

К.т.н. И.Н. Бычков, И.А. Молчанов, д.т.н., проф. Ю.С. Рябцев
(ПАО «ИНЭУМ им.И.С. Брука», АО «МЦСТ»)

I. Bychkov, I. Molchanov, Y. Ryabtsev

РАЗВИТИЕ КОНСТРУКЦИЙ МНОГОПРОЦЕССОРНЫХ СИСТЕМ

DEVELOPMENT OF DESIGNS FOR MULTI PROCESSOR SYSTEMS

Приведены результаты исследования завершённых разработок многопроцессорных систем высокой производительности на основе многоядерных процессоров архитектуры «Эльбрус». Выбраны направления улучшения технических характеристик при проектировании многопроцессорных систем общего и специального назначения. Предлагается проведение одновременного конструкторско-технологического проектирования многоядерных процессоров, вычислительных модулей и базовых конструкций, сокращающее время выхода на рынок многопроцессорных систем при повышении их конкурентоспособности. Рассмотрены разработки импортозамещающих многопроцессорных систем на основе новых процессоров архитектуры «Эльбрус». Предложены вычислительные модули многопроцессорной крейтовой системы унифицированной конструкции и многопроцессорная система из этих модулей, которая масштабируема по количеству вычислительных ядер, оперативной памяти и периферийных интерфейсов.

In this work, results of investigation of existing high-performance multiprocessor systems based on multi-core CPUs with «Elbrus» architecture are discussed. Parameters to improve while designing general purpose and specific purpose multiprocessor systems are chosen. The co-design process of technological and mechanical features for multi-core CPUs, modules and computing systems is proposed. Such co-design allows shortening design time of final system and enhancing its economical properties. Additionally, an overview of multiprocessor systems based on CPUs with «Elbrus» architecture is presented. Uniform modules for multiprocessor systems based on industry standards are proposed. Slot-based systems consisting of these modules, with ability to be scaled according to requirements of CPU cores, RAM size, peripheral interface types and quantity, are presented.

Ключевые слова: многоядерный процессор, многопроцессорная система, крейтовые системы, масштабируемость, унификация.

Keywords: multi-core processor, multiprocessor system, slot-based systems, scalability, unification.

Введение

В современных условиях для развития импортозамещающей вычислительной техники необходимо появление многопроцессорных систем (в т.ч. серверного типа) с конкурентоспособными технико-экономическими характеристиками. Эта необходимость обусловлена следующими факторами:

- ограничением поставок импортной высокопроизводительной вычислительной техники, предполагаемой к применению в ряде ответственных областей;
- возможностью выполнения широкого класса приложений;
- высокой прибыльностью поставок серверов;
- практикой продажи персональных компьютеров ниже себестоимости при их совместной поставке с серверами;
- перспективностью применения при создании высокопроизводительных кластеров, центров обработки данных, суперкомпьютеров.

При создании многопроцессорных систем необходимо рассматривать значительное число технических характеристик, определяющих выбор направления разработки [1, 2]. Основными из них являются производительность и наличие распространенных интерфейсов пользователя. Немаловажными являются также характеристики, специфичные для модернизации, проводимой с целью адаптации к новым условиям использования, а также возможность построения новой системы на базе компонентов уже выполненной разработки. В качестве таких характеристик при конструкторско-технологическом проектировании используются показатели унификации и масштабируемости. Проектирование, ставящее целью получить необходимое качество этих характеристик, применительно к многопроцессорным системам включает создание базовых конструкций составных вычислительных модулей и периферийных устройств.

Для многопроцессорных систем типичным является горизонтальное масштабирование – возможность добавлять к системе новые узлы, модули, многоядерные процессоры для увеличения функциональности и общей производительности. Можно выделить следующие варианты такого масштабирования по количеству:

- многоядерных процессоров и оперативной памяти в системе с неравномерным доступом к общей памяти (NUMA – Non-Uniform Memory Access);
- многопроцессорных модулей в кластере, объединенных высокоскоростной сетью;
- периферийных устройств.

