

## Сибирский Суперкомпьютерный Центр (ССКЦ), перспективы его развития

А.С. Алексеев, Б.М. Глинский, С.П. Котелевский,  
Н.В. Кучин, В.Э. Малышкин, А.В. Селихов\*

Информационно-вычислительные технологии интенсивно развиваются в современном мире, особенно в науке, где они были созданы в середине прошлого века и непрерывно стимулируются быстро растущими потребностями математического моделирования в различных фундаментальных и прикладных науках.

В Сибирском отделении АН с момента его образования академики М. А. Лаврентьев [1] и Г. И. Марчук отдавали вычислительной технике и методам математического моделирования в различных науках высший приоритет. Они всегда обеспечивали Сибирское отделение и, в частности, Вычислительный центр ННЦ (в настоящее время ИВМиМГ) самой передовой отечественной вычислительной техникой. Большую помощь Отделению оказал председатель Государственной комиссии по вычислительной технике академик А. А. Дородницын.

В 1975 г. в ННЦ была организована первая сеть дистанционного доступа институтов к крупным ЭВМ, расположенным в ВЦ. Тогда по телефонным каналам городской телефонной станции с помощью самодельных модемов была реализована первая сетевая система коллективного пользования тремя машинами БЭСМ-6 с объединенной памятью. В 1979 г. был разработан проект создания корпоративной скоростной кабельной сети, связывающей практически все институты ННЦ с базовыми вычислительными комплексами БЭСМ-6 (3 шт.) и тремя ЕС-1060 в ВЦ (проект ВЦ КП). Реализация в 1984 г. этой корпоративной сети СО РАН с траншейной прокладкой кабелей между институтами сыграла важную роль при восстановлении ВЦКП в 1996 г. и при переходе на волоконно-оптические системы связи между ЭВМ и с развитием Интернета на скоростных оптических каналах [2].

Интенсивную работу по созданию вычислительной базы СО АН СССР и развитию сетевых технологий математического моделирования выполняли кроме ВЦ в ННЦ (ВЦН) – Вычислительные центры в Красноярске (ВЦК) и Иркутске (ВЦИ). Эти три института выполняли функции региональных центров коллективного пользования информационно-вычислительными ресурсами. Эти ВЦ были первыми в Сибири институтами в области информационно-вычислительных технологий [3].

Следующим шагом в развитии идеи вычислительного центра коллективного пользования было создание Сибирского Суперкомпьютерного Центