

СРАВНИТЕЛЬНОЕ ТЕСТИРОВАНИЕ ВЫЧИСЛИТЕЛЬНЫХ УЗЛОВ ДВУХ СУПЕРКОМПЬЮТЕРОВ

А.А. Соколов, Д.В. Андрюшин

ФГУП «НИИ «Квант», Москва

Приводятся результаты исследования производительности кластерного суперкомпьютера, развернутого для Нижегородского государственного университета им. Н.И. Лобачевского, с использованием новых процессоров Intel Sandy Bridge. Выполнено сравнение с кластерным суперкомпьютером К-100, развернутым в ИПМ им. М.В. Келдыша и созданным с использованием процессоров предыдущего поколения Intel Westmere. При проведении исследования использовались элементы оригинальной многоуровневой методики тестирования, позволяющей проводить комплексный анализ производительности вычислительных систем. Особое внимание уделено сравнению вычислительных узлов, базовых блоков суперкомпьютеров.

Введение

В 2012 году для Нижегородского государственного университета им. Н.И. Лобачевского развернут вычислительный кластер, базовым блоком которого является двух-сокетный узел с новыми 8-ми ядерными процессорами Intel Xeon E5-2660 (Sandy Bridge-EP) 2.2 ГГц. Десять узлов оснащены тремя и пятьдесят узлов оснащены двумя графическими ускорителями NVIDIA M2090 Tesla. Каждый узел снабжен 64 Гбайт оперативной памяти DDR3-1333 МГц. Все шестьдесят узлов кластера объединены вычислительной коммуникационной сетью Mellanox QDR InfiniBand, служебной сетью Gigabit Ethernet и работают под управлением ОС MS Windows Server 2008 HPC Edition.

Микропроцессор Intel Sandy Bridge-EP (далее - Sandy Bridge) содержит принципиально новые решения по микроархитектуре [1], среди которых необходимо выделить поддержку новых векторных команд AVX, четырехканальный контроллер памяти, а также наличие встроенного в кристалл контроллера PCI-E 3.0, который позволяет значительно повысить эффективность работы с GPU и коммуникационными сетями, например, InfiniBand.

В докладе представлены результаты исследования производительности кластера ННГУ с использованием элементов оригинальной базовой многоуровневой методики оценочного тестирования [2] – тесты 1, 3 и 5-го уровня. В рамках методики выполняется комплексная оценка производительности, от отдельных элементов до системы в целом. В приведенных результатах на специальных тестах уровня 1 и 3 показывается, как влияют на производительность особенности обращений к памяти при выполнении вычислений. Тесты 5-го уровня – комплексные, ядра прикладных задач. В них сочетание режимов работы оборудования определяется решаемой задачей. На этих задачах представляет интерес результат совместного влияния особенностей работы оборудования в разных режимах – достигаемая реальная производительность на тесте.

Для сравнения приведены результаты тестирования суперкомпьютера К-100, расположенного в ИПМ им. М.В. Келдыша [3]. Суперкомпьютер К-100 состоит из шестидесяти четырёх двухсокетных вычислительных узлов с 6-ти ядерными процессорами предыдущего поколения Intel Xeon X5670 Westmere 2,93 ГГц, снабженных тремя графическими ускорителями NVIDIA C2050 Fermi. Каждый узел снабжен 96 Гбайт оперативной памяти DDR3-1333 МГц. Узлы объединены вычислительными сетями

Qlogic QDR InfiniBand, МВС-Экспресс gen 2, служебной сетью Gigabit Ethernet и находятся под управлением ОС SUSE Linux Enterprise Server 11 SP1 (2.6.32.12-0.7).

При проведении исследования использовался компилятор icc версии 12.1 (кластер ННГУ) и 11.0 (К-100).

Результаты исследований

В тесте АРЕХ-МАР [4] (уровень 1) измеряется среднее время доступа вычислительных процессов к памяти в тактах или микросекундах в зависимости от пространственно-временной локализации обращений. В процессе работы выполняется последовательное считывание с шагом 1 элементов некоторого массива L , начало которого в области памяти M (ее размер сопоставим с доступной физической памятью исследуемого оборудования) определяется случайным образом и управляется параметром A . Размер L – задает пространственную локализацию обращений к памяти, меняется от 1 (худшая) до 65536 (лучшая). Параметр A задает временную локализацию, изменяется от 0.001 (лучшая), до 1.0 (худшая). Для каждого сочетания L и A тест измеряет среднее время выполнения обращения к памяти в тактах процессора. По этим измерениям строится АРЕХ-поверхность. Для анализа представляет интерес форма этой поверхности и точки сочетания предельных пространственно-временных локализаций – G , L , T и F (рис. 1а).