Наряду с этим, при использовании стандартных базовых конструкций вычислительных модулей, периферийных устройств и шасси для их установки допускается вертикальное масштабирование – возможность по мере роста требований или развития новых поколений процессоров и соответствующих им периферийных контроллеров заменять компоненты в существующей многопроцессорной системе более производительными и функциональными.

В последнее время широко распространены микросхемы с интегрированными периферийными контроллерами (South Bridge). Эти микросхемы содержат наборы контроллеров для распространенных интерфейсов. Применение таких микросхем и интерфейсных разъемов в составе вычислительного модуля придает ему функциональность компьютера на плате (on-board computer), что экономически целесообразно при создании современных многопроцессорных систем.

1. Результаты исследования завершенных разработок

При исследовании ставилась цель проанализировать достоинства и недостатки конструкций, в которых выпускаются основные классы многопроцессорных систем: стоечные системы (Rackmount Servers), системы типа «Евромеханика» (Eurocard slot-based systems) и системы лезвий (Blade servers). В качестве представителей определенного класса были выбраны, соответственно, изделия MB4C-CPB разработки ПАО «ИНЭУМ им. И.С. Бру-

ка», БВ631 разработки АО «МЦСТ» и «Лезвие-БНК» разработки АО «Т-Платформы».

Многопроцессорная система MB4C-CPB в стандартном корпусе для установки в стойку (Rackmount Server) представлена на рис. 1. Данный модуль позволяет конструировать на его основе многопроцессорные системы высотой от 1U для обеспечения возможности их компактного размещения в стойке. Тем не менее, с целью использования всех периферийных разъёмов для системы был выбран стандартный корпус высотой 3U. На рис. 1 слева представлен одноименный четырёхпроцессорный модуль форм-фактора SWTX с двумя периферийными разъёмами PCI 2.3 32 бита, 33/66 МГц и двумя разъёмами PCI Express 1.0 x8, один из которых позволяет устанавливать платы расширения форм-фактора до x16.

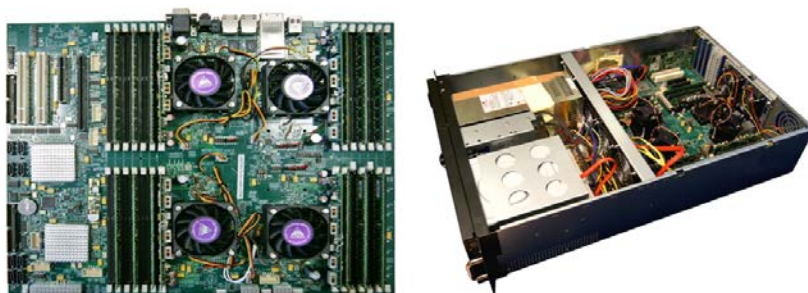


Рисунок 1. Многопроцессорная система MB4C-CPB (справа) и вычислительный модуль (слева) на основе четырех процессоров «Эльбрус-4С»

Важным достоинством модуля и многопроцессорной системы MB4C-CPB является низкая себестоимость за счет применения коммерческих компонентов, в т.ч. корпуса с резервируемыми блоками питания. Многопроцессорные системы на основе таких модулей могут включать популярные на рынке периферийные устройства и иметь при этом различные назначения в составе кластера или автоматизированной системы. Недостатком такой конструкции является необходимость ее доработки для специальных применений с обеспечением работоспособности в условиях внешних факторов, включая выпадение осадков, вибрацию, механические воздействия и т.д. Учитывая все варианты применения вычислительного модуля, показатель N_s (число реализаций многопроцессорных систем

на основе модуля) равен трем. Количество систем этого класса также равно трем. В этих системах применены одинаковые вычислительные модули.

