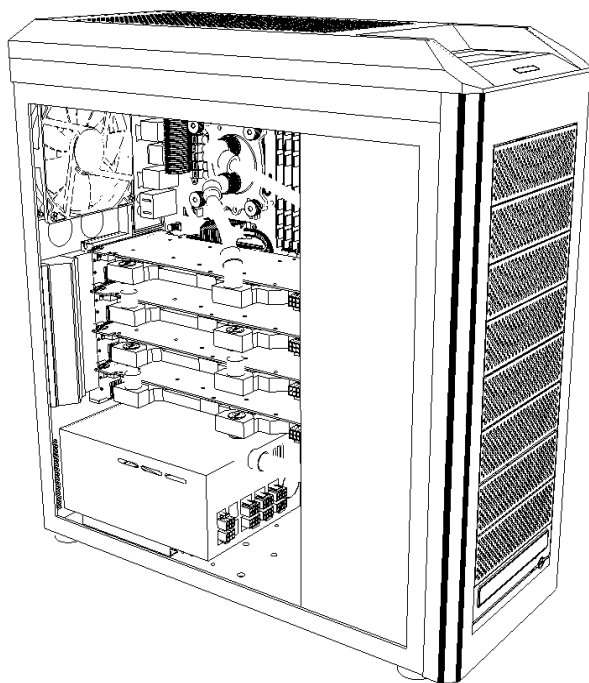


Рис. 3. 3D модель специализированной компактной суперЭВМ

В настоящее время ведется разработка нового, более производительного образца – вычислительной системы ГВС-20 производительностью до 10 ТФлоп/с.

Некоторые архитектурные особенности специализированных компактных суперЭВМ

Одной из уникальных особенностей разработанных компактных вычислительных систем является подсистема жидкостно-воздушного охлаждения, которая позволяет:

- снизить температуру основных компонент вычислительной системы и поддерживать в рабочем режиме, тем самым повышая их надежность;
- снизить уровень акустического шума от вычислительной системы и повысить комфортность работы;
- повысить плотность компоновки арифметических ускорителей, что в условиях ограниченного физического объема повышает общую производительность вычислительной системы в целом.

Внутренняя структура конструкции представлена массивом микроребер толщиной 0,5 мм каждое, что обеспечивает малое гидросопротивление потоку и высокую эффективность решения.

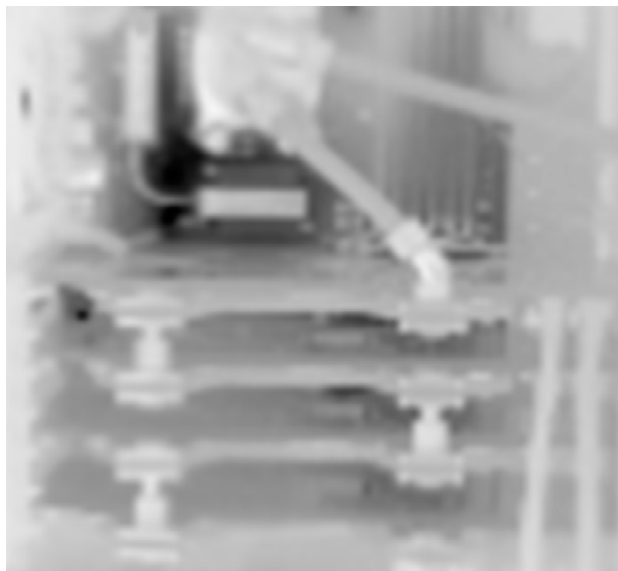
Используемый водоблок арифметического ускорителя не только охлаждает графический процессор и видеопамять, а также силовые модули. Он изготовлен из чистой меди с никелевым покрытием, что обеспечивает высокую эффективность охлаждения.

В процессе наладки и диагностики выполняется контроль за температурными характеристиками вычислительной системы. На рис. 4 приведены снимки

вычислительных систем с воздушным и жидкостным охлаждением, выполненные тепловизором. Обе системы обладают одинаковой архитектурой, ключевым отличием является система охлаждения.



