

Издается с 2004 года

**Председатель
редакционного совета**
Г.Г. Себряков, чл.-корр. РАН

Главный редактор
И.А. Каляев, чл.-корр. РАН

**Заместители председателя
редакционного совета:**
С.Ю. Желтов, чл.-корр. РАН
М.Н. Красильщиков, д.т.н.

Редакционный совет:
В.И. Аверченков, д.т.н.
(зам. гл. редактора)
А.И. Башмаков, к.т.н.
С.Н. Васильев, акад. РАН
Ю.В. Визильтер, д.ф.-м.н.
А.И. Кибзун, д.ф.-м.н.
П.Е. Клейзер (зам. гл. редактора)
Ю.Н. Кофанов, д.т.н.
В.В. Лебедев, чл.-корр. РАН
Е.А. Микрин, акад. РАН
В.В. Попов, д.т.н.
А.В. Рыбаков, к.т.н.
Р.Л. Смелянский, чл.-корр. РАН

Региональные редсоветы:
Белгород О.И. Константинов, д.т.н.
Орел В.Т. Еременко, д.т.н.
Волгоград В.А. Камаев, д.т.н.
Самара В.А. Виттих, д.т.н.
Воронеж В.П. Смоленцев, д.т.н.
Санкт-Петербург Ю.А. Гатчин, д.т.н.
Иркутск И.В. Бычков, акад. РАН
Ставрополь П.А. Аверичкин, д.т.н.
Курск О.И. Атакищев, д.т.н.
Таганрог В.В. Курейчик, д.т.н.
Лима (Республика Перу) А.П. Риос, д-р
Тула А.Н. Иноземцев, д.т.н.
Минск (Республика Беларусь) С.В. Абламейко, акад. НАНБ
Уфа Б.Г. Ильясов, д.т.н.
Нижний Новгород Р.Я. Вакуленко, д.э.н.
Ярославль А.И. Яманин, д.т.н.
С.И. Ротков, д.т.н.

Редакция:
Н.Н. Морозова
О.В. Мочалина
Н.В. Пантина

Журнал зарегистрирован в
Министерстве Российской Феде-
рации по делам печати, телерадио-
вещания и средств массовых
коммуникаций.
Свидетельство о регистрации
ПИ № ФС77-46365 от 26.08.2011 г.

Тел. редакции: (495) 589-56-41,
514-76-50

Адрес редакции: 119048, г. Москва,
ул. Усачева, д. 35, стр. 1

E-mail: vkit@idspektr.ru;
vkitpost@rambler.ru
Http://www.vkit.ru

СОДЕРЖАНИЕ

КОМПЬЮТЕРНОЕ ЗРЕНИЕ. ВИРТУАЛЬНАЯ РЕАЛЬНОСТЬ

- Инсаров В.В., Обросов К.В., Ким В.Я., Лисицын В.М.** Формирование комплексированных телевизионно-тепловизионных изображений в системах переднего обзора летательных аппаратов 3
- Тихомирова Т.А., Федоренко Г.Т., Кириллова Л.Н.** Флис-трассировка контурных изображений с помощью «гибких» связей 10

КОМПЬЮТЕРНАЯ ГРАФИКА И ГЕОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Чобану М.К., Кургускин О.С.** Применение сеточного метода для сжатия графической информации. Основной алгоритм 13

РАСПРЕДЕЛЕННЫЕ ВЫЧИСЛИТЕЛЬНЫЕ И ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

- Горобец В.В.** Облачная модель транзакционной системы 19
- Мельник Э.В., Иванов Д.Я.** Принципы организации децентрализованных сетевых информационных-управляющих систем 25

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

- Никонов В.И., Щерба Е.В., Щерба М.В.** Методика защиты информации в беспроводных сетях на основе динамической маршрутизации трафика 31
- Стародубцев Ю.И., Бухарин В.В., Кирьянов А.В., Баленко О.А.** Метод оценки защищенности информационно-телекоммуникационной сети от деструктивных программных воздействий 37
- Щеглов К.А., Щеглов А.Ю.** Практическая реализация дискреционного метода контроля доступа к создаваемым файловым объектам 43

