

# ОТРАСЛЕВЫЕ СУПЕРКОМПЬЮТЕРЫ СЕМЕЙСТВА «СКИФ» В РЕСПУБЛИКЕ БЕЛАРУСЬ: СОСТОЯНИЕ И ПЕРСПЕКТИВЫ РАЗВИТИЯ

С.М. Абрамов, В.В. Анищенко, Н.Н. Парамонов, А.Г. Рымарчук, О.П. Чиж

## Введение

Суперкомпьютерное направление «СКИФ» развивается в рамках программ Союзного государства Беларуси и России (СГ) уже более десяти лет, начиная с 2000 года. За это время созданы четыре поколения суперкомпьютеров «СКИФ»: Ряда 1, Ряда 2, Ряда 3 и Ряда 4 (рис. 1) [1]. Главные исполнители этих программ – Объединенный институт проблем информатики НАН Беларуси и Институт программных систем РАН.

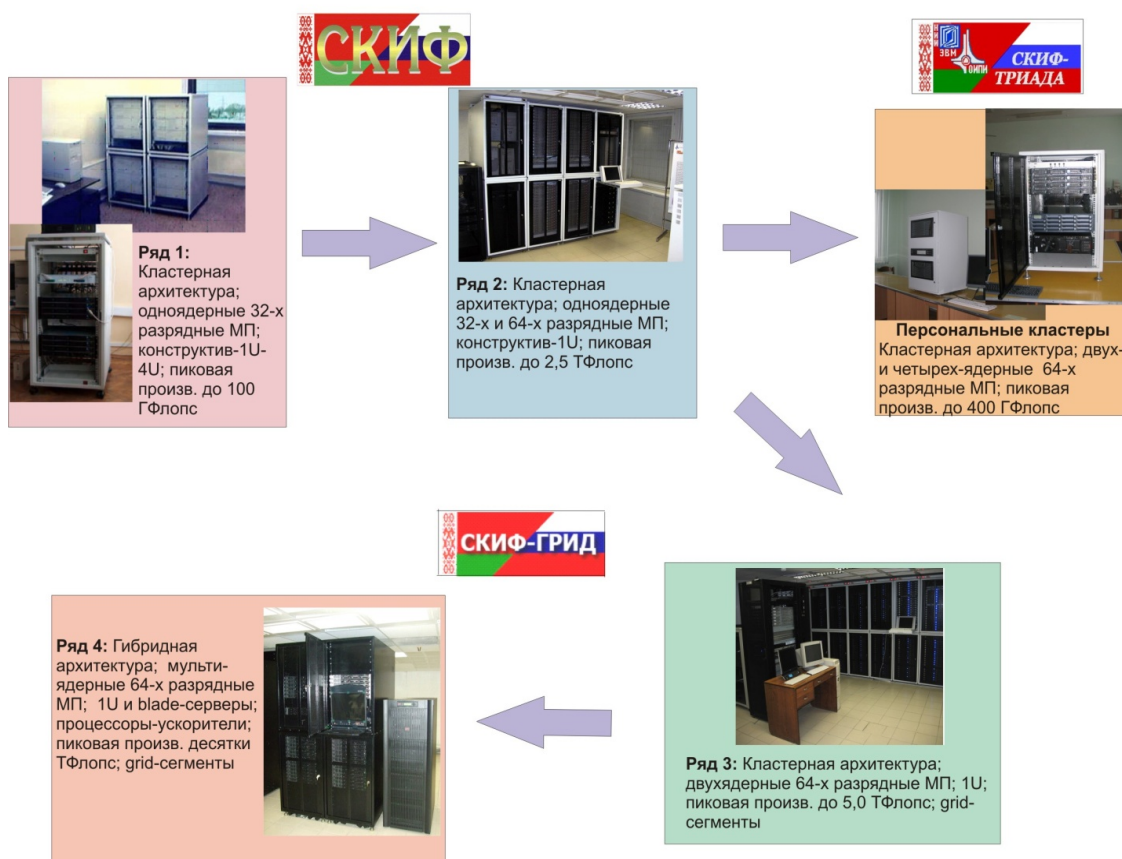


Рис. 1. Развитие суперкомпьютерного направления «СКИФ»