а



б

Рис. 4. Воздушное и жидкостное охлаждение

В целом система с жидкостным охлаждением обладает более низкой температурой (в среднем на 20°). На рис. 5 проиллюстрированы значения температуры на графических процессорах на примере высокопроизводительного теста перемножения матриц.

Другой не менее важной подсистемой, обеспечивающей надежную работу нескольких арифметических ускорителей и вычислительного модуля, является масштабируемая подсистема питания. Она выполнена на основе платформы Channel Well, в виде реализации блока питания общей мощностью 1500 Вт, который состоит из двух субблоков, каждый из которых можно нагрузить до 750 Вт.

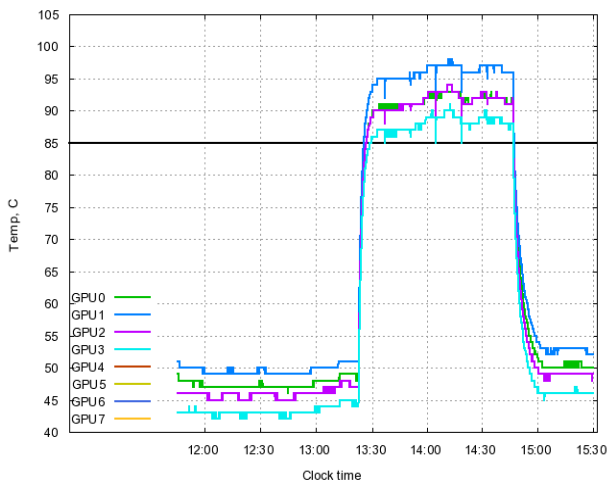


Рис. 5. Зависимость температуры графических процессоров на воздушном охлаждении

Между собой субблоки связаны только схемой включения. Каждый субблок имеет собственный: активный PFC, ШИМ-контроллер, силовой трансформатор и выходные выпрямители.

Каждый субблок имеет две независимые шины +12 В, каждая из которых дополнительно разделена на две «виртуальные» линии. Собраны они по совершенно одинаковым схемам и параметры имеют очень близкие. Напряжение +12 В обеспечивается во всем диапазоне нагрузок, +5 В укладывается в 4 %-й допуск при разрешенном стандартом 5 %-ом.

Эффективность работы блока до 86 % на нагрузке порядка сотен ватт с падением ниже 80 % на большой нагрузке.

В разрабатываемой ГВС-20 планируется использовать шесть субблоков для обеспечения 4 кВт энергии.

Численное моделирование с использованием специализированных компактных высокопроизводительных вычислительных систем

Создание приложений для гетерогенных параллельных вычислительных платформ является относительно несложным процессом, поскольку классические подходы к программированию для многоядерных универсальных и графических процессоров похожи. NVIDIA предлагает масштабируемую программно-аппаратную архитектуру для параллельных вычислений Compute Unified Device Architecture (CUDA). Разработанная компанией NVIDIA технология позволила использовать вычислительные мощности видеокарт в любых программах при помощи специального интерфейса прикладного программирования для языка C. AMD предлагает аналогичную среду разработки Accelerated Parallel Processing (APP), основанную на едином стандарте для программ, исполняющихся в гетерогенной среде, – OpenCL, который позволяет эффективно использовать все преимущества гетерогенных вычислительных платформ.

В мире растет количество прикладных программ, способных использовать в полной мере возможности специализированных компактных вычислительных машин с арифметическими ускорителями.

В 2010 году сотрудниками РФЯЦ-ВНИИЭФ в рамках НИР по государственному контракту № 02.514.12.4008 в интересах атомной промышленности были разработаны и адаптированы программные комплексы для моделирования большого класса физических процессов [4] методами Монте-Карло [5–7] и молекулярной динамики [8–11].

Метод Монте-Карло (ММК) является одним из основных численных методов, используемых при математическом моделировании различных физических процессов. Область его применения весьма широка и многогранна. Он применяется как в медицине, так и в промышленности, например, для задач ядерной безопасности, защиты от проникающих излучений, геологоразведке, нейтронно-, гамма-, протонографии и т. д.

Особо следует выделить класс реакторных задач, в которых роль метода возрастает с каждым годом. Расширяются как классы решаемых задач, так и точность описания процессов, протекающих в реакторах. Это связано и со стремительным развитием вычислительной техники, и с преимуществами данного метода перед остальными. В настоящее время ММК используется для расчета эффективного коэффициента размножения нейтронов активных зон ядерных реакторов, обоснования радиационной и ядерной безопасности транспортных упаковочных комплексов для перевозки и долговременного хранения отработанного ядерного топлива АЭС.

Преимущества метода Монте-Карло заключаются, во-первых, в использовании прямого моделирования (вероятностей реально протекающих физических процессов) при решении уравнений переноса. При этом, для описания взаимодействия частиц с веществом используются оцененные спектральные константы, у которых число точек в сечениях приближается к 100 тыс. Такое описание позволяет избегать неточностей различного рода математических и физических приближений и точно моделировать резонансные области для каждого изотопа. Второе преимущество состоит в том, что метод позволяет моделировать особенности различных процессов, протекающих в реакторах, и их конструктивные особенности до мельчайших подробностей (потвельное описание), не используя гомогенизацию и другие упрощения. Например, длительность численных исследований активной зоны реактора ВВЭР-1000 на ГВС-14 уменьшена по сравнению с 12-ядерным универсальным процессором AMD Opteron 6180 SE в семь раз.

Другим перспективным быстроразвивающимся методом математического моделирования является молекулярная динамика. Основное направление – это моделирование на микроуровне фундаментальных процессов, определяющих свойства вещества. Для моделирования методом классической молеку-

лярной динамики доступны объемы вещества микрометрового масштаба. Результаты моделирования могут представлять интерес для задач материаловедения, конденсированного состояния вещества и физики твердого тела. В настоящее время проводится большое количество исследований свойств твердых тел. Одним из перспективных направлений, представляющих интерес для задач атомной промышленности и энергетики, является молекулярно-динамическое моделирование влияния облучения на механические свойства металлов и их сплавов. На данном классе задач также ускорение вычислений составляет 10 раз.

Также на вычислительной системе ГВС-14 выполняются расчеты по программному комплексу MASTER Professional [12–14], который уже давно зарекомендовал себя как удобный современный компьютерный инструмент для проведения расчетно-теоретического моделирования быстропротекающих нестационарных многомерных импульсных газо-, гидродинамических и упругопластических процессов.

MASTER Professional представляет собой многозадачную, многооконную интегрированную среду, позволяющую оперативно проводить 1–3D расчеты. Однако выполнение 3D расчетов задач с миллионами счетных точек на универсальных процессорах занимает, как правило, длительное время, что снижает оперативность и лишает программный комплекс свойств оперативного инструмента исследователя.

Выполнена большая работа по разработке и адаптации алгоритмов трехмерного моделирования на гибридных вычислительных системах. Получено ускорение вычислительного процесса от 40 до 60 раз на одном арифметическом ускорителе по сравнению с одним ядром универсального процессора. Эксперименты проводились на ГВС-14, размер задачи варьировался от 200 тыс. до 4 млн точек. Данные размерности являются типичными для этого класса задач. Для ускорителя реализованы наиболее востребованные уравнения состояний.

Заключение

Получение на рабочем месте исследователя или разработчика возможностей суперкомпьютера становится возможным благодаря специализированным компактным суперЭВМ. Основной особенностью таких машин является использование арифметических ускорителей, которые обладают на порядок большей производительностью, по сравнению с самыми современными универсальными процессорами. Но для эффективного использования гибридного решения необходимо адаптировать прикладное программное обеспечение, способное использовать все аппаратные возможности.

