

О развитии инфраструктуры суперкомпьютерных и распределенных вычислений в СО РАН

Ю.И. Шокин, М.П. Федорук, Д.Л. Чубаров, А.В. Юрченко

Аннотация. В работе представлен обзор современного состояния инфраструктуры для высокопроизводительных и распределенных вычислений в Сибирском отделении Российской академии наук. Создание подобной инфраструктуры представляет собой последовательный процесс, включающий в себя создание ресурсных центров и компьютерных сетей, а также изучение потребностей настоящих и потенциальных пользователей инфраструктуры. В статье перечислены основные вычислительные ресурсы, имеющиеся в ресурсных центрах и организациях СО РАН, представлено состояние сетевой инфраструктуры и обсуждаются перспективы её дальнейшего развития.

Ключевые слова: киберинфраструктура науки, высокопроизводительные и распределенные вычисления, суперкомпьютерный центр коллективного пользования, сеть передачи данных.

Введение

Сибирское отделение РАН является региональным объединением 87 научно-исследовательских, опытно-конструкторских, производственных организаций РАН, работающих во всех областях науки, а также подразделений, обеспечивающих функционирование инфраструктуры отделения и его научных центров, расположенных на территории Сибири в 7 областях, 2 краях и 4 республиках. Обеспечение их высокопроизводительными вычислительными системами, системами хранения и обработки данных является одним из приоритетных направлений развития инфраструктуры науки в Сибири.

Потребности учреждений СО РАН в современных инструментах для высокопроизводительных вычислений, а также в средствах для обработки и размещения данных весьма велики и постоянно растут. В отделении функционирует большое число институтов, деятельность которых связана с накоплением данных как путем мониторинга различных природных, технологических, социальных и экономических

процессов, так и в результате численного моделирования, а также вследствие сбора разнообразной информации, например, библиотечного характера. Далеко не для всех таких институтов создание и содержание центров обработки данных (ЦОД) является профильным видом деятельности, особенно с учетом ее высокой трудоемкости и финансовой затратности.

Практика показывает, что консолидация усилий заинтересованных организаций является эффективным подходом для решения столь масштабных задач, как удовлетворение потребностей в высокопроизводительных средствах для вычислений, накопления, обработки и хранения данных. Соответствующие тенденции в мировой IT-индустрии называются «облачными» (cloud) технологиями, и все большее количество информационных и вычислительных услуг переводятся в режим работы «по требованию» (on-demand) в крупные ЦОД коллективного пользования. При этом концентрация всех ресурсов в одном месте не является предпочтительной схемой, вместо этого создается и развивается множество ЦОД, объединенных общей концепцией, соединенных

специализированной организационной и технологической инфраструктурой, а их ресурсы разделяются между заинтересованными потребителями. Таким образом достигается максимальная эффективность использования финансовых и других возможностей с обеспечением высокого уровня доступности ресурсов и качества обслуживания пользователей.

Усилия СО РАН по обеспечению своих сотрудников высокопроизводительной компьютерной техникой, развитию инфраструктуры суперкомпьютерных вычислений и распределенного хранения и обработки данных направлены на создание и развитие крупных ЦОД, обеспечение условий для успешного взаимодействия потребителей и поставщиков высокопроизводительных ресурсов, а также на поддержание взаимовыгодных партнерских отношений с профильными центрами ВУЗов региона. Выделенным направлением является развитие сетевой инфраструктуры научных центров и отделения в целом для повышения доступности ресурсов и качества обслуживания их пользователей. Кроме того, развитая сетевая инфраструктура является необходимым условием для успешного внедрения технологий распределенных вычислений, хранения и обработки данных.

1. Центры обработки данных СО РАН и сибирского региона и их ресурсы

Крупные ЦОД созданы практически в каждом научном центре СО РАН, что позволяет научным центрам вести относительно независимую политику обслуживания потребителей. Однако можно выделить некоторые ЦОД, обладающие наиболее мощными информационно-вычислительными ресурсами и предоставляющие к ним доступ для сотрудников всех институтов отделения.

1.1. Новосибирский научный центр СО РАН и НГУ

В Новосибирске расположен Сибирский суперкомпьютерный центр (ССКЦ), функционирующий с 2001 года на базе Института вычислительной математики и математической геофизики СО РАН. Его материально-техническая база за последний год существенно обновилась, и в настоящее время основными ресурсами ССКЦ являются:

1. вычислительный кластер (192 вычислительных узла) на 4-ядерных процессорах Intel Xeon™ (всего 1536 ядер) архитектуры x86-64 (Clovertown, Nehalem) с пиковой производительностью порядка 16,5 TFLOPS, кластерной системой хранения данных на 48 ТБ и интерконнектом Infiniband QDR/DDR;

2. вычислительный кластер на процессорах Intel Itanium™ (архитектура IA64) с вычислительной мощностью порядка 1 TFLOPS;

3. система для длительного хранения данных емкостью 6 ТБ.

ССКЦ предоставляет свои ресурсы по требованию тех или иных организаций СО РАН и развивается в соответствии с проектом суперкомпьютерных вычислений СО РАН и Президентской программой Интеграция (совместно с НГУ) при поддержке РФФИ.