ПРОГРАММНОЕ И АППАРАТНОЕ ОБЕСПЕЧЕНИЕ ПРОЦЕССОВ И СИСТЕМ

- Березко А.Е., Бодунов Н.В., Фролов С.М.** Высокопроизводительная многопроцессорная вычислительная система с водяным охлаждением как основа для построения мультимедийных суперкомпьютеров 50
- НОВОСТИ В МИРЕ КОМПЬЮТЕРНЫХ И ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ** 55

Журнал распространяется по подписке, которую можно оформить в любом почтовом отделении, или непосредственно в издательстве.

Индексы по каталогам: «Роспечать» – 84197, «Пресса России» – 39244, «Почта России» – 60263

Журнал входит в перечень изданий, утвержденных ВАК РФ, для публикации трудов соискателей ученых степеней

Перепечатка, все виды копирования и воспроизведения материалов, публикуемых в журнале «Вестник компьютерных и информационных технологий», допускаются со ссылкой на источник информации и только с разрешения редакции



УДК 004.272.43

А.Е. Березко, канд. техн. наук, Н.В. Бодунов, С.М. Фролов, канд. техн. наук
(ОАО «Т-Платформы», Москва);
e-mail: Sergey.Frolov@t-platforms.ru

ВЫСОКОПРОИЗВОДИТЕЛЬНАЯ МНОГОПРОЦЕССОРНАЯ ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА С ВОДЯНЫМ ОХЛАЖДЕНИЕМ КАК ОСНОВА ДЛЯ ПОСТРОЕНИЯ МУЛЬТИПЕТАФЛОПСНЫХ СУПЕРКОМПЬЮТЕРОВ*

Представлены практические результаты решения задач повышения производительности многопроцессорной вычислительной системы с использованием системы водяного охлаждения.

This paper presents the results of practical solution of the problem aimed to increase the performance of a multiprocessor computer system based on a water cooling system.

Ключевые слова: водяное охлаждение; высокопроизводительные вычисления; вычислительный узел; многопроцессорная вычислительная система; энергоэффективность.

Keywords: Water-cooled; High performance computing; Computing node; Multiprocessor computer system; Energy efficiency.

Введение

Предоставленные самим себе события имеют тенденцию развиваться от плохого к худшему.

Закон Мерфи, следствие 5

Технологии создания высокопроизводительных вычислительных комплексов на базе стандартных стоечных серверов и блейд-платформ подошли к рубежу, за которым увеличение вычислительной мощности ведет к лавинообразному росту стоимости, энергопотребления и веса систем, а также снижению эффективности обмена данными между узлами. Воздушная система охлаждения стандартных серверных платформ не позволяет отводить от стойки более 30 кВт тепла, сильно ограничивая возможности наращивания производительности. Стандартные серверные платформы не обеспечивают нужной степени плотности размещения компонентов, а также надежности, – чем больше компонентов содержит вычислительная система, тем чаще происходят сбои из-за выхода какого-либо компонента из строя.

В ОАО «Т-Платформы» ведется разработка вычислительной системы в стоечном исполнении с высокой плотностью компоновки вычислительных узлов (ВУ) и водяным охлаждением. Система построена на базе стандартных, общедоступных центральных процессоров (ЦП) архитектуры x86, а также ВУ гибридной архитектуры с поддержкой новейших ускорителей вычислений и предназначена для использования в составе высокопроизводительных суперкомпьютеров (СК) петафлопсного уровня производительности.

Предлагаемое решение позволяет обойти проблему отвода тепла и связанное с этим ограничение наращивания производительности за счет применения оригинальной системы водяного охлаждения ВУ. Такое решение принципиально изменяет отношение к проблеме охлаждения, позволяя отвести от стойки 150...200 кВт тепла. Это делает возможным достижение петафлопсного и мультипетафлопсного (десятки петафлопс) уровня производительности. Дополнительным преимуществом предлагаемого решения является его стоечный форм-фактор.

* Работа выполнена при финансовой поддержке Министерства образования и науки РФ, проект № 07.524.12.4010 «Разработка программно-аппаратного многопроцессорного вычислительного комплекса (МВК) для построения высокопроизводительных вычислительных систем мультипетафлопсного диапазона».

В отличие от наиболее распространенного на данный момент подхода, базовым элементом системы становится не отдельная серверная платформа, а целая стойка, где все ВУ консолидированы с помощью единой объединительной платы. Такое решение увеличивает плотность расположения компонентов в стойке, повышает эффективность межузловых коммуникаций, надежность за счет высокой интеграции компонентов и сокращения межузловой коммутации и позволяет создавать более простую и безопасную систему охлаждения.

Общее описание вычислительной системы

Если вам непонятно какое-то слово в техническом тексте, не обращайтесь на него внимания. Текст полностью сохраняет смысл и без него.

Закон Купера

Разрабатываемый *многопроцессорный вычислительный комплекс (МВК)* будет являться основой для построения мультипетафлопсных суперкомпьютерных установок, обеспечивая размещение до 512 ЦП в одной вычислительной стойке и пиковую производительность не менее 100 Тфлопс с использованием ЦП архитектуры x86 при текущем на сегодняшний день максимальном уровне порядка 20...30 Тфлопс.

Отличительными особенностями конструкции МВК, обеспечивающими улучшение энергоэффективности, вычислительной плотности и надежности, являются стоечное исполнение и применение жидкостного охлаждения. Объединяющий конструктивный элемент МВК – *вычислительная стойка*, представленная на рис. 1, и состоящая из шкафа, подсистем вычислений, жидкостного охлаждения и электропитания.

Подсистема вычислений состоит:

- из 256 ВУ, с восемью модулями *оперативной памяти (МОП)* каждый;
- двух модулей управления, с четырьмя модулями памяти каждый;
- 42 коммутаторов InfiniBand коммуникационной и вспомогательной сетей;
- 18 коммутаторов Ethernet управляющей и сервисной сетей;
- 10 объединительных плат.

Каждый ВУ МВК содержит два ЦП Intel Sandy/Ivy Bridge стандартной архитектуры x86-64, которые обеспечивают максимальную степень совместимости МВК с широко используемыми прикладными приложениями. Многопроцессорный вычислительный комплекс будет являться одной из

