

Особенности новых средств модернизации вычислительного комплекса СКИТ

Комухаев Э.И., Черепинец В.В.

Институт кибернетики им. В.М. Глушкова НАН Украины, пр. Ак. Глушкова, 40, Киев, Украина

Аннотация. *Анализируются новые hard-средства для модернизации суперкомпьютеров, выделена эффективность многоядерного сопроцессора Intel Xeon Phi, ускорителей Nvidia, IBM Cell, узлов IBM Netezza FPGA, обсуждаются варианты их интеграции при модернизации комплекса СКИТ.*

Ключевые слова

Многоядерные чипы, FPGA, твердотельные накопители, обработка данных.

1 Перспективность многоядерных структур

Институт кибернетики им. В.М. Глушкова НАН Украины развивает свои вычислительные системы высокой производительности и расширяет их применение для решения актуальных задач науки, техники, экономики, экологии, медицины [1].

В 2012 году успешно выполнена значительная модернизация комплекса СКИТ путем дополнения его новым кластером СКИТ-4, реализованным на передовой высокоинтегрированной элементной базе.

Приведем характеристики кластера СКИТ-4 [1,2]. Кластер включает 12 узлов на центральных процессорах Intel Xeon E5-2600 с частотой 2.6 ГГц, образуя 192 вычислительных ядра, 36 ускорителей NVidia Tesla M2075 и 768 ГБ оперативной памяти, достигнута пиковая производительность 25,6 ТФлопс, реальная производительность 11,8 ТФлопс.

СКИТ-4 вдвое производительней, чем СКИТ-3, меньше последнего по габаритам, имеет уровень энергопотребления 15 кВт (СКИТ-3 - 60 кВт), обеспечивая сравнительно высокую энергоэффективность - 788 МФлопс/Вт.

В СКИТ-4 применен передовой интерконнект связи узлов Infiniband FDR 56 Гбит/с, подключено хранилище данных объемом 100 ТБ.

Сегодня исследуются варианты ввода в комплекс СКИТ ряда инновационных hard-средств для обеспечения прироста производительности до уровня показателей рейтинговых суперкомпьютеров.

Сложилась весьма благоприятная ситуация для достижения большого прироста производительности, снижения удельного уровня энергозатрат в связи с доступностью инновационного 50-60 ядерного сопроцессора Intel Xeon Phi, который в тандеме с 6-8 ядерным процессором Intel Xeon E5 обеспечивает скачек роста производительности.

Intel Xeon E5 стал основным процессором узлов СКИТ-4, что упрощает подключение эффективнейшего сопроцессора Intel Xeon Phi и позволяет использовать предлагаемые компанией Intel программные инструменты Intel Parallel Studio XE и Intel Cluster Studio XE для оптимизации кодов за счет использования одинаковых языков и моделей программирования для сопроцессоров Xeon Phi и процессоров Xeon E5.

В последнем 40-вом списке всемирного рейтинга суперкомпьютеров Top500 [3] уже появился ряд систем, эффективно использующих упомянутые связки E5 с Xeon Phi. В таблице 1 приведены данные таких систем в 40-вом списке рейтинга Top500.

Таблица 1. Системы рейтинга Top500, использующие сопроцессоры Intel Xeon Phi

Позиция рейтинга, страна, компания	Название суперкомпьютера	Кол-во ядер	Типы процессоров	Типы коммутаторов	Реальная производительность, ТФлопс	Энергопотребление, кВт
7. США Dell	Stamped	204900	Xeon E5 – 2680 BC Intel Xeon Phi	Infiniband FDR	2660,3	
53. США IBM	Discover	35568	Xeon E5 – 2670 Intel Xeon Phi 5110 P	Infiniband QDR	628,8	215,6
58. США Intel	Endeavor	27489	Xeon E5 – 2670 Intel Xeon Phi	Infiniband FDR	379,3	299,9
59. Россия RSC Group	MVS-10P-RSC Tornado	28704	Xeon E5 – 2690 Intel Xeon Phi	Infiniband FDR	375,7	222,7
117. США SGI	Maia	17408	Xeon E5 – 2670 BC Intel Xeon Phi	Infiniband FDR	212,9	132
170. Россия RSC Group	RSC Tornado SUSU	14016	Xeon X5 – 680 Intel Xeon Phi	Infiniband QDR	146,8	147,5
253. США Appro Intenational	Beacon	921,6	Xeon E5 – 2670 BC Intel Xeon Phi 5110 P	Infiniband FDR	110,5	48,1

Приведенная таблица 1 строк списка 40-го рейтинга Top500, фиксирующих начало растущих использований E5 и Xeon Phi, содержит на позиции 253 также данные лидера последнего 17-того списка рейтинга Green500 [4] (такие списки составляются каждые полгода с ноября 2007 года с перечнем наиболее энергосберегающих систем мира).