Специализированный ЦОД расположен в Институте вычислительных технологий (ИВТ) СО РАН. Он ориентирован, в первую очередь, на хранение и обработку данных и обеспечивает основные потребности ИВТ СО РАН и функционирующего на его базе Центра мониторинга социально-экономических и природных процессов СО РАН. Развитие этого ЦОД происходит, в основном, за счет средств, привлекаемых ИВТ СО РАН, и направлено на создание крупного хранилища данных с возможностью их локальной или удаленной обработки – Интегрированной системы обработки данных (ИСОД). В настоящее время ИСОД состоит из:

- высокопроизводительной расширяемой системы хранения данных (СХД) Enterprise уровня EMS Clarion CX4-120, емкость которой сейчас составляет порядка 130 ТБ (введена в эксплуатацию в 2010 году);

- трех нерасширяемых СХД Proware SB-3164 с суммарной емкостью 48 ТБ (эксплуатируются с 2008 года);

- системы хранения типа SAN – HP MSA1500 с емкостью 30 ТБ (закуплена в 2005 году, неоднократно расширялась, в настоящее время ее эксплуатация завершается);

- вычислительного кластера и сервера с общей памятью на основе двухъядерных процессоров AMD (эксплуатируется с 2007 года);

- вычислителя на основе GPGPU ускорителя NVidia Tesla s1070 с пиковой производительностью более 4 TFLOPS (эксплуатируется с 2009 года, с новым хост-сервером – с 2010 года);

- восьмиузловой кластера на основе шестиядерных процессоров AMD с высокопроизводительным интерконнектом Infiniband DDR и пиковой производительностью порядка 1 TFLOPS (введен в эксплуатацию в 2010 году).

Ведутся работы по интеграции этих ресурсов в распределенную информационно-вычислительную среду СО РАН.

Для решения задач аэрогидродинамики в Институте теоретической и прикладной механики (ИТПМ) СО РАН установлен вычислительный кластер T-платформы T-Edge-32 с пиковой производительностью 2,5 TFLOPS. Данные, получаемые с датчиков ускорителей частиц в Институте ядерной физики (ИЯФ) СО РАН, накапливаются в созданном на его базе ЦОД с СХД емкостью 25 TB и небольшим вычислительным кластером для их обработки.

Мощный импульс возрождению вычислительной отрасли России в XXI веке дали суперкомпьютерные программы ВУЗов. В Новосибирске это, в первую очередь, государственный университет (НГУ), расположенный в Академгородке. НГУ имеет давние и тесные связи с СО РАН. В рамках этого сотрудничества институтами отделения поддерживается деятельность большинства кафедр университета, что позволяет готовить высококвалифицированные исследовательские кадры. Очередным этапом успешного взаимодействия НГУ и институтов СО РАН стало возрождение вычислительного центра НГУ, который, в новых условиях, существенно расширил номенклатуру решаемых задач и предоставляемых услуг.

В конце 2007 года по государственной программе поддержки инновационных проектов ВУЗов для НГУ была закуплена вычислительная система, которая к моменту ввода в эксплуатацию была самой мощной в Новосибирске и третьей по производительности в регионе. В подготовке этой закупки – анализе, сопоставлении потребностей и возможностей, подборе вариантов, разработке технических требований и конкурсной документации – принимал активное участие ИВТ СО РАН. В результате был

создан крупный некоммерческий ЦОД. К настоящему времени основными ресурсами ИВЦ НГУ являются:

- вычислительный кластер на базе блэйд-серверов Hewlett-Packard с процессорами Intel Xeon Nehalem (96 вычислительных узлов) и Clovertown (64 вычислительных узла), объединенных сетью Infiniband QDR/DDR с суммарным объемом оперативной памяти более 2,5 TB и пиковой производительностью 13,2 TFLOPS;

- кластерная СХД Hewlett-Packard SFS на технологии Lustre с объемом дисковой подсистемы 24 TB;

- СХД на быстрых дисках HP EVA 4100 емкостью 4 TB;

- 16-ядерный вычислительный сервер с общей памятью объемом 128 GB.

Введен в эксплуатацию новый 64-ядерный (8 процессоров Intel Xeon X7560) вычислительный сервер с 2 TB общей памяти для решения ресурсоемких задач с высокими требованиями к объему оперативной памяти.

Совместными усилиями специалистов НГУ и ИВТ СО РАН осуществляется мониторинг инженерной инфраструктуры ЦОД и серверного оборудования, разработана автоматическая подсистема останова компонентов ЦОД в случае форс-мажорных обстоятельств, например, отключения электроснабжения; создана специализированная подсистема и осуществляется мониторинг приложений, учет использования ресурсов, статистика и мониторинг загрузки вычислительных компонентов комплекса; осуществляется поддержка пользователей как по вопросам освоения системы, так и по частным проблемам создания эффективных параллельных вычислительных кодов.

1.2. Томск, Красноярск, Иркутск и другие

В Томском научном центре введена в строй новая вычислительная система, которая представляет собой объединение специального сервера управления, надежного сетевого хранилища данных (NAS) с мощным вычислительным сервером на основе графических процессоров NVidia и высокопроизводительных процессоров Intel. Сервер управления используется для предоставления удаленного доступа к высоко-

производительным ресурсам, мониторинга различных параметров состояния вычислительно-го сервера (загрузки CPU, температуры, занятости дисковых ресурсов, количества свободной оперативной памяти и т.д.), постановки вычислительных задач в очередь. Основной используемый вычислитель - Nvidia Tesla C1060 - представляет собой систему на основе графического процессора (GPU) с поддержкой архитектуры CUDA (Compute Unified Device Architecture). По своим вычислительным возможностям он сопоставим с небольшим кластером – пиковая производительность C1060 составляет примерно 1 TFLOPS.