При рассмотрении АРЕХ-поверхностей одного ядра (рис. 1а и б) можно отметить, что Sandy Bridge демонстрирует лучшие, чем Westmere показатели работы с памятью при наличии хорошей временной или пространственной локализации (точки F , T и L), что говорит о лучшей оптимизации алгоритмов работы с кэш памятью в новой микроархитектуре. Однако в самой плохой точке – G – результат Sandy Bridge немного хуже, чем у Westmere. При рассмотрении АРЕХ-поверхностей варианта максимальной занятости вычислительных ядер (рис. 1в и г) видно общее снижение количества тактов на одно обращение, что говорит о наличии толерантности микропроцессоров к задержкам выполнения операций с памятью за счет увеличения одновременно выполняемого их числа. При этом результаты Sandy Bridge лучше результатов Westmere.

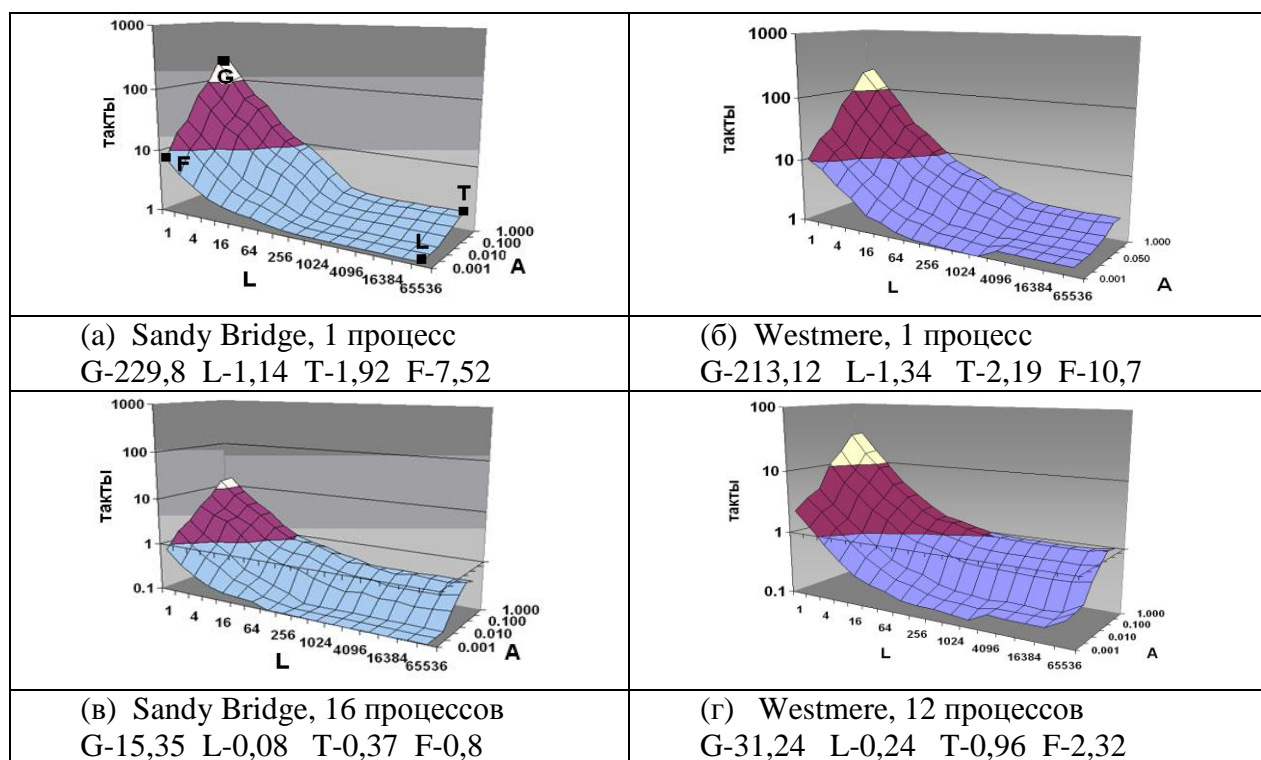


Рис. 1. Тест АРЕХ-МАР

Тесты EuroBen Benchmark [5] (уровень 3) – это 31 тест в виде операций разной сложности над векторами. Собранные специальным образом в группы они используются для испытания процессоров в условиях целенаправленно изменяемой нагрузки. В табл. 1 приведены две группы тестов. В группе 1 выполняется операция перемножения векторов с изменением режима доступа к элементам массива. В группе 2 изменяется соотношение между количеством вычислительных операций и операций обращения в память. Результаты испытаний на этих группах даны на рис.2-3.