Для лидирующей в этом “зеленом” рейтинге системы Beacon достигнут показатель 2,499.44 MFLOPS/W. Третью позицию 17-того списка Green500 принадлежит сегодняшнему лидеру 40-го списка рейтинга Top500 системе Titan-Cray XK7 с показателем 2,142.77 MFLOPS/W.

Эксперты уже определили лидера следующего 41-го списка рейтинга Top500 - им станет вероятнее всего суперкомпьютер XC30 компании Cray с очень большим показателем производительности за счет использования множества чипов Intel Xeon E5, Intel Xeon Phi и чипов графических ускорителей NVIDIA.

Отметим стремление компании Intel расширить в своих чипах графические функции, создав конкурентоспособную альтернативу чипам NVIDIA. К важнейшим преимуществам компании Intel относится ее обладание передовыми технологическими линиями изготовления чипов.

Переход к изготовлению высокоинтегрированных чипов по минимальным нанометровым нормам технологического процесса обеспечивает чипам прирост быстродействия и энергоэффективности, снижение габаритов и стоимости.

Компания Intel производит на своих технологических линиях сопроцессоры Xeon Phi по нормам 22 нанометра. NVIDIA является компанией типа Fabless без собственных технологических линий, она заказывает

изготовление своих передовых графических чипов тайваньской компании TSMC. Последняя добилась высокого процента выхода годных чипов NVIDIA по нормам 40 нанометров, теперь осваивает уровень их достаточно конкурентоспособного выхода по нормам 28 нанометров.

Intel уже анонсировала начало выпуска в 2013 году ряда своих чипов по нормам 14 нанометров.

2 Инновационные типы вычислительных структур

Вычислительные узлы суперкомпьютеров все чаще насыщают различными инновационными типами высокоинтегрированной элементной базы (ВЭБ) с их гибким подключением для решения различных задач, обеспечивая достижение высоких показателей производительности и энергосбережения. Ниже в таблице 2 приведены основные типы современной ВЭБ, их поставщики, передовые образцы.

Таблица 2. Типы ВЭБ

Тип ВЭБ	Ведущие поставщики	Передовые образцы	Примечания
Многоядерные процессоры	Intel, IBM, AMD, Oracle-Sun	Intel Xeon E5-2600	СКИТ-4 содержит процессоры E5
Коммутаторы вычислительных узлов	Cray, Mellanox	Чипы Infiniband FDR	СКИТ-4 содержит узлы Infiniband FDR
Ускорители вычислений	Nvidia, AMD, Intel, IBM	Nvidia Kepler, Intel Xeon Phi	СКИТ-4 содержит ускорители Nvidia
Твердотельные флэш-накопители	Intel, Toshiba, Samsung	Intel 710, 910, 335	Чипы 710 содержат серийные суперкомпьютеры Gordon (США)
ПЛИС типа FPGA	Xilinx, Altera	Virtex 5, Virtex 6, Virtex 7, Stratix 4, Stratix 5	СКИФ-Аврора (РФ) содержит чипы Stratix 4 Altera
Вычислители типа "Processor-in-memory"	IBM, Oracle, HP, SAP	IBM Netezza	В Netezza использованы чипы FPGA и CPU

Выдающимся примером суперкомпьютера с использованием большинства приведенных в таблице 2 типов ВЭБ стала новая итальянская система Euroga, нацеленная на решение задач материаловедения, астрофизики, бионаук, земледования. Эта система недавно признана наиболее энергоэффективным суперкомпьютером.

Гибкая гибридизация большинства вышеуказанных типов ВЭБ таблицы 2 в каждом узле обеспечила суперкомпьютеру Euroga компании Eurotech рекордный на сегодня в мире уровень энергоэффективности – 3150 Мфлопс на ватт. Возглавляющий на текущие полгода рейтинг Green500 суперкомпьютер Beason (см. таблицу 1) обеспечивает только 2499 Мфлопс на ватт.

Суперкомпьютер Euroga состоит из 64 крупных узлов, каждый имеет размер ноутбука с производительностью узла 1700 Гфлопс.

Узел Euroga включает: два процессора Intel Xeon E5 – 2687 W, две карты – ускорителя Nvidia Tesla K20, или два ускорителя Intel Xeon Phi, чип FPGA Altera Stratix5, 16 Гбайт ОЗУ типа DDR3 и твердотельный накопитель на 160 Гбайт [5]. Здесь сравнительные испытания эффективности приведенных вариантов ускорителей Nvidia, Intel еще не завершены.

