Б. М. Шабанов, А. П. Овсянников, А. В. Баранов, С. А. Лещев,
Б. В. Долгов, Д. Ю. Дербышев

Проект распределенной сети суперкомпьютерных центров коллективного пользования

Аннотация. Одним из ключевых методов повышения доступности и эффективности использования ресурсов суперкомпьютерных центров коллективного пользования (СКЦ) является их объединение в единую распределенную сеть.

Подобное объединение дает возможность оперативного перераспределения нагрузки между ресурсами путем перенаправления пользовательских заданий из очереди одного СКЦ в очередь другого, менее загруженного в определенный момент времени. В статье рассматриваются цель, задачи и ожидаемые результаты проекта по созданию распределенной сети СКЦ.

В статье приводятся результаты экспериментов по исследованию характеристик защищенного канала связи между СКЦ пилотной зоны проекта распределенной сети — МСЦ РАН (г. Москва) и ССКЦ СО РАН (г. Новосибирск).

Ключевые слова и фразы: центры коллективного пользования, суперкомпьютерные центры, защищенные каналы связи, распределенные вычисления.

Введение

Задача развития и повышения эффективности центров коллективного пользования (ЦКП) научным оборудованием является одной из актуальных задач научно-технологического развития России, вынесенной на государственный уровень [1]. В полной мере эта задача относится и к суперкомпьютерным научным ЦКП. Одним из ключевых методов повышения доступности и эффективности использования ресурсов суперкомпьютерных ЦКП (далее — СКЦ) является их объединение в единую распределенную сеть.

© Б. М. Шабанов, А. П. Овсянников, А. В. Баранов, С. А. Лещев, Б. В. Долгов, Д. Ю. Дербышев, 2017

© Межведомственный суперкомпьютерный центр РАН, 2017

© Программные системы: теория и приложения, 2017

DOI: 10.25209/2079-3316-2017-8-4-245-262

Задача объединения территориально распределенных высокопроизводительных ресурсов в единую систему известна с начала 2000-х годов [2], когда для ее решения применялись грид-технологии. При всей важности и востребованности грид-технологий построенные с их помощью распределенные системы обладали следующими недостатками.

- (1) На тот момент времени отсутствовали выделенные высокоскоростные защищенные сети передачи данных, объединение суперкомпьютерных центров приходилось производить посредством открытой и небезопасной сети Интернет. По этой причине значительную часть промежуточного программного обеспечения для построения грид-систем (например, Globus Toolkit) занимала подсистема обеспечения безопасности. Для авторизации пользователей использовался надежный, но, как показал опыт эксплуатации, крайне громоздкий и неудобный в использовании протокол X.509. На практике использование этого протокола безопасности создавало лишний барьер для пользователя, который предпочитал работе с грид-системой индивидуальные соединения с каждым суперкомпьютерным центром.
- (2) Низкая переносимость параллельных программ. Даже если в двух разных суперкомпьютерных центрах использовалось совместимое оборудование, незначительные различия в стеке системного и инструментального ПО неизбежно приводили к невозможности переноса исполняемых программных модулей из одного центра в другой и, как следствие, — существенному усложнению систем управления заданиями в грид.
- (3) Нестабильность развития промежуточного ПО. Используемый как фактический стандарт известный пакет Globus Toolkit несколько раз за свою историю менял идеологию построения, что заставляло разработчиков грид-систем либо непрерывно модифицировать схемы управления заданиями, либо использовать устаревшие версии пакета.

Особенно болезненно подобные изменения сказывались на информационной подсистеме, в качестве которой выступала MDS (Monitoring and Discovery System).

Сегодняшний уровень развития сетевых технологий позволяет организовать надежно защищенную распределенную сеть СКЦ, не прибегая к использованию таких «тяжелых» протоколов, как X.509. Современные средства контейнерной виртуализации способны решить

проблему бинарной переносимости параллельных программ между СКЦ [3]. Сравнительно недавно появившиеся распределенные системы управления базами данных позволяют организовать надежную единую информационную систему.

1. Цель и задачи проекта

Основной целью проекта является создание условий для повышения уровня применения высокопроизводительных вычислений при реализации научно-технических проектов по приоритетным направлениям научно-технологического развития России, за счет чего будет обеспечен рост показателей результативности научной деятельности. В рамках проекта создается и внедряется инновационная инфраструктура — распределенная сеть суперкомпьютерных центров коллективного пользования, объединяющая вычислительные ресурсы СКЦ в единую систему.