В табл. 1 (группа 2) тест 9th DegrPolyn – это вычисление полинома 9-й степени по схеме Горнера: $y(i) = ((...(a9*x(i) + a8)*x(i) + ...) * x(i) + a0$.

Таблица 1. Группы тестов со специально изменяемой нагрузкой оборудования (область точки Т арх-мар)

Группа 1 (усложнение доступа к памяти)	Группа 2 (повышение интенсивности вычислений)
$y(i) = x1(i) * x2(i), i = 1, n$	$y(i) = x1(i) + x2(i), i = 1, n$
$y(i) = x1(i) * x2(i), i = 1,3*n,3$	$y(i) = x1(i)*x2(i) + x3(i)*x4(i)$
$y(i) = x1(i) * x2(i), i = 1,4*n,4$	$y(i) = x1(i+1) - 2*x1(i) + x1(i-1)$
$y(i) = x1(ind(i)) * x2(ind(i)), i = 1, n$	9th DegrPolyn

Анализ результатов показывает, что при выполнении операции умножения векторов при меньшей тактовой частоте процессор Sandy Bridge показывает сравнимые с Westmere показатели производительности при работе в кэш памяти L1. При усложнении режима доступа к массиву, расположенному в кэш памяти, процессор Westmere, за счет большей тактовой частоты, незначительно опережает Sandy Bridge. При работе с оперативной памятью показатели производительности процессоров выравниваются. При этом полученные результаты показывают резкое падение реальной производительности при усложнении доступа к данным из памяти (группа 1).

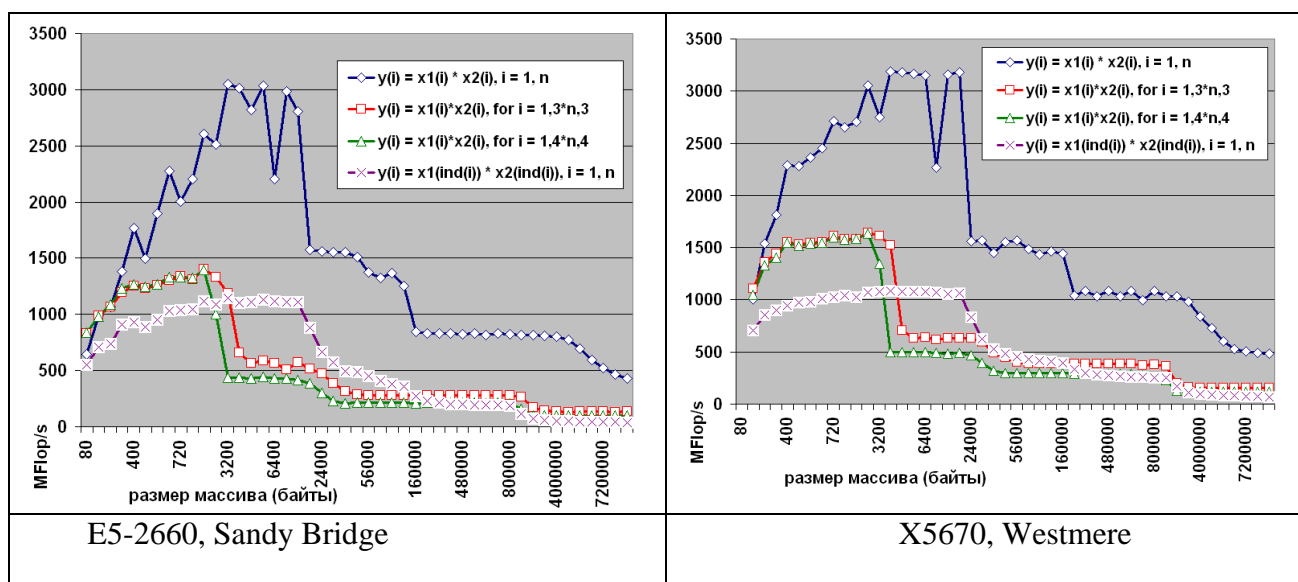


Рис. 2. Результаты измерений Euroben (группа 1)

На тесте вычислений по схеме Горнера процессор Sandy Bridge лучше процессора Westmere за счет наличия 256-разрядных устройств типа SIMD (команды AVX). На этом тесте на два обращения к памяти приходится двадцать семь вычислительных операций, и показатели производительности процессоров приближаются к пиковым значениям. Увеличение операций доступа к памяти по отношению к вычислительным опера-

циям и разбалансировка операций умножения и сложения приводит серьезному ухудшению демонстрируемой производительности для процессора Sandy Bridge, а большая тактовая частота процессора Westmere позволяет ему обогнать Sandy Bridge при работе с кэш памятью.

