



УДК 681.3

В.П. КЛИМЕНКО, Э.И. КОМУХАЕВ

СУПЕРКОМПЬЮТЕРЫ: ОСОБЕННОСТИ СОВРЕМЕННОЙ МОДЕРНИЗАЦИИ

Abstract. Features of modernization of High-performance computer (HPC) are analyzed in the article. A considerable growth of productivity and energy efficiency for standardized clusters, disaster-resistance decisions for Data-processing centre (DPC), workings out of specialized microcircuits for strategic supercomputers with common access to the memory are noted. Cases of introduction of accelerators of specialized calculations for 12 systems of new list Top500 are allocated in the table. We consider that introducing accelerators in structures of Ukrainian clusters and DPC is of current importance.

Key words: expansion of applications, modernization, petaflop, exaflop, power savings, many-core, parallelization, Grid, accelerators of calculations, common memory.

Анотація. У статті проаналізовані особливості модернізації НРС (High Performance Computer). Відмічено значний ріст продуктивності та показників енергозбереження для стандартизованих кластерів, рішення катастрофостійкості для Центру обробки даних (ЦОД), розробки спеціалізованих мікросхем для стратегічних суперкомп'ютерів зі спільною пам'яттю. Виділені в таблицю випадки застосування прискорювачів спеціалізованих обчислень для 12 систем нового списку Top500. Вважаємо актуальним упровадження прискорювачів і в структури українських кластерів та ЦОД.

Ключові слова: розширення застосувань, модернізація, петафлоп, ексафлоп, енергозбереження, багатоядерність, паралелізація, Грід, прискорювачі обчислень, спільна пам'ять.

Аннотация. В статье проанализированы особенности модернизации НРС (High Performance Computer). Отмечен значительный рост производительности и показателей энергосбережения для стандартизованных кластеров, решения катастрофоустойчивости для Центра обработки данных (ЦОД), разработки специализированных микросхем для стратегических суперкомпьютеров с общей памятью. Выделены в таблицу случаи внедрения ускорителей специализированных вычислений для 12 систем нового списка Top500. Считаем актуальным внедрение ускорителей и в структуры украинских кластеров и ЦОД.

Ключевые слова: расширение применений, модернизация, петафлоп, эксафлоп, энергосбережение, многоядерность, параллелизация, Грід, ускорители вычислений, общая память.

1. Ускоренная модернизация НРС

Кризис незначительно снизил высокие темпы модернизации НРС с целью повышения их продуктивности, особенно в США, Германии, Китае, России, Японии, где суперкомпьютеры все шире используются при решении актуальных задач науки, экономики, обороны, экологии [1]. Из долговременных тенденций ускорения, рассмотренных нами в [2], выделяются замена гонок мегагерц процессоров на использование многоядерности последних с разработкой новых алгоритмов, знаний в области параллельных вычислений и растущая роль энергосбережения для НРС. Теперь в мире ранжируют ведущие НРС не только списком рейтинга Top500 по уровню производительности, но и списком Top Green500 по уровню энергосбережения.

Петафлопный уровень производительности НРС (квадрильон операций с плавающей запятой в секунду) для теста Linpack был превзойден в середине 2008 года системой IBM Roadrunner вопреки ряду прогнозов, относивших такое событие на более поздний период. Для преодоления следующего, эксафлопного, барьера (тысяча петафлопс) мировому сообществу потребуются огромные усилия по созданию новых языков программирования, алгоритмов, средств разработки, намного более эффективных средств вычислений, памяти, коммутации, периферии. По оценкам Джека Донгарры, разработчика теста Linpack и одного из авторов рейтинга Top500, появление эксафлопного суперкомпьютера, вероятно, в 2019 году, следует соответственно ожидать через 11

лет после покорения петафлопного рубежа. Донгарра при этом учитывает сложившийся четкий 11-летний цикл роста быстродействия суперкомпьютеров. Действительно, гигафлопный рубеж был преодолен в 1986 году, терафлопный – в 1997 году и т.д.