Конструкция многопроцессорной системы БВ631 в корпусе типа «Евромеханика» высотой 6U на основе соединительной панели, вычислительного модуля с двумя многоядерными процессорами и возможностью увеличения периферийных устройств представлена на рис. 2. Вычислительные модули этой конструкции являются взаимозаменяемыми в случае соответствия одинаковому стандарту крейтовой системы. Представленный вычислительный модуль соответствует стандарту PICMG 2.30 CompactPCI PlusIO, который не оказался популярным на рынке.

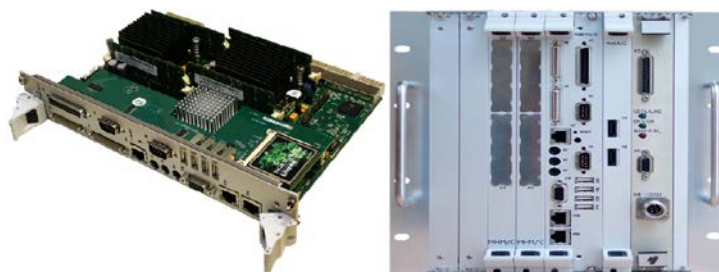


Рисунок 2. Многопроцессорная система БВ631 (справа) и вычислительный модуль (слева) на основе двух процессоров «Эльбрус-2С+»

Основными достоинствами этой конструкции является стойкость к механическим воздействиям, установка системы в стойку с односторонним обслуживанием, удобство замены или модернизации вычислительного модуля, возможность увеличения периферийных устройств. Основным недостатком следует считать невозможность исполнения вычислительного модуля для кондуктивного типа охлаждения с теплоотводом на корпус. Существенными недостатками также являются отсутствие горизонтального масштабирования по количеству многоядерных процессоров и двойное место для установки вычислительного модуля (8 HP) с принудительным охлаждением воздушным потоком. Учитывая все варианты применений вычислительного модуля, показатель N_s равен 1. Количество систем такого класса значительно превышает этот показатель и равно пяти. В этих систе-

мах применены пять различных вычислительных модулей.

Оригинальная конструкция «Лезвие-БНК» в корпусе высотой 5U на основе соединительной панели, вычислительного модуля с четырьмя многоядерными процессорами и разъемами периферийных интерфейсов представлена на рис. 3. Достоинство этой конструкции – возможность компактного размещения десяти вычислительных модулей в корпусе высотой 5U при установке в стойку, что обеспечивает эффективное горизонтальное масштабирование по количеству многопроцессорных модулей в кластере суперкомпьютера. Другим достоинством является низкая себестоимость готового изделия за счет единой системы питания, охлаждения и управления десятью вычислительными модулями.

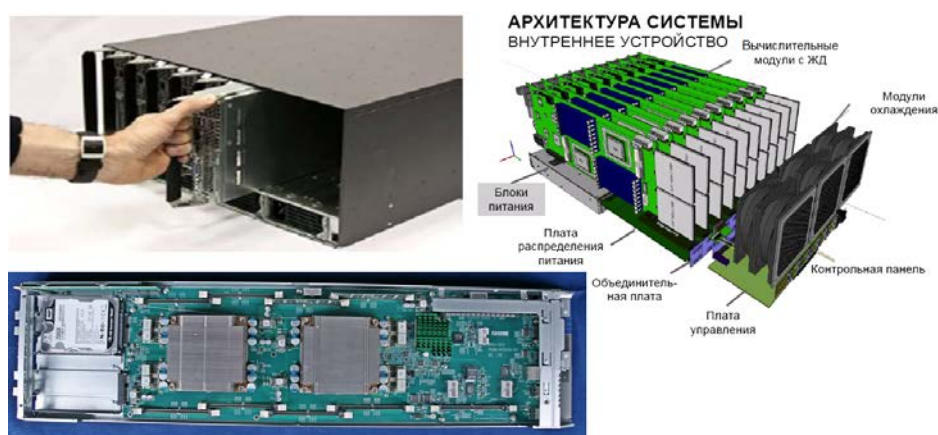


Рисунок 3. Многопроцессорная система «Лезвие-БНК» (вверху) и вычислительный модуль (внизу) на основе четырех процессоров «Эльбрус-2С+»

Основным недостатком такой конструкции является необходимость ее доработки для специальных применений, включая использование шасси с обеспечением работоспособности в условиях внешних воздействий. Другими недостатками являются небольшое количество разъемов периферийных интерфейсов при отсутствии возможности добавления более одного периферийного устройства (карты или модуля) и небольшое количество слотов оперативной памяти при четырех процессорах на модуле. Учитывая все варианты применений вычислительного модуля, показатель N_s равен 1. Количество систем этого класса также равно 1.

