

УДК 05.13.00

## **Направления развития высокоскоростной коммуникационной сети для многопроцессорных вычислительных систем**

**Симонов А.С.\*, Семенов А.С.\*\*, Макагон Д.В.\*\*\***

*Научно-исследовательский центр электронной вычислительной техники,*

*Варшавское шоссе, 125, Москва, 117587, Россия*

*\*e-mail: [simonov@nicevt.ru](mailto:simonov@nicevt.ru)*

*\*\*e-mail: [semenov@nicevt.ru](mailto:semenov@nicevt.ru)*

*\*\*\*e-mail: [makagond@nicevt.ru](mailto:makagond@nicevt.ru)*

*Статья поступила 18.10.2019*

### **Аннотация**

Для современных массово-параллельных суперкомпьютеров межузловая коммуникационная сеть является ключевым компонентом. Её характеристики и функциональные возможности во многом определяют показатели масштабируемости производительности при решении прикладных задач. В статье рассматриваются подходы, направления совершенствования и требования к перспективным коммуникационным сетям для многопроцессорных вычислительных систем с высокой масштабируемостью производительности, выработанные на основе опыта, полученного при разработке высокоскоростной коммуникационной сети Ангара.

**Ключевые слова:** многопроцессорные вычислительные системы, коммуникационная сеть, масштабируемость производительности.

Для современных массово-параллельных суперкомпьютеров межузловая коммуникационная сеть является ключевым компонентом. Её характеристики и функциональные возможности во многом определяют показатели масштабируемости производительности при решении прикладных задач. В АО «НИЦЭВТ» разработана высокоскоростная коммуникационная сеть (ВКС) Ангара первого поколения [1-4], предназначенная для использования в составе многопроцессорных вычислительных систем (МВС) петафлопсного класса и обеспечивающая высокую масштабируемость производительности при решении реальных задач.

Основные тенденции в развитии ВКС направлены на достижение характеристик, позволяющих обеспечить эффективную масштабируемость производительности МВС субэкзафлопсного и экзафлопсного, а в перспективе и трансэкзафлопсного диапазонов производительности. Для этого необходимо обеспечить последовательное улучшение основных характеристик ВКС по пропускной способности и коммуникационной задержке таким образом, чтобы поддерживать рост производительности процессоров и обеспечить возможность создания сбалансированных МВС. В этой связи, в дальнейшем развитии ВКС можно выделить следующие основные тенденции:

- увеличение связности;
- применение оптических соединений;
- интеграция сетевого адаптера (NIC) ВКС с процессором;

– универсализация высокоскоростных коммуникационных сетей для расширения применения на массовом рынке: поддержка протоколов 10G, 40G, 100G, 200G, 400G Ethernet.

Актуальными направлениями исследований в области архитектуры высокоскоростных коммуникационных сетей являются:

1. Разработка топологий (модифицированные сети Кэли [5], dragonfly [6], slim fly [7]) и механизмов маршрутизации, хорошо подходящих для решения целевых задач и обладающих достаточным потенциалом масштабируемости для построения систем субэкзафлопсного, экзафлопсного и трансэкзафлопсного уровня производительности.
2. Максимальная интеграция NIC с процессором и контроллером памяти, а в перспективе, со встроенным в процессор контроллером кэш-памяти, обеспечение аппаратных возможностей поддержки интеграции с сетями Ethernet.
3. Разработка эффективных алгоритмов отказоустойчивости, поддержки сохранения и восстановления состояния параллельно выполняющихся задач из контрольных точек.
4. Разработка гибких алгоритмов адаптивной маршрутизации с механизмом предотвращения перегрузки каналов связи.
5. Разработка эффективных алгоритмов поддержки коллективных и синхронизационных операций, активных сообщений.
6. Поддержка средств аппаратной виртуализации и интеграции с другими устройствами (GP GPU, устройствами хранения данных).

Для создания передового сетевого оборудования, ориентированного на пропускные способности каждого лейна 50 Гбит/с и выше, в перспективе будут активно использоваться оптические коммуникационных технологий. Ключевыми преимуществами сетевого оборудования, построенного на их основе, являются:

- возможность использования кабелей произвольной длины;
- высокая пропускная способность каналов передачи данных;
- высокая надёжность и безопасность;
- малые габариты и масса кабельной сети.