В Томске расположена одна из самых известных сибирских вычислительных систем – кластер СКИФ-Cyberia, установленный в Томском государственном университете (ТГУ). Хотя комплекс и не является сейчас самым мощным компьютером в России, пробыв на этом пьедестале менее года, тем не менее он продолжает оставаться одной из крупнейших некоммерческих вычислительных систем, обладая после обновления пиковой производительностью порядка 30 TFLOPS. В Томском политехническом университете (ТПУ) также идет наращивание возможностей вычислительного центра, в 2010 году здесь был установлен новый кластер на процессорах Intel Xeon Nealem с пиковой производительностью 3,7 TFLOPS.

В Красноярске, на базе Института вычислительного моделирования (ИВМ) СО РАН функционирует Красноярский суперкомпьютерный центр (КСКЦ) СО РАН. Здесь, совместно с Сибирским Федеральным университетом (СФУ), эксплуатируется вычислительная система кластерного типа на процессорах Intel Xeon Quad Core (архитектура x86-64) с пиковой производительностью более 1 TFLOPS и системой хранения данных емкостью 3,5 ТВ. Кроме того в ИВМ СО РАН активно развиваются методы гибридной параллелизации с применением ускорителей вычислений на базе GPGPU, для чего закуплены рабочие станции на основе NVidia Fermi, планируется установка мощного вычислительного сервера с такими вычислительными модулями. Здесь же установлен 16-ядерный SMP сервер на основе процессоров AMD Opteron. Вычислительный комплекс СФУ

включает в себя кластер с пиковой производительностью порядка 17 TFLOPS и две СХД емкостью 50 ТВ и 20 ТВ.

Мощная кластерная вычислительная система (20 вычислительных узлов с процессорами Intel Xeon Clovertown, пиковая производительность порядка 1,5 TFLOPS) установлена в Институте динамики систем и теории управления (ИДСТУ) СО РАН, на базе которого функционирует Иркутский суперкомпьютерный центр СО РАН. Здесь также осваиваются технологии гибридной параллелизации с использованием многопоточковых ускорителей вычислений на основе GPGPU – уже эксплуатируются два сервера с платами NVidia Tesla c1060, запланирована закупка гибридного кластера на базе NVidia Fermi и серверов плотной упаковки.

Нужно отметить, что гибридные вычисления с использованием серверов и кластеров, оснащенных ускорителями вычислений на GPGPU, являются главным трендом 2010 года, который неизбежно будет продолжен в будущем. Кроме уже названных вычислительных подразделений СО РАН, закупивших необходимое оборудование и осваивающих соответствующую технологию, по достоинству она была оценена и в менее крупных научных центрах и подразделениях СО РАН. Так, в Омский суперкомпьютерный центр на базе Омского филиала Института математики СО РАН закуплен двухмодульный кластер Tesla Meijin на базе ускорителей NVidia Tesla c1060.

Все владельцы крупных академических вычислительных и информационных ресурсов в сибирском регионе имеют серьезные планы по расширению своих систем, наращиванию их мощности и функционала, чтобы обеспечить сибирскую науку необходимыми для решения сложных задач возможностями.

2. Высокоскоростной сегмент СПД СО РАН

ИВТ СО РАН практически с самого своего основания координирует работы по развитию и поддержанию крупнейшей некоммерческой корпоративной сети передач данных (СПД) за Уралом [1]. Сеть объединяет между собой институты научных центров, для чего прокла-

дываются оптоволоконные линии связи, закупается современное телекоммуникационное оборудование, осуществляется постоянный мониторинг. Кроме того, централизованно арендуются линии связи для объединения СПД научных центров между собой и обеспечения выхода в глобальную сеть Интернет.

Развитие сетей связи и компьютерных сетей позволило обеспечить пользователей суперкомпьютеров и высокопроизводительных систем обработки и хранения данных возможностью удаленной работы с ними. Такой режим позволяет, в частности, для решения различных специфических задач использовать системы, расположенные в ЦОД, удаленных и от пользователя, и друг от друга. Эффективное решение задач распределенной обработки и хранения данных требует высокоскоростных соединений между системами и ЦОД, участвующими в этом процессе. Каналы общего пользования, предназначенные для организации доступа в сеть Интернет, не всегда отвечают предъявляемым требованиям. Например, для решения различного рода интерактивных задач необходимы соединения между пользователем и системами, с которыми он взаимодействует, обеспечивающие малые времена отклика, что невозможно совместить с мерами, предпринимаемыми в сетях общего пользования для обеспечения безопасности и защиты от сетевых атак.

Важной специфической особенностью сибирского региона и СО РАН является существенная территориальная разнесенность крупных городов и расположенных в них научных центров (НЦ). Так, в СО РАН можно выделить четыре крупных НЦ, концентрирующих основные информационные и вычислительные ресурсы: самый большой – Новосибирский (ННЦ), менее крупные – Томский (ТНЦ), Красноярский (КНЦ) и Иркутский (ИНЦ).

Одно из следствий низкой плотности населения и значительных расстояний между центрами научных исследований в Сибири – географически низкая плотность информационно-вычислительных ресурсов и, как дальнейшее следствие, трудности с развитием объединяющей их сетевой инфраструктуры. В настоящее время локальные СПД различных НЦ СО РАН объединены выделенными каналами связи, ко-

торые арендуются на коммерческой основе. Высокая стоимость аренды заставляет искать пути оптимизации расходов на эти каналы связи, которые загружены сейчас только передачей небольшого объема служебной информации.