2. Определяющие характеристики

По результатам анализа завершенных разработок были определены требования, направленные на сокращение отставания от импортной вычислительной техники, т.е. определена необходимость улучшения характеристик многопроцессорных систем, таких как:

- общая производительность;
- пропускная способность каналов обмена с оперативной памятью;
- пропускная способность каналов межпроцессорного обмена;
- количество и быстродействие ядер процессора, объём кэш-памяти;
- масштабируемость по количеству многоядерных процессоров, оперативной памяти и периферийных интерфейсов;
- уровень унификации вычислительных модулей.

Существенно, что масштабирование по количеству многоядерных процессоров и оперативной памяти особенно важно для востребованных многопроцессорных систем типа «Евромеханика». При этом добавление процессоров и оперативной памяти не означает пропорциональное увеличение производительности, которое в отношении ряда пользовательских задач либо невозможно, либо возможно только при изменении кода программ на языке высокого уровня [3].

Важной проблемой конструкторско-технологической разработки новых многопроцессорных систем, во многом определяющей их конкурентоспособность, является сокращение времени проектирования. Многоядерный процессор нового типа создается несколько лет, и появление многопроцессорных систем на его основе крайне желательно сразу после выпуска первой партии микросхем. Это обязывает начинать проектирование многопроцессорной системы до момента готовности микросхемы, по возможности сократив его продолжительность относительно момента завершения разработки кристалла. Опыт совместного конструкторско-технологического проектирования процессора, вычислительного

модуля и многопроцессорной системы показал реальность временной диаграммы, представленной на рис. 4. Важно отметить и то, что параллельная организация конструкторско-технологического проектирования и подготовки производства изделия позволяет значительно повысить конкурентоспособность проектной организации [4].

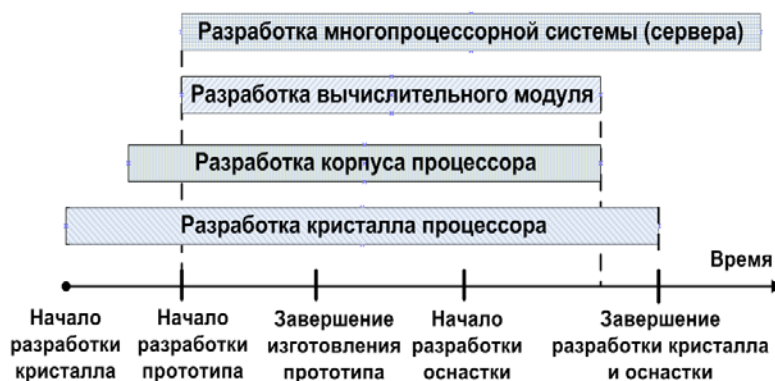


Рисунок 4. Диаграмма Ганта для реализации многопроцессорной системы

В результате анализа завершенных разработок импортозамещающих многопроцессорных систем выявлена необходимость унификации вычислительных модулей для системы типа «Евромеханика». Целесообразность унификации обосновывается следующим соотношением:

$$-R_u + \sum_{i=1}^{N_s} R_i + \Delta e \cdot \left(\sum_{i=1}^{N_s} Q_i \right)^a + \sum_{i=1}^{N_s} (Q_i^b \cdot \Delta P_i + Q_i^c \cdot \Delta C_i) > 0,$$