Использование оптических модулей, монтируемых непосредственно на плату, обеспечит возможность создания каналов передачи данных с пропускными способностями 300 Гбит/с и выше. Развитие данного направления и использование активных оптических кабелей обеспечит возможность расположения серверов на значительном расстоянии как друг от друга, что позволит сократить диаметр сети и повысить плотность компоновки оборудования, а так же может быть эффективно использовано при создании распределенных информационно управляющих систем.

Основным недостатком использования волоконно-оптических каналов передачи данных является весьма высокая стоимость трансиверов. Активное внедрение оптических технологий 25/40/100/200/400 Гбит/с Ethernet должно привести повышению доступности перечисленных технологий и позволит снизить стоимость коммуникационного оборудования.

Дальнейшее развитие ВКС Ангара направлено на достижение конкурентоспособных характеристик, позволяющих обеспечить эффективную масштабируемость производительности МВС субэкзафлопсного и экзафлопсного диапазона производительности, а также на получение технического задела, необходимого для создания ВКС для МВС трансэкзафлопсного диапазона производительности.

Основные характеристики каждого поколения на момент вывода в серию должны быть оптимизированы под создание МВС среднего/высшего диапазона производительности, сбалансированных по показателям пиковой производительности и пропускной способности сети. При этом создание элементной базы для каждого поколения должно осуществляться с темпом не реже, чем один раз в 3-4 года на современных на тот момент и отработанных технологических нормах производства СБИС.

Для построения МВС субэкзафлопсного уровня производительности, сбалансированной по производительности вычислительных узлов и пропускной способности коммуникационной сети, необходима разработка сетевого оборудования с характеристиками, соответствующими показателям производительности перспективных образцов процессоров вычислительных узлов.

Оценку показателя сбалансированности между производительностью вычислительных узлов и пропускной способностью коммуникационной сети целесообразно провести путём сопоставления соответствующих характеристик наиболее производительных вычислительных систем из первой десятки списка

TOP500 (табл. 1) [8-10], а также динамики роста производительности микропроцессоров.

Таблица 1.

Оценка показателей сбалансированности характеристик наиболее производительных вычислительных систем

	<b>CRAY XK7</b>	<b>CRAY XC30</b>	<b>CRAY XC40</b>	<b>IBM Sequoia BG/Q</b>	<b>Fujitsu K computer</b>
Interconnect	Gemini [11]	Aries [12]	Aries	Custom torus [13]	Tofu [14]
Nnodes	18688	5586	1248+565	98304	88128
Processor	AMD Opteron 6274, 16 cores, 2.2 GHz	2x Intel Xeon E5-2695v2, 12 cores, 2.40 GHz	2x Intel E5-2695v4, 18 cores, 2.1 GHz	PowerBQC, 16 cores, 1.6 GHz	SPARC64 VIIIfx 8C 2 GHz
L3 cache size, MB	2 x 8	30	45	32 MB (L2)	6
Node memory, GB	32 GB + 6 GB	64	128 (+64 GB)	16	16
CPU peak performance (CPU Pp), Gflops	141	461	604,8	204,8	128
Peak performance, Tflops	27113	2574	2193	20132659,2	11280
Linpack Performance, TFlop/s	17590	1655	1896	17173,2	10510
BW Memory peak, GB/s	102,4	119,4	153,6	42,7	64
BW Interconnect (BW ic), GB/s	9,6	15,4	15,4	2	10
Balance (BW ic / CPU Pp)	0,068	0,033	0,025	0,010	0,078

Оценку требуемых значений показателя сбалансированности между производительностью вычислительных узлов и пропускной способностью ВКС целесообразно провести путём сопоставления соответствующих характеристик наиболее производительных вычислительных систем из первой десятки списка TOP500 [5-7], а также динамики роста производительности микропроцессоров. Проведённый анализ показал, что для поддержания приемлемого масштабирования производительности необходимо обеспечить значение показателя сбалансированности  $\gamma \in (0,01 \div 0,08)$  ГБайт/с на ГФлоп/с.

