

Предварительные результаты оценочного тестирования отечественной высокоскоростной коммуникационной сети Ангара

А.А. Агарков, Т.Ф. Исмагилов, Д.В. Макагон, А.С. Семенов, А.С. Симонов

АО «НИЦЭВТ»

В статье представлены результаты сравнительного оценочного тестирования 36-узловой вычислительного кластера «Ангара-К1», оснащенного адаптерами коммуникационной сети Ангара, и суперкомпьютера МВС-10П с сетью Infiniband 4x FDR, установленного в МСЦ РАН.

Ключевые слова: высокоскоростная сеть, интерконнект, Ангара, Infiniband FDR, HPCG, HPL, NPВ, ПЛАВ

1. Введение

По статистике списка TOP500 самых мощных суперкомпьютеров мира ([1], ноябрь 2015) можно заметить, что большинство представленных в нем систем используют коммерческие сети Infiniband и Gigabit Ethernet. Однако суперкомпьютеры из первой десятки — Tianhe-2 (Китай), Cray (США), IBM Blue Gene/Q (США), K Computer (Япония) — используют собственные («заказные») коммуникационные сети, и только одна система использует коммерческую сеть Infiniband. «Заказные» сети не поставляются отдельно от вычислительной системы, а коммерческие сети далеко не всегда подходят для эффективной реализации систем с высокими требованиями по масштабируемости, надежности и производительности.

Сеть «Ангара» [2–7] — первый в России проект высокоскоростной сети с маршрутизаторами на основе СВИС ЕС8430 отечественной разработки. Микросхема ЕС8430 стала итогом семилетней работы подразделения АО «НИЦЭВТ» — разработчика высокоскоростной сети Ангара. СВИС выпущена на фабрике TSMC с использованием технологии 65 нм. Размер кристалла — 13,0×10,5 мм, количество транзисторов — 180 миллионов. СВИС работает на частоте 500 МГц и потребляет 20 Вт. Поддерживается топология сети «многомерный тор» (возможны варианты от 1D до 4D-тор), каждый сетевой узел может иметь до 8 соединений с соседними узлами.

В настоящий момент в АО «НИЦЭВТ» собран 36-узловой кластер «Ангара-К1» с коммуникационной сетью Ангара. Данная работа посвящена оценочному тестированию сети Ангара в составе данного кластера и сравнению результатов с суперкомпьютером МВС-10П, оснащенного сетью Mellanox Infiniband 4x FDR.

Статья построена следующим образом. В разделе 2 дано описание использованного оборудования и программного обеспечения. В разделе 3 приведены результаты тестирования простых коммуникационных операций. В разделе 4 представлены результаты производительности тестов HPL и HPCG, как наиболее распространенных тестов суперкомпьютеров. В разделе 5 приведены результаты тестов NPВ, представляющих собой набор часто встречающихся в практике задач и охватывающих широкий диапазон требований к коммуникационной сети. В разделе 6 приведен анализ производительности большого суперкомпьютерного приложения — модели предсказания погоды ПЛАВ.

2. Вычислительные системы

В АО «НИЦЭВТ» собран кластер «Ангара-К1», состоящий из 36 узлов, объединенных сетью Ангара. Кластер состоит из 24 узлов с двумя процессорами Intel Xeon E5-2630 (по 6 ядер, 2.3 ГГц) и 12 узлов с одним процессором Intel Xeon E5-2660 (8 ядер, 2.2 ГГц). Память

каждого узла — 64 ГБ. Узлы объединены сетью Ангара с топологией 3D-тор 3×3×4. В исследовании использовалась библиотека MPI, основанная на MPICH версии 3.0.4, а также собственная реализация библиотеки SHMEM, соответствующая стандарту 1.0 OpenSHMEM [8].

Сопоставление результатов проводилось с суперкомпьютером МВС-10П, установленным в МСЦ РАН и включающем 207 узлов, в каждом узле по два процессора Intel Xeon E5-2690 (по 8 ядер, 2.9 ГГц) и 64 ГБ памяти. Узлы объединены сетью Infiniband 4x FDR. Во время тестирования использовалась библиотека Intel MPI 14.1.0.030.

В таблице 1 приведена сводная характеристика узлов кластеров «Ангара-К1» и МВС-10П. В таблице 2 приведены основные характеристики используемых суперкомпьютеров.

Необходимо отметить, что используемые в обеих вычислительных системах процессоры фирмы Intel относятся к одному поколению Sandy Bridge, однако в узле МВС-10П находятся два процессора с частотой 2.9 ГГц, которая значительно выше, чем частота процессоров в кластере «Ангара-К1». Поэтому узел МВС-10П значительно мощнее любого из двух типов узлов кластера «Ангара-К1», что необходимо учитывать особенно при сравнении систем на прикладных тестах. Для адекватного сравнения прикладных тестов на обеих системах используется по 8 ядер; на каждом узле это значение соответствует максимальному количеству ядер в узле типа В кластера «Ангара-К1». При использовании большего числа ядер на В-узле режим работы задачи может меняться из-за использования технологии Hyper-Threading, адекватное сравнение в этом случае, например, с кластером МВС-10П вряд ли возможно.

Таблица 1. Параметры вычислительных узлов, используемых в кластере с сетью Ангара и суперкомпьютере МВС-10П.

Параметр	«Ангара-К1»		МВС-10П
	Узел типа А	Узел типа В	
Процессор	2×Intel Xeon E5-2630	Intel Xeon E5-2660	2×Intel Xeon E5-2690
Тактовая частота процессора, ГГц	2.3	2.2	2.9
Количество ядер в узле	12	8	8
Размер кэша L3, МБ	15	20	20
Память узла, ГБ	64	64	64
Пиковая производительность узла, Гфлопс	221	141	371

3. Коммуникационная задержка

Одной из важнейших характеристик сети является задержка на передачу сообщений между двумя соседними узлами. Задержка измерялась при помощи теста `osu_latency` из пакета OSU Micro-Benchmarks, версия 5.1 [9]. Результаты измерений приведены на рисунке 1.

Таблица 2. Параметры вычислительных систем

Параметр	«Ангара-К1»	МВС-10П
Количество узлов	24×А, 12×В	207 (36)
Общая пиковая производительность, Тфлопс	6.988	76.838 (13.356)
Сеть	Ангара 3D-тор 3×3×4	Infiniband 4x FDR Fat Tree

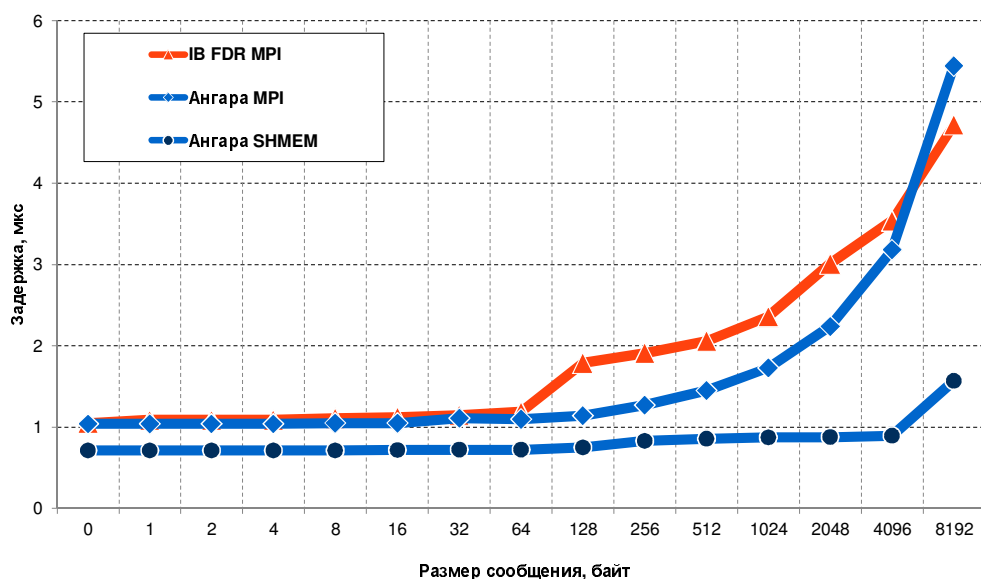


Рис. 1. Задержка передачи сообщения в сетях Ангара и Infiniband FDR в зависимости от размера сообщения.

При использовании библиотеки MPI задержка на сети Ангара меньше, чем на сети Infiniband FDR при размерах сообщений от 8 байт до 4 Кбайт. Использование библиотеки SHMEM позволяет уменьшить задержку на сети Ангара на 31% в сравнении с результатом, полученным при использовании библиотеки MPI. При этом задержка на передачу коротких сообщений с использованием SHMEM составляет 685 наносекунд, из которых 40 наносекунд уходят на обработку пакета в маршрутизаторе, 89 наносекунд — на передачу по линку, а оставшиеся 556 наносекунд — это суммарная задержка на инъекцию и эжекцию сообщения через PCI Express. Сеть Ангара имеет торовую топологию, важным параметром является задержка на каждый шаг (хоп), которая составляет 129 наносекунд.

Результаты, полученные на сети Ангара с использованием библиотеки SHMEM, значительно лучше, чем при использовании библиотеки MPI. С одной стороны, это говорит о том, что библиотека SHMEM значительно ближе к аппаратуре и вносит меньше накладных расходов чем библиотека MPI; с другой стороны, это показывает возможность оптимизации библиотеки MPI для сети Ангара.

4. Тесты HPL и HPCG

Тест HPL (High-Performance LINPACK) [11] используется для ранжирования суперкомпьютеров в списке TOP500. Тест LINPACK представляет собой решение СЛАУ $Ax = f$ методом LU -разложения, где A — плотнозаполненная матрица. Тест HPL — это реализация LINPACK на языке C для суперкомпьютеров с распределенной памятью.

Запуск теста HPL проводился с использованием 8 MPI-процессов на узел кластера. На кластере «Ангара-К1» на HPL получено 85% от пиковой производительности в расчете на 8 используемых в узле ядер, см. таблицу 3. Тест HPL содержит мало коммуникационных обменов и не представляет интереса с точки зрения сети, поэтому сравнительного измерения на МВС-10П не проводились.

Тест HPCG [12] — относительно новый тест, предназначенный для дополнения теста HPL. В основе HPCG лежит решение линейных уравнений с разреженной матрицей большой размерности при помощи итерационного метода сопряженных градиентов с многосеточным предобуславливателем. В отличие от HPL тест HPCG обеспечивает стрессовую нагрузку подсистемы памяти вычислительных узлов и коммуникационной сети, представляя значительный класс современных суперкомпьютерных приложений.

Для тестирования на обеих вычислительных системах использовалась оптимизированная реализация HPCG [13], разработанная в АО «НИЦЭВТ». Все проведенные оптимизации являются допустимыми с точки зрения спецификации теста HPCG. Данная реализация включает оптимизации уровня вычислительного узла, в том числе изменение формата хранения разреженной матрицы и векторизацию. Выполнена также оптимизация межпроцессных обменов для сети Ангара при помощи использования библиотеки SHMEM вместо библиотеки MPI. По сравнению с базовой версией теста HPCG 2.4 оптимизированная версия дает выигрыш на 36 узлах кластера практически в 2 раза.

Результаты выполнения теста приведены в таблице 3; пиковая производительность рассчитывалась исходя из вычислительной мощности 8 используемых ядер на каждой вычислительной системе. В HPCG межузловой обмен используется в двух функциях: обмен между соседними процессами при вычислении произведения разреженной матрицы на вектор и вычислении предобуславливателя, а также в функции вычисления скалярного произведения (редукция). За счет того, что размер передаваемых в тесте сообщений между соседними процессами небольшой, а задержка передачи коротких сообщений с использованием библиотеки SHMEM в сети Ангара значительно меньше, чем с использованием библиотеки MPI (см. рисунок 1), получен значительный выигрыш в производительности теста при замене библиотеки MPI на библиотеку SHMEM в обеих функциях. В итоге, с использованием SHMEM на кластере «Ангара-К1» удалось получить значительно более высокую производительность по отношению к пиковой в сравнении с достигнутой на МВС-10П.

5. Тесты NPВ

Набор тестов NAS Parallel Benchmarks (NPB) [14] является одним из самых распространенных тестов вычислительных систем. Тесты NPВ включают в себя ряд синтетических задач и псевдоприложений, эмулирующих реальные приложения в области гидро и аэродинамики. Существуют параллельные реализации тестов NPВ с использованием MPI, OpenMP, MPI+OpenMP. Сравнение вычислительных систем проводилось на MPI версии теста 3.3.1. Для всех тестов NPВ размеры обрабатываемых данных разбиваются на классы, расположенные в порядке увеличения расчетных данных: S, A, B, C, D, E. Для тестирования был выбран класс C, так как задачи этого класса уже достаточно большие, чтобы обеспечить необходимый параллелизм, но при этом влияние сети является заметным. На узле использовалось 8 MPI-процессов, максимальное количество используемых узлов — 32, оно связано с требованием, что количество используемых процессов должно быть степенью двойки.

Таблица 3. Результаты выполнения тестов HPL и HPCG

		«Ангара-К1»	МВС-10П
HPL	Тфлопс	4.44	–
	% пиковой	85 %	–
HPCG MPI	Гфлопс	279	363
	% пиковой	5.3 %	5.4 %
HPCG SHMEM	Гфлопс	342	–
	% пиковой	6.5 %	–

На рисунках 2-6 показано сравнение производительности и ускорения тестов LU, MG, CG, FT и IS на кластере «Ангара-К1» и суперкомпьютере МВС-10П в зависимости от количества вычислительных узлов. На кластере «Ангара-К1» использовалось следующее правило выбора узлов: для заданного числа узлов выбирались узлы типа В, в случае их нехватки (для конфигураций от 16 узлов) добавлялись узлы типа А.

Графики расположены в порядке убывания реальной производительности. Каждый следующий тест более требователен к коммуникационной сети, чем предыдущий. Так как узлы МВС-10П более производительны по сравнению с узлами кластера «Ангара-К1», то практически на всех тестах производительность МВС-10П выше. Поэтому для лучшего сравнения сетей кроме реальной производительности используется характеристика полученного ускорения выполнения задачи с увеличением количества вычислительных узлов.

В тесте LU решается система уравнений с равномерно разреженной блочной треугольной матрицей методом симметричной последовательной верхней релаксации к которой приводят трехмерные уравнения Навье — Стокса. Как можно видеть на рисунке 2, ускорение на кластере «Ангара-К1» совпадает с полученным на МВС-10П. Как можно видеть на рисунке 2, ускорение на кластере с сетью Ангара совпадает с полученным на МВС-10П. Это вызвано тем, что доля коммуникаций теста LU мала в сравнении с долей локальных вычислений.

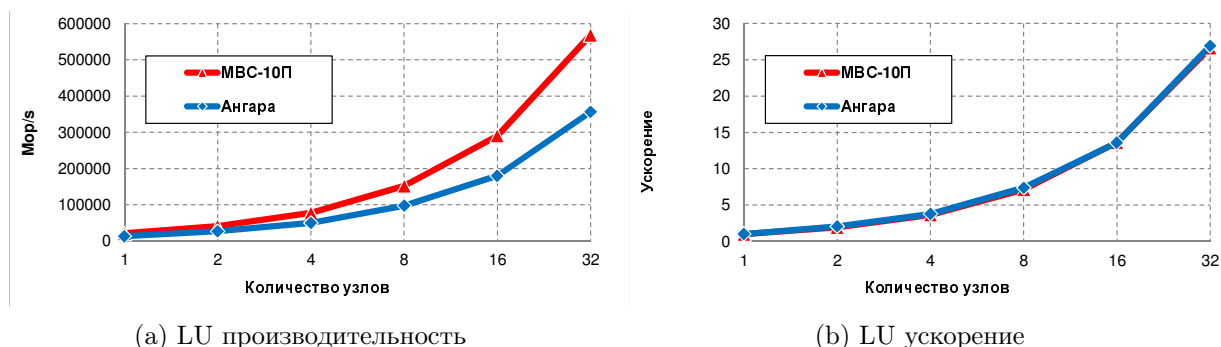


Рис. 2. Результаты выполнения теста NPV LU на кластере с сетью Ангара в сравнении с суперкомпьютером МВС-10П.

Тест MG представляет собой приближенное решение трехмерного уравнения Пуассона в частных производных на заданной сетке с периодическими граничными условиями. В этом тесте объем и интенсивность коммуникаций выше, чем в тесте LU. На 8-ми, 16-ти и

32-х узлах ускорение на кластере с сетью Ангара выше, чем на МВС-10П (см. рисунок 3).

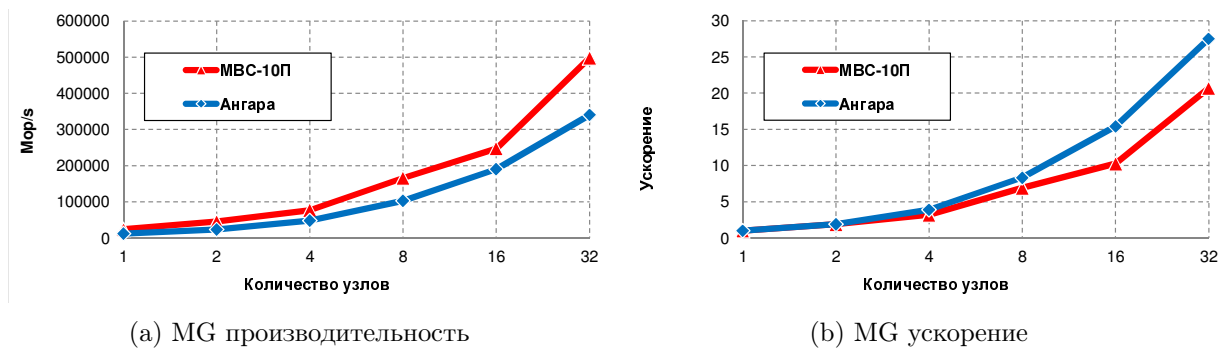


Рис. 3. Результаты выполнения теста NPV MG на кластере с сетью Ангара в сравнении с суперкомпьютером МВС-10П.

Тест FT заключается в нахождении решения уравнения в частных производных $\frac{\partial u(x,t)}{\partial t} = \alpha \nabla^2 u(x,t)$ при помощи быстрого прямого и обратного преобразования Фурье. В этом тесте появляются передачи значительных объемов данных от каждого MPI-процесса каждому. На 16-ти и 32-х узлах (см. рисунок 4) кластер «Ангара-К1» показывает лучшее ускорение, чем МВС-10П.

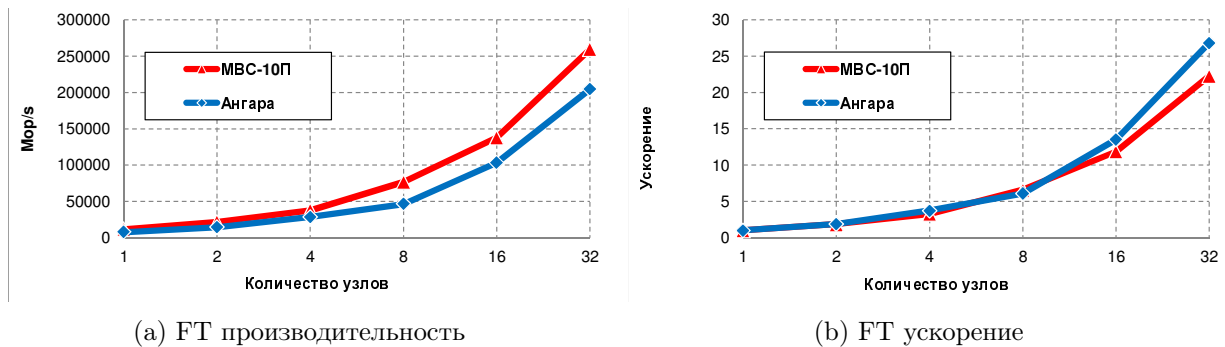


Рис. 4. Результаты выполнения теста NPV FT на кластере «Ангара-К1» в сравнении с суперкомпьютером МВС-10П.

В тесте CG вычисляется методом сопряженных градиентов наименьшее собственное значение разреженной матрицы. Основная часть обменов проходит внутри небольших групп MPI-процессов. На 8-ми и 16-ти узлах (см. рисунок 5) кластер «Ангара-К1» показывает немного худшее ускорение, чем МВС-10П, а на 32-х узлах — немного лучшее.

Тест IS выполняет распределенную сортировку N целых чисел. Доля локальных вычислений в этом тесте мала в сравнении с долей обменов. Производительность теста ограничена временем выполнения обменов типа «все-всем». Это единственный тест, на котором благодаря сети Ангара удастся получить лучшую реальную производительность на 32-х узлах, чем на кластере МВС-10П (см. рисунок 6).

6. Тестирование на модели ПЛАВ

Модель атмосферы ПЛАВ (ПолуЛагранжева, основана на уравнении Абсолютной завихренности) [15, 16] является основной моделью глобального среднесрочного прогноза погоды в России с 2010 года. Она включает в себя блок решения уравнений динамики атмосферы, алгоритмы параметризации процессов подсеточного масштаба (крупномасштабных осадков, глубокой конвекции, коротковолновой и длинноволновой радиации, пограничного

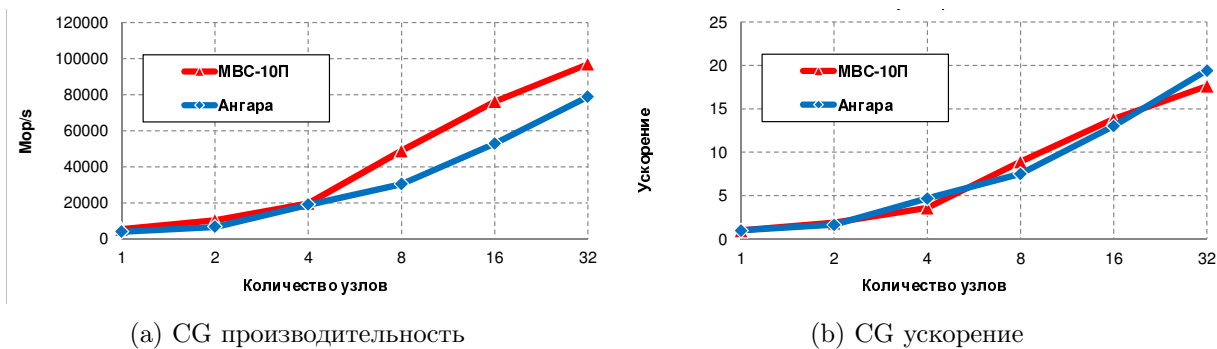


Рис. 5. Результаты выполнения теста NPV CG на кластере «Ангара-К1» в сравнении с суперкомпьютером МВС-10П.

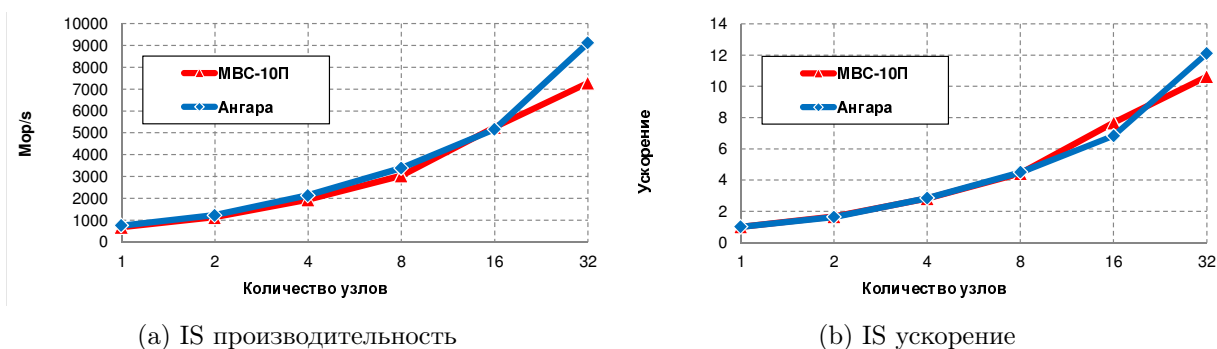


Рис. 6. Результаты выполнения теста NPV IS на кластере «Ангара-К1» в сравнении с суперкомпьютером МВС-10П.

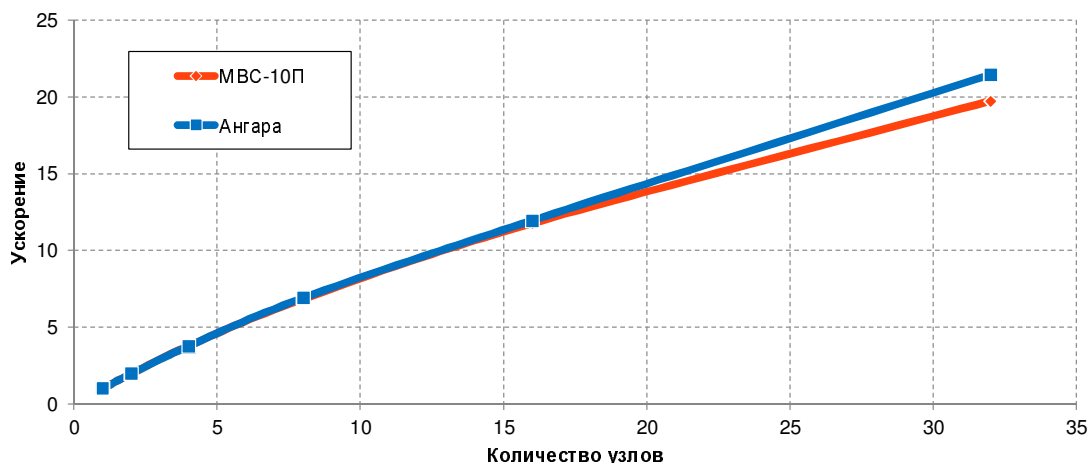


Рис. 7. Сравнение кластера «Ангара-К1» и суперкомпьютера МВС-10П на модели ПЛАВ.

слоя атмосферы, торможения гравитационных волн рельефом, модель многослойной почвы). Для распараллеливания ПЛАВ применяется сочетание библиотеки MPI и технологии OpenMP.

Модель ПЛАВ с разрешением 0.5625 градусов по долготе, переменным шагом по широте

от 0.26 до 0.9 градуса, 50 уровнями по вертикали (размерности сетки $640 \times 400 \times 50$) запущена на кластере «Ангара-К1» и суперкомпьютере МВС-10П. Данное разрешение выбрано значительно меньшим применяемого в реальных расчетах для демонстрации возможностей сетей на требуемом числе узлов. Для оценки ее производительности использовалось время, затраченное на получение прогноза на 400 часов вперед. При запуске на каждом узле кластера задействовалось 8 ядер: 4 MPI-процесса и 2 OpenMP тредра на каждый MPI-процесс. При получении графика ускорения выполнения модели ПЛАВ на кластере «Ангара-К1» использовалось то же правило выбора узлов, что и для тестов NPВ: для заданного числа узлов выбирались узлы типа В, в случае их нехватки (для конфигураций от 16 узлов) добавлялись узлы типа А.

На рисунке 7 показано сравнение ускорения, достигнутого при выполнении модели ПЛАВ, на кластере «Ангара-К1» и на суперкомпьютере МВС-10П. Время расчета прогноза ПЛАВ на одном узле кластера «Ангара-К1» составило 4697 секунд, на 32-х узлах — 217 секунд, на одном узле суперкомпьютера МВС-10П — 3411 секунд, на 32-х узлах — 173 секунды. Один из факторов, ограничивающих производительность модели ПЛАВ, — транспонирование матрицы, которое ведет к сложному для сети коммуникационному шаблону обменов «все со всеми». Таким образом, сеть Ангара обеспечивает на 9.6% более высокое ускорение при расчете ПЛАВ по сравнению с Infiniband: 21.6 раз против 19.7 на МВС-10П. При заданном разрешении модель перестает масштабироваться после 32-х узлов на обеих вычислительных системах из-за отсутствия достаточного параллелизма.

7. Заключение

В статье представлены результаты сравнительного оценочного тестирования 36-узлового вычислительного кластера «Ангара-К1», оснащенного адаптерами коммуникационной сети Ангара, и суперкомпьютера МВС-10П с сетью Infiniband 4x FDR, установленного в МСЦ РАН.

Оценочное тестирование проведено при помощи тестов разных уровней: простых коммуникационных операций, широкораспространенных тестов оценки производительности суперкомпьютеров HPL и HPCG, набора тестов NPВ уровня прикладных задач, охватывающих широкий диапазон требований к коммуникационной сети, а также на модели предсказания погоды ПЛАВ.

Тест измерения задержки передачи сообщения с использованием библиотеки MPI показывает превосходство сети Ангара над сетью Infiniband FDR при размерах сообщения от 8 байт до 4 Кбайт. Использование библиотеки SHMEM на сети Ангара позволяет получить значительно лучшие характеристики.

Для теста HPL показана возможность получения на кластере «Ангара-К1» необходимой реальной производительности. Тест HPCG предъявляет значительно более высокие требования к подсистеме памяти и коммуникационной сети, чем тест HPL. Для исследования использовалась оптимизированная авторами данной работы версия теста. Применение библиотеки MPI на кластере «Ангара-К1» в сравнении с МВС-10П позволило получить на данном тесте одинаковый уровень производительности по отношению к теоретической пиковой, а использование библиотеки SHMEM позволило значительно улучшить результат на кластере с сетью Ангара.

Рассматриваемый набор тестов NPВ включает тесты LU, MG, FT, CG, IS, рассматривается класс С. Каждый узел МВС-10П более производителен по сравнению с узлом кластера «Ангара-К1», поэтому общая производительность МВС-10П на тестах NPВ опережает «Ангара-К1». Однако по ускорению кластер «Ангара-К1» опережает на всех тестах МВС-10П с сетью Infiniband 4x FDR. Кроме того, на тесте сортировки целых числе IS за счет преимуществ сети Ангара кластер «Ангара-К1» достиг на 32-х узлах более высокой производительности, чем суперкомпьютер МВС-10П с сетью Infiniband 4x FDR.

Модель прогноза погоды ПЛАВ на небольшой расчетной сетке на кластере «Ангара-К1» показала ускорение, превышающее на 9.6% полученное на суперкомпьютере МВС-10П.

Дальнейшие работы включают в себя детальный анализ производительности тестов NPВ, оптимизацию тестов NPВ и прикладных задач при помощи библиотеки SHMEM, а также оптимизацию библиотеки MPI для сети Ангара.

Применение сети Ангара полностью охватывает суперкомпьютерный рынок, а также рынок ЦОДов. В настоящее время ведется разработка второго поколения высокоскоростной коммуникационной сети Ангара, что показывает важный для пользователей факт, что пользователям при смене поколений оборудования не придется переучиваться и привыкать к новой технологии, оптимизированные под сеть Ангара программы будут также эффективнее работать при использовании сети Ангара-2.

Авторы статьи выражают благодарность Михаилу Андреевичу Толстых и Ростиславу Фадееву за помощь в исследовании производительности модели прогноза погоды ПЛАВ.

Литература

1. Top500 Supercomputing Sites. URL: [Top500.org](http://top500.org) (дата обращения: 21.02.2016).
2. Макагон Д.В., Сыромятников Е.Л. Сети для суперкомпьютеров // Открытые системы. — 2011. — №7.
3. Корж А.А., Макагон Д.В., Жабин И.А., Сыромятников Е.Л. Отечественная коммуникационная сеть 3D-тор с поддержкой глобально адресуемой памяти для суперкомпьютеров трансгигафлопсного уровня производительности. // Параллельные вычислительные технологии (ПаВТ'2010): Труды международной научной конференции (Уфа, 29 марта-2 апреля 2010 г.): С. 227-237, Челябинск: Издательский центр ЮУрГУ, ISBN 978-5-696-03987-9, 2010.
URL: <http://omega.sp.susu.ac.ru/books/conference/PaVT2010/full/134.pdf> (дата обращения: 29.04.2015).
4. Симонов А.С., Жабин И.А., Макагон Д.В. Разработка межзвонковой коммуникационной сети с топологией «многомерный тор» и поддержкой глобально адресуемой памяти для перспективных отечественных суперкомпьютеров. // Научно-техническая конференция «Перспективные направления развития вычислительной техники», ОАО «НИЦЭВТ», 2011.
5. Симонов А.С., Макагон Д.В., Жабин И.А., Щербак А.Н., Сыромятников Е.Л., Поляков Д.А. Первое поколение высокоскоростной коммуникационной сети «Ангара» // Научные технологии. — 2014. — Т. 15, №1. — С. 21-28.
6. Слуцкий А.И., Симонов А.С., Жабин И.А., Макагон Д.В., Сыромятников Е.Л. Разработка межзвонковой коммуникационной сети EC8430 «Ангара» для перспективных суперкомпьютеров // Успехи современной радиоэлектроники. — 2012. — №1.
7. Жабин И.А., Макагон Д.В., Симонов А.С. Кристалл для Ангара // Суперкомпьютеры. — Зима-2013. — С. 46-49.
8. OpenSHMEM Application Programming Interface, Version 1.0, 31 January 2012.
URL: http://openshmem.org/site/sites/default/site_files/openshmem_specification-1.0.pdf (дата обращения: 29.11.2015).
9. OSU Micro-Benchmarks. URL: <http://mvapich.cse.ohio-state.edu/benchmarks/> (дата обращения: 29.11.2015).

10. Intel MPI Benchmarks.
URL: <https://software.intel.com/en-us/articles/intel-mpi-benchmarks> (дата обращения: 29.11.2015).
11. High-Performance LINPACK. URL: <http://www.netlib.org/benchmark/hpl/> (дата обращения: 29.11.2015).
12. M. Heroux, J. Dongarra, P. Luszczek. HPCG Technical Specification. Sandia Report SAND2013-8752. Printed October 2013.
URL: <https://software.sandia.gov/hpcg/doc/HPCG-Specification.pdf> (дата обращения: 10.06.2015).
13. А.А. Агарков, А.С. Семенов, А.С. Симонов. Оптимизация теста HPCG для суперкомпьютеров с сетью «Ангара» // Труды международной конференции (28–29 сентября 2015 г., г. Москва). — 2015. — С. 294–302.
14. NAS Parallel Benchmarks. URL: <https://www.nas.nasa.gov/publications/npb.html> (дата обращения: 29.11.2015).
15. Толстых М.А. Глобальная полулагранжева модель численного прогноза погоды. М, Обнинск: ОАО ФОП, 2010. 111 стр.
16. Толстых М.А., Мизяк В.Г. Параллельная версия полулагранжевой модели ПЛАВ с горизонтальным разрешением порядка 20 км // Труды Гидрометеорологического научно-исследовательского центра Российской Федерации. — 2011. — No 346. — С. 181–190.

Performance Evaluation of the «Angara» Interconnect.

A.A. Agarkov, T.F. Ismagilov, D.V. Makagon, A.S. Semenov, A.S. Simonov

AO «NICEVT»

The paper presents performance evaluation results of 36-nodes cluster with «Angara» interconnect compared with MVS-10P supercomputer of JSCC RAS with Infiniband 4x FDR interconnect.

Keywords: interconnect, «Angara», Infiniband FDR, NPB, HPCG, HPL, SLAV

References

1. Top500 Supercomputing Sites. URL: [Top500.org](http://top500.org) (accessed: 21.02.2016).
2. Makagon D.V., Syromyatnikov E.L. Seti dlya superkomp'yuterov [Supercomputers Interconnect]. Otkrytyye sistemy. SUBD. [Open Systems. DBMS]. — 2011. — N 7. — P. 33–37.
3. Korzh A.A., Makagon D.V., Zhabin I.A., Syromyatnikov E.L. Otechestvennaya kommunikatsionnaya set' 3D-tor s podderzhkoy global'no adresuyemoy pamyati dlya superkomp'yuterov transpetaflopsnogo urovnya proizvoditel'nosti [Russian 3D-torus Interconnect with Support of Global Address Space Memory]. Parallelnye vychislitelnye tekhnologii (PaVT'2010): Trudy mezhdunarodnoj nauchnoj konferentsii (Ufa, 29 marta – 2 aprelya 2010) [Parallel Computational Technologies (PCT'2010): Proceedings of the International Scientific Conference (Ufa, Russia, March, 29 – April, 2, 2010)]. Chelyabinsk, Publishing of the South Ural State University, 2010. P. 527–237.
URL: <http://omega.sp.susu.ac.ru/books/conference/PaVT2010/full/134.pdf> (accessed: 29.04.2015).
4. Simonov A.S., Zhabin I.A., Makagon D.V. Razrabotka mezhuzlovoy kommunikatsionnoy seti s topologiyey «mnogomernyy tor» i podderzhkoy global'no adresuyemoy pamyati dlya perspektivnykh otechestvennykh superkomp'yuterov [Development of the Multi-Dimensional Torus Topology Interconnect with Support of Global Address Space Memory for Advanced National Supercomputers]. Nauchno-tekhnicheskaya konferentsiya «Perspektivnyye napravleniya razvitiya vychislitel'noy tekhniki» (Moskva, 28 iyunya) [Scientific and Technical Conference «Advanced Directions of the Computers Development Technology». — Moscow: JSC «Concern «Vega», 2011. — P. 17–19
5. Simonov A.S., Makagon D.V., Zhabin I.A., Shcherbak A.N., Syromyatnikov E.L., Polyakov D.A. Pervoye pokoleniye vysokoskorostnoy kommunikatsionnoy seti «Angara» [The First Generation of Angara High-Speed Interconnect]. Naukoyemkiye tekhnologii [Science Technologies]. — 2014. — V. 15, N 1. — P. 21–28.
6. Slutskin A.I., Simonov A.S., Zhabin I.A., Makagon D.V., Syromyatnikov E.L. Razrabotka mezhuzlovoy kommunikatsionnoy seti YES8430 «Angara» dlya perspektivnykh superkomp'yuterov [Development of ES8430 Angara Interconnect for Future Russian Supercomputers]. Uspekhi sovremennoy radioelektroniki [Progress of the Modern Radioelectronics]. — 2012. — N 1. — P. 6–10.
7. Zhabin I.A., Makagon D.V., Simonov A.S. Kristall dlya Angary [Angara Chip] // Superkomp'yutery [Supercomputers]. — Winter-2013. — P. 46–49.

8. OpenSHMEM Application Programming Interface, Version 1.0, 31 January 2012.
URL: http://openshmem.org/site/sites/default/site_files/openshmem_specification-1.0.pdf (accessed: 29.11.2015)
9. OSU Micro-Benchmarks. URL: <http://mvapich.cse.ohio-state.edu/benchmarks/> (accessed: 29.11.2015).
10. Intel MPI Benchmarks.
URL: <https://software.intel.com/en-us/articles/intel-mpi-benchmarks> (accessed: 29.11.2015).
11. High-Performance LINPACK. URL: <http://www.netlib.org/benchmark/hpl/> (accessed: 29.11.2015).
12. M. Heroux, J. Dongarra, P. Luszczek. HPCG Technical Specification. Sandia Report SAND2013-8752. Printed October 2013.
URL: <https://software.sandia.gov/hpcg/doc/HPCG-Specification.pdf> (accessed: 10.06.2015).
13. Agarkov A.A., Semenov A.S., Simonov A.S. Optimizaciya testa HPCG dlya superkomp'yuteroj s set'yu «Angara» [Optimized Implementation of HPCG Benchmark on Supercomputer with "Angara" Interconnect]. // Superkomp'yuternye dni v Rossii: Trudy mezhdunarodnoj konferencii (28-29 sentyabrya 2015 g., g. Moskva) [Russian Supercomputing Days: Proceedings of the International Conference (Moscow, Russia, September 28-29, 2015.)]. — 2015. — С. 294-302.
14. NAS Parallel Benchmarks. URL: <https://www.nas.nasa.gov/publications/npb.html> (accessed: 29.11.2015).
15. Tolstykh M.A. Global'naya polulagranzheva model' chislennogo prognoza pogody [Global Semi-Lagrangian Model Numerical Weather Prediction Model]. M, Obninsk: OAO FOP, 2010. P. 111.
16. Tolstykh M.A., Mizyak V.G. Parallel'naya versiya polulagranzhevoj modeli PLAV s gorizontalmym razresheniem poryadka 20 km [Parallel Implementation of the Semi-Lagrangian Model SLAV with Resolution about 20 km] // Trudy Gidrometeorologicheskogo nauchno-issledovatel'skogo centra Rossijskoj Federacii [Proceedings of the Hydrometeorological Center of Russian Federation]. 2011. No 346. P. 181-190.