где R_u – затраты на разработку унифицированных модулей; N_s – число многопроцессорных систем, для которых предназначаются унифицированные модули; R_i – затраты на разработку специализированных модулей, которые заменяются унифицированными модулями; Δe – эффект, обусловленный повышением обращаемости унифицированных модулей (например, повышение производительности и качества при специализации производства, снижение стоимости комплектующих при повышении количества изготавливаемых модулей); Q_i – обращаемость унифицированных модулей в i -й системе; ΔP_i – положительный или отрицательный эффект, обусловленный сборкой унифицированных модулей

вместо сборки специализированных модулей в i -й системе; ΔC_i – положительный или отрицательный эффект, обусловленный приобретением комплектации унифицированных модулей вместо комплектации специализированных модулей в i -й системе; a , b и c и – числа ≤ 1 (определяются опытным путем).

Предлагается также учитывать показатель, демонстрирующий сокращение номенклатуры вычислительных модулей, на которые приходятся основные затраты по разработке и производству многопроцессорных систем:

$$U = (N_s / N_m)_2 - (N_s / N_m)_1,$$

где N_s – число реализаций многопроцессорных систем, N_m – число вычислительных модулей этих систем, числа 1 и 2 – моменты времени до и после проведения унификации.

3. Конструкции многопроцессорных систем общего применения

Для систем общего применения, которые не подлежат адаптации к специальным условиям эксплуатации, характерно требование низкой себестоимости. При этом востребованы разработки вычислительных модулей в виде панели с размерами форм-факторов SSI EEB или SWTX, которые содержат четыре многоядерных процессора и контроллеры периферийных интерфейсов. Для реализации межсетевых экранов, систем хранения данных или другого назначения может быть целесообразно использование двухпроцессорных вычислительных модулей с размерами форм-фактора EATX.

Показательны российские многопроцессорные системы общего применения, а именно четырёхпроцессорные системы на основе процессоров «Эльбрус-4С» и «Эльбрус-8С». В первом случае выполнена разработка четырехпроцессорного вычислительного модуля для системы (сервера) «Эльбрус-4.4» в корпусе высотой 1U. Предполагается также его использование для многопроцессорных систем в корпусах под установку в стойку высотой 2U или 3U. Достигнуты следующие характеристики модуля:

- скорость передачи данных между процессорами – 128 Гб/с (дуплекс) на канал;

- тепловыделение процессора при максимальной загрузке – 70 Вт;
- три канала памяти типа DDR3 у каждого процессора, по одному слоту DIMM на канал;
- пропускная способность памяти – 1333 МТ/с;
- в качестве основы панели использована 12-слойная печатная плата с применением доступных и широко используемых материалов группы FR-4 производства фирмы «Isola» (Германия).

Вычислительный модуль, являющийся развитием модуля MB4C-CPB, и многопроцессорная система на основе процессора «Эльбрус-4С» показаны на рис. 5.

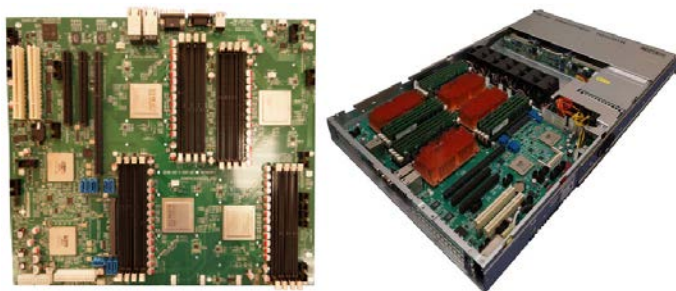


Рисунок 5. Многопроцессорная система «Эльбрус-4.4» (справа) вычислительный модуль на основе процессоров «Эльбрус-4С» (слева)

На основе процессора «Эльбрус-8С» выполнена разработка четырехпроцессорного вычислительного модуля для системы «Эльбрус-8.4» в корпусе высотой 2U. Предполагается также его использование для многопроцессорных систем в корпусах под установку в стойку высотой 1U или 3U. Достигнуты следующие характеристики вычислительного модуля:

- скорость передачи данных между процессорами – 384 Гб/с (дуплекс) на канал;
- тепловыделение процессора при максимальной загрузке – 90 Вт;
- четыре канала памяти типа DDR3 у каждого процессора, по два слота DIMM на канал;
- пропускная способность памяти – 1600 МТ/с;

- в качестве основы панели использована 14-слойная печатная плата на базе доступных и широко используемых материалов группы IT-150DA производства фирмы «Iteq» (Тайвань).