Развитие информационных и вычислительных сервисов приводит к необходимости передачи все больших объемов данных. В ситуации, когда емкость каналов связи между НЦ невелика, нуждающимся в таких сервисах приходится ограничиваться ресурсами в пределах своего НЦ или города. Чтобы изменить ситуацию коренным образом, необходимо развивать СПД в направлении качественного роста емкости каналов связи между НЦ. Параллельно необходимо активизировать пользователей, привлекать их к использованию уникальных информационных и вычислительных сервисов, предоставляемых ресурсными центрами других НЦ, а также к созданию и публикации новых сервисов с использованием соответствующих аппаратных и программных ресурсов СО РАН.

Примером такого уникального сервиса является доступ к данным дистанционного (спутникового) зондирования и результатам их обработки, накапливаемых Центром мониторинга в ЦОД ИВТ СО РАН [2]. Востребованность этих данных высока, но в настоящий момент основной их объем, требуемый для деятельности институтов СО РАН за пределами ННЦ, передается по публичным каналам сети Интернет. Чтобы обеспечить доступ академических организаций к столь большим объемам данных без использования арендуемых коммерческих каналов связи, необходимо, следуя примеру Евросоюза и США, реализовать возможность безвозмездной аренды у операторов дальней связи неиспользуемых оптоволоконных линий (так называемого «темного волокна») для соединения крупных научных центров и информационно-вычислительных ресурсов как Академии наук, так и ВУЗов. Это, в свою очередь, требует принятия и скорейшей реализации соответствующих федеральных программ поддержки развития научно-образовательных компьютерных сетей.

В рамках своих возможностей институты и подразделения СО РАН, ответственные за развитие СПД, решают обозначенные проблемы путем развития специализированных локаль-

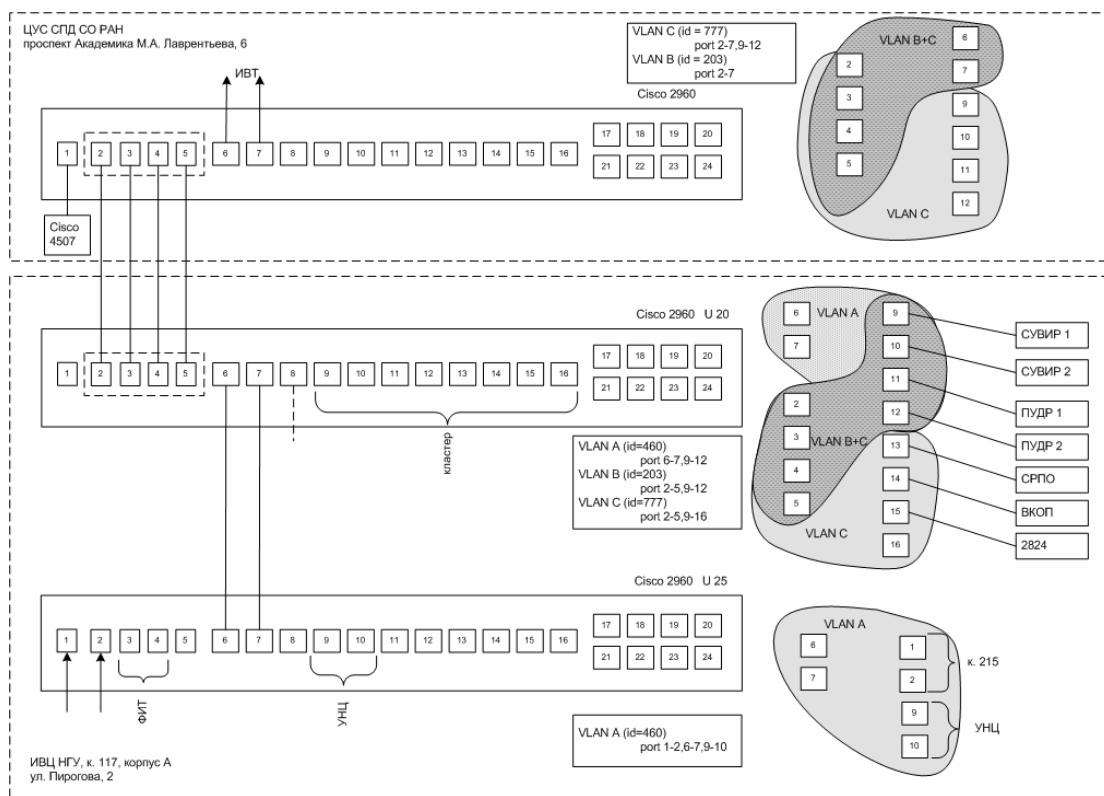


Рис. 1. Схема первой очереди сегмента ИВТ–НГУ высокоскоростной технологической сети

ных сетевых структур и подключения к ним как поставщиков высокопроизводительных ресурсов, так и их наиболее активных потребителей. Одним из первых этапов создания такой сетевой структуры в ННЦ СО РАН стало создание высокоскоростного соединения между ИВЦ НГУ и ЦУС СПД СО РАН. В ИВТ СО РАН была разработана схема подключения к СПД СО РАН вычислительного кластера НГУ с предоставлением удаленного доступа к ресурсам для всех заинтересованных организаций СО РАН и ВУЗов Сибирского региона. Схема подключения основана на создании виртуальной локальной сети с использованием протокола IEEE 802.1Q, позволяющей в рамках одного физического канала связи и одного физического интерфейса разделять доступ для нескольких виртуальных сетей. Первоначально подключение было осуществлено путем агрегирования четырех гигабитных линий (Рис. 1), впоследствии была произведена установка оборудования с 10 Gb портами, а гигабитные линии оставлены для обеспечения отказоустойчивости.

