

СУПЕРКОМПЬЮТЕРЫ: ТЕНДЕНЦИИ И ТЕХНОЛОГИИ ПОСЛЕДНЕГО ДВАДЦАТИЛЕТИЯ

Abstract: *The problems of tendencies, architectures, technology of the developments of the supercomputers of the last twentieth anniversary are discussed in the paper. There are adduced data about new abilities of the element base of the supercomputers, about broadening of using. The prospects of the development of cluster systems, the ways of perfection of interaction of the cluster nodes are accepted. The significant contribution of the leading specialists of the Institutes of Cybernetic Centre by Glushkov V.M. of NASU to the developments of decisions and technologies of supercomputers realisation is marked.*

Key words: *growth of productivity, cluster, multinuclear ness, parallelization, Computer nets, reconfigurable accelerators*

Анотація: *В статті розглянуто питання тенденцій, архітектур, технологій розробок суперкомп'ютерів останнього двадцятиліття (1986 – 2006 рр.). Наведено дані про нові можливості елементної бази суперкомп'ютерів та про розширення застосувань. Акцентовано перспективи розвитку кластерних систем, шляхи вдосконалення взаємодії вузлів кластерів. Відмічено значний внесок провідних фахівців інститутів Кібцентру ім. В.М. Глушкова НАНУ в розробки рішень та технологій реалізації суперкомп'ютерів.*

Ключові слова: *зростання продуктивності, кластери, багатоядерність, паралелізація, обчислювальні мережі, реконфігураційні прискорювачі.*

Аннотация: *В статье рассмотрены вопросы тенденций, архитектур, технологий разработок суперкомпьютеров последнего двадцатилетия (1986 – 2006 гг.). Приведены данные о новых возможностях элементной базы суперкомпьютеров, о расширении применений. Акцентированы перспективы развития кластерных систем, пути совершенствования взаимодействия узлов кластеров. Отмечен значительный вклад ведущих специалистов институтов Кибцентра им. В.М. Глушкова НАНУ в разработки решений и технологий реализации суперкомпьютеров.*

Ключевые слова: *рост производительности, кластеры, многоядерность, параллелизация, вычислительные сети, реконфигурируемые ускорители.*

1. Введение

Новые задачи и вызовы значительно повысили требования к более широкой доступности, надежности и особенно к повышению производительности суперкомпьютерных систем. Именно рост производительности значительно расширяет круг эффективно решаемых задач.

В последнем двадцатилетии (1986 – 2006 гг.) важнейшей общей тенденцией в сфере развития суперкомпьютерных систем можно считать реальный переход к международному сотрудничеству, открытости после прежнего периода жесткой конфронтации Запада и СССР. Пока еще сохраняется эмбарго на поставку в Россию и Китай мощных суперкомпьютеров со стороны США и Японии с целью затруднить им работы в области моделирования процессов ядерного оружия. Зато сняты барьеры при приобретении совершенных микросхем для суперкомпьютеров, лицензионных пакетов ПО и др. Подчеркнем также усиливающуюся тенденцию конвергенции вычислительных средств и коммуникационных, сетевых средств. Теперь запросы к высокопроизводительным серверам и ответы последних широко транспортируются с помощью коммуникационных средств Интернета. Эта всемирная сеть – главный источник информации и ряда услуг. Сложный комплекс высокоскоростных оптоволоконных линий Интернета подсоединен к множеству компаний, организаций, к домам индивидуальных пользователей. Очень перспективны международные разработки Grid – систем, о чем подробнее ниже.

Выдающаяся роль в становлении суперкомпьютерных систем принадлежит крупнейшим идеологам – академикам С.А. Лебедеву, В.М. Глушкову и Сеймуру Крею. В 1985 году в мире действовало уже более 150 таких систем стоимостью около \$10 млн. каждая [1]. Сегодня число систем

высокой производительности превышает несколько десятков тысяч, среди них особенно быстро растет количество коммерческих систем. Растет также коммерческое использование бюджетных комплексов. Пик разработок в СССР суперкомпьютерных систем пришелся примерно на 1986 год. В достаточной мере тогда финансировались и координировались через оборонные ведомства следующие разработки многопроцессорных систем:

- ереванский матричный спецпроцессор ЕС2700 (этот спецпроцессор выполнял только узкий набор операций над матрицами и векторами, но с очень большой скоростью);
- киевский макроконвейер ЕС2701 (ведущие разработчики В.М. Глушков, С.Б. Погребинский, А.Г. Кухарчук, В.П. Клименко, Ю.В. Капитонова и др.);
- ленинградский мультипроцессор с динамической архитектурой ЕС2704;
- таганрогский мультипроцессор ЕС2706;
- семейство мультипроцессоров ПС ИПУ АН СССР;
- Электроника СС – БИС;
- московские комплексы «Эльбрус-1», «Эльбрус-2» (ведущие разработчики С.А. Лебедев, В.С. Бурцев, Б.А. Бабаян и др.);
- киевский коллективный интеллектуальный терминал для «Эльбрус-2» (ведущие разработчики З.Л. Рабинович, А.А. Якуба и др.);
- московские системы НИИ «Квант» и ряд других.

В 1989 году Г. Беллом и Д. Нельсоном была разработана шуточная классификация, предлагающая любой компьютер, весящий более тонны, считать суперкомпьютером. Из ряда предложенных определений суперкомпьютеров безоговорочно сегодня признается статус суперкомпьютеров лишь для систем, включенных в последнюю версию выпуска рейтинга Top 500.

Государства и компании ожесточенно борются за высокие места во всемирном списке Top500. Высокое место в этом списке свидетельствует о владении эффективным интеллектуальным инструментарием, который привлекает соответственно выгодные заказы, лучших специалистов, повышает престиж страны.

Список Top500, утвержденный в ноябре 2005 года, свидетельствует:

- главным потребителем суперкомпьютеров становится индустрия, 226 систем из списка принадлежат коммерческим организациям, остальными владеют научно-исследовательские фирмы и академические организации;
- 305 систем из последнего списка Top500 установлены в США;
- самый мощный в мире суперкомпьютер установлен IBM в ядерной Национальной лаборатории Лоуренса в Ливерморе, он может выполнять 281 триллион вычислений в секунду после недавнего удвоения числа процессоров (теперь 131076 процессоров);
- главными европейскими претендентами на места в «десятке» являются Испания (суперкомпьютер в Барселоне с производительностью 28 Тфлопс), Германия (суперкомпьютер JUBL с производительностью 46 Тфлопс), Франция (суперкомпьютер Комиссии по атомной энергии с производительностью 50 Тфлопс).

На германском суперкомпьютере можно вести до пяти исследовательских проектов одновременно. Выражается надежда, что благодаря JUBL «мозги потекут в Германию, а не в США», на

нем будет эффективно моделироваться синтез белков, поведение жидкостей, газов и др. Крупнейший азиатский суперкомпьютер создают в Токио совместно компании NEC, AMD, Sun. Его производительность превысит 100 триллионов операций в секунду. Предполагают, что этот кластер станет крупнейшим в мире по количеству ядер процессоров.

В Украине в настоящее время действует ряд суперкомпьютерных кластеров [2]. Два крупнейших из них установлены в Институте кибернетики им. В. М. Глушкова НАНУ. Они включены под номерами 21, 35, соответственно, в обновленный список Top50 суперкомпьютеров СНГ, имея производительность на тестах Linpack 278,3 гигафлопс и 189,3 гигафлопс [3].

2. Актуальные архитектурные решения

Чаще всего аналитики выделяют как актуальные следующие четыре направления архитектурных решений: векторно-конвейерные системы, массивно-параллельные системы с распределенной памятью (MPP массивно-параллельная архитектура – massive parallel processing), параллельные системы с общей памятью, кластерную архитектуру.

После 2002 года кластерные системы стали лидировать по количеству применений, вытеснив в значительной мере векторно-конвейерные машины, доминировавшие с середины 70-х годов. Векторно-конвейерные системы, оснащенные специализированными процессорами для обработки сразу наборов (векторов) данных, сохранили преимущества в быстродействии и простоте программирования по сравнению с кластерами. Зато кластеры, использующие стандартные процессоры в своих узлах, значительно дешевле, хотя имеют значительно более сложное ПО.