Вычислительный модуль и соответствующая ему многопроцессорная система «Эльбрус-8.4» показаны на рис. 6.

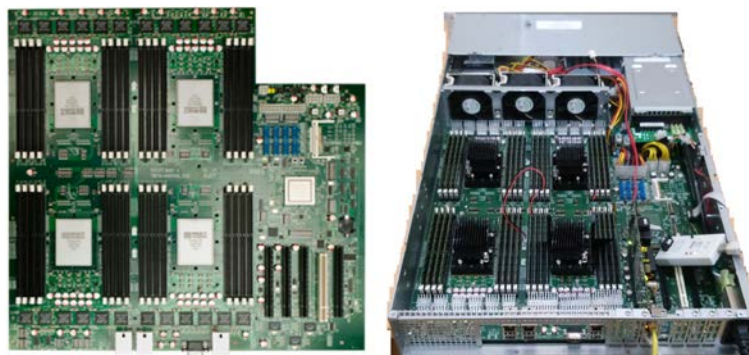


Рис 6. Многопроцессорная система «Эльбрус-8.4» (справа) вычислительный модуль на основе процессоров «Эльбрус-8С» (слева).

4. Конструкции многопроцессорных систем специального применения

Для специального применения возможно построение системы типа «Евромеханика» на основе стандартов CompactPCI Serial (PICMG CPCI-S.0), AdvancedTCA (PICMG 3.x) или VPX (ANSI/VITA 46, ANSI/VITA 48). Наиболее востребованным и перспективным направлением развития многопроцессорных систем является разработка вычислительных модулей и периферийных устройств (карт) в соответствии со стандартами VPX и СМС (PMC)/ХМС (ХМС.3).

На основе многоядерных процессоров «Эльбрус» реализованы вычислительные модули для многопроцессорных систем специального применения, устанавливаемых в кейт или корпус со слотами под установку модулей. При этом достигаются те же характеристики системы, которые отмечены ранее для систем общего применения. В условиях выпуска небольших серий целесообразна реализация многопроцессорной системы с применением нескольких вычислительных модулей. В настоящее время широко представлены вычислительные модули двух типов:

1) с набором периферийных интерфейсов и без каналов для построения многопроцессорной системы;

2) с каналами для построения многопроцессорной системы и без набора периферийных интерфейсов.

Модули первого типа для промышленных систем предлагаются различными производителями телекоммуникационного оборудования, в то время как модули второго типа предлагаются несколькими производителями суперкомпьютеров для воспроизводительных вычислений.

С учетом этой конъюнктуры новизна предлагаемой нами многопроцессорной системы состоит в использовании нескольких одинаковых вычислительных модулей, которые не относятся к известным типам. Вычислительным модулем является компьютер на плате под установку в крейт с выводом межпроцессорных каналов для кэш-когерентного доступа к памяти (сNUMA) на высокочастотные соединители. Эти соединители классифицируются для различных стандартов и имеют модификации [5]. В составе предлагаемого модуля имеется процессор, оперативная память, интегрированный периферийный контроллер (южный мост), соединители и разъемы пользователя, носитель данных и другая аппаратура. Несколько таких модулей в крейте образуют вычислительную систему на общей памяти. Этим предлагаемая многопроцессорная система и отличается от распространенных систем с одним вычислительным модулем в корпусе, соответствующих стандартам VPX (VITA 46, VITA 48 выпуска 2007–2013 годов) или Advanced®TCA (PICMG 3.x выпуска 2008–2012 годов).