Растущее быстроедействие суперкомпьютеров обеспечивает значительное расширение круга эффективно и своевременно решаемых задач. Например, вычисления истинных показателей погоды на конкретные сутки теряют смысл после наступления этих суток.

Важность оперативного доступа к хранилищам данных суперкомпьютеров постоянно возрастает, все чаще необходима обработка данных большого объема в реальном масштабе времени. Большие возможности для ускоренной обработки данных хранилищ предоставляют гибкие технологии IBM Netezza FPGA.

3 Гибкие технологии IBM Netezza FPGA

Лавинообразный рост объема новых данных (проблемы Big Data) повысил актуальность совершенствования разработок вычислений внутри памяти. Такого рода решения разработали и совершенствуют такие гиганты, как IBM, SAP, Oracle, HP.

По нашим оценкам, весьма перспективным стало бы подключение к системе хранилища данных вычислительного комплекса СКИТ-4 устройств IBM Netezza FPGA [6,7].

На рис.1 представлена архитектура Netezza на базе фирменных сочетаний Snipped чипов FPGA и CPU, обеспечивающих значительные ускорения обработки данных.

Архитектура Netezza AMPP™

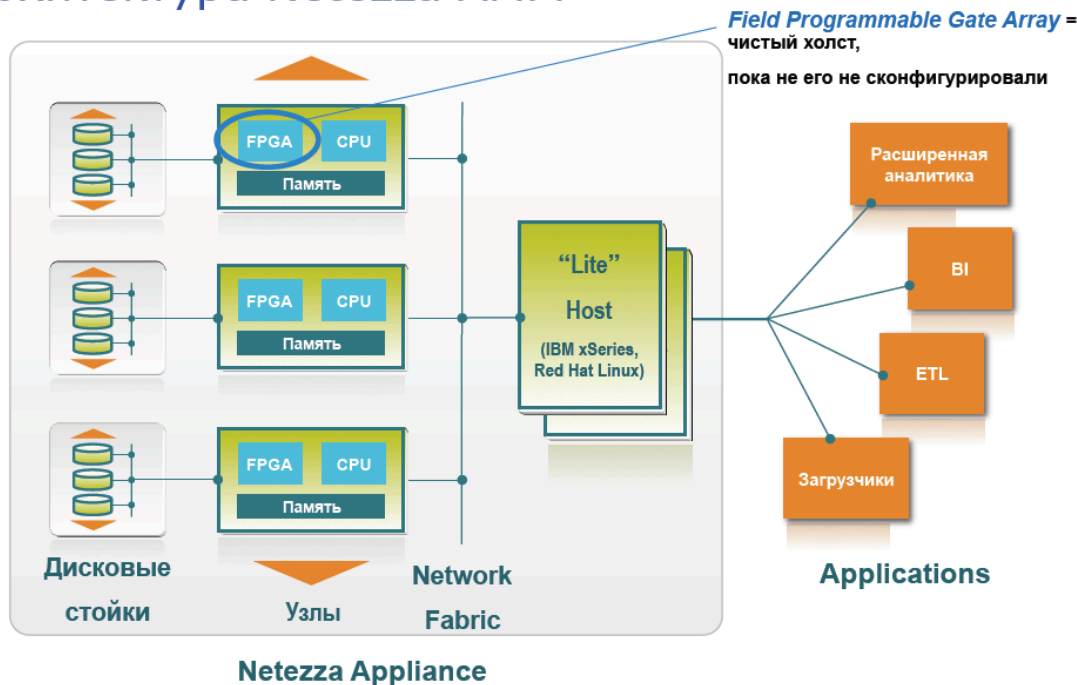


Рис.1 Архитектура Netezza

Прямое взаимодействие здесь блоков хранения и обработки данных позволяет обрабатывать данные в сжатом (архивированном) виде без предварительной декомпрессии. При этом запрос раскладывается на части, выполняемые в параллельном режиме. Эти технологии способствуют расширению режимов обработки данных в реальном масштабе времени, ускоряют решения оптимизации транспортных потоков, социальных сетей, выделения ключевых шаблонов в обслуживании клиентов, прогнозов, оценок банковских рисков.

Соединяя системы хранения данных с аналитическими функциями внутри баз данных, передовые технологии IBM Netezza Analytics реализуют встроенную высокопроизводительную аналитическую платформу с массовым параллелизмом, предназначенную для обработки петабайтных объемов данных. IBM Netezza Analytics служит для быстрого и эффективного предоставления более оптимальных и правильных ответов на сложные запросы.