Узлы кластера обычно имеют близкое географическое расположение и коммутируются средствами высокоскоростных соединений. Значительное внедрение в системах высокой производительности СНГ получили недорогие сетевые технологии SCI (Scalable Coherent Interface) норвежских фирм Dolphin (аппаратура) и Scali (ПО). Здесь аппаратно поддерживаются операции узла кластера над памятью других узлов без обращения к ОС и библиотекам. Суперкомпьютеры с сетью SCI уже не попадают по быстродействию в новые версии Top500, несколько систем, использующих SCI, пока входят в список Top50.

Более эффективной является технология InfiniBand, используемая в кластерах в качестве межузлового соединения (медного или оптоволоконного), и основы для интерфейса передачи сообщений MPI. Внедрено немало продуктов InfiniBand: платы или микросхемы хост-адаптеров и канальных адаптеров, коммутаторы, драйверы устройств, реализации MPI. В обзорах кластерных вариантов межсоединений узлов выделяют такие достоинства InfiniBand: наличие стандарта IBTA, доступные варианты пропускной способности 2, 10, 30 Гбит/с, наличие сдвоенных адаптеров 2x10 Гбит/с, небольшое время задержки микросхем, гибкость сочетаний с другими типами сетей. Например, занимающий в Top50 7-ое место кластер НИВЦ МГУ имеет коммутирующую сеть типа InfiniBand с использованием микросхем компании Mellanox, транспортную сеть типа Gigabit Ethernet, сервисную сеть Fast Ethernet. Доля систем с InfiniBand имеет устойчивую тенденцию к росту.

Под управлением Linux работает примерно 75% из 500 суперкомпьютеров списка Top500.org, доля Unix – 20%, MacOSX – 1% Windows – 0%. Билл Гейтс недавно объявил о разработках подключения к мощным кластерам специализированных “настольных Windows – суперкомпьютеров”. Среди систем параллельного программирования особенно распространены решения с

передачей сообщений MPI (Message Passing Interface). MPI – основное средство программирования высокопроизводительных мультикомпьютеров, включая Silicon Graphics Origin 2000, Cray T3E, IBM SP2, MBC-1000 и много других. Использование стандарта MPI, принятого в 1994 году, обеспечивает полную независимость приложений от архитектуры многопроцессорной системы. Особенно распространена библиотека функций MPI с интерфейсом для языков C и FORTRAN [4, 5].

3. Инновационные технологии. Grid-проекты

Стремительно развиваются инновационные технологии, продукты, принимаются решения для высокопроизводительных систем. Начался процесс массовой модернизации кластеров путем перехода на двухъядерные процессоры производства компаний AMD, Intel, IBM. Разрабатываются четырехъядерные процессоры и варианты с большим числом ядер.

За небольшими компаниями, имеющими эффективные инновационные технологии, ведут постоянную “охоту” компании – гиганты, стараясь их выкупить вместе с соответствующими интеллектуальными активами, присоединив к себе. Так, компания Cray в 2004 году приобрела небольшую канадскую компанию OctigaBay, которая создала эффективные акселераторы приложений на базе ПЛИС VirtexII Pro. Компания Cray внедрила эти акселераторы в свой суперкомпьютер XD1, используя уже новые ПЛИС Xilinx Virtex-4 по 6 штук в каждой стойке системы XD1. Cray в новых суперкомпьютерах планирует подключать вспомогательный DRC FPGA чип непосредственно в сокет Opteron-процессора. Это создаст ряд преимуществ для пользователей разъема CraySeaStar.