Не способствуют развитию сотрудничества на этом пути практика эмбарго на распространение некоторых перспективных микросхем, пакетов ПО, устройств, засекречивание новых перспективных технологий и проектов НРС. С ностальгией упоминают в [3] существовавшее в СССР многообразие собственных перспективных суперкомпьютерных проектов, когда, в частности, “...Институт кибернетики АН Украины разрабатывал машину с макроконвейерной архитектурой ЕС2701 и целое семейство языков программирования для него...” наряду с проектами других организаций при соответствующей координации и взаимоподдержке решений [4]. Отметим, что заказываемые государством высокопроизводительные вычислительные системы нередко ориентированы на отдельные приоритетные применения, где необходимо особое сочетание высоких производительности и надежности. Например, спроектированные в прошлом веке российские высокопроизводительные средства вычислений для противовоздушной обороны в составе комплекса С-300 до сих пор эффективны благодаря инновационным высоконадежным архитектурно-алгоритмическим решениям параллелизации вычислений, даже при использовании устаревших микросхем, которые позже были обновлены в комплексах С-400. Главный конструктор указанных вычислительных средств академик В. Бурцев [2] охарактеризовал, с учетом своего опыта, супер-ЭВМ как “ЭВМ с наивысшей среди других производительностью, разработанную с предельным интеллектуальным напряжением проектировщиков и построенную с предельным технологическим напряжением заводов-изготовителей”. Нередко высокой интеллектуальностью решений, применением специализированных ускорителей удается компенсировать невысокую производительность доступных стандартных компонентов. Соответствующие супер-ЭВМ с предельно достижимыми специализированными параметрами востребованы для ряда оборонных применений, среди них и упомянутая петафлопная военная система Roadrunner с ускорителями вычислений, используемая, в частности, для моделирования ядерных арсеналов, лазеров высокой мощности и др. В [5] отмечается архитектурное влияние работ группы академика В. Бурцева по машине, управляемой потоками данных, на современное развитие российского проекта суперкомпьютера с глобально адресуемой памятью, выполняемого в московском ОАО “НИЦЭВТ”.

Эксперты утверждают, что в условиях растущего сокращения сроков и усложнения решаемых задач для существенного повышения продуктивности коренная модернизация суперкомпьютеров необходима не реже чем через 5 лет. Часто намного целесообразнее установить новый суперкомпьютер с повышенными показателями производительности и энергосбережения, новым ПО, чем продолжать выборочно обновлять устаревшие средства “старого” суперкомпьютера. Один из мощных суперкомпьютеров немецкого университета Штутгарта, несколько лет назад еще входивший в список Top500, на этапе обновления состава суперкомпьютеров этого немецкого университета недавно подарен Донецкому техническому университету. Ввод в эксплуатацию этого “научного подарка” планируют приурочить к открытию в Донецке Генерального консульства Германии в середине 2010 года. До этого Украина использовала три суперкомпьютера уровня 6–7 Тфлопс производительности (в ИК НАНУ, ГАО НАНУ, КПИ) и десятки значительно менее производительных сис-

тем. Предусматривается, что донецкий суперкомпьютер станет частью европейской сети суперкомпьютерных вычислений. Отметим высокую квалификацию специалистов суперкомпьютерного центра упомянутого университета Штутгарта (HLRS), которые совместно со специалистами компании Cray разработали новую линию суперкомпьютеров “среднего уровня” Cray XT5m для широкого круга потребителей. Причем, первым заказчиком системы Cray XT5m стал именно HLRS. Суперкомпьютеры серии Cray XT5m ориентированы на гибкую модернизацию и расширение блоков, например, на базе использования шестиядерных процессоров AMD с возможностью доведения продуктивности модернизируемой системы даже до уровня Cray XT5 – теперешнего лидера.