Для достижения поставленной цели необходимо решение задач по созданию:

- (1) Децентрализованной автоматизированной системы управления заданиями и ресурсами. Система управления будет поддерживать глобальную очередь пользовательских заданий и обеспечивать за счет этого оперативное перераспределение вычислительной нагрузки в сети СКЦ.
- (2) Единой системы мониторинга, которая позволит оперативно получать информацию о текущих состоянии и загрузке суперкомпьютерных ресурсов распределенной сети.
- (3) Единой системы доступа на основе удостоверяющей федерации суперкомпьютерных центров. Система доступа обеспечит пользователей унифицированным интерфейсом к суперкомпьютерным ресурсам распределенной сети, причем авторизоваться в сети можно будет с использованием только личной учетной записи в своей организации.
- (4) Общей облачной системы хранения данных, обеспечивающей единое файловое пространство для всех СКЦ сети.

В деятельности любого центра коллективного пользования научным оборудованием можно выделить три основные категории участников научно-организационного процесса: пользователей ЦКП (потребителей услуг), владельцев научного оборудования (поставщиков услуг), руководителей и представителей учредителей — государственных органов исполнительной власти (регуляторов процесса предоставления

услуг). Успешная реализация проекта по созданию распределенной сети СКЦ для указанных категорий позволит:

- (1) Повысить эффективность использования суперкомпьютерных ресурсов сети за счет возможности оперативного перераспределения вычислительной нагрузки. В случае недоступности, неисправности или перегруженности ресурсов какого-либо СКЦ пользовательские задания будут перенаправлены на ресурсы других СКЦ, что сократит время ожидания в очереди и повысит качество обслуживания пользователей, как потребителей услуг СКЦ. При недогруженности какого-либо СКЦ в нем могут быть выполнены задания других, перегруженных в этот момент времени СКЦ, за счет чего удастся избежать простоев ресурсов и повысить показатели эффективности работы СКЦ, как поставщиков услуг и владельцев суперкомпьютерного оборудования. Все это позволит увеличить пропускную способность совокупности ресурсов за счет исключения простоя одних ресурсов при перегруженности других ресурсов распределенной сети, что в свою очередь повысит эффективность расхода выделенных учредителями на развитие и содержание СКЦ бюджетных средств.
- (2) Упростить управление и мониторинг использования суперкомпьютерных ресурсов. Доступность для регуляторов информации единой системы мониторинга упростит процесс и сократит сроки принятия управленческих решений, облегчит обоснование финансирования для развития и поддержки суперкомпьютерных ЦКП.
- (3) Повысить качество доступа к суперкомпьютерным ресурсам различных архитектур для научных и образовательных организаций. В составе оборудования СКЦ, как правило, представлены суперкомпьютерные ресурсы либо уникальной архитектуры, либо уникальной конфигурации.

За счет реализации единой системы доступа и создания удостоверяющей федерации СКЦ пользователям будет предоставлена возможность использовать одну учетную запись для доступа ко всем суперкомпьютерным ресурсам сети.

Рассмотрим подробнее планируемые решения поставленных задач.

2. Управление заданиями и ресурсами в сети СКЦ

Неделимой единицей оборудования, включаемой в состав распределенной сети СКЦ, будет являться высокопроизводительная вычислительная система (ВС) — суперкомпьютер. По решению руководства

Рис. 1. Двухуровневая система управления сети СКЦ

СКЦ, только часть из имеющихся в центре ВС может войти в состав распределенной сети, оставшаяся часть оборудования может использоваться исключительно под нужды и задачи центра. Для управления отдельной ВС используется локальная система управления ресурсами (ЛСУР), действия всех ВС в сети координирует глобальная система управления ресурсами (ГСУР). В качестве ЛСУР может выступать любая система управления заданиями (СУППЗ, SLURM, PBS и т.п.).

После включения ВС в состав сети ее вычислительные ресурсы не отчуждаются от их владельца и продолжают использоваться для выполнения локальных заданий, которые образуют локальный поток заданий. Задания этого потока поступают на вход определенной ВС и могут выполняться на ресурсах только этой ВС.