Учеными Калифорнийского университета в Беркли на основе перепрограммируемых логических интегральных схем (FPGA) инициирована разработка вычислительного комплекса RAMP (Research Accelerator for Multiple Processors) – перестраиваемого имитатора – ускорителя разработок ПО. Этот комплекс нацелен на решение злободневной проблемы организации четкого взаимодействия между конструкторами аппаратной части системы и программистами. Часто разработка оборудования занимает несколько лет, а написание кода без готовой аппаратной части вызывает большие затруднения. Матрицы FPGA в этом комплексе будут конфигурироваться для адаптации всей системы для отработки решений очередных задач программистов. Предполагается, что проект комплекса RAMP в 1000 узлов обойдется примерно в \$100 тысяч, будет потреблять всего 1,5 кВт энергии. Разработчики считают, что использование обычного кластера для аналогичных задач потребовало около \$2 млн. и 120 кВт энергии. К проекту RAMP подключились Массачусетский технологический институт, Университеты Стэнфорда, Техаса, Карнеги-Меллона, Вашингтона. Недавно IBM объявила об открытии бесплатного доступа к спецификации процессорного ядра IBM Power PC405 для исследовательских и образовательных учреждений. Это решение существенно поддерживает RAMP. В табл. 1 отмечено наличие указанного ядра IBM в высокоинтегрированных ПЛИС фирмы Xilinx, которые достаточно широко используются и в Украине рядом организаций Киева, Чернигова, Харькова и других городов. Таким образом. Это решение IBM расширяет возможности и наших разработчиков при проведении исследований перспективных микросхем, включая исследования многоядерных технологий [6].

Энергосберегающий проект суперкомпьютера с производительностью в один терафлопс на базе FPGA разрабатывают сотрудники Центра параллельных вычислений Эдинбургского универ-

ситета. Энергопотребление этой системы будет в 100 раз ниже, чем существующие системы с производительностью 1 терафлопс.

Встречаются и другие варианты интеграции активов, сотрудничества. Например, упомянутый в разделе 1 Б.А. Бабаян, удостоенный в 1987 году звания лауреата Ленинской премии за разработку и внедрение многопроцессорной системы “Эльбрус-2”, сохраняет, с одной стороны, пост директора Института микропроцессорных вычислительных систем РАН, с другой стороны, Б.А. Бабаян с августа 2004 года занимает должность директора по архитектуре в подразделении программных решений корпорации Intel и руководит глобальным проектом по технологии безопасных вычислений. Он стал первым европейским ученым, получившим звание заслуженного инженера – исследователя Intel, ряд его изобретений внедряются корпорацией.

В таблице приведены типы и примеры интеллектуальных активов по тематике высокопроизводительных систем.

Таблица. Типы и примеры интеллектуальных активов для HPC

N	Тип	Примеры	Примечания
1	Изобретения представляют переход на новый изобретательный уровень	US Patent 6973559 Scalable hypercube multiprocessors US Patent 6629267 Способ регистрации и анализа программных ошибок	Дается оригинальное решение масштабируемых межсоединений мультипроцессорных систем, патентодержатель – компания SGI. Объектом этого программного патента является метод передачи сообщения об ошибке, сначала определяется программный модуль, где произошла ошибка, затем выявляется место сбоя в модуле. Патентодержатель – компания Microsoft
2	Лицензия на технологии	IBM предоставила лицензию на ядро процессоров Power PC компании Xilinx	ПЛИС компании Xilinx (особенно Virtex 4, 5), содержащие указанное ядро, широко внедряются в ряд систем HPC
3	Лицензия на технологии	Московская компания “Т-Платформы” приобрела у фирмы GDA лицензию на использование технологии шины HyperTransport	Т-Платформы разрабатывают ПЛИС – компоненты для HPC, использующие шину HyperTransport
4	Полезная модель, не имеет изобретательского уровня, но ее патент дает право на гонорар при внедрениях внутри страны. Экспертиза на требование новизны и промышленной применимости не проводится	RU 38092 U1 Серверная платформа, автор Бречалов А.В. компания “Т-Платформы”	Предложена серверная платформа, отличающаяся дополнительными контроллером серверной сети и портами USB, доступ к которым осуществляется через отверстие в передней стенке шасси, контроллер закреплен винтами или болтовым соединением. Документацию предлагает “Т-Платформы”