Различные реализации такого типа многопроцессорной системы представлены на рис. 7. Предлагаемые вычислительные модули обеспечивают постоянное соотношение количества вычислительных ядер, объема когерентной памяти и пользовательских интерфейсов. Поскольку модули полностью идентичны, то просто обеспечивается масштабируемость, настройка и диагностика неисправности в процессе эксплуатации многопроцес-

сорной системы. Важно отметить, что в случае двухстороннего обслуживания крейта для каждого вычислительного модуля возможна установка дополнительного интерфейсного модуля расширения (Rear Transition Module). Основным недостатком предлагаемого типа многопроцессорной системы является повышенная себестоимость по сравнению с компактными стоечными системами (Rackmount Servers), которая обусловлена использованием соединителей и приемопередатчиков данных, отнесенных к каждому межпроцессорному каналу, а также избыточностью интегрированных контроллеров периферийных интерфейсов.



Рисунок 7. Многопроцессорные системы С-1, С-2 и С-3 на основе унифицированного однопроцессорного вычислительного модуля, соответствующего стандартам VPX

Учитывая все варианты применений вычислительного модуля на основе многоядерного процессора, показатель N_s равен пяти. Общее количество систем этого класса также равняется пяти. В этих системах применены одинаковые вычислительные модули.

Заключение

На основе конструкторско-технологических проектных решений для повышения уровня унификации вычислительных модулей и масштабируемости вычислительных систем реализованы многопроцессорные системы общего и специального применения, являющиеся развитием аппаратно-программной платформы «Эльбрус». Для многопроцессорных систем специального назначения более чем в два раза сокращена номенклатура вычислительных модулей путем унификации их конструкции.

Унификация вычислительного модуля многопроцессорных систем специального применения позволила на 20% уменьшить его себестоимость. При росте популярности крейтовых систем на основе современных стандартов прогнозируется существенное уменьшение комплектующих в их составе, а следовательно, дальнейшее снижение их себестоимости и повышение доступности для широкого круга потребителей.

Литература

1. Мистюков В.Б., Бурковский В.Л. Анализ специфики технологического маршрута проектирования вычислительных кластеров специального назначения // Вестник ВГТУ. – 2010. – №11. – С. 55-58.

2. Афонин И.Н. Решение для обеспечения постоянной доступности информационных систем // Современные технологии автоматизации. – 2014. – №3. – С. 6-14.

3. Majo Z., Gross T.R. (Mis)Understanding the NUMA Memory System Performance of Multithreaded Workloads // Proceedings of the IEEE International Symposium on Workload Characterization, IISWC 2013, Portland, OR, USA, September 22-24, 2013. – pp. 11-22.

4. Банкрутенко В.В., Долбунов Д.А., Комиссаров К.В., Павлин В.Н., Штарев В.В. Организация параллельной конструкторско-технологической подготовки производства на ФГУП «ОКБМ» // САПР и Графика. – 2006. – №7.

5. Д. Гаманюк. Электрические соединители для встраиваемых систем специального назначения. // Современная электроника. – 2012. – №3. – С. 22-24.

Бычков Игнат Николаевич, к.т.н., начальник отдела

ПАО «Институт электронных управляющих машин им. И.С. Брука», АО «МЦСТ»

119334, г. Москва, ул. Вавилова, д. 24

(499)135-42-02

bychkov_i@ineum.ru

Рябцев Юрий Степанович, д.т.н., проф., главный научный сотрудник

ПАО «Институт электронных управляющих машин им. И.С. Брука», АО «МЦСТ»

119334, г. Москва, ул. Вавилова, д. 24

(499)135-20-16

Yury.S.Ryabtsev@mcst.ru

Молчанов Игорь Анатольевич, инженер

ПАО «Институт электронных управляющих машин им. И.С. Брука», АО «МЦСТ»

119334, г. Москва, ул. Вавилова, д. 24

(499)135-50-32

Igor.A.Molchanov@mcst.ru