Создаваемые суперкомпьютерные модели семейства «СКИФ» соответствовали мировому (на текущий момент) уровню. Так, например, старшая модель семейства СКИФ Ряда 2 – кластер СКИФ К-1000 с пиковой производительностью 2,5 триллиона операций в секунду 9 ноября 2004 года был включен в очередной 24-й выпуск списка Top500 под номером 98. Этот кластер на 1 января 2005 г. являлся самым мощным компьютером на территориях СНГ и Восточной Европы и входил в четыре подряд редакции списка Top500. Белорусские кластеры семейства «СКИФ» Ряда 4 соответствуют современным базовым мировым трендам развития суперкомпьютерных технологий: кластерная архитектура, операционная система Linux, сетевые интерфейсами – GigabitEthernet и InfiniBand, вычислительные узлы на основе многоядерных процессоров с архитектурой x86 и спецпроцессоров.

Суперкомпьютерные модели семейства «СКИФ» Ряда 1 – Ряда 4 создавались как универсальные вычислительные системы. Однако в последнее время появилась потребность создания суперкомпьютерных конфигураций, оптимально реализующих перспективные математические модели и алгоритмы, ориентированные на решение наиболее сложных и наукоемких задач в ключевых экономически значимых секторах экономики. Суперкомпьютеры, оптимизированные (архитектурно, функционально, конструктивно-технологически и т. п.) для решения определенных задач конкретной отрасли экономики, называют предметно-ориентированными суперкомпьютерными системами отраслевого назначения (сокращенно – отраслевыми суперкомпьютерами) [2]. Использование отраслевых суперкомпьютеров может обеспечить наибольшую экономическую отдачу вложенных отраслью средств и мультипликативный инновационный эффект. В связи с

этим особое значение приобретают обоснование и выбор базовых параметров суперкомпьютеров, проектируемых для решения задач конкретной отрасли экономики.

Развитие суперкомпьютерного направления «СКИФ» в части создания перспективных отраслевых кластеров (предметно-ориентированная серия суперкомпьютеров «СКИФ-ГЕО» для решения геолого-геофизических задач) планируется в рамках формируемой программы Союзного Государства (СГ) «СКИФ-НЕДРА» [3]. В последние годы в стратегически важном для государств-участников СГ топливно-энергетическом секторе экономики резко возрастают потребности в скорости и качестве высокопроизводительных вычислений при поисках, разведке и разработке месторождений природных ресурсов. Такое положение обусловлено необходимостью решения задач, возникающих в связи с истощением разрабатываемых месторождений, снижением эффективности добычи, экономическими и технологическими сложностями в поисках и освоении новых месторождений из-за усложнения условий проведения геологоразведочных работ. Повышение точности и скорости расчетов и качества моделирования процессов в технологических задачах поисков, разведки, разработки и использования месторождений углеводородного и углеродного сырья является наиболее важной и перспективной задачей топливно-энергетического сервиса, решение которой существенно влияет на прирост запасов, экономию ресурсов и использование простаивающего фонда скважин, рост уровня отдачи месторождений. Развитие предметно-ориентированных программно-аппаратных суперкомпьютерных систем с соответствующими техническими параметрами и прикладным программным обеспечением для решения наиболее актуальных и перспективных вычислительных задач нефтегазового сервиса является одной из стратегических задач инновационного развития и модернизации экономики государств-участников СГ [3].

В настоящей статье отражены основные возможные направления создания в РБ опытных образцов отраслевых суперкомпьютеров серии «СКИФ-ГЕО» с учетом концепции формируемой программы СГ «СКИФ-НЕДРА», целей создания суперкомпьютеров, специфики приложений, мировых тенденций развития суперкомпьютерных технологий и других факторов.