2. Перспективные системы и проекты

Итак, мировой рейтинг, согласно 34-му списку Top500, возглавил кластер Jaguar Cray XT5 с производительностью на тесте Linpack 1759 TFlop/s. Третье место в списке принадлежит суперкомпьютеру Cray XT5 Kraken с производительностью 831,7 TFlop/s. Обе эти системы компании Cray, установленные в национальной лаборатории Оак-Ридж, США, используют возможности шестиядерных процессоров AMD Opteron с тактовой частотой 2,6 ГГц, маршрутизаторы Seastar 2+ с пропускной способностью 57,6 Гб/с. Компания Cray также выиграла конкурсы на поставки в 2010 году трех суперкомпьютерных систем повышенного быстродействия для МО США. Эти системы Baker будут иметь новый чипсет Gemini для межсоединений и усовершенствованную архитектуру Cray XT. Применение этих систем будут обеспечивать фундаментальные и прикладные научные исследования, разработки новых материалов, горючего, вооружений, разработки долговременных прогнозов погоды.

Главным конкурентом компании Cray при создании систем петафлопсной производительности остается компания IBM, системы которой занимают второе и четвертое места в рейтинге Top500. Причем военный суперкомпьютер IBM Roadrunner имеет производительность 1042 TFlop/s, занимая второе место. На четвертом месте суперкомпьютер IBM Blue Gene/P JUGENE, установленный в Германии, с производительностью 825,5 TFlop/s на тесте Linpack.

Пятую позицию рейтинга занял суперкомпьютер Китая Tianhe-1 с производительностью 563,1 TFlop/s на тесте Linpack.

Неожиданным стало появление на двенадцатой позиции российского суперкомпьютера “Ломоносов” с производительностью 350 TFlop/s на тесте Linpack после невысоких позиций России в прежних рейтингах. Ключевую роль для достижения высокой производительности системы “Ломоносов” обеспечили запатентованные специалистами московской компании “Т-Платформы” вычислительные узлы на базе блейд-структур T-Blade2. Рекордная вычислительная плотность на квадратный метр площади структур T-Blade2 открывает реальные возможности создавать системы с показателем 1 петафлопс и выше.

Основные технические характеристики суперкомпьютера “Ломоносов” следующие:

– пиковая производительность	420Тфлопс;
– реальная производительность	350Тфлопс;
– эффективность (соотношение пиковой и реальной производительности)	83%;

– число вычислительных узлов	4 446 ;
– число процессоров	8 892;
– число процессорных ядер	35 776;
– число типов вычислительных узлов	3 (T-Blade2, T-Blade1.1, платформа на базе процессора PowerXCell 8i);
– основной тип вычислительных узлов	T-Blade2;
– процессор основного типа	
вычислительных узлов	Intel Xeon X5570;
– оперативная память	56 576ГБ;
– общий объем дисковой памяти вычислителя	166 400 ГБ;
– занимаемая площадь	252 кв.м;
– энергопотребление вычислителя	1,5 МВт;
– интерконнект	QDR Infiniband;
– система хранения данных	трехуровневая с параллельной файловой системой;
– объем системы хранения данных	до 1 350ТБ;
– операционная система	Clustrx T-Platforms Edition.

Данную систему относят к гибриднему типу, поскольку здесь взаимодействуют универсальные четырехъядерные процессоры Intel Xeon, выполняющие системные функции, и ускоряющие вычисления спецпроцессоры PowerXCell. Напомним, что первый петафлопный кластер был также гибридным, использующим сочетания ускорителей PowerXCell и универсальных процессоров AMD Opteron.

Наибольшие известность и применение получили следующие разновидности ускорителей для HPC:

- Cell-процессоры, разработка IBM совместно с компаниями Sony, Toshiba;
- платы компании ClearSpeed (Великобритания – США);
- платы на базе FPGA разработок компаний Celoxica, Nallatech, DRC и др.;
- блоки и платы GPGPU (General Purpose Graphical Processing Units) разработок компаний NVIDIA, AMD (ATI);
- платы и чипы GRAPE, разработанные университетом Токио. Все указанные разновидности ускорителей уже представлены в новом списке Top 500.