Учитывая резкий рост производительности процессоров, достигнутый в 2016-2017 гг., с последующей стабилизацией в 2018 г., разумно предположить, что на рубеже 2019-2020 гг. ситуация существенным образом не изменится. Из этого следует, что для поддержания приемлемого масштабирования производительности МВС субэкзафлопсного класса пропускная способность линков ВКС  $V_L \in (100 \frac{\text{Гбит}}{\text{с}} \div 800 \frac{\text{Гбит}}{\text{с}})$ .

План развития высокоскоростной коммуникационной сети Ангара с топологией многомерный тор составлен с учётом сложившихся тенденций в развитии микропроцессоров.

Второе поколение высокоскоростной коммуникационной сети с топологией многомерный тор, разработка которого планируется к завершению в 2019 г., предназначено для создания МВС субэкзафлопсного диапазона производительности и характеризуется поддержкой топологии модифицированный тор и существенно более высокими характеристиками NIC, позволяющего подключать до 4-х процессоров на каждом узле.

Третье поколение высокоскоростной коммуникационной сети Ангара, разработку которого планируется осуществить в 2021-2023 гг., ориентировано на создание МВС экзафлопсного диапазона производительности и характеризуется еще большей пропускной способностью линка, обеспечиваемой с использованием новых технологий в создании высокоскоростных приёмопередатчиков, поддержкой произвольных топологий и дальнейшим улучшением характеристик NIC.

Опыт, полученный при разработке первого поколения коммуникационной сети, позволил определить направления дальнейшего эволюционного развития и совершенствования аппаратуры и программного обеспечения для коммуникационной сети Ангара следующего поколения.

Так, в ходе разработки спецификации коммуникационной сети Ангара второго поколения, помимо улучшения эксплуатационных параметров, акцент был сделан на существенное расширение функциональных возможностей, направленных на обеспечение эффективной работы не только в сегменте высокопроизводительных МВС, но и в МВС, адаптированных для решения смежных задач.

Не секрет, что в России рынок сетевого оборудования для сегмента High Performance Computing (HPC) составляет всего 1% от общего объёма рынка сетевого оборудования для вычислительных систем, и этот рынок пока ещё зависит от государственных инвестиций в создание высокопроизводительных вычислительных систем. По этой причине продолжение разработки линейки высокоскоростного сетевого оборудования, применение которого ограничено лишь сегментом HPC, по экономическим соображениям бесперспективно. Именно поэтому расширение областей применения создаваемого сетевого оборудования столь важно для обеспечения дальнейшего устойчивого развития направления в целом.

В ходе внедрения первого поколения высокоскоростной коммуникационной сети с топологией многомерный тор был разработан целый пул технологий,

обеспечивающих возможность создания на её основе прикладных систем и эффективное решение задач в смежных областях с учётом преимуществ, достигнутых в сегменте НРС.

Несмотря на значительную избыточность, ставший привычным в сегменте Ethernet стек протоколов TCP/IP остаётся обязательным элементом МВС, применяемых в центрах обработки данных (ЦОД), АСУ различного назначения и специализированных МВС. Его поддержка необходима как для реализации основных функций прикладных систем, построенных на базе данного стека протоколов, так и для решения задач интеграции между сегментами АСУ, использующими технологии НРС и Ethernet.

В этой связи в ходе работ по расширению областей применения первого поколения высокоскоростной коммуникационной сети с топологией многомерный тор были проведены работы по разработке бэкэнда для стека протоколов TCP/IP, с учётом ограничений, позволившего обеспечить возможность параллельного выполнения приложений на МВС, построенной на основе коммуникационной сети Ангара. Достигнутые характеристики показали, что даже с учётом существенного увеличения накладных расходов, формируемого уровнями стека протоколов, реальная пропускная способность, достигнутая с использованием коммуникационной сети Ангара, была существенно выше, чем у сопоставимых решений, предлагаемых поставщиками сетевого оборудования Ethernet.

Одним из наиболее востребованных сегментов, где применение высокоскоростной коммуникационной сети может дать значительный

синергетический эффект, являются системы хранения и обработки Больших данных. Для наиболее популярного фреймворка обработки Больших Данных Apache Spark в рамках работ по расширению сфер применения первого поколения коммуникационной сети Ангара разработан проект бэкэнда, цель которого – обеспечить достижение конкурентоспособных показателей по скорости работы по сравнению с типовой реализацией поверх стека протоколов TCP/IP. Ключевая особенность сети Ангара, позволяющая реализацию такого бэкэнда – технология прямой записи в память удаленного узла RDMA. Поддержка этой технологии необходима и втором поколении коммуникационной сети Ангара.

Для применений как в сегменте НРС, так и в ЦОД, важно обеспечение максимально эффективного использования ресурсов МВС при её использовании в режиме разделения между множеством пользователей. Изначально, с целью упрощения аппаратной части NIC, в микроархитектурную спецификацию СБИС первого поколения коммуникационной сети Ангара было заложено приемлемое для суперкомпьютеров высшего диапазона производительности допущение о том, что вычислительные узлы МВС передаются задаче полностью. Это допущение позволило значительно упростить аппаратуру трансляции адресов NIC, ведь в этом случае NIC должен был обрабатывать лишь обращения в два адресных пространства – решаемой задачи и операционной системы.

Однако рост числа ядер, доступных на одном процессоре, привел к тому, что данное допущение привело к неэффективному использованию доступных ядер и МВС в целом. Очевидным решением данной проблемы является, во-первых,

поддержка нескольких адресных пространств приложений, а во-вторых, наличие эффективной поддержки виртуализации [15]. По этой причине в спецификацию второго поколения внесены дополнения, обеспечивающие возможность эффективной реализации механизмов виртуализации сетевого оборудования Ангара в исполняемых на вычислительных узлах виртуальных машинах (поддержка режима SR-IOV стандарта PCI Express [16]). Указанные дополнения включают поддержку одновременного выполнения до 16 виртуальных функций с динамическим разделением ресурсов СБИС, а также механизмов синхронизации виртуальных функций и физической. Таким образом, каждый узел с сетью Ангара может поддерживать до 16 гостевых операционных систем.

Одним из наиболее сложных классов задач, решаемых на МВС, являются задачи параллельной обработки графов [17-19]. Известно, что даже при использовании различных оптимизаций данный класс задач является одним из наиболее трудных для получения высоких показателей производительности. В силу оптимальной поддержки в архитектуре сети Ангара передачи коротких и среднего размера сообщений, представляется перспективным её применение при решении сложных информационно-аналитических задач, в частности для обработки Больших Данных, обладающих сложной семантической структурой, представляемой в виде графа.

Построение параллельных перспективных библиотек и языков программирования, учитывающих особенности как архитектуры, так и программного обеспечения сети Ангара, является задачей, которая решается в

настоящее время разработчиками сети. В частности, были реализованы набор библиотек для работы с графами с использованием различных средств распараллеливания (MPI [20], SHMEM [21], Charm++ [22]), а также предложена оптимизированная версия предметно-ориентированного языка программирования Green-Marl [23].

Разработка второго поколения коммуникационной сети Ангара выполняется путём эволюционного развития созданного ранее первого поколения путём введения улучшений, необходимых для обеспечения конкурентоспособных характеристик и расширения областей применения. Ключевыми нововведениями в микроархитектуре маршрутизатора коммуникационной сети второго поколения являются:

- повышенная пропускная способность линков, обеспечиваемая применением высокоскоростных приёмопередатчиков, работающих в режиме NRZI и обеспечивающих бюджет потерь не менее 35 dB с использованием блока коррекции ошибок FEC;

- реализация в рамках одного маршрутизатора двух архитектурно независимых блоков маршрутизации Net\_IF, что позволило обеспечить баланс по пропускным способностям линков и внутренних каналов связи без увеличения ширины внутренних шин данных до 256 бит;

- поддержка 4x NIC и, соответственно, интерфейсов PCI Express последнего поколения в одной СБИС, что необходимо для эффективной поддержки узлов, включающих 2 или 4 многоядерных процессора, а также различных вариантов

вычислительных узлов с гибридной архитектурой, включающих суперскалярные процессоры, массово-параллельные (графические) процессоры, а также вычислительные поля на основе FPGA;

– аппаратная поддержка виртуализации, обеспечивающая незначительную деградацию характеристик коммуникационной сети при её использовании в среде виртуальных машин, что даёт возможность построения высокопроизводительных МВС с высоким коэффициентом использования;

– более эффективные алгоритмы маршрутизации, обеспечивающие возможности построения сетей с топологией «модифицированный тор» и задания произвольного порядка направлений. Кроме того, нововведения коснулись и алгоритмов адаптивной маршрутизации, обеспечивающих зональную адаптивную маршрутизацию, необходимую для более эффективного обхода перегруженных участков;

– улучшены механизмы справедливого арбитража, введены новые механизмы предотвращения голодания (credit aggregation) на основе анализа данных о свободном месте в нескольких виртуальных каналах;

– для улучшения поддержки стека протоколов TCP/IP и исключения дублирования пакетов на уровне аппаратуры и стека протоколов во втором поколении введен механизм негарантированной доставки пакетов (unreliable delivery), обеспечивающий их автоматическое удаление в случае, если время доставки превысит заданный лимит;

- для поддержки современных и перспективных многоядерных процессоров в состав маршрутизатора введена аппаратная поддержка 128 независимых инжекционных конвейеров;

- для обеспечения возможности использования узлов МВС в режиме разделения ресурсов между несколькими приложениями и запуска нескольких задач на пересекающихся множествах узлов в аппаратную часть NIC введен механизм защиты памяти от пакетов сторонних задач (jobID);

- для повышения эффективности работы бэкэнда синхронных операций MPI и стека протоколов TCP/IP в маршрутизатор была добавлена аппаратная поддержка механизма уведомления об успешном выполнении операций на основе суффиксов;

- для более эффективной поддержки параллельных приложений были введены атомарные операции с возвратом значений;

- для обеспечения эффективной поддержки библиотеки MPI в маршрутизаторе реализована аппаратная поддержка кольцевых буферов приёма пакетов в памяти узла;

- поддержана более современная и эффективная технология прерываний MSI-X.

Кроме того, в рамках проработки и уточнения спецификации второго поколения коммуникационной сети Ангара были проведены дополнительные работы, целью которых являлось уточнение спецификации и функциональных возможностей. Среди них наиболее значимыми являются:

- анализ сценариев использования технологии виртуализации SR-IOV;

- анализ сценариев отказов оборудования и механизмов их обработки;
- исследование возможности работы под ОС на базе QNX;
- анализ вариантов реализации параллельной файловой системы Lustre [24];
- разработка системы бездисковой загрузки узлов кластера (для макетирования и тестирования под ОС с разными конфигурациями);
- разработка системы автоматизированного пакетирования системного ПО под дистрибутивы различных ОС;
- разработка системы автоматизированного тестирования успешности сборки системного ПО под различные ОС;
- портирование библиотеки параллельной передачи данных libfabric [25].

Микроархитектура ядра маршрутизатора коммуникационной сети Ангара второго поколения представляет собой реализацию разработанной RTL-модели маршрутизатора, представляющей собой хорошо параметризованную блочную структуру, которая может быть конфигурирована под другие области применения и обеспечивает варьирование числа и ширины линков, числа интерфейсов с процессором.

Для поддержки нескольких интерфейсов с процессорами в микроархитектуру введены 4 NIC, включающие EndPoint-контроллеры интерфейса PCI Express. Вариант с одним интерфейсом PCI Express планируется задействовать при использовании СБИС в составе встраиваемого в вычислительные узлы

адаптера коммуникационной сети в форм-факторе стандартной карты расширения PCI Express. Варианты с двумя и четырьмя интерфейсами планируется задействовать при использовании СБИС в составе заказных двух- или четырёхпроцессорных серверных вычислительных платформ.

Кроме того, для поддержки использования СБИС в составе коммутаторов реализована работа маршрутизатора в режиме транзитного узла без задействования интерфейсов PCI Express. Конфигурирование СБИС в этом случае осуществляется встроенным сервисным процессором.

Учитывая необходимость обеспечения взаимодействия между процессорами, расположенными в одном узле, например, при отправке сообщения от одного процессора другому, ядро маршрутизатора обеспечивает возможность перехвата управления инжекционными конвейерами одного PCIe другим PCIe.

Для обеспечения масштабируемости по пропускной способности линков без изменения формата пакетов в состав маршрутизатора введена новая структурная единица Net\_IF, включающая линки и кроссбар. Масштабирование пропускной способности линков может быть обеспечено простым увеличением числа данных блоков без изменения форматов пакетов.

Реализация механизма уведомлений и подтверждений с использованием суффиксов потребовало введение в структуру маршрутизатора специального виртуального канала vcs0, в дополнение к реализованным в первом поколении виртуальным каналам запросов vcd0 и ответов vcd1 и vca0. Таким образом, в рамках коммуникационной сети обеспечивается разделение соответствующих

потоков данных и исключением взаимной блокировки на уровне библиотек параллельного программирования и прикладного программного обеспечения.

В отличие от первого поколения, в котором была обеспечена возможность подключения внешнего сервисного процессора, во втором поколении соответствующее процессорное ядро интегрировано в СБИС и, в зависимости от варианта встроенного программного обеспечения (firmware), может выполнять функции тестирования, конфигурирования, отладки, поддержания целостности и обеспечения отказоустойчивости коммуникационной сети.

С целью профилирования работы маршрутизатора в ходе эксплуатации в наиболее важных точках прохождения трафика установлены программно доступные счётчики производительности. Кроме того, для работы в составе систем специального назначения, функционирующих в расширенном диапазоне температур, введена функция настройки частоты работы ядра маршрутизатора.

Перспективные образцы сетевого оборудования, разрабатываемые на основе СБИС коммуникационной сети второго поколения, включают следующие варианты исполнения:

- адаптер в форм-факторе карты расширения PCI Express;
- коммутатор на 24/48/72/96 портов для установки в стандартную 19-дюймовую стойку;
- заказные варианты исполнения мезонинных модулей расширения для использования в составе blade-серверов крупных компаний-интеграторов.

За счёт перечисленных нововведений планируется расширение сфер применения сетевого оборудования второго поколения, в том числе для создания:

- кластеров высокой доступности, в том числе на основе ПК СВ Брест;
- отказоустойчивых кластеров;
- систем хранения данных на основе технологии NVMe over Ангара;
- вычислительных полей в центрах обработки данных.

Следующее, третье поколение коммуникационной сети Ангара, планируемое к разработке в 2020-2023 гг., будет продуктом дальнейшего эволюционного развития второго поколения. При этом с целью значительного повышения связности сети и поддержки произвольной, в том числе иерархической, топологии, предполагается существенно увеличить число внешних линков в каждом маршрутизаторе вплоть до 24-36 шт.

Варианты исполнения сетевого оборудования предполагается расширить за объединения на единой подложке кристаллов микропроцессоров, памяти и коммуникационной сети.

Таким образом, разработка последующих поколений высокоскоростной коммуникационной сети Ангара выполняется путём эволюционного развития созданного ранее первого поколения путём введения улучшений, необходимых для обеспечения конкурентоспособных характеристик и расширения функциональных возможностей и областей применения.

**Библиографический список**

1. Симонов А.С., Жабин И.А., Макагон Д.В. Разработка междузловой коммуникационной сети с топологией «многомерный тор» и поддержкой глобально адресуемой памяти для перспективных отечественных суперкомпьютеров // Научно-техническая конференция «Перспективные направления развития средств вычислительной техники»: тезисы докладов. – М.: НИЦЭВТ, 2011. С. 17 - 19.
2. Слущкин А.И., Симонов А.С., Жабин И.А., Макагон Д.В., Сыромятников Е.Л. Разработка междузловой коммуникационной сети ЕС8430 «Ангара» для перспективных российских суперкомпьютеров // Успехи современной радиоэлектроники. 2012. № 1. С. 6 - 10.
3. Жабин И.А., Макагон Д.В., Поляков Д.А., Симонов А.С., Сыромятников Е.Л., Щербак А.Н. Первое поколение высокоскоростной коммуникационной сети «Ангара» // Научноёмкие технологии. 2014. Т. 15. № 1. С. 21 - 27.
4. Жабин И.А., Макагон Д.В., Симонов А.С., Сыромятников Е.Л., Фролов А.С., Щербак А.Н. Кристалл для «Ангары» // Суперкомпьютеры. 2013. № 4 (16). С. 46 - 49.
5. Heydemann M.C. Cayley graphs and interconnection networks, Graph symmetry, Springer, Dordrecht, 1997, pp. 167 - 224.
6. Kim J. et al. Technology-driven, highly-scalable dragonfly topology // International Symposium on Computer Architecture, IEEE, 2008. pp. 77 - 88.

7. Besta M., Hoefler T. Slim fly: A cost effective low-diameter network topology // Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, IEEE Press, 2014, pp. 348 - 359.
8. Rogers J. Power efficiency and performance with ORNL's cray XK7 Titan // High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion, IEEE, 2012, pp. 1040 - 1050.
9. Cordery M.J. et al. Analysis of Cray XC30 performance using Trinity-NERSC-8 benchmarks and comparison with Cray XE6 and IBM BG/Q // International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems, Springer, Cham, 2013, pp. 52 - 72.
10. Список TOP500, июнь 2018 г. URL: <https://www.top500.org/list/2018/06/>
11. Vishnu A., ten Bruggencate M., Olson R. Evaluating the potential of Cray Gemini interconnect for PGAS communication runtime systems // 19th Annual Symposium on High Performance Interconnects, IEEE, 2011, pp. 70 - 77.
12. Faanes G. et al. Cray cascade: a scalable HPC system based on a Dragonfly network // Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, IEEE, 2012, pp. 1 - 9.
13. Chen D. et al. The IBM Blue Gene/Q interconnection network and message unit // Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis, IEEE, 2011, pp. 1 - 10.
14. Ajima Y., Sumimoto S., Shimizu T. Tofu: A 6d Mesh/Torus interconnect for exascale computers // IEEE Computer, 2009, no 11 (42), pp. 36 - 40.

15. Кондрашин М.А., Арсенов О.Ю., Козлов И.В. Применение технологии виртуализации и облачных вычислений при построении сложных распределенных моделирующих систем // Труды МАИ. 2016. № 89. URL: <http://trudymai.ru/published.php?ID=73411>
16. PCI Special Interest Group, available at: <http://www.pcisig.com/home>
17. Lumsdaine A. et al. Challenges in parallel graph processing // Parallel Processing Letters, 2007, no. 1 (17), pp. 5 - 20.
18. Sengupta D. et al. Graphin: An online high performance incremental graph processing framework //European Conference on Parallel Processing, Springer, Cham, 2016, pp. 319 - 333.
19. Mazeev A., Semenov A., Simonov A. A Distributed Parallel Algorithm for the Minimum Spanning Tree Problem // International Conference on Parallel Computational Technologies, Springer, Cham, 2017, pp. 101 - 113.
20. MPI: A Message-Passing Interface Standard, Version 3.1, available at: <https://www.mpi-forum.org/docs/mpi-3.1/mpi31-report.pdf>
21. OpenSHMEM Application Programming Interface, Version 1.4, available at: [http://www.openshmem.org/site/sites/default/site\\_files/OpenSHMEM-1.4.pdf](http://www.openshmem.org/site/sites/default/site_files/OpenSHMEM-1.4.pdf)
22. Charm++ Parallel Programming Framework, available at: <http://charmplusplus.org/>
23. Hong S. et al. Green-Marl: a DSL for easy and efficient graph analysis // ACM SIGARCH Computer Architecture News, 2012, no. 1 (40), pp. 349 - 362.
24. Lustre parallel filesystem, available at: <http://lustre.org/>
25. Libfabric OpenFabrics, available at: <https://ofiwg.github.io/libfabric/>