

## **Предварительные результаты оценочного тестирования отечественной высокоскоростной коммуникационной сети «Ангара»**

Михеев В.А., Изгалин С.П., Слуцкий А.И., Симонов А.С., Леонова А.Е., Куштанов Е.Р.,  
Жабин И.А., Щербак А.Н., Макагон Д.В., Сыромятников Е.Л., Казаков Д.В., Фролов А.С.,  
Мошкин Д.В., Хахаев А.В., Десяткова Т.А., Семовских В.Г., Бирюков С.А., Семенов А.С.  
ОАО «НИЦЭВТ»

В 2013 году ОАО «НИЦЭВТ» был представлен СБИС [1-2] отечественной высокоскоростной сети «Ангара», предназначенной для объединения узлов в вычислительных системах среднего и высшего диапазона производительности. Основным режимом параллельного программирования для данной сети является совместное использование MPI, OpenMP и SHMEM. Сеть поддерживает топологию «многомерный тор» с числом измерений до четырёх включительно, детерминированную и адаптивную передачу пакетов, односторонние операции с удалённой памятью (запись, чтение, атомарные операции). Аппаратно поддерживаются механизмы синхронизации и коллективные операции. На канальном и сетевом уровнях реализованы отказоустойчивые протоколы передачи данных.

СБИС EC8430 выпущена на фабрике TSMC с использованием технологии 65 нм. Размер кристалла – 13,0×10,5 мм, количество транзисторов – 180 миллионов. Кристалл размещён в корпусе flip-chip BGA, имеет 1521 вывод в виде массива 39×39 контактов с шагом 1 мм, размер подложки – 40×40 мм. СБИС работает на частоте 500 МГц, пиковое потребление – 36 Вт. В рамках поддерживаемой топологии 4D-тор каждый сетевой узел может иметь до 8-ми соединений с соседними узлами, взаимодействие СБИС с вычислительным узлом осуществляется через интерфейс PCI Express Gen2 x16.

В сентябре 2014 года был собран тестовый кластер из 12 узлов, оснащённый одновременно двумя сетями: сетью Infiniband FDR 4x (IB FDR) (однопортовые адаптеры ConnectX-3 VPI) и ВКС «Ангара» (EC8430). Пиковая производительность кластера составляет 1689,6 Гфлопс. В каждом узле находится 8-ядерный процессор Intel Xeon E5-2660, 2,2 ГГц и 32 ГБ памяти. Результаты запуска синтетических тестов и некоторых прикладных задач с использованием этих двух сетей представлены в настоящем докладе.

На рис. 1 представлена задержка (мкс) передачи сообщения в зависимости от длины сообщения для сети IB FDR и сети «Ангара» (EC8430). Видно, что при использовании библиотеки SHMEM получены значительно лучшие результаты для сети «Ангара» в сравнении с использованием MPI на IB FDR. Задержка с использованием SHMEM на сети «Ангара» составляет 693 нс.

На рис. 2 приведены результаты выполнения теста IMB Barrier на сетях IB FDR и «Ангара» в сравнении с выполнением барьерной операции при использовании библиотеки SHMEM на сети «Ангара». Сеть «Ангара» показывает себя лучше, чем IB FDR, в особенности – при использовании библиотеки SHMEM.

В табл. 1 показаны результаты выполнения тестов HPL и HPCG на сетях «Ангара» (EC8430) и IB FDR. HPCG использовался в стандартной комплектации, без дополнительных

оптимизаций. Сети «Ангара» и IB FDR показывают близкие результаты. Увеличение размера кластера позволит лучше оценить разницу в производительности сетей на этих тестах.

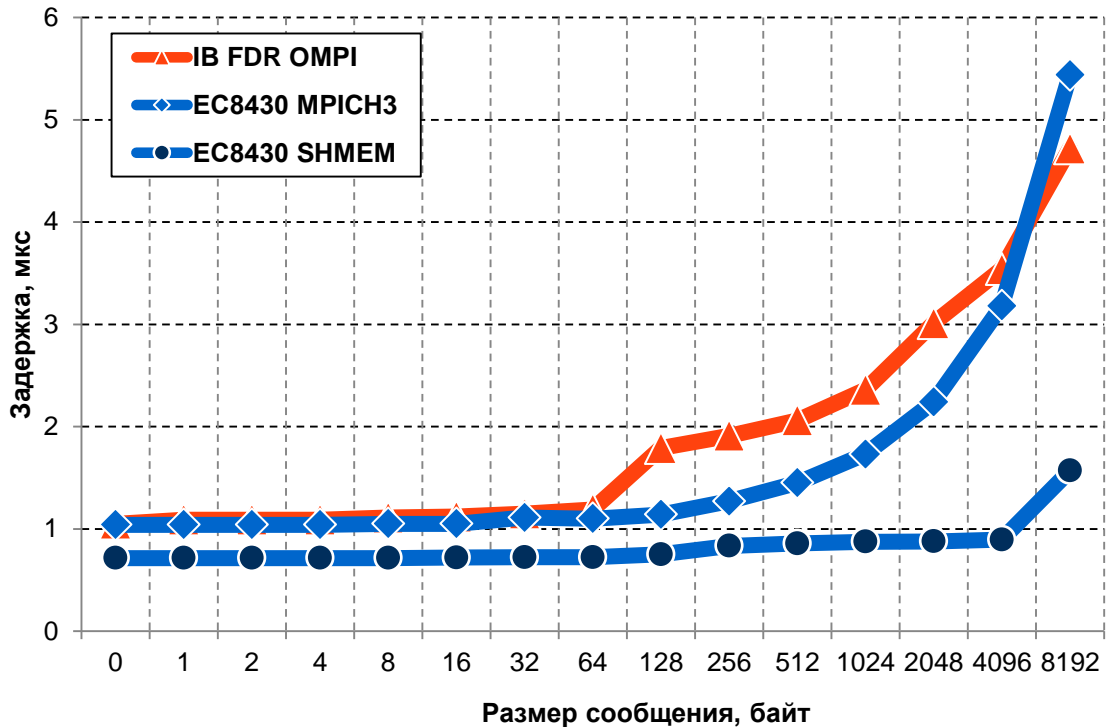


Рис. 1. Задержка (мкс) передачи сообщения в зависимости от длины сообщения для сети IB FDR и сети «Ангара» (EC8430) для библиотек MPI-1 и SHMEM.

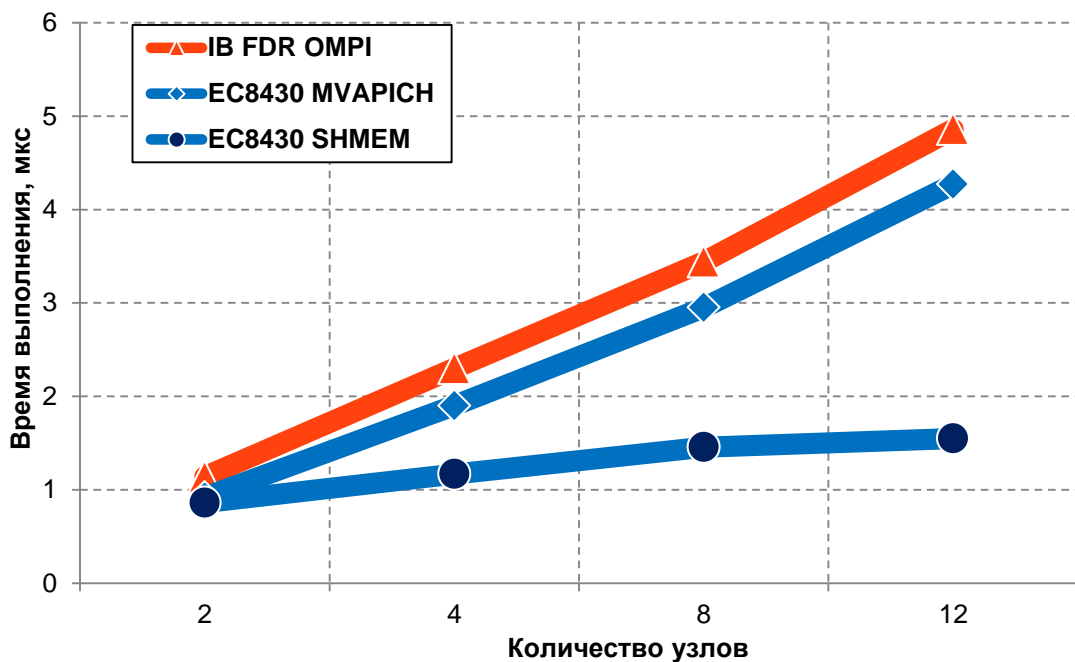


Рис. 2. Результаты выполнения теста IMB Barrier на сетях IB FDR и «Ангара» в сравнении с выполнением барьерной операции при использовании библиотеки SHMEM на сети «Ангара».

На рис. 3 показано ускорение времени счета модели ПЛАВ [3-4] в зависимости от количества ядер кластера для сетей IB FDR и «Ангара» (EC8430). Размерность задачи 640×400×50. Сеть «Ангара» показывает немного лучшее ускорение, однако в кластере не так много вычислительных узлов, чтобы можно было делать более общие выводы. С учетом полученных результатов с использованием библиотеки SHMEM (см. рис. 1, рис. 2) можно сделать предположение о возможности получения значительно лучших результатов производительности на модели ПЛАВ на сети «Ангара», так как в модели ПЛАВ есть вариант реализации на SHMEM.

Таблица 1. Результаты выполнения тестов HPL и HPCG на сетях «Ангара» (EC8430) и IB FDR.

Тест		EC8430	IB FDR
HPL	Тфлопс	1,46	1,44
	% пиковой	86,8 %	85,3 %
HPCG	Гфлопс	39	39
	% пиковой	2,3 %	2,3 %

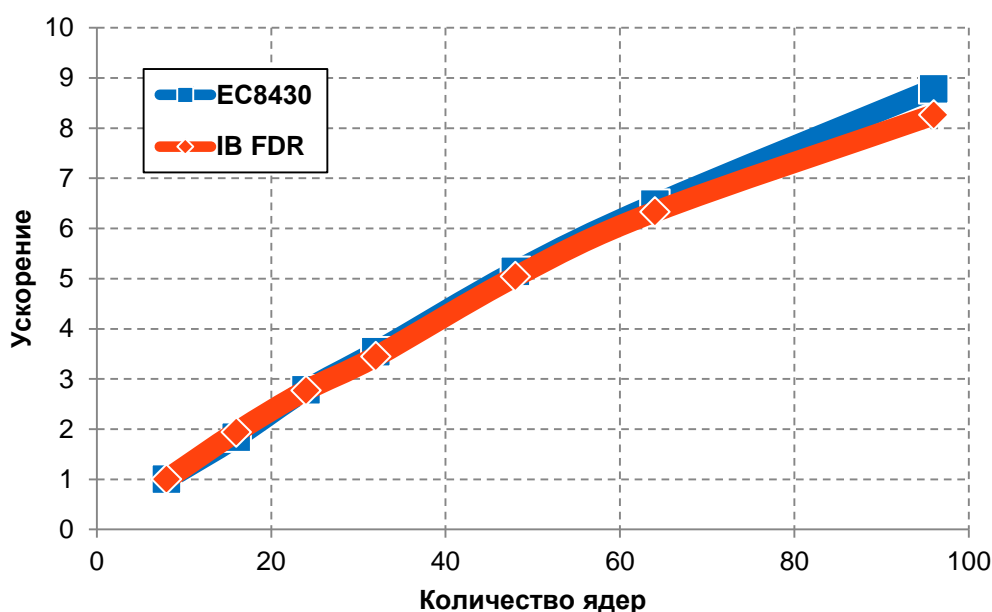


Рис. 3. Ускорение времени счета модели ПЛАВ в зависимости от количества ядер кластера для сетей IB FDR и «Ангара» (EC8430).

Сеть «Ангара» показывает характеристики производительности не хуже, чем сеть Infiniband FDR, а на некоторых тестах и реальных приложениях ведет себя даже лучше. Показанные результаты производительности являются предварительными, так как предстоит работа по оптимизации и настройке как системного ПО сети «Ангара», так и тестов-бенчмарков и используемых приложений. Стоит ожидать, что с использованием библиотеки SHMEM сеть «Ангара» будет позволять задачам показывать лучшую производительность в сравнении с сетью Infiniband FDR.

## Литература

1. И. Жабин, Д. Макагон, А. Симонов, Е. Сыромятников, А. Фролов, А. Щербак. Кристалл для «Ангары» // Суперкомпьютеры. – 2007. – №4 (16). – С. 46-49.
2. А. А. Корж, Д. В. Макагон, И. А. Жабин, Е. Л. Сыромятников. Отечественная коммуникационная сеть 3D-тор с поддержкой глобально адресуемой памяти для суперкомпьютеров транспетафлопсного уровня производительности // Параллельные вычислительные технологии (ПаВТ'2010): Труды международной конференции (Уфа, 29 марта 2 апреля 2010 г.). Челябинск: Издательский центр ЮУрГУ. – 2010. – С. 227-237.
3. Володин Е.М., Толстых М.А. Параллельные вычисления в задачах моделирования климата и прогноза погоды // Вычислительные методы и программирование. – 2007. – Т. 8. – С. 113-122.
4. Толстых М.А. Глобальная полулагранжева модель численного прогноза погоды. – М.-Обнинск: ОАО ФОРС, 2011. – 111 с.