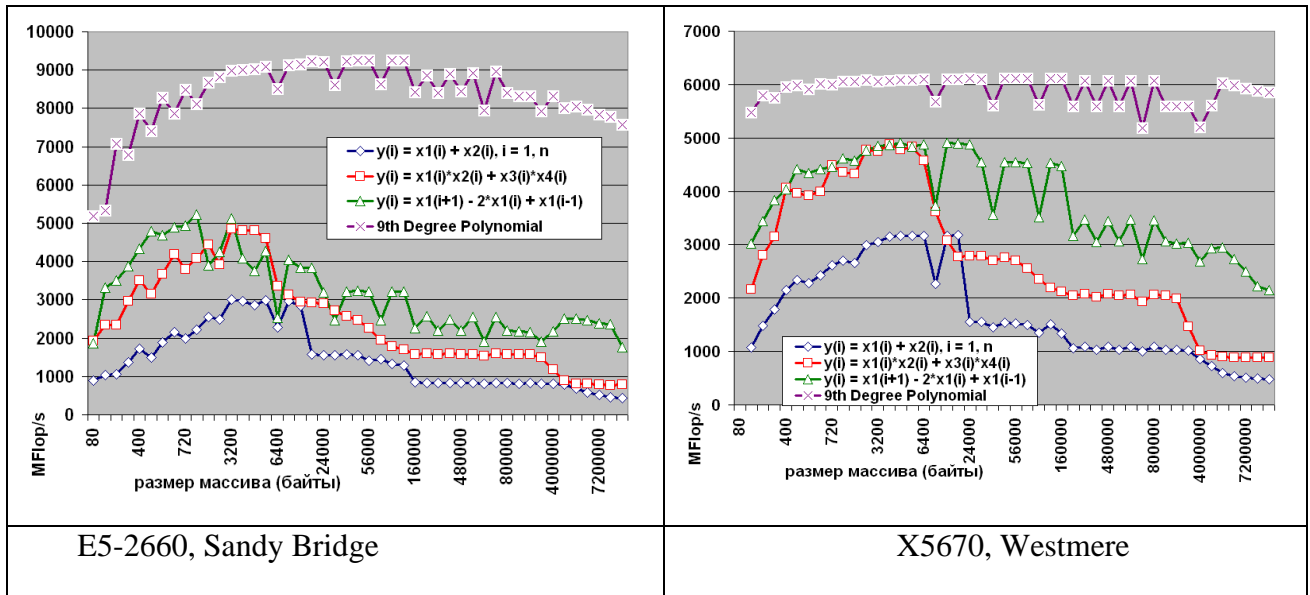


Рис. 3. Результаты измерений Euroben (группа 2)

Тесты пакета NAS Parallel Benchmark [6] (уровень 5) используются для комплексной оценки производительности. Приведены результаты для OpenMP-вариантов тестов EP, CG (класс C).

Тест EP реализует алгоритм накопления двумерной статистики для псевдослучайных чисел с распределением Гаусса, является типичным для задач, решаемых методом Монте-Карло, и позволяет оценивать производительность на задачах с интенсивными вычислениями с плавающей точкой двойной точности.

Тест CG реализует метод сопряженных градиентов для вычисления приближения наименьшего собственного значения большой разреженной симметричной, положительно определенной матрицы. Для теста CG характерно попадание в область точки G APEX-поверхности с плохой пространственно-временной локализацией обращений к памяти.

Результаты измерений на тесте EP показывают, что, несмотря на меньшую тактовую частоту, процессор Sandy Bridge демонстрирует сравнимые результаты при интенсивных вычислениях с плавающей точкой, а за счет большего количества вычислительных ядер, общая производительность узла с процессорами Sandy Bridge в 1,5 раза выше производительности узла с процессорами Westmere. Результаты измерений на тесте CG показывают, что для процессора Westmere при переходе рубежа в 10 вычислительных процессов на узел, подсистема памяти перестает справляться с потоком обращений в память, что ограничивает масштабируемость приложения. Однако при любом количестве процессов в диапазоне от 1 до 12 производительность узла с процессорами Westmere выше производительности узла с процессорами Sandy Bridge. При использовании узла с процессорами Sandy Bridge, за счет большей ширины каналов памяти и кэш, масштабируемость с ростом вычислительных узлов сохраняется, и, как следствие, общая производительность узла в 2,2 раза выше производительности узла с процессорами Westmere.