Открытие доступа к высокопроизводительным ресурсам НГУ существенно расширило возможности сотрудников институтов СО РАН по использованию вычислений на многопроцессорных гибридных системах. Вычислительными ресурсами НГУ активно пользуются институты СО РАН: ИВТ, ИМ, ИГиЛ, ИФП, ИЦиГ, ИВМиМГ, ИАиЭ, ИТПМ, ИЛФ, ИК, ИТ, ОИГТМ, ИНГ, ИЯФ, ИВМ и другие организации академического профиля.

На базе соединения НГУ – СО РАН создан специализированный технологический высокоскоростной сегмент СПД СО РАН. В настоящий момент он устроен по схеме типа «звезда» (Рис. 2), в центре которой расположен мощный коммутационный узел на базе коммутатора Cisco 4900M, оснащенного 10GbE и SFP10G модулями. «Лучи звезды» уходят в ЦОД, расположенные в НГУ, ССКЦ СО РАН, ИВТ СО РАН и ИЯФ СО РАН, где установлены коммутирующие устройства Cisco 3560E, в каждом из которых по два SFP10G модуля, наряду с достаточным количеством 1GbE портов.

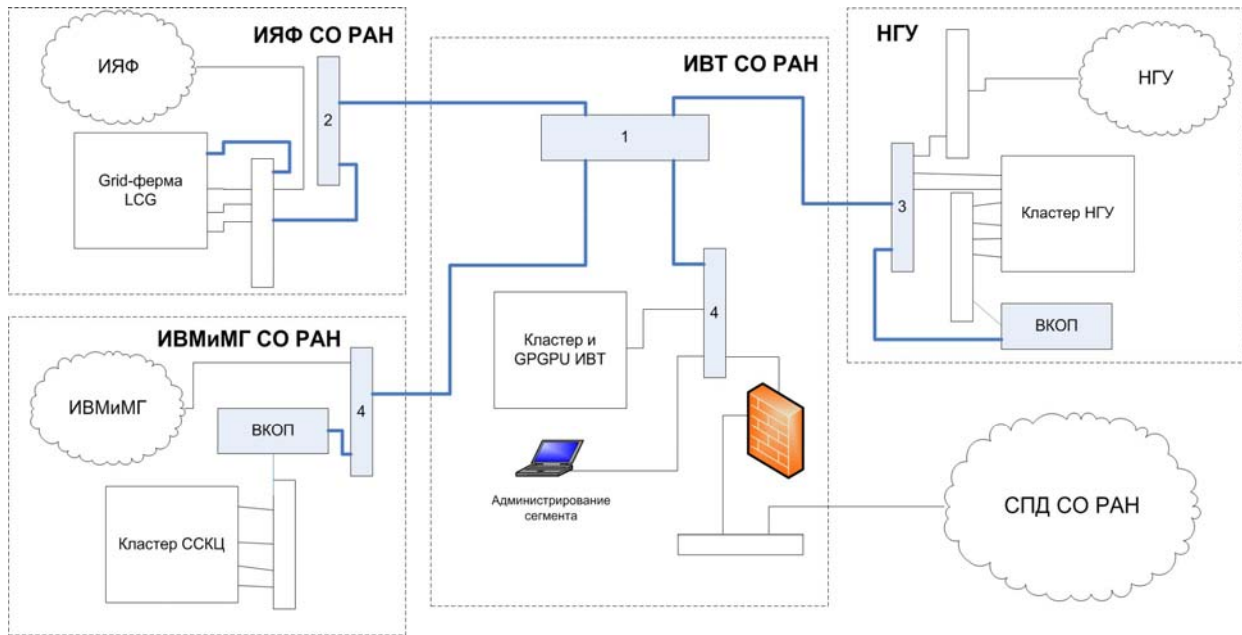


Рис. 2. Общая схема первой очереди высокоскоростного сегмента СПД СО РАН

Такая схема позволила подключить к сегменту со стороны клиентов по одному устройству на скорости 10 Gb и дополнительные устройства на скоростях 1 Gb, а также обеспечить резервные 1 Gb каналы. В вычислительных центрах коллективного пользования – ССКЦ СО РАН и ИВЦ НГУ – к высокоскоростным портам в настоящий момент подключены высокопроизводительные вычислительные серверы с 16-процессорными ядрами и большим объемом оперативной памяти (от 128 GB) и через них имеется выход в высокопроизводительные кластерные сети Infiniband. Соединение высокоскоростной сетью производительных вычислительных ресурсов ССКЦ СО РАН и ИВЦ НГУ расширяет возможности их пользователей, позволяя объединять вычислительные мощности на основе Grid-технологий [3,4].

ИВТ СО РАН совместно с ИЯФ СО РАН проведены эксперименты по передаче данных внутри созданного высокоскоростного сегмента в интенсивном режиме. В рамках эксперимента были достигнуты скорости обмена данными между двумя узлами в сегменте на уровне 6 Gb и выше. В экспериментах по насыщению канала была достигнута степень использования ресурсов свыше 90%.