IBM Netezza Analytics – мощная платформа расширенной аналитики IBM Netezza, которая поддерживает технологическую инфраструктуру для развертывания аналитических функций внутри баз данных.

Аналитическая платформа объединяет полноценный набор встроенных аналитических средств с аналитическими инструментами, такими как R and Hadoop™, а также с основными решениями для управления хранилищами данных IBM Netezza.

IBM Netezza положила начало созданию современных решений для управления хранилищами данных; ее пользователи получают различные выгоды от объединения системы хранения данных и аналитики в единую высокопроизводительную интегрированную систему.

При работе с IBM Netezza аналитические модели можно создавать и развертывать непосредственно в самих хранилищах данных. При этом существенно сокращается время, затрачиваемое на создание и развертывание аналитических моделей.

Уменьшив время от начала работы над моделью до ее развертывания, пользователи могут оперативно перейти к принятию опирающихся на факты решений, шире используя в них аналитику по запросу, предоставляющую ценные знания.

Каждое хранилище данных IBM Netezza поставляется с библиотекой аналитических функций внутри баз данных, доступ к которым предоставляется через любой SQL- совместимый интерфейс.

Заказчики могут также сами разрабатывать новые возможности при помощи определяемых расширений платформы. Это простая и открытая платформа, обеспечивающая поддержку многих инструментов, языков и структур.

IBM Netezza Analytics служит для:

- исследования и поиска данных;
- преобразования данных;
- построения моделей;
- диагностики моделей;
- оценки моделей.

Для быстрой аналитической обработки с возможностями масштабирования в IBM Netezza Analytics используется мощная платформа параллельных вычислений – решение для управления хранилищами данных IBM Netezza.

Лежащая в его основе архитектура асимметричной обработки данных с массовым параллелизмом (AMPP), обеспечивающая высокое быстродействие и производительность, позволяет выполнять аналитическую обработку внутри базы данных быстро и эффективно. Архитектура AMPP представляет собой потоковую архитектуру на основе blade-серверов.

Использование расширенной аналитики в решениях по управлению хранилищами данных IBM Netezza нацелено на решения многих запросов финансовых групп, тематических отделов, отделов продаж, служб ИТ, исполнительного руководства, работающих в различных отраслях.

Решение для управления хранилищами данных IBM Netezza ускоряет весь аналитический процесс. Программные интерфейсы и параметры распараллеливания позволяют гибко переносить в решение большую часть аналитики – как выполняемой при помощи инструментов типа IBM SPSS, SAS и R, так и написанной на языках Java™, Python™ и Fortran.

Кроме того, решения для управления хранилищами данных IBM Netezza поставляются со встроенной библиотекой аналитических функций с параллелизмом, созданных специально для больших объемов данных, которые позволяют быстро запустить и ускорить разработку и развертывание любого аналитического приложения.

Решения IBM Netezza отличает от аналогичных продуктов удобство и простота расширений. Это решение, в котором мощность и масштабируемость сотен процессоров заложена в архитектуру, подходящую для параллельной аналитической обработки.

В отличие от традиционных решений с их фрагментированной инфраструктурой аналитики, имеющей в составе множество систем, в которых выполняется репликация данных, IBM Netezza Analytics объединяет все аналитические операции в одном мощном программно-аппаратном комплексе. Он легко развертывается и не требует особого текущего администрирования, – и потому имеет низкую совокупную стоимость владения в целом.

Возможности AMPP решений IBM Netezza Analytics снимают многие традиционные ограничения аналитики. Для прогнозирования и оценок данные не нужно перемещать на другое место. Результаты прогнозных оценок формируются часто в режиме реального времени.

В решении для управления хранилищами данных IBM Netezza используются программируемые логические матрицы типа FPGA, запрограммированные для эффективной обработки больших объемов данных. Эти матрицы отсеивают посторонние данные еще на выходе с диска. Тем самым устраняются узкие места ввода-вывода и разгружаются компоненты в нисходящем направлении, такие как процессор, системы хранения и сеть, повышая производительность системы.

Сложные аналитические операции баз данных выполняются над отфильтрованным потоком данных. Аналитические задачи запускаются как независимые процессы, работающие с потоками данных, на каждом сервере S-Blade.

Некоторые решения, алгоритмы раскрываются пользователю лишь в процессе эксплуатации после приобретения систем IBM Netezza 100, IBM Netezza 1000. Эта непрозрачность связана с продолжающимися процессами патентования алгоритмов (в режимах прохождения заявок).