Упомянутая в таблице московская компания “Т – платформы” установила и настроила ряд высокопроизводительных систем в СНГ, является организатором и спонсором проведения рейтингов Top50 СНГ. Главным интеграционным направлением сегодня для всей суперкомпьютерной отрасли становятся вычислительные сети grid computing. Они реализуют трехуровневую модель: узел (например, сервер), кластер (набор связанных между собой узлов), сеть (набор кластеров, других систем). Территориально-распределенные Grid-сети объединяют компьютеры с различными аппаратными и программными системами. На базе высокопроизводительных GRID-систем планируется решать множество задач, включая идентификацию личности по биометрическим параметрам в реальном масштабе времени и др. Во многих странах развернуты национальные GRID-проекты с целью не допустить отставания в этом перспективном направлении. Компании Dell, EMC, Intel, Oracle совместно проектируют систему интеграции MegaGrid для обеспечения оперативного информационного сервиса по запросам бизнес-структур [7]. Мощный суперкомпьютер Linux компа-

нии HP по заказу Министерства энергетики США уже подключен к системе DOE Science Grid. Число публикаций о проектах, достижениях, технологиях Grid постоянно возрастает. Например, сообщается, что один из банков США, внедривший технологию Grid, вместо суток сообщает клиенту через минуту после запроса о своем решении предоставлять ли ему значительную ссуду. Институт теорфизики НАНУ на базе своего кластера планирует сотрудничество через систему Alien Grid с центром ядерных исследований CERN, где создается уникальный ускоритель элементарных частиц.

4. Выводы

1. За последнее двадцатилетие суперкомпьютеры во многом превратились из засекреченного средства для оборонных приложений в мощный высокоинтеллектуальный ресурс для открытых массовых применений. Суперкомпьютеры – лидеры по производительности вычислений становятся предметом престижа государств, мощным инструментарием, привлекающим лучших исследователей приоритетных отраслей.
2. Развивается международное сотрудничество в области разработок и применения высокопроизводительных систем. В центре внимания здесь технологии Grid Computing. Ряд компаний участвуют в проекте MegaGrid. В СНГ ряд организаций Москвы, Санкт-Петербурга, Киева, Харькова и других городов стремятся активнее участвовать в применениях технологий Grid при разработках ускорителя элементарных частиц в европейском центре CERN.
3. Большинство современных суперкомпьютеров использует кластерную архитектуру, например, в последнем списке TOP50 СНГ из пятидесяти лауреатов 49 имеют кластерную архитектуру. В Украине функционирует ряд кластеров, два самых высокопроизводительных из них установлены в Институте кибернетики им. В.М. Глушкова НАНУ, они включены в список Top50 под номерами 21, 35.
4. Расширяются разработки и приложения ПО для мультипроцессоров: ОС Linux (75% охвата систем), технологии передачи сообщений (стандарты MP/1.1, MP/1.2 и др.), протоколы InfiniBand, HyperTransport и др. Прежние заделы, опыт разработчиков параллельного программирования мультипроцессоров СССР [8] нередко получают эффективное развитие.
5. В числе перспективных решений элементной базы мультипроцессоров: двух- и многоядерные процессоры, ускорители приложений на реконфигурируемых ПЛИС, адаптеры и коммутаторы межсоединений узлов кластеров и др.

СПИСОК ЛИТЕРАТУРЫ

1. СуперЭВМ / Под ред. С. Фернбаха. – М.: Радио и связь, 1991. – 320 с.
2. Коваль В.Н., Савьяк В.В., Сергиенко И.В. Тенденции развития современных высокопроизводительных систем // УСиМ. – 2004. – № 6. – С. 31–43.
3. www.supercomputers.ru.
4. Корнеев В.Д. Параллельное программирование в MPI. – М.: МПИ, 2003. – 304 с.
5. Воеводин В.В., Воеводин В.В. Параллельные вычисления. – СПб.: БХВ – Петербург, 2004. – 608 с.
6. www.bm.com/news/ru/ru/2005/12/1216.htm.l.
7. Савватеев И. Dell, EMC, Intel и Oracle запускают проект MegaGrid // PC Week. – М., 2004. – № 47. – С. 9.
8. Михалевич В.С., Капитонова Ю.В., Кухарчук А.Г., Клименко В.П. и др. Структурно – архитектурные решения многопроцессорных ЭВМ с макроконвейерной обработкой данных // Электронная вычислительная техника: Сб. статей. – 1988. – Вып. 2. – С. 151–160.