#### **Научно-технический задел**

Существенный научно-технический задел для разработки опытных образцов серии «СКИФ-ГЕО» сформирован в РБ при создании моделей Ряда 4 семейства «СКИФ» (кластеры «СКИФ ОИПИ», «СКИФ-GPU» и «СКИФ-ГРИД») в рамках программы СГ «СКИФ-ГРИД» (2007 – 2010 г.г.) [1].

При создании в РБ кластерных конфигураций «СКИФ» Ряда-4 учитывались следующие концептуальные положения:

- Диапазон производительности – терафлопсный.
- Архитектура – кластерная.
- Вычислительные узлы (ВУ) – на базе классических мультиядерных процессоров с архитектурой x86-64 и/или специализированных процессоров (сопроцессоров) типа Cell, графических процессоров-ускорителей (GPU) и др.
- Базовые конструктивно-технологические решения: форм-факторы – 1U, 2U, blade-серверы; стойки – 19” высотой до 44 U; система охлаждения – воздушная.

Структурно суперкомпьютер «СКИФ ОИПИ» («первенец» конфигураций «СКИФ» Ряда 4) представляет собой метаcluster, состоящий из двух кластеров - blade-кластер (52 узла на blade-серверах на базе 4-х ядерных процессоров Intel XEON E5472 3.0GHz) и Cell-кластер (8 двухпроцессорных ВУ на базе 9-ядерных процессоров PowerXCell 8i 3.2Ghz). Пиковая производительность blade-кластера суперкомпьютера «СКИФ ОИПИ» - 5,0 Тфлопс. Показатель эффективности на тесте Linpack для 52 узлов на blade-серверах - 80,67%.

В кластере «СКИФ-GPU» 34 вычислительных и 2 управляющих узла. Каждый ВУ кластера содержит два 4-х ядерных процессора с архитектурой x86-64 и видеокарту NVIDIA GeForce GTX 295 с двумя 240-ядерными графическими потоковыми процессорами. Пиковая производительность кластера «СКИФ-GPU» (36 узлов) без учёта ускорения с помощью графических процессоров – не менее 3375.36 Гфлопс при показателе эффективности на тесте Linpack 89,62%.

В кластере «СКИФ-ГРИД» используются 12-ядерные процессоры AMD Opteron с архитектурой x86-64 и графические процессоры-ускорители. Пиковая производительность кластера «СКИФ-ГРИД» (36 ВУ и 2 управляющих узла) без учёта ускорения с помощью графических процессоров составляет 8,0 Тфлопс при показателе эффективности на тесте Linpack 82,15%.

Опытные образцы кластеров «СКИФ» Ряда-4 эксплуатируются в суперкомпьютерном центре ОИПИ НАН Беларуси под управлением Операционного центра национальной grid-сети Республики Беларусь [4]. Наличие записи пользователя в едином реестре grid-ресурсов гарантирует возможность запуска пользовательской задачи на любой вычислительной установке, входящей в национальную grid-сеть. Тем не менее, для большинства вычислительных установок можно выделять классы прикладных задач, на которых достигается наиболее эффективная работа. В процессе эксплуатации суперкомпьютеров проводились исследования и оценка функциональных возможностей вычислительных узлов с использованием спецпроцессоров. В частности, для Cell-кластера суперкомпьютера «СКИФ ОИПИ» можно отметить высокую эффективность целочисленных вычислений при использовании векторизованного кода с применением

регистров большой длины (например, работа с арифметикой конечного поля для модулей большой длины). Вычислительные установки с графическими процессорами GPU демонстрируют хорошие результаты в параллельной обработке данных, когда с помощью одной и той же последовательности действий обрабатывается большой объем данных. На кластерах «СКИФ-GPU» и «СКИФ-GRID» с использованием графических процессоров было успешно произведено разложение на множители числа из конкурса RSA в 190 десятичных знаков. Данная факторизация была впервые произведена на суперкомпьютерах МГУ. Более подробная сравнительная информация о решении данной задачи RSA-190 на кластерах Чебышев, Ломоносов, SKIF\_GPU и SKIF\_GRID имеется в [5].

На основании практического опыта, полученного авторами в процессе создания и эксплуатации моделей суперкомпьютеров семейства «СКИФ» Ряда 4, были сформулированы предварительные требования к параметрам перспективных отраслевых кластеров для решения геолого-геофизических задач, разработка которых планируется в рамках формируемой программы СГ «СКИФ-НЕДРА». Технические требования формировались с учетом финансовых возможностей отраслевой структуры, степени развития ее вычислительной инфраструктуры, специфики решаемых задач, мировых тенденций развития суперкомпьютерных технологий и других факторов.

### **Предварительные требования к базовым параметрам белорусских отраслевых кластеров серии «СКИФ-ГЕО»**

При разработке проекта суперкомпьютера для отраслевой структуры того или иного уровня (отраслевой центр обработки данных, исследовательское или учебное заведение, отраслевые структуры типа офисов и т.п.), как правило, возникает естественный конфликт желаний и возможностей. Поэтому на начальном этапе жизненного цикла отраслевого суперкомпьютера необходимо уделить особое внимание совокупности ряда определяющих действий, таких как оценка реального объема финансирования проекта создания кластерной конфигурации (бюджет проекта), анализ назначения и целей создания отраслевой кластерной конфигурации, анализ специфики приложений и др.

Существенного эффекта в повышении производительности вычислений, то есть в увеличении скорости расчетов, качества получаемых результатов, сокращения затрат и снижения рисков, можно добиться только при оптимальном сочетании характеристик аппаратной и программной составляющих используемых высокопроизводительных информационно-вычислительных технологий. Поэтому в формируемой программе СГ «СКИФ-НЕДРА» предусмотрены два взаимозависимые базовые мероприятия: исследование и разработка программного обеспечения (ПО «СКИФ-НЕДРА») и опытных образцов отраслевых суперкомпьютеров серии «СКИФ-ГЕО» [3]. Проведение комплексного анализа специфики геолого-геофизических приложений планируется на первом этапе формируемой программы СГ «СКИФ-НЕДРА». Поэтому в настоящей статье акцент сделан в основном на предварительном выборе параметров непосредственно опытных образцов белорусских суперкомпьютеров серии «СКИФ-ГЕО».

В соответствии с утвержденной концепцией формируемой программы СГ «СКИФ-НЕДРА» в РБ планируется создание двух опытных образцов суперкомпьютеров серии «СКИФ-ГЕО»:

- опытный образец стационарной модели для решения ресурсоемких геолого-геофизических задач в центрах обработки данных - кластер «СКИФ-ГЕО-ЦОД РБ»;
- опытный образец стационарной модели для структурных подразделений (офисов) геолого-геофизической отрасли - кластер «СКИФ-ГЕО-Офис РБ».

В статье также сделана оценка перспективы создания опытного образца мобильной модели кластера серии «СКИФ-ГЕО» при поисках, разведке и разработке месторождений (кластер «СКИФ-ГЕО-Мобил»).

#### **Кластер «СКИФ-ГЕО-ЦОД РБ»:**

- **Процессоры:** Многоядерные процессоры с системой команд x 86 микроархитектуры Intel Haswell-EP E5-2600 V3 или AMD Opteron версий Streamroller/Excavator. Для ускорения выполнения потоковых операций - процессоры Xeon Phi, GPU NVIDIA типа Maxwell или GPU AMD Radeon.
- **Вычислительные узлы:** Количество и типы ВУ определяются после выбора процессоров с архитектурой X 86 и GPU. Ориентировочно 20-24 двухпроцессорных узлов с архитектурой X 86 и часть (или все) вычислительные сервера со спец-вычислителями. Два управляющих узла. Серверы покупные, сборка РБ по ТУ РБ.
- **Файл-сервер:** 36-40 Тбайт, 2U.
- **Сетевые интерфейсы:** Системная сеть – Infiniband последнего поколения (~ 80 Gbit/s); вспомогательная сеть – 10 Gigabit Ethernet.
- **Пиковая производительность:** 30-40 Тфлопс (без учета GPU).
- **Стойка:** Одна 19” стойка высотой (42-44) U. Воздушное охлаждение. Возможно использование отечественных стоек (типа стоек белорусских кластеров «СКИФ» Ряда 4) или импортных (стойки фирм RITTAL, APC или Conteg).
- **Потребляемая кластером мощность:** не более 15кВА.