Таблица 1. Ускорители в системах Top500

Тип ускорителя	Система, организация, страна установки	Ранг системы в Top500/ производит. в TFlops по Linpack	Ранг системы в Top Green500 List/показатель энергоэффект. в MFlops	Ключевые аппаратные средства
IBM Power XCell 8i	Roadrunner, DOE/NNSA/LANL, США	2/1042	4/458.33	Opteron DC, InfiniBand
AMD ATI Radeon HD 4870	NUDT TH-1 National SuperComputer Center in Tianjin, Китай	5/563.1	8/379.24	Xeon E5540/E5450
IBM Power XCell 8i	“Ломоносов”, МГУ, Россия	12.350.1	56/230.42	T-Platforms T-Blade2, Xeon 5570, InfiniBand QDR

IBM Power XCell 8i	DOE/NNSA, LANL, США	29/126.5	4/458.33	Opteron DC, InfiniBand
IBM Power XCell 8i	IBM Ponghkeepsie Benchmarking Center, США	78/63.25	4/458.33	Opteron DC, InfiniBand
IBM Power XCell 8i	Forschungszentrum Juelich, Германия	110/43.01	1/722.98	3D Torus IBM
IBM Power XCell 8i	Universitaet Regensburg, Германия	111/43.01	1/722.98	3D Torus IBM
IBM Power XCell 8i	Universitaet Wuppertal, Германия	112/43.01	1/722.98	3D Torus IBM
ClearSpeed CSX600	Институт технологий, Токио, Япония	56/87.01	290/78.88	Opteron, Xeon E5440, InfiniBand
nVidia GT200	Институт технологий, Токио, Япония	56/87.01	290/78.88	Opteron, Xeon E5440, InfiniBand
GRAPE-DR	Национальная астрономическая обсерватория, Япония	445/21.96	7/428.91	InfiniBand
FPGA Stratics III	Skif Aurora, Южноуральский университет, Россия	450/21.84	47/248.18	Xeon E55XX

Специфические задачи некоторых ЦОД внесли свой вклад в расширение разновидностей ускорителей на чипах FPGA. В частности, здесь отметим возможности аппаратных платформ Crescendo Networks CN-5504 и CN-5510 на базе чипов FPGA, внедренные в ЦОД Франции, Кореи, России. Платформы эффективно устраняют избыточные соединения TCP, мультиплексируют и ускоряют веб-приложения. Продукты компании Crescendo Networks получили название ADC-Application Delivery Controller. ADC обеспечивают интеллектуальную балансировку нагрузки отдельных серверов и географически распределенных ЦОД, значительно снижают количество серверов приложений, распознают тип трафика и оперативно управляют им. Аналитики Gartner уже представляют компанию в качестве безусловного лидера рынка ADC. Крупнейший хостинг Франции Интернет Fr

повысил эффективность работы более чем 15000 web-сайтов средствами ADC, которые по главным тестам Fr превосходят конкурентов на 300%. Крупный бизнес, особенно банковский, Интернет-профилей, убедившись в преимуществах ЦОД, все чаще внедряет именно географически распределенные системы для обеспечения их максимальной катастрофоустойчивости. Соответственно совершенствуются и расширяются волоконно-оптические сети, ускорители специализированных функций для оперативных смен режимов географически распределенных кластерных средств (рис. 1), включая функции обеспечения информационной безопасности таких средств.

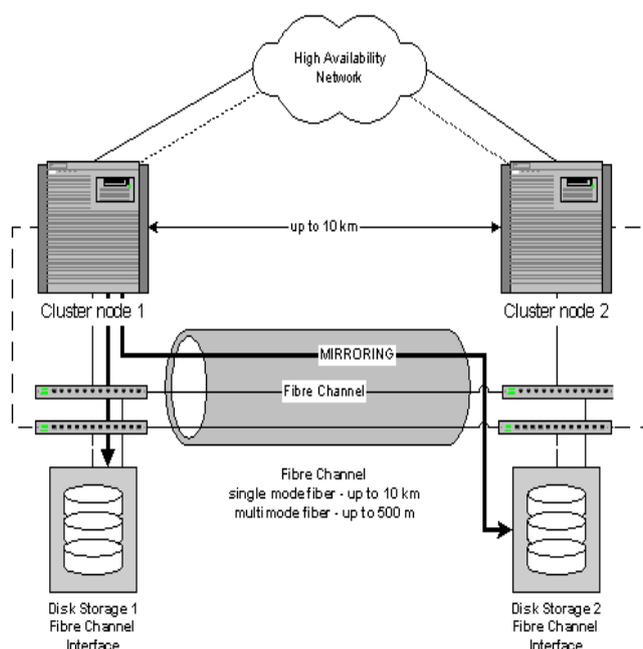


Рис. 1. Географически распределенный кластер

Развитие перспективных Грид-систем, где интегрирует большой объем географически удаленных ресурсов вычислений и хранения данных, также во многом базируется на совершенствовании