Важным следующим этапом развития высокоскоростного сегмента СПД СО РАН должно стать обеспечение соединениями с максимальной пропускной способностью от 1 Gb настоящих и потенциальных пользователей подключенных к нему вычислительных ресурсов, систем обработки и хранения данных. Для этого необходимо вести совместную работу с организациями, сотрудники которых являются активными пользователями высокопроизводительных ресурсов. Такая работа должна включать разъяснение новых технологических возможностей, повышение скорости подключения организаций СО РАН к корпоративной СПД, проектирование локальных сетей с учетом требований по обеспечению скоростными (от 1 Gb) подключениями конечных пользователей этих сетей и созданию внутренних высокоскоростных (10 Gb) магистралей.

В ТНЦ, КНЦ и ИНЦ СО РАН в настоящее время развиваются сети на основе 1 Gb каналов, организуется подключение к этим сетям ВУЗов региона. Так, запущены 1 Gb каналы связи ТНЦ СО РАН – ТГУ для предоставления сотрудникам СО РАН прямого доступа к вычислительным ресурсам кластера Cyberia, установленного в ТГУ, аналогичные работы выпол-

нены в КНЦ СО РАН для объединения ресурсов СФУ и КСКЦ.

3. Проекты распределенной обработки данных и метавычислений

3.1. Распределенные (мета-) вычисления с использованием RASX-MPI

Проведен ряд экспериментов по выполнению параллельных программ с одновременным использованием ресурсов нескольких географически разделенных вычислительных кластеров, расположенных в различных организациях с использованием библиотеки RASX-MPI. Библиотека RASX-MPI создана в Штуттгартском суперкомпьютерном центре и с технической точки зрения является прослойкой между MPI-приложением и нижележащей MPI-библиотекой и обладает следующими преимуществами:

- не требуется изменения кода MPI-приложения, достаточно его откомпилировать с библиотекой RASX;
- осуществляется сжатие данных в межкластерных сообщениях перед отправкой;
- производится буферизация передаваемых данных;
- имеется поддержка аппаратно-программной гетерогенности кластеров;
- для внешней коммуникации используются стандартные протоколы: TCP, ATM, SSL;
- внутрикластерные сообщения используют стандартные оптимизированные MPI-вызовы нижележащей MPI библиотеки.

Концептуальная схема взаимодействия кластеров с использованием библиотеки RASX проиллюстрирована на Рис. 3. На каждом кластере для коммуникации с другими кластерами используется два процесса. Один – RASX In-Server – фоновый процесс для отправки исходящих сообщений, другой – RASX Out-Server – фоновый процесс для приема входящих сообщений. RASX In-Server каждого кластера устанавливает связь с RASX Out-Server-ами остальных кластеров.

Расчеты в пределах высокоскоростного сегмента СПД СО РАН, на базе ресурсов ИВТ СО РАН и НГУ, показали принципиальную работоспособность технологии и возможность ее эффективного использования для решения широкого класса задач.

На основе указанной выше технологии проводились работы по объединению вычислительных ресурсов ТНЦ СО РАН, ТГУ и ТПУ. Для этого произведено необходимое согласование специализированного программного обеспечения, даны подробные рекомендации по настройкам программного обеспечения для метавычислительных приложений. Для тестирования созданного метакомпьютера Томскими специалистами разработана программа решения трехмерного нестационарного адвективно-диффузионного уравнения конечно-разностными методами с использованием явных разностных схем [5]. Параллельная реализация этого метода с использованием одномерной декомпозиции применена для исследования влияния способа организации межпроцессорных обменов на основе стандарта MPI. Выполнены теоретические оценки ускорения разработанной параллельной программы и дано сравнение с результатами вычислительных экспериментов на кластере ИОА СО РАН.

3.2. Удаленные расчеты на CAE-системах

С использованием высокопроизводительных вычислительных ресурсов НГУ и КСКЦ СО РАН в ИВТ СО РАН начат цикл исследований по сравнению трехмерных и двумерных моделей тонкостенных многослойных композитных элементов конструкций. В рамках этого исследования требуется проводить тестовые расчеты задач трехмерной теории упругости с высокой степенью детализации и повышенными требованиями к точности.

Задачи трехмерной теории упругости решаются с помощью пакета ANSYSTM Mechanical методом конечных элементов, в котором высокая детализация и точность приводят к существенному росту требуемого числа элементов, а наличие обширных контактных областей – к большому числу сложных элементов высокой размерности. Вследствие этого растут и требования к вычислительным ресурсам. Так, попытки выполнить расчет на современной персональной рабочей станции с 4-ядерным процессором и 4GB оперативной памяти привели к необходимости применения модели расчета с использованием дискового пространства для размещения временных данных. Поскольку