- **ИБП:** встроенный, 12U.
- **Операционная система:** Linux.
- **Программное обеспечение для сопроцессоров:** должна быть обеспечена поддержка OpenCL.
- **КД и ПД:** Проведение приемочных (государственных) испытаний. Литера О1.
- **Стоимость вычислительной мощности кластера:** не более 40 USD за 1 GFlops.
- **Энергоэффективность кластера** (отношение пиковой производительности кластера к потребляемой мощности): не менее 0,7 Gflops/VA.
- **Показатель эффективности кластера на тесте Linpack** (отношение реальной производительности кластера к пиковой без учета спецпроцессоров): не менее 85%.
- **Условия эксплуатации:** круглосуточный режим работы в нормальных климатических условиях при напряжении сети 220 В. Нормальными климатическими условиями эксплуатации кластера являются:
  - температура окружающего воздуха (20±5)°С;
  - относительная влажность окружающего воздуха (60±15)%;
  - атмосферное давление от 84 до 107 кПа.

#### **Кластер «СКИФ-ГЕО-Офис РБ»:**

- **Процессоры:** Многоядерные процессоры APU AMD типа Kaveri с элементами HAS, мобильные процессоры Intel Haswell с графическим ядром GT3 или серверные Intel Haswell серии E3-1200 V3 с графическим ядром.
- **Вычислительные узлы:** Количество и типы ВУ определяются после выбора процессоров APU. ВУ однопроцессорные на платах Mini-ITX/Micro-ATX – 6-8 узлов. Один управляющий узел. Сборка серверов из покупных узлов. Сборка РБ по ТУ РБ. ВУ без вентиляторов. Использование **отечественных тепловых трубок**. Диски – SSD.
- **Файл-сервер:** 8-12 Тбайт, 1U(2U).
- **Сетевые интерфейсы:** Системная сеть/вспомогательная сеть – **10 Gigabit Ethernet** с возможностью коммуникаций по технологии **RDMA**.
- **Подключение к Internet:** **беспроводная связь** (наличие модулей WLAN и т.п.).
- **Пиковая производительность:** **5-8 Тфлопс** (без учета GPU).
- **Стойка:** Одна 19” стойка высотой (12-18) U фирм APC или **Rittal**. Воздушное охлаждение.
- **Уровень шума:** не более **55 ДБ**.
- **Вес:** не более **80 кг**.
- **Потребляемая кластером мощность:** не более **2 кВА**.
- **ИБП:** встроенный, 1U (2U).
- **Операционная система:** Linux, MS Windows.
- **Программное обеспечение для сопроцессоров:** должна быть обеспечена поддержка **OpenCL**.
- **КД и ПД:** Проведение приемочных (государственных) испытаний. **Литера О1**.
- **Стоимость вычислительной мощности кластера:** не более **40 USD** за 1 GFlops.
- **Энергоэффективность кластера:** не менее **0,7 Gflops/VA**.
- **Показатель эффективности кластера на тесте Linpack:** не менее **85%**.
- **Условия эксплуатации:** круглосуточный режим работы в нормальных климатических условиях при напряжении сети 220 В.

#### **Кластер «СКИФ-ГЕО-Мобил РФ».**

Адресной аудиторией мобильных кластеров «СКИФ-ГЕО-Мобил» являются исследовательские поисковые партии (группы) и другие низовые структуры геолого-геофизической отрасли. Беларусь – небольшая страна, территория которой (практически вся) охвачена скоростной 3G-мобильной связью. В связи с этим в РБ возможно использование технологии обработки и обмена данными между структурными подразделениями геолого-геофизической отрасли без мобильных кластеров. В принципе, в качестве псевдо мобильных возможно использование кластеров «СКИФ-ГЕО-Офис РБ», которые должны иметь небольшие габариты и вес, а также средства беспроводной связи, в том числе с Internet. Для обеспечения возможности оперативных перевозок для этих изделий должна быть разработана специальная упаковка многократного пользования. Возможен выпуск модификации кластера «СКИФ-ГЕО-Офис РБ» с обеспечением длительной (несколько часов) автономной работы от аккумуляторов без их подзарядки. При использовании в мобильном режиме кластеры «СКИФ-ГЕО-Офис РБ» должны эксплуатироваться в помещениях с нормальными климатическими условиями.

В России (с учетом ее огромной территории) мобильные кластеры при поисках, разведке и разработке месторождений, безусловно, будут востребованы. Для создания в рамках формируемой программы СГ «СКИФ-НЕДРА» мобильного кластера «СКИФ-ГЕО-Мобил РФ» в ИПС РАН создан существенный научно-технический задел. В ИПС РАН разработан, изготовлен и проходит опытную эксплуатацию экспериментальный образец

кластера (рис. 2) со специфическим способом охлаждения. Специфика системы охлаждения заключается в том, что базовые компоненты кластера - серверы погружены в диэлектрическую жидкость. Технология жидкостного охлаждения с погружением компонентов представляет удачное решение в области высокопроизводительных вычислений для оборудования с высокой плотностью размещения (30кВт на стойку и выше). При этом обеспечивается практически бесшумная работа, так как в серверах отсутствуют вентиляторы. Диэлектрические жидкостные хладагенты обладают существенно более высокой (в 1000 раз) теплоаккумулирующей способностью, чем воздух такого же объема. Данная технология позволяет обеспечить высокую экономию электроэнергии, затрачиваемой на охлаждение (до 90-95%). ИПС РАН не является изобретателем технологии охлаждения с погруженными в диэлектрическую жидкость компонентами, но специалисты института нашли способ внедрения данной технологии в суперкомпьютерную индустрию. Ниже приведены основные характеристики экспериментального образца кластера ИПС РАН с погруженными в диэлектрическую жидкость компонентами. Этот принцип может быть также эффективен и для создания стационарных кластерных конфигураций (например, для создания кластера «СКИФ-ГЕО-Офис РФ»).

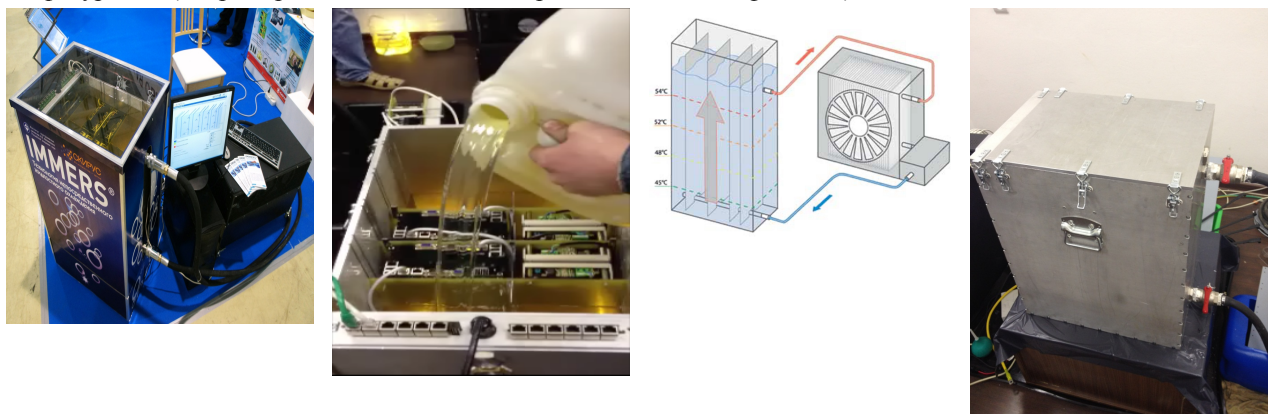


Рис. 2. Опытный кластер (ИПС РАН и группа компаний «СТОПУС»), охлаждаемый диэлектрической жидкостью по технологии IMMERS® (слева–направо): установка на выставке «Открытые инновации» 02.11.2012; заливка жидкости в систему; общая схема системы; система IMMERS–10

Технические параметры:

- **IMMERS–10** — Высокопроизводительная система, персональный суперкомпьютер. Стационарная установка. Размер: 60x60x35 см. 10 двухпроцессорных узлов. Хот-свап блоки питания и жесткие диски(по 2 на узел). FDR, 10G Eth. GPU Ready.
- **IMMERS–2M** (поставлена первому заказчику) — Мобильная отказоустойчивая система. Два 4-х процессорных узла с одним GPU каждый. Виброзащита, герметичность корпуса. Электропитание: 12/24/48V DC, 100—260V AC. Рабочая температура: –40...+40°C. Влажность: до 100%

### Заключение

Главная цель создания любой вычислительной конфигурации – обеспечение эффективного выполнения приложений. Задачи, решаемые в различных отраслях экономики, помимо специфики назначения отличаются количеством вычислительных операций, размерностью, объёмами входных данных, формами параллелизма, требованиями к времени решения задачи и т.п. Необходимо адекватно отобразить свойства прикладных задач в создаваемой инструментальной среде. Эта цель достигается путем выбора соответствующих тактико-технических параметров создаваемых суперкомпьютерных конфигураций с учетом специфики отраслевых приложений. Внедрение предметно-ориентированных отраслевых суперкомпьютеров «СКИФ-ГЕО» позволит использовать решения, оптимизированные именно для геолого-геофизических задач.

Приведенные в статье предварительные оценки параметров перспективных белорусских отраслевых кластеров «СКИФ-ГЕО» могут быть использованы при формировании технических требований к этим изделиям с учетом анализа специфики геолого-геофизических приложений, проведение которого планируется на первом этапе формируемой программы СГ «СКИФ-НЕДРА».

Состав авторов настоящей статьи подчеркивает согласованность действий белорусских и российских специалистов уже на этапе формирования программы СГ «СКИФ-НЕДРА».

### ЛИТЕРАТУРА:

1. Белорусские кластеры семейства «СКИФ»: состояние и перспективы развития / В.В. Анищенко, А.М. Криштофик, Н.Н. Парамонов, О.П. Чиж // Новороссийск: Труды Всероссийской суперкомпьютерной конференции Научный сервис в сети Интернет: эксафлопсное будущее. – 19-24 сентября 2011. – С. 23-27.

2. Принципы создания базовых конфигураций суперкомпьютерных систем отраслевого назначения / В.В. Анищенко, В.В. Мурашко, Н.Н. Парамонов, О.П. Чиж // Минск: Ежеквартальный журнал «Информатика» № 1 (33), январь-март 2012 г. Мн: Изд-во ОИПИ НАН Беларуси, 2012. – С. 97-105.
3. Исследования и разработка высокопроизводительных информационно-вычислительных технологий для увеличения и эффективного использования ресурсного потенциала углеводородного сырья Союзного государства. Концепция научно-технической программы Союзного государства. Постановление Совета Министров Союзного государства № 42 от 12.12.2012.
4. Операционный центр национальной грид-сети Республики Беларусь [Электронный ресурс]. - Режим доступа: <http://noc.grid.by/> - Дата доступа: 23.04.2013.
5. RSA-190 factored [Electronic resource]. - Mode of access: <http://mersenneforum.org/showthread.php?t=14177> — Date of access: 23.04.2013.