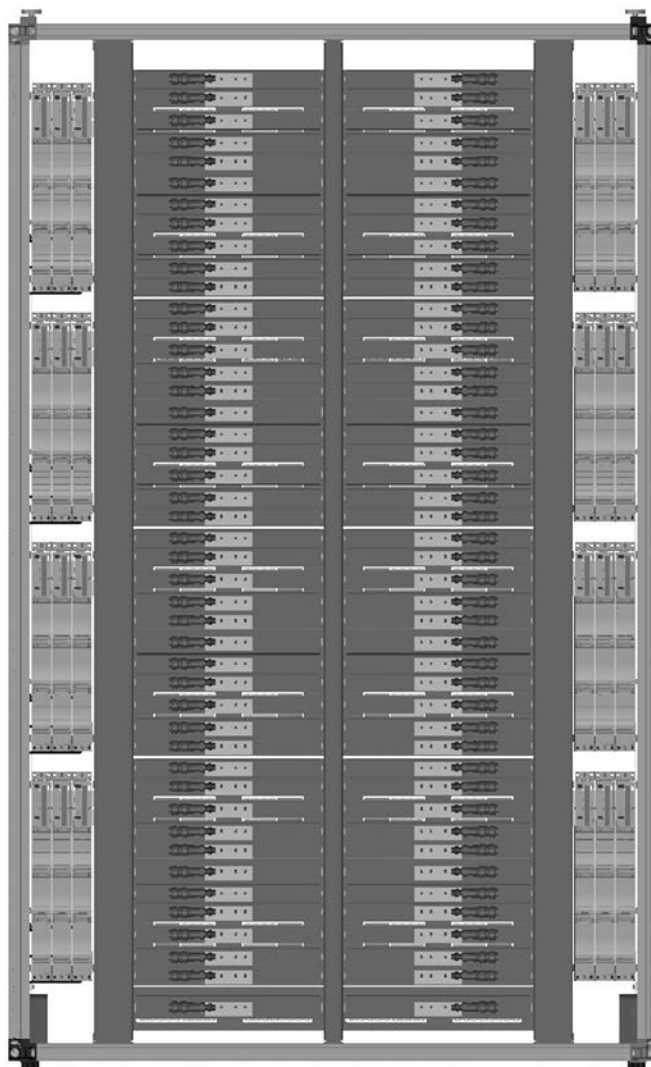


Рис. 1. Вычислительная стойка

наиболее плотных из имеющихся на рынке вычислительных систем на базе такой архитектуры. Предусмотрена возможность использования ВУ гибридной архитектуры, оснащенных ускорителями вычислений архитектуры GPU (*Graphics Processing Unit*) или MIC (*Intel Many Integrated Core Architecture*), что в перспективе позволит увеличить производительность МВК до 300...400 Тфлопс.

Подсистема жидкостного охлаждения обеспечивает нормальную рабочую температуру электронных компонентов ВУ и содержит в гидравлическом тракте быстроразъемные соединения, которые многократно повышают надежность и улучшают ремонтпригодность МВК. Для того чтобы произвести «горячую замену» любого узла, достаточно быстро отсоединить его вместе с радиатором от подсистемы охлаждения, при этом исключается утечка охлаждающей жидкости и попадание ее на активные компоненты МВК.

**Конструктивные особенности
вычислительного узла
и модуля оперативной памяти**

При проектировании МВК разработаны принципиально новые технические решения, касающиеся ВУ и МОП.

Конструкция платы ВУ рассчитана на водяное охлаждение в сборке с радиатором подсистемы жидкостного охлаждения. Для обеспечения контакта с поверхностью радиатора все электронные компоненты платы располагаются на верхней стороне и имеют монтажную высоту, не превышающую монтажную высоту ЦП. Сигнальные и силовые разъемы располагаются на заднем торце платы, пример такой платы приведен на рис. 2 (см. обложку).

Вычислительный узел имеет следующие основные параметры:

- *оперативное запоминающее устройство (ОЗУ)*: восемь модулей памяти объемом 4...16 ГБ, до 64 ГБ DDR3 SDRAM на ЦП;

- коммуникационная/вспомогательная сеть: FDR4X InfiniBand (56 Гбит/с);

- управляющая/сервисная сеть: 1G/10G Ethernet;

- *постоянное запоминающее устройство (ПЗУ)*: твердотельный накопитель *SSD (Solid State Drive)*, емкостью 256 ГБ;

- размер печатной платы 500×210 мм.

Разработанный в ОАО «Т-Платформы» МОП объемом 4...16 ГБ реализован в соответствии со стандартом DDR3-1600. Особенностью разработанного МОП является оригинальная конструкция мезанинного типа с информационным разъемом, находящимся с одной стороны печатной платы МОП, и микросхемой памяти – с другой (рис. 3, см. обложку). Это позволяет применить общий радиатор жидкостного охлаждения как для электронных компонентов платы ВУ, так и для МОП. Для распределения усилия при использовании общего охлаждающего радиатора для ВУ и МОП используются специализированные крепления, расположенные по краям платы МОП.

Модуль оперативной памяти имеет следующие основные параметры:

- одноканальный режим работы;

- регистр между микросхемами памяти и системным контроллером памяти;

- поддержку кода коррекции ошибок *ECC (Error Correcting Code)*;

- максимальный объем до 16 ГБ;

- максимальная тактовая частота шины 1866 МГц;

- размер печатной платы 25×130 мм.

Подсистема жидкостного охлаждения

Все не так легко, как кажется.