После включения ВС в состав сети совместно с заданиями локального потока на вычислительные ресурсы ВС начинают поступать задания с других ВС — из глобального потока заданий (рис. 1). В отличие от локального, задания глобального потока допускают обработку на вычислительных ресурсах любой ВС сети (либо ВС из заданного списка), ГСУР распределяет задание глобального потока в ту ВС сети, время обработки задания в которой будет минимальным. Важно отметить, что ГСУР не планирует задания на локальные ресурсы ВС, а только выбирает т.н. целевую ВС для размещения задания. Размещенное задание глобального потока поступает под управление ЛСУР целевой ВС и планируется наряду с заданиями локального потока.

Рис. 2. Организация глобальной очереди заданий

Существенной характеристикой распределенной сети СКЦ является ее динамически изменяющийся состав, допускающий в любой момент времени как включение в состав сети новых ВС, так и исключение из состава сети отдельных ВС. Выключение из состава сети (включение в состав) некоторой ВС может быть связано как с неисправностью (возвращением в строй) самой ВС, так и с неисправностью (возвращением в строй) коммуникационных каналов.

ГСУР имеет децентрализованную архитектуру, представленную в виде коллектива равноправных диспетчеров. Каждый диспетчер обслуживает свою ВС из состава сети, обеспечивая две основные функции:

- помещение в глобальную очередь заданий, поступивших в подконтрольную диспетчеру ВС;
- выборка из глобальной очереди заданий, наиболее подходящих для выполнения на подконтрольной диспетчеру ВС.

СКЦ объединяются в сеть с помощью закрытых каналов связи, например, с помощью сертифицированной аппаратуры шифрования трафика. За счет этого появляется возможность размещения глобальной очереди в единой децентрализованной информационной системе, построенной на основе распределенной СУБД (рис. 2). Такой подход обеспечивает следующие возможности:

- в каждом СКЦ обрабатываются локальные потоки пользовательских заданий, часть из которых могут быть перемещены в единую глобальную очередь;
- в каждом СКЦ проводится своя политика предоставления ресурсов пользователям и планирования их заданий, определяется квота для заданий из единой очереди, виды и характеристики заданий, которые будет обрабатывать СКЦ;
- в сети отсутствует единый центр принятия решения, поэтому предполагаются динамические подключения (отключения) СКЦ к сети в произвольные моменты времени.

3. Информационная система сети СКЦ на базе распределенной СУБД

Противоречие между надежностью и масштабируемостью децентрализованной схемы управления и необходимостью ведения централизованной глобальной очереди заданий в сети СКЦ может быть разрешено за счет применения распределенной СУБД [4]. Идея состоит в том, что каждый диспетчер сети СКЦ хранит часть общих данных, данные при этом физически распределены по всем СКЦ, однако логически они являются одним целым. Применяемые в современных распределенных СУБД механизмы сегментирования и дублирования позволяют при отказе одной из точек хранения восстановить часть ее данных, тем самым не допустив их потери.

Современные распределенные СУБД, как правило, вместо традиционной реляционной модели хранения данных используют одну из трех, так называемых NoSQL-моделей: столбцово-ориентированную, ассоциативную или документо-ориентированную.

Столбцово-ориентированная модель [5] подразумевает хранение данных в таблицах не по строкам, а по столбцам, что существенно ускоряет выполнение поисковых запросов, но осложняет изменение и удаление записей. Столбцово-ориентированные СУБД поддерживают высокую скорость добавления данных, что в совокупности с быстрым поиском делает их привлекательными для расчета статистических показателей при больших объемах данных.

Ассоциативные распределенные СУБД [6] используют модель хранения данных типа «ключ-значение». Добавляемые в базу данные ассоциируются с идентификатором (ключом), используемым в дальнейшем для выполнения поисковых запросов. Ассоциативные СУБД отличаются высокой производительностью и часто используются для

организации промежуточного хранения (кэширования) данных с целью снижения нагрузки на основное хранилище.

Документо-ориентированные СУБД [7] используют для хранения иерархические структуры данных (документов) в виде дерева или леса. Данные в виде полуструктурированных документов хранятся в листьях дерева. При добавлении данные заносятся в индексы базы, благодаря которым можно осуществлять быстрый поиск даже при достаточно сложной общей структуре хранилища.

В работе [4] произведено сравнение указанных моделей распределенных СУБД, и сделан выбор в пользу документо-ориентированной модели хранения и представления данных, которая позволяют сделать универсальным решение по построению распределенной информационной системы для хранения глобальной очереди заданий пользователей сети СКЦ. В качестве основы для создания такой информационной системы может быть применена известная распределенная СУБД Elasticsearch [8].

4. Системы доступа к ресурсам и мониторинга сети СКЦ

Единая система мониторинга предназначена для оперативного отображения состояния и загрузки сети СКЦ. Предполагаются следующие примерные функции системы:

- просмотр состава сети СКЦ, характеристик входящий в состав суперкомпьютерных ресурсов;
- отображение текущей загрузки суперкомпьютерных ресурсов;
- просмотр содержания выполняемой в СКЦ вычислительной работы (какие пользователи, где, и какие задачи решают);
- отображение объемов израсходованных ресурсов, текущего состояния бюджета, выделенных квот ресурсов, настроек приоритетов, графиков изменения этих характеристик по всем научным проектам;
- просмотр состояния глобальной и локальных очередей заданий;
- просмотр состояния вычислительных узлов суперкомпьютеров (занят, неисправен, свободен и т.п.).

Единая система доступа к ресурсам сети СКЦ предоставит пользователю унифицированный интерфейс со следующими возможностями:

- просмотр списка, характеристик и состояния (текущей загрузки) доступных суперкомпьютерных ресурсов;

- направление задания в глобальную очередь для выполнения в наименее загруженном СКЦ из заданного пользователем списка;
- просмотр состояния глобальной и локальных очередей заданий;
- отслеживание статуса своих заданий;
- получение результатов расчетов.

Для возможности проведения простой и однократной авторизации пользователя в сети будет организована удостоверяющая федерация суперкомпьютерных центров [9], являющаяся расширением российской инфраструктуры аутентификации и авторизации проекта FEDURUS. Удостоверяющая федерация обеспечит технологию единого входа для доступа ко всем суперкомпьютерным ресурсам федерации на основе единственной учетной записи пользователя, поддерживаемой его организацией.

Процедура авторизации на суперкомпьютерном ресурсе с использованием удостоверяющей федерации будет выглядеть следующим образом. При входе на суперкомпьютер пользователь для подтверждения своих полномочий по использованию вычислительных ресурсов выбирает организацию — входящий в удостоверяющую федерацию институт (вуз), в котором он работает. Далее пользовательские данные перенаправляют на сервер аутентификации этого института, который проверяет логин и пароль пользователя с помощью локальной базы и отправляет службе авторизации на входе суперкомпьютерного ресурса подтверждение, что пользователю разрешено использование этого ресурса (например, в заданном объеме). На основании полученных от института данных службы авторизации и учета суперкомпьютерного ресурса принимают решение о допуске пользователя, а по завершении его заданий информируют институт пользователя об объемах израсходованных ресурсов. Сообщения, которыми обмениваются службы авторизации и учета суперкомпьютерного ресурса и сервер идентификации института, шифруются с использованием SSL-сертификатов.

Таким образом, пользователь получит возможность использования единой учетной записи для доступа на все суперкомпьютерные ресурсы, а его институт — гибкую систему управления выделенного ему суперкомпьютерного ресурса и распределения его между своими сотрудниками-пользователями. Причем распределение ресурса осуществляется исключительно внутри института и не требует взаимодействия с системами учета суперкомпьютерного центра, что снижает сложность управления системой в целом.

5. Облачная система хранения пользовательских данных

Системы хранения данных в дата-центрах, в т.ч. суперкомпьютерных, строятся по иерархическому принципу. Применительно к СКЦ информация первого уровня хранения представляет собой промежуточные результаты расчетов (временные файлы), размещаемые на локальных дисках узлов суперкомпьютера. К этому же уровню часто относят массивы на быстрых дисках (например, SSD) относительно небольшой емкости, но с высоким быстродействием. В таких массивах могут размещаться контрольные точки заданий, временно снятых с выполнения и возвращенных в очередь. Второй уровень образуют блочные дисковые устройства, обеспечивающие долговременное хранение пользовательских данных. Третий уровень представлен магнитными лентами, служащими для долговременного хранения редко используемых и/или архивных данных.

Для создаваемой сети СКЦ можно выделить отдельный, четвертый уровень, обеспечивающий отказоустойчивое хранение и глобальный доступ к данным. На этом уровне пользователям будет предоставлено общее пространство хранения данных, доступное из любого СКЦ. Для организации общего пространства часть ресурсов систем хранения данных СКЦ будет использовано для развертывания программно-определяемого хранилища данных (узла облачной системы хранения), реализующего распространенные протоколы доступа к облачным хранилищам (S3, WebDAV, SMB 3.0, HTTPS и пр.). На объединяемых узлах облачной системы хранения реализуется единое пространство имен файлов [10]. Доступ пользователей к данным будет организован через шлюз облачной системы хранения [11], преобразующий используемые в СКЦ файловые протоколы доступа к данным (NFS, SMB, FTP/SFTP и пр.) в распространенные протоколы доступа к облачным хранилищам. На уровне шлюза могут быть реализованы механизмы кэширования данных и политики автоматического перемещения «остывших» данных в облачную систему хранения.

Надежность хранения данных в облачной системе хранения обеспечивается либо репликацией данных, либо использованием алгоритмов рассредоточения информации (Information Dispersal Algorithm) [12] между узлами облачной системы хранения.

Использование единого пространства имен облачной системы хранения пользовательских данных позволит пользователю одинаковым образом работать со своими данными независимо от того, ресурсами какого СКЦ он воспользуется, а также упрощает формирование паспортов заданий глобальной очереди.

Рис. 3. Пилотный проект на базе МСЦ РАН и ССКЦ СО РАН

6. Результаты исследования характеристик канала связи для пилотной зоны сети СКЦ

На первом этапе (рис. 3) реализации проекта построения распределенной сети СКЦ предполагается объединение вычислительных ресурсов Межведомственного суперкомпьютерного центра РАН (МСЦ РАН, г. Москва) и Сибирского суперкомпьютерного центра на базе Института вычислительной математики и математической геофизики Сибирского отделения РАН (ССКЦ СО РАН, г. Новосибирск). Предпосылками такого объединения являются:

- близость архитектур суперкомпьютеров, поставленных в эти СКЦ группой компаний РСК;
- наличие канала связи пропускной способностью 10 Гбит/с, закрытие канала обеспечивается с помощью средств криптографической защиты информации от компании «С-Терра СиЭсПи».

Для реализации пилотного проекта распределенной сети СКЦ необходимым представлялось проведение экспериментального исследования характеристик и свойств предоставленного защищенного коммуникационного канала. Канал непосредственно связывал сети передачи данных высокопроизводительных вычислительных кластеров, обеспечивая при этом автоматическое шифрование передаваемых данных. Расстояние между Москвой и Новосибирском по прямой свыше 2800 км, что обусловило значительные задержки в линии связи (свыше 28 мс в одну сторону).

Рис. 4. Скорость передачи в прямом (нижняя ось абсцисс, сплошные линии на графике) и обратном (верхняя ось абсцисс, пунктирные линии на графике) направлениях в зависимости от размера файла и числа потоков

В эксперименте производилась передача файлов с узлов ВС МСП РАН на узлы ВС ССКЦ и обратно. Передача проводилась в несколько параллельных потоков, каждый из которых передавал множество файлов разного размера. Параметры эксперимента принимали следующие значения:

- размеры передаваемых файлов: 8 КБ, 16 КБ, 32 КБ, 64 КБ, 128 КБ, 256 КБ, 512 КБ, 1 МБ, 4 МБ, 16 МБ, 64 МБ, 256 МБ, 512 МБ, 1 ГБ;
- количество передаваемых файлов в одном потоке: 2, 10;
- количество потоков: 1, 2, 4, 8, 16, 32, 48, 64.

На рис. 4 приведены результаты проведенных в ходе эксперимента измерений. Видно, что предоставленный канал несимметричен — характеристики при передаче из Москвы в Новосибирск кратно лучше, чем в обратном направлении. Латентность оказалась равна времени передачи файлов объема 10 КБ.

Рост утилизации пропускной способности канала до момента насыщения происходит как по мере увеличения размеров передаваемых файлов, так и по мере роста числа потоков. Полная утилизация пропускной способности канала в прямом направлении происходит при

передаче файлов размером не менее 256 МБ при числе потоков 32 и выше. При передаче файлов в одном потоке утилизация падает на порядок.

После насыщения канала при дальнейшем увеличении количества потоков наблюдается падение скорости передачи. При асимметричном канале связи между СКЦ становится важным направление передачи данных, и именно меньшая пропускная способность в одном из направлений будет ограничивать производительность объединенной системы.

Полученные результаты позволяют говорить о целесообразности использования защищенного коммуникационного канала с высокой латентностью для организации общей для объединяемых СКЦ облачной системы хранения данных. С одной стороны, падение производительности канала на порядок при использовании одного потока (т.е. в условиях сети СКЦ при передаче данных одним пользователем) не даст возможности использовать общую СХД в качестве блочной системы хранения второго уровня иерархии. Общее файловое пространство распределенной сети СКЦ должно быть выделено в отдельный, четвертый, уровень иерархии систем хранения. Пользователи должны помещать данные на этот уровень только при перемещении данных между СКЦ или для заданий глобальной очереди. С другой стороны, канал обладает достаточно хорошей масштабируемостью (до 32 одновременных потоков), что является важным при многопользовательском режиме работы.

Заключение

В МСЦ РАН подготовлен проект распределенной сети суперкомпьютерных центров коллективного пользования, реализация которого позволит объединить вычислительные ресурсы суперкомпьютерных центров. Для системы управления заданиями и ресурсами сети СКЦ предложена децентрализованная архитектура на базе коллектива равноправных диспетчеров, поддерживающих единую глобальную очередь пользовательских заданий. Для хранения глобальной очереди выбрана документо-ориентированная распределенная база данных, позволяющая сочетать достоинства централизованного и децентрализованного подходов. Для реализации системы доступа к объединенным ресурсам сети СКЦ предлагается создать удостоверяющую федерацию суперкомпьютерных центров, которая позволит использовать единственную учетную запись пользователя для получения доступа ко

всем суперкомпьютерным ресурсам распределенной сети. Для возможности организации свободного перемещения пользовательских данных между СКЦ сети предлагается организовать облачное хранилище с единым пространством имён файлов.

Коллектив авторов приступил к реализации проекта, на начальном этапе которого была создана пилотная зона из двух ЦКП: МСЦ РАН и ССКЦ СО РАН. В рамках работ по реализации проекта было проведено экспериментальное исследование защищенного канала связи между МСЦ РАН и ССКЦ СО РАН, результаты которого показали целесообразность его использования для нужд проекта.

Список литературы

- [1] *О требованиях к центрам коллективного пользования научным оборудованием и уникальным научным установкам, которые созданы и (или) функционирование которых обеспечивается с привлечением бюджетных средств, и правилах их функционирования*, Постановление Правительства РФ от 17 мая 2016 г. № 429. ↑ ²⁴⁵
- [2] Г. И. Савин, В. В. Корнеев, Б. М. Шабанов, П. Н. Телегин, Д. В. Семенов, А. В. Киселёв, А. В. Кузнецов, О. И. Вдовикин, О. С. Аладышев, А. П. Овсянников. «Создание распределенной инфраструктуры для суперкомпьютерных приложений», *Программные продукты и системы*, 2008, №2, с. 2–7. ↑ ²⁴⁶
- [3] А. В. Баранов, Д. С. Николаев. «Использование контейнерной виртуализации в организации высокопроизводительных вычислений», *Программные системы: теория и приложения*, 7:1 (2016), с. 117–134, URL: http://psta.psiras.ru/read/psta2016_1_117-134.pdf ↑ ²⁴⁷
- [4] А. В. Баранов, А. И. Тихомиров. «Методы и средства организации глобальной очереди заданий в территориально распределенной вычислительной системе», *Суперкомпьютерные дни в России*, Труды международной конференции (25–26 сентября 2017 г., г. Москва), Изд-во МГУ, М., 2017, с. 776–787. ↑ ^{251, 252}
- [5] M. Stonebraker, U. Cetintemel. ““One size fits all”: An idea whose time has come and gone”, *21st International Conference on Data Engineering (ICDE 2005)*, 2005, pp. 2–11. ↑ ²⁵¹
- [6] D. Florencio, D. Oliveira, E. Freitas, F. Souza. “Which Fits Better? A Comparative Analysis about NoSQL Key-Value Databases”, *IEEE Latin America Transactions*, 15:11 (2017), pp. 2251–2256. ↑ ²⁵¹
- [7] S. Gupta, R. Rani. “A Comparative Study of Elasticsearch and CouchDB Document Oriented Databases”, *International Conference on Inventive Computation Technologies (ICICT)*. V. 1, 2016, pp. 211–214. ↑ ²⁵²
- [8] *Elastic Stack and Product Documentation*, <https://www.elastic.co/guide/index.html>, 2016 (дата обращения 13.11.2017). ↑ ²⁵²