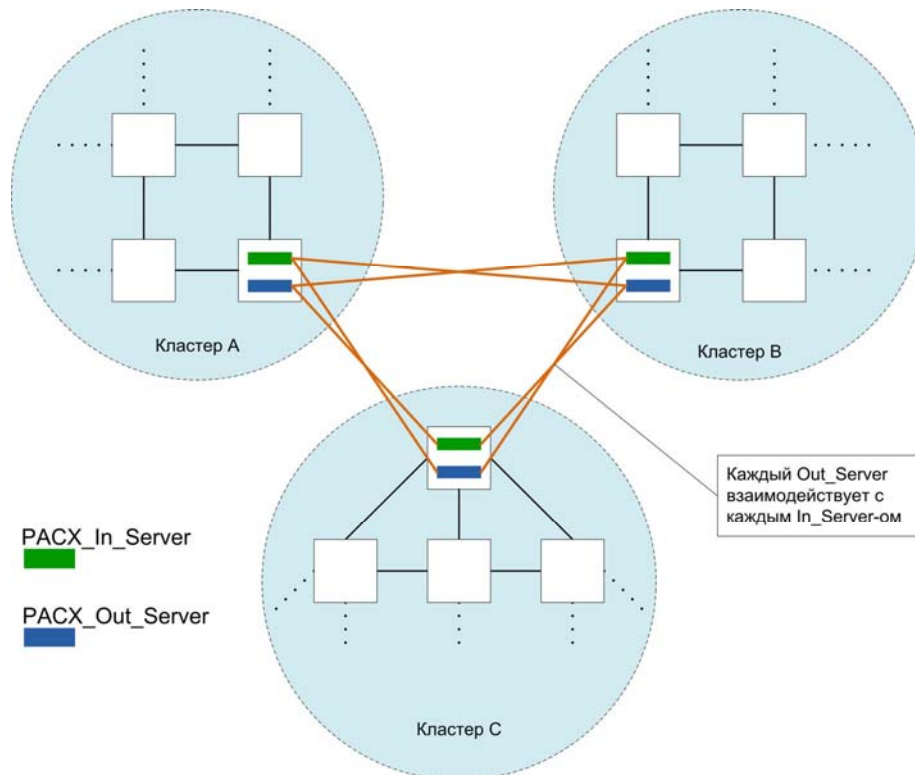


Рис. 3. Концептуальная схема взаимодействия трех кластеров при использовании PACX

операции с жестким диском являются очень медленными, то это привело к более чем 100-кратному увеличению времени расчета, а в ряде случаев просто не позволило его завершить, так как самые большие задачи требовали для своего решения до 120 GB памяти.

Для разрешения проблемы были привлечены ресурсы с большим объемом общей памяти – высокопроизводительные 16-ядерные серверы в НГУ и ССКЦ СО РАН. Это позволило не только разрешить проблему использования жестких дисков при выполнении интенсивных вычислений и обменов данными, но и существенно сократить время вычислений за счет их параллелизации. Необходимый инструментарий (параллельный и распределенный решатель, управление удаленным решателем – RSM) встроен в пакет программ ANSYS, понадобилась лишь незначительная его адаптация к существующей конфигурации сети и вычислительных устройств. При этом наличие высокоскоростных соединений существенно повысило комфортность работ по передаче исходных расчетных данных, мониторингу и кон-

тролю выполнения расчетов, передаче результатов конечному пользователю в ИВТ СО РАН.

Так, типичными объемами исходных данных и результатов расчетов в вышеназванных задачах анализа трехмерного упругого состояния тонкостенных многослойных элементов конструкций были 100MB и 1GB. Для их передачи с использованием высокоскоростных соединений требовалось 2 и 20 секунд соответственно, при том, что скорость ограничивалась не возможностями сети, а дисковой подсистемой конечной пользовательской рабочей станции. При использовании обычных каналов связи для передачи данных требуется в 10 – 100 раз больше времени, из-за чего работа в пакетном режиме становится существенно менее удобной, а в интерактивном – практически невозможной.

Таким образом, наличие высокоскоростных каналов связи с центрами коллективного пользования вычислительными и информационными ресурсами позволяет существенно расширить возможности последних, повышая их привлекательность с точки зрения доступности и удобства использования. А широкий охват и

наличие большого числа пользователей, в свою очередь, позволяют поддерживать высокий уровень эффективности использования вычислительных ресурсов.

3.3 Распределенная обработка данных с датчиков ускорителей частиц

Совместными усилиями ИВТ СО РАН, ИВЦ НГУ и ИЯФ СО РАН разработана, создана и отлажена схема использования высокопроизводительных ресурсов НГУ для распределенной обработки данных с датчиков ускорителей частиц, расположенных в ИЯФ СО РАН, а также других организациях, занимающихся экспериментальными исследованиями в области физики элементарных частиц и высоких энергий, таких как CERN (Швейцария) с большим адронным коллайдером (БАК).

Для распределенной обработки данных используются технологии виртуализации и так называемых «облачных» вычислений (cloud computing). В конкретной реализации высокоскоростные соединения используются для передачи образов виртуальных машин с расчетными программами, данных, необходимых для проведения расчетов, и результатов обработки. Схема использования распределенных ресурсов выдержала несколько модификаций и в настоящее время выглядит следующим образом.

- Программа обработки сконфигурирована в виде виртуальной машины (VM), VM могут быть конвертированы в форматы, специфичные для конкретных серверов виртуальных машин (СВМ).

- Образы клонируются, распределяются по СВМ, в качестве которых выступают выделенные узлы кластера НГУ, и запускаются.

- При запуске VM подключают тома сетевой файловой системы, которая размещена физически на СХД в ИЯФ СО РАН, и исполняют заложенный вычислительный код.

Обмен данными между VM и СХД является основным генератором трафика и может запросто поглотить всю пропускную способность 10Gb канала, что заставляет задуматься над оптимизацией использования сетевого ресурса и его дальнейшего наращивания. В настоящее время для целей обработки данных с датчиков ускорителей частиц выделяется до 16 вычисли-

тельных узлов кластера НГУ. Решается вопрос о выделении ресурсов для включения созданной инфраструктуры в RDIG.