Закон Мерфи, следствие 1

Вычислительная стойка МВК оборудована подсистемой прямого жидкостного охлаждения. В качестве жидкого теплоносителя используется вода с номинальной температурой теплоносителя 44 °С на входе, допустимый диапазон температуры воды на входе 30...45 °С. Подсистема жидкостного охлаждения применяется для охлаждения большей части оборудования вычислительной стойки, в том числе ВУ, модулей управления и мониторинга, коммутаторов InfiniBand коммуникационной и вспомогательной сетей, коммутаторов Ethernet управляющей и сервисной сетей. Охлаждающая жидкость подводится к вычислительной стойке через располагающийся в нижней части шкафа гидрораспределительный узел. В состав подсистемы жидкостного охлаждения входят:

- два «холодных» трубопровода;

- два «горячих» трубопровода;

- 90 радиаторов;

- гибкие шланги;

- быстроразъемные соединители.

Теплоноситель через систему трубопроводов, гибких шлангов и быстроразъемных соединителей подводится ко всем охлаждаемым компонентам системы. Отвод тепла от отдельных компонентов оборудования вычислительной стойки осуществляется с помощью радиаторов. Теплоноситель подводится к радиаторам и отводится гибкими шлангами. Для возможности горячей замены оборудования в вычислительной стойке радиаторы и гибкие шланги оборудованы быстроразъемными непротекающими соединителями. Для минимизации числа соединителей используется один радиатор для охлаждения группы ВУ или коммутаторов, образец которого представлен на рис. 4 (см. обложку). Максимальный общий расход воды для одной стойки составляет 8,6 л/с, разница температур на входе и выходе из подсистемы охлаждения – 5 °С.

Безопасность МВК в случае возможных протечек воды обеспечивается путем изолированного расположения всей инфраструктуры жидкостного охлаждения в передней части стойки и герметизации ВУ, модулей управления и мониторинга, коммутаторов. Все трубопроводы находятся в обслуживаемых зонах стойки и отделены от остальных зон герметичными кожухами, соединенными с дренажной системой в гидрораспределительном узле.

Сетевая инфраструктура

Сетевая инфраструктура МВК включает: коммуникационную сеть на базе сетевой технологии FDR4X InfiniBand для обмена вычислительными данными между ВУ; вспомогательную сеть на основе сетевой технологии FDR InfiniBand для хранения, загрузки PXE и удаленного доступа; управляющую сеть, использующую сетевую технологию 1G/10G Ethernet для загрузки ВУ; сервисную сеть для контроля и мониторинга ВУ с использованием контроллера BMC по вспомогательному каналу управляющей сети.

Технология InfiniBand реализована на основе топологии n -мерный тор. На рис. 5 показана возможность гибкого конфигурирования с использованием медных кабелей внутри вычислительной стойки.

При реализации технологий Ethernet и InfiniBand предусмотрено двойное дублирование.

В управляющей/сервисной сети Ethernet используются два уровня коммутаторов: первый уровень – агрегирует 40x1G нисходящих каналов в 2x1G восходящих канала; второй уровень – агрегирует 16x1G нисходящих каналов в 2x10G восходящих канала.

В коммуникационной и вспомогательной сетях применяются три типа коммутаторов InfiniBand. Во вспомогательной сети используются: коммутатор первого типа, имеющий 32 нисходящих и 4 восходящих каналов FDR4X; коммутатор второго типа, имеющий 16 нисходящих, 18 восходящих и 2 управляющих каналов FDR4X. В коммуникационной сети используется коммутатор третьего типа, который имеет 8 нисходящих и 28 восходящих каналов FDR4X.

Во всех коммутаторах встроены контроллеры для сбора данных с датчиков: температуры, обнаружения протечек охлаждающей жидкости, параметров электропитания и т.п.

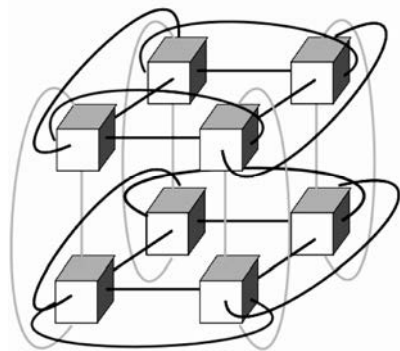


Рис. 5. Топология InfiniBand

Программное обеспечение многопроцессорного вычислительного комплекса

Всякая работа требует больше времени, чем вы думаете.