- [9] А. П. Овсянников, Г. И. Савин, Б. М. Шабанов. «Удостоверяющие федерации научно-образовательных сетей», *Программные продукты и системы*, 2012, №4, с. 3–7. [↑] [253](#)
- [10] S. Koulouzis, A. Belloum, M. Bubak, P. Lamata, D. Nolte, D. Vasyunin, C. de Laat. “Distributed Data Management Service for VPH Applications”, *IEEE Internet Computing*, **20**:2 (2016), pp. 34–41. [↑] [254](#)
- [11] A. Kapadia, S. Varma, K. Rajana. *Implementing Cloud Storage with OpenStack Swift*, Packt Publishing, 2014, ISBN: 1782168052. [↑] [254](#)
- [12] М. Джонс. *Анатомия облачной инфраструктуры хранения данных. Модели, функции и внутренние детали*, <https://www.ibm.com/developerworks/ru/library/cl-cloudstorage/cl-cloudstorage-pdf.pdf>, 2016 (дата обращения 24.11.2017). [↑] [254](#)

Рекомендовал к публикации

Программный комитет

Шестого национального суперкомпьютерного форума *НСКФ-2017*

Пример ссылки на эту публикацию:

Б.М. Шабанов, А.П. Овсянников, А.В. Баранов и др. «Проект распределенной сети суперкомпьютерных центров коллективного пользования», *Программные системы: теория и приложения*, 2017, 8:4(35), с. 245–262.

URL: http://psta.psiras.ru/read/psta2017_4_245-262.pdf

Об авторах:

Борис Михайлович Шабанов

Директор МСЦ РАН, к.т.н., доцент, лауреат Государственной премии Российской Федерации в области науки и техники. Области научных интересов: высокопроизводительные вычисления, архитектура суперкомпьютеров и процессоров, САПР в электронике, параллельное программирование

e-mail: shabanov@jscs.ru

Алексей Павлович Овсянников

Заместитель директора МСЦ РАН. Области научных интересов: высокопроизводительные вычисления, научные телекоммуникации, развитие высокоскоростной сетевой инфраструктуры для суперкомпьютерных ресурсов

e-mail: ovsiannikov@jscs.ru

Антон Викторович Баранов

Ведущий научный сотрудник МСЦ РАН, к.т.н., доцент. Области научных интересов: организация высокопроизводительных вычислений, планирование заданий и управление вычислительными ресурсами в суперкомпьютерах, технологии виртуализации и облачных вычислений

e-mail:

antbar@mail.ru

Сергей Алексеевич Лещев

Научный сотрудник МСЦ РАН. Области научных интересов: организация хранения и доступа к данным в суперкомпьютерах, параллельные файловые системы

e-mail:

sergey.leshchev@jscs.ru

Борис Вячеславович Долгов

Научный сотрудник МСЦ РАН. Области научных интересов: организация и развитие сетевой инфраструктуры суперкомпьютерных центров коллективного пользования

e-mail:

borisdol@jscs.ru

Дмитрий Юрьевич Дербышев

Научный сотрудник МСЦ РАН. Области научных интересов: планирование заданий и управление вычислительными ресурсами в суперкомпьютерах, параллельное программирование

e-mail:

kirikus@jscs.ru

Boris Shabanov, Aleksey Ovsiannikov, Anton Baranov, Sergey Leshchev, Boris Dolgov, Dmitry Derbyshev. *The distributed network of the supercomputer centers for collaborative research.*

ABSTRACT. The article is devoted to the project of the creation of the distributed network of the supercomputer centers for collaborative research. The key technique for improving the availability and efficiency of the supercomputer centers is the integration of the centers into a joint distributed network. This technique allows to redistribute the workload between resources by the redirecting of the user's jobs from one center job queue to another depending on the instant resource usage.