Очень привлекательна для пользователя предусмотренная возможность изменять алгоритмы в чипах FPGA в составе сотен специализированных лезвий Snippet Processing Unit. Компания IBM традиционно развивает гибкие, интеллектуальные решения на базе FPGA, дополняет их своими структурами, теперь становятся весьма перспективными продукты IBM Netezza 100, IBM Netezza 1000. В этих продуктах в узлах Snippet с чипами FPGA взаимодействуют процессоры IBM Power PC.

Напомним, что компания IBM уже 18 лет сохраняет статус мирового лидера по наибольшему числу ежегодно полученных патентов.

Активно развиваются и применяются все новые сочетания FPGA и CPU. В журнале Xcell [8] компании Xilinx «новой эрой» называют использование сочетаний FPGA и CPU на одном кристалле при начальных управляющих функциях CPU, которые реализованы в новейших 28 нм чипах семейства Xilinx Zynq-7000. Здесь

двухядерный процессор ARM Cortex – A9 MP Core взаимодействует с FPGA Kintex-7, причем исходные управляющие функции выполняет CPU, оптимально подключая ресурсы FPGA.

Указанная функциональная интеграция повышает быстродействие и упрощает проектирование для пользователя. Появились сообщения, что по такому же пути с доминированием CPU идут корпорация Intel и FPGA-компания Altera, создавая совместный чип, но уже на базе технологических норм 14-20 нм.

4 Выводы

1. На статус основного средства модернизации узлов кластера СКИТ-4 несомненно претендуют сопроцессоры Intel Xeon Phi, которые обладают фирменной согласованностью с процессорами Intel Xeon E5, уже введенными в узлы СКИТ-4.
2. О растущем признании эффективности сопроцессоров Intel Xeon Phi свидетельствует их оперативное использование в семи суперкомпьютерах последнего 40-го списка рейтинга суперкомпьютеров Top500, а также ввод этих сопроцессоров во все узлы системы XC30 компании Cray, которая, по признанию экспертов, будет вскоре иметь наивысший в мире показатель производительности.
3. Инновационные типы высокоинтегрированной элементной базы (ВЭБ) включают: многоядерные процессоры, коммутаторы вычислительных узлов, ускорители вычислений, твердотельные накопители, реконфигурируемые FPGA, узлы вычислений in-memory. Большинство этих типов ВЭБ эффективно взаимодействуют в узлах нового суперкомпьютера Euroga, которому теперь принадлежит мировой рекорд энергоэффективности (3150 МФлопс на ватт).
4. Ускорители вычислений все шире используют многоядерные структуры, например, в составе IBM Cell, графических чипов, чипов FPGA и др.
5. Важным критерием отбора высокоинтегрированных чипов при модернизациях систем является учет освоенных для производства их кристаллов передовых норм технологического процесса, которые обеспечивают конкурентоспособный уровень быстродействия, энергоэффективности, снижения габаритов и стоимости.
6. Упомянутые сопроцессоры Intel Xeon Phi уже производятся по передовым нормам 22 нм. Большинство высокоинтегрированных чипов пока производятся по нормам 28-45 нм.
7. Весьма перспективны для эффективных решений задач Big Data узлы вычислений in-memory, в частности, блоки IBM Netezza FPGA, которые целесообразно интегрировать в состав системы хранилища данных комплекса СКИТ.

Литература

1. А.Л. Головинський, А.Л. Маленко, І.В. Сергієнко, В.Г. Тульчинський / Энергоэффективный суперкомпьютер СКИТ-4 // Вісник Національної академії наук України, №2, 2013, с. 50-59.
2. Головинський А.Л., Маленко А.Л., Бандура О.Ю., Горенко С.О., Черепинець В.В., Ємченко І.О. / «СКИТ-4 – суперкомпьютер Інституту кібернетики імені В.М. Глушкова НАН України» // Матеріали міжнародної конференції «Високопродуктивні обчислення HPC-UA 2012», с. 149-151.
3. Рейтинг Top500: <http://www.top500.org/>
4. Рейтинг Green500: <http://www.green500.org/>
5. Intel и NVidia конкурируют в области «зеленых» суперкомпьютеров <http://servernews.ru/news/597070>
6. Выходцев А. / Платформа для Больших Данных // Открытые системы, №6, 2012, с.12-13.
7. IBM Netezza FPGA <http://www.ibm.com/>
8. Mike Santarini / Zynq 7000 EPP Sets Stage for New Era of Innovations // Xcell Journal, №2, 2011.