Достоинствами подобных гибридных вычислительных систем являются:

- высокая производительность;
- низкая стоимость;

- низкое удельное энергопотребление;
- возможность использовать в обычном помещении.

Литература

1. Стрюков В. Н., Бартенев Ю. Г., Басалов В. Г. и др. Универсальная компактная супер-ЭВМ // Сб. трудов XII межд. семинара «Супервычисления и математическое моделирование» / Под ред. Р. М. Шагалиева. Саров: РФЯЦ-ВНИИЭФ, 2011. С. 329–331.

2. Специализированные компактные суперЭВМ ГВС-18А и ГВС-14 [Электронный ресурс] <http://www.vniief.ru/directions/grazrab/catalog/infprod/razrabotki/super-evm.html>.

3. Super Micro Computer, Inc. [Электронный ресурс] <http://www.supermicro.com>.

4. Воронин Б. Л., Грушин С. А., Житник А. К. и др. Программно-аппаратные комплексы на базе вычислительных систем с арифметическими ускорителями для моделирования методом Монте-Карло и методом молекулярной динамики // Сб. трудов XII межд. семинара «Супервычисления и математическое моделирование» / Под ред. Р. М. Шагалиева. Саров: РФЯЦ-ВНИИЭФ, 2011. С. 88–93.

5. СМК-У. Свидетельство о регистрации № 2010614638 от 14.07.2010.

6. Программный комплекс расчета критических параметров методом Монте-Карло для вычислительных систем с арифметическими ускорителями. Свидетельство о регистрации № 2010617162 от 27.10.2010.

7. Рыбкин А. С., Залялов А. Н., Малькин А. Г. и др. Программный комплекс на базе гибридных вычислительных систем для расчета критических параметров методом Монте-Карло // Сб. трудов XII межд. семинара «Супервычисления и математическое моделирование» / Под ред. Р. М. Шагалиева. Саров: РФЯЦ-ВНИИЭФ, 2011. С. 310–315.

8. Комплекс программ молекулярно-динамического моделирования (MoDyS). Свидетельство о регистрации № 2010614974 от 29.07.2010.

9. Программный комплекс моделирования свойств материалов методом молекулярной динамики (МД) для вычислительных систем с арифметическими ускорителями. Свидетельство о регистрации № 2010617163 от 27.10.2010.

10. Крючков И. А., Копкин С. В. Программный комплекс моделирования методом молекулярной динамики для гибридных вычислительных систем // Сб. трудов XII межд. семинара «Супервычисления и математическое моделирование» / Под ред. Р. М. Шагалиева. Саров: РФЯЦ-ВНИИЭФ, 2011. С. 234–242.

11. Копкин С. В., Крючков И. А. Алгоритм модернизированного многочастичного потенциала для молекулярно-динамического моделирования на графическом арифметическом ускорителе // Вопросы атомной науки и техники. Сер. Математическое моделирование физических процессов. 2010. Вып. 3.

12. Tchekhounov E., Roudenko V., Chabourov M. Tutorial and Research Package MASTER on Continuum Mechanics. Proc. Second European Conference // Physics Teaching in Engineering Education, Budapest, Hungary, 2000.

13. Давыдов И. А., Пискунов В. Н., Мешков Е. Е. и др. Программный комплекс MASTER – интегрированная среда визуального компьютерного моделирования процессов физики сплошных сред // Сб. тезисов XVII Всероссийской конф. «Теоретические основы и конструирование численных алгоритмов и решение задач математической физики с приложением к многопроцессорным системам». Абрау-Дюрсо, 2008.

14. Башуров В. В., Крючков И. А., Павлов С. В. и др. Визуальное численное моделирование процессов физики взрыва и удара в программном комплексе MASTER Professional // Сб. материалов докладов Международной конференции «Седьмые Окуневские чтения», Санкт-Петербург, 2011.