The pilot project of the supercomputer centers distributed network is based on Joint Supercomputer Center in Moscow and Siberian Supercomputer Center in Novosibirsk. The centers communicate by the secure high-performance link. The results of the communication environment research are presented. (*In Russian*).

Key words and phrases: supercomputer center, center for collaborative research, secure communications, distributed computing.

References

- [1] *On the requirements for the centers of collective use of scientific equipment and unique scientific installations that are created and (or) whose operation is provided with the involvement of budgetary funds, and the rules for their functioning* (in Russian).
- [2] G. I. Savin, V. V. Korneyev, B. M. Shabanov, P. N. Telegin, D. V. Semenov, A. V. Kisel'ev, A. V. Kuznetsov, O. I. Vdovikin, O. S. Aladyshev, A. P. Ovsiannikov. "Creation of distributed infrastructure for supercomputer applications", *Programmnyye produkty i sistemy*, 2008, no.2, pp. 2–7 (in Russian).
- [3] A. V. Baranov, D. S. Nikolayev. "The use of container virtualization in the organization of high-performance computing", *Program Systems: Theory and Applications*, 7:1 (2016), pp. 117–134 (in Russian), URL: http://psta.psiras.ru/read/psta2016_1_117-134.pdf
- [4] A. V. Baranov, A. I. Tikhomirov. "Methods and tools for organizing the global job queue in the geographically distributed computing system", *Superkomp'yuternyye dni v Rossii*, Trudy mezhdunarodnoy konferentsii (25–26 sentyabrya 2017 g., g. Moskva), Izd-vo MGU, M., 2017, pp. 776–787 (in Russian).
- [5] M. Stonebraker, U. Cetintemel. "One size fits all": An idea whose time has come and gone", *21st International Conference on Data Engineering (ICDE 2005)*, 2005, pp. 2–11.
- [6] D. Florencio, D. Oliveira, E. Freitas, F. Souza. "Which Fits Better? A Comparative Analysis about NoSQL Key-Value Databases", *IEEE Latin America Transactions*, 15:11 (2017), pp. 2251–2256.
- [7] S. Gupta, R. Rani. "A Comparative Study of Elasticsearch and CouchDB Document Oriented Databases", *International Conference on Inventive Computation Technologies (ICICT)*. V. 1, 2016, pp. 211–214.

- [8] *Elastic Stack and Product Documentation*, <https://www.elastic.co/guide/index.html>, 2016 (data obrashcheniya 13.11.2017).
- [9] A. P. Ovsyannikov, G. I. Savin, B. M. Shabanov. “Identity federation of the research and educational networks”, *Programmnyye produkty i sistemy*, 2012, no.4, pp. 3–7 (in Russian).
- [10] S. Koulouzis, A. Belloum, M. Bubak, P. Lamata, D. Nolte, D. Vasyunin, C. de Laat. “Distributed Data Management Service for VPH Applications”, *IEEE Internet Computing*, **20**:2 (2016), pp. 34–41.
- [11] A. Kapadia, S. Varma, K. Rajana. *Implementing Cloud Storage with OpenStack Swift*, Packt Publishing, 2014, ISBN: 1782168052.
- [12] M. Dzhons. *Anatomiya oblačnoy infrastruktury khraneniya dannykh. Modeli, funktsii i vnutrenniye detali*, <https://www.ibm.com/developerworks/ru/library/cl-cloudstorage/cl-cloudstorage-pdf.pdf>, 2016 (data obrashcheniya 24.11.2017); M. Jones. *Anatomy of the cloud storage infrastructure. Models, functions and internal parts*, <https://www.ibm.com/developerworks/cloud/library/cl-cloudstorage/>, 2010.

Sample citation of this publication:

Boris Shabanov, Aleksey Ovsiannikov, Anton Baranov, Sergey Leshchev, Boris Dolgov, Dmitriy Derbyshev. “The distributed network of the super-computer centers for collaborative research”, *Program systems: Theory and applications*, 2017, **8**:4(35), pp. 245–262. (*In Russian*).

URL: http://psta.psiras.ru/read/psta2017_4_245-262.pdf