нии скоростных оптоволоконных линий связи и средств информационной безопасности. Хотя еще не решен ряд юридических вопросов по оплате Грид-сервисов, существенно расширились приложения Грид. Так, с помощью Грид-систем успешно обрабатывают большие объемы экспериментальных данных большого адронного коллайдера ЦЕРН, ряда масштабных медицинских исследований для долговременных прогнозов синоптиков. Программу Грид в Украине поддерживают Национальная академия наук (координатор – Институт теорфизики) и Министерство науки и образования (координатор – Национальный технический университет «КПИ»).

3. Выводы

1. Широкая модернизация суперкомпьютеров, в первую очередь, вызвана требованиями весомого прироста производительности для решения многих актуальных задач. Новый список Top500 включает две системы США с превышением петафлопного уровня, анонсирован предстоящий ввод систем с показателями 10–20 петафлопс.
2. Для ряда применений критичен уровень энергозатрат суперкомпьютеров, теперь введен обновляемый каждые полгода список рейтинга Top Green500. Сейчас его возглавляют три системы, установленные в Германии, с уровнем 722,98 MFLOPS/W (производительность на ват).
3. Особенно значимо снижение уровня энергозатрат для ЦОД, число которых быстро растет в Украине, России. Низкие энергозатраты для ЦОД обеспечивают приемлемый уровень стоимости, улучшают катастрофоустойчивость систем, заказываемых бизнес-структурами телекоммуникационной, банковской, транспортной, нефтегазовой и других отраслей.
4. В последнем списке Top500 уже 12 систем используют различные типы ускорителей специализированных операций. Представляется, что украинские разработчики имеют достаточную квалификацию для активного участия в реализации перспективных проектов создания ускорителей для внутренних и внешних HPC.
5. Совершенствуются технологии и применения Грид-систем с использованием многих кластеров различных стран, включая Украину. Эффективно используются возможности сервисов Грид для обработки данных коллайдера ЦЕРН, медицинских данных, особенно в онкологии. Недавно украинская сеть «УАРНЕТ» была подключена через польскую систему «Пионер» к общеевропейской академической сети Грид.
6. Лидером модернизационных разработок для HPC в РФ стала компания «Т-Платформы», создавшая самый быстродействующий в РФ суперкомпьютер «Ломоносов», эффективные блейд-конструкции T-Blade2, собственные решения ПО для пета-, экса-суперкомпьютеров, установившая HPC в ряде стран, существенно модернизировав два академических кластера Украины [6].

СПИСОК ЛИТЕРАТУРЫ

1. Суперкомпьютерные технологии в науке, образовании и промышленности. – М.: Изд. Московского университета, 2009. – 232 с.
2. Клименко В.П. Суперкомпьютеры: тенденции и технологии последнего десятилетия / В.П. Клименко, Э.И. Комухаев // Математичні машини і системи. – 2006. – № 3. – С. 146 – 151.
3. www.delyagin.ru/citation/9468.html.
4. Структурно-архитектурные решения многопроцессорных ЭВМ с микроконвейерной обработкой данных / В.С. Михалевич, Ю.В. Капитонова, А.Г. Кухарчук [и др.] // Электронная вычислительная техника: сб. статей. – 1988. – Вып. 2. – С. 151 – 160.
5. Слуцкий А. Российский суперкомпьютер с глобально адресуемой памятью / А. Слуцкий, Л. Эйсымонт // Открытые системы. – 2007. – № 9. – С. 42 – 51.
6. www.t-platforms.ru.

Стаття надійшла до редакції 09.06.2010