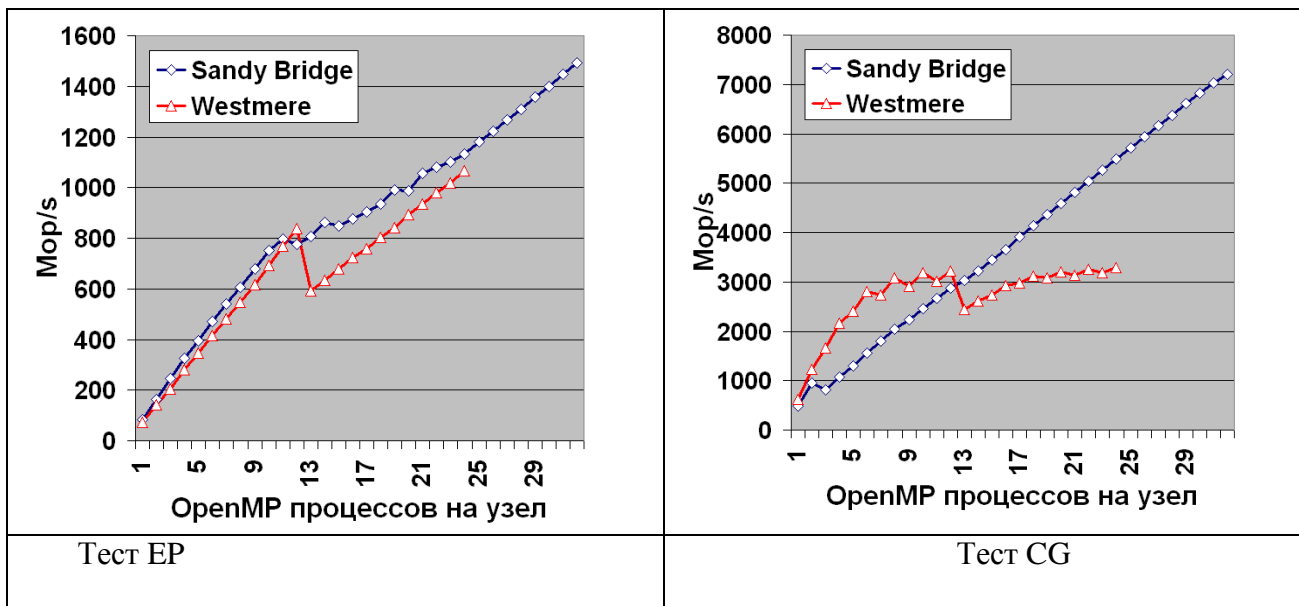


Рис. 4. Тесты пакета NAS Parallel Benchmark (сравнительные результаты)

Результаты сравнительного тестирования показывают, что новые архитектурные решения обеспечивают процессору Sandy Bridge превосходство над Westmere. Использование 256-битных блоков SIMD обработки и более широких трактов передачи данных позволяет процессору Sandy Bridge эффективнее, по сравнению с Westmere, выполнять операции с плавающей точкой, особенно в случае размещения данных в кэш памяти. Наличие 4-х канального контролера памяти в процессоре Sandy Bridge (вместо 3-х канального у Westmere) позволяет обеспечивать большую толерантность узла по обращениям к памяти и, как следствие масштабируемость приложений с ростом числа вычислительных процессов.

В настоящее время заканчиваются работы по тестированию коммуникационной сети InfiniBand и графических ускорителей кластера ННГУ. Результаты тестирования планируются привести в докладе на конференции.

Литература

1. Sandy Bridge for Servers D. Kanter, 07/08/2011, сайт <http://www.realworldtech.com/>
2. Горбунов В.С., Эйсымонт Л.К., Речинский А.В., Заборовский В.С., Забеднов П.В. Суперкомпьютеры для промышленности – вопросы тестирования, анализа и разработки // Материалы Второй Всероссийской конференции по суперкомпьютерным технологиям (СКТ2012). Дивноморское, октябрь 2012.
3. <http://www.kiam.ru/MVS/resources/k100.html>. Сетевой ресурс.
4. Strohmaier E., Shan H. Architecture independent performance characterization and benchmarking for scientific applications. 2004.
5. EuroBen Benchmark Homepage – [<http://www.hpcresearch.nl/euroben/index.php>].
6. NAS Parallel Benchmark Homepage – [<http://www.nas.nasa.gov/publications/npb.html>].