Закон Мерфи, следствие 2

Для описываемого МВК разработано специализированное системное программное обеспечение, основными особенностями которого являются:

- построение на базе операционной системы Linux;
- наличие набора оптимизированных математических библиотек;
- наличие комплекта средств разработки.

В число базовых систем программного обеспечения входят системы:

- мониторинга и управления;
- автоматического отключения оборудования;
- управления ресурсами (менеджер ресурсов). Функции мониторинга реализованы на трех уровнях.

- на уровне вычислительной стойки:
 - обнаружение протечек охлаждающей жидкости;
 - мониторинг входного и выходного давлений в подсистеме жидкостного охлаждения;
- на уровне вычислительного блока (32 ВУ):
 - мониторинг параметров электропитания;
 - мониторинг температуры компонентов объединительной платы;
- на уровне платы электронных компонентов:
 - мониторинг температуры электронных компонентов;
 - обнаружение протечек охлаждающей жидкости;
 - мониторинг параметров электропитания.

Заключение

Мерфи был оптимистом.

Следствие из закона Мерфи Каллагана

Техническое решение описываемого МВК аккумулирует в себе большинство преимуществ аналогичных разработок. По сравнению с аналогами, описываемый комплекс имеет особенности:

- одновременное сочетание поддержки высокоплотной архитектуры с современным «горячим» водяным охлаждением;
- агрессивные показатели производительности из расчета на одну стойку и на 1 Вт потребляемой энергии;
- поддержка гетерогенной архитектуры ВУ;
- модульность, упрощающая ремонтпригодность (включая смену модулей памяти);

– мощность, потребляемая МВК на максимальной вычислительной производительности 106 Тфлопс, будет составлять не более 120 кВт;

– максимальный размер вычислительного кластера на базе МВК может достигать 128 вычислительных стоек.

При проектировании МВК были разработаны принципиально новые технические решения, касающиеся ВУ и МОП.

Применение в описываемом МВК универсальных процессоров архитектуры x86 в сочетании со стандартными сетями коммутации позволяет получить хорошую совместимость с приложениями. В то же время стоечное исполнение обеспечивает оптимальную коммуникацию между узлами с минимальными задержками и, следовательно, создает

предпосылки для хорошего ускорения работы приложений. Важнейшими преимуществами разработки являются низкое энергопотребление и высокая энергоэффективность системы, обеспечивающая до 50 % экономии расходов на электроэнергию.

Библиографический список

1. **Начало** совещания с членами Совета Безопасности по вопросам создания и применения супер-компьютеров [Электронный ресурс]. URL: <http://www.kremlin.ru/text/appears/2009/07/220213.shtml> (дата обращения 01.02.2013).

2. **Текущий рейтинг:** 17-я редакция от 18.09.2012 [Электронный ресурс]. URL: <http://top50.supercom-puters.ru/?page=rating> (дата обращения 04.02.2013).

3. **Суперкомпьютерные** технологии в науке, образовании и промышленности / под ред. В.А. Садовниченко, Г.И. Савина, Вл.В. Воеводина. М.: Изд-во Моск. ун-та, 2009. 232 с.

Статья поступила в редакцию 10.01.2013 г.

◆◆◆◆◆

Общество с ограниченной ответственностью "Спектр"

ОПТОВАЯ ТОРГОВЛЯ КОМПЬЮТЕРАМИ И ПЕРИФЕРИЙНЫМИ УСТРОЙСТВАМИ

Вы хотели бы купить в Вашу компанию компьютерную технику и не знаете, где предложат самые выгодные условия?

Наша компания готова поставить Вам абсолютно любые аппаратные и программные решения с максимально выгодным и интересным предложением для Вас.

Тел.: (495)795-87-70 ,8-903-795-87-70

E-mail: ooo-spektr2012@yandex.ru

[Http://ooo-spektr.ru/](http://ooo-spektr.ru/)