Заключение

Наряду с ростом потребностей в суперкомпьютерных вычислениях и высокопроизводительной обработке данных растут и требования к качеству предоставления соответствующих услуг. Неотъемлемым компонентом качественного обслуживания пользователей высокопроизводительных вычислительных систем и систем обработки и хранения данных является обеспечение доступности ресурсов, удобства их использования в различных режимах. Для выполнения таких требований необходимо создание достаточного количества открытых ЦОД, концентрирующих ресурсы коллективного пользования, и соединение их надежными, высокоскоростными сетями. Задачу обеспечения науки необходимыми высокопроизводительными ресурсами позволит решить наличие развитой, качественной инфраструктуры, включающей как сами высокопроизводительные ресурсы, так и сетевую инфраструктуру, объединяющую их между собой и с пользователями.

Инфраструктура суперкомпьютерных и распределенных вычислений, созданная в сибирском регионе к настоящему времени, обеспечивает доступность вычислительных ресурсов высокой производительности для сотрудников академических организаций региона, в первую очередь – институтов СО РАН. Даже в рамках существенных ресурсных ограничений, в научных центрах СО РАН активно развивается сетевая инфраструктура и ее специализированные сегменты, ориентированные на задачи высокопроизводительных вычислений и обработки данных.

Основные академические ресурсы сибирского региона для высокопроизводительных вычислений и обработки данных сконцентрированы в ССКЦ СО РАН и крупнейших ВУЗах – СФУ, ТГУ, НГУ. При этом суперкомпьютерные центры функционируют в КНЦ, ИНЦ, ТНЦ и Омском НЦ СО РАН. Ряд институтов СО РАН, особо нуждающихся в высокопроизводительных ресурсах, развивают соответствующую материально-техническую базу самостоятельно.

Суперкомпьютерная инфраструктура региона используется научно-исследовательскими организациями для распределенного решения в удаленном режиме различных прикладных и фундаментальных задач в области нанотехнологий, исследования геологических процессов, геномики и протеомики, защиты от последствий природных катастроф, нанофотоники и волоконной оптики, томографического исследования быстропротекающих процессов, разработки новых технических устройств в самолетостроении и др.

При составлении обзора использованы материалы отчетов ИВТ СО РАН, отчетов по проекту «Распределенная информационно-вычислительная среда СО РАН», по Целевым программам СО РАН «Суперкомпьютер» и «Сеть передачи данных СО РАН», а также информация, опубликованная в открытом виде на веб-сайтах организаций: ИВТ СО РАН, ССКЦ СО РАН, ИВМ СО РАН, ИДСТУ СО РАН, НГУ, ТГУ, ТПУ, СФУ.

Литература

1. Важнейший инфраструктурный компонент // Наука в Сибири № 6, 2011
2. Шокин Ю.И. Об информационном обеспечении задач мониторинга регионов Сибири и Дальнего Востока с использованием данных дистанционного зондирования // Материалы Сибирского научно-практического семинара (2-3 июня 2009 г.). – Барнаул: изд-во ААЭП. – 2009. – С. 3-8.
3. Шокин Ю.И., Федорук М.П., Чубаров Д.Л., Юрченко А.В. О перспективах Grid в Сибирском регионе // Труды Шестого Совещания Российско-казахстанской рабочей группы по вычислительным и информационным технологиям (16 - 18 марта 2009 г.) / Под общ. ред. академика Б.Т. Жумагулова. – Алматы: Казак университеті, 2009. – С. 324–338.
4. Шокин Ю.И., Федорук М.П., Чубаров Д.Л., Юрченко А.В. Об организации деятельности ресурсных центров распределенной информационно-вычислительной среды // Математические и информационные технологии: Тр. международной конф. Копоник, Сербия, 27 - 31 августа 2009 г.; Будва, Черногория, 31 августа - 5 сентября 2009 г. – С. 377–380.
5. Старченко А.В., Панасенко Е.А. Параллельная реализация численного метода решения обратных задач переноса примеси // Сборник материалов "Пятая Сибирская конференция по параллельным и высокопроизводительным вычислениям". – Томск: Том. ун-та, 2010. – С. 25-35.

Шокин Юрий Иванович. Директор Института вычислительных технологий Сибирского отделения РАН. Окончил Новосибирский государственный университет в 1966 году. Академик РАН (1994 г.), доктор физико-математических наук, профессор. Награжден орденами «Знак почета» (1982 г.), «Дружбы» (1999 г.), «Почета» (2004 г.). Автор более 300 научных работ, из них 24-х монографий. Специалист в области прикладной математики и информатики. Область научных интересов: математическое моделирование, в том числе катастрофических явлений, методы интервальной математики, информатика и телекоммуникации. E-mail: dir@ict.nsc.ru.

Федорук Михаил Петрович. Заместитель директора по научной работе Института вычислительных технологий Сибирского отделения РАН. Окончил Новосибирский государственный университет в 1982 году. Доктор физико-математических наук. Автор и соавтор более 150 научных работ. Область научных интересов: вычислительное моделирование нелинейных задач математической физики. E-mail: mife@ict.nsc.ru.

Чубаров Дмитрий Леонидович. Младший научный сотрудник Института вычислительных технологий Сибирского отделения РАН. Окончил Новосибирский государственный университет в 2000 году. Автор 8 печатных работ. Область научных интересов: моделирование параллельных вычислительных процессов, организация параллельных и распределенных вычислений. E-mail: dchubarov@ict.nsc.ru.

Юрченко Андрей Васильевич. Ученый секретарь Института вычислительных технологий Сибирского отделения РАН. Окончил Новосибирский государственный университет в 2002 году. Кандидат физико-математических наук. Автор 29 печатных работ. Область научных интересов: математическое и численное моделирование, организация параллельных и распределенных вычислений, механика композитных конструкций и материалов. E-mail: yurchenko@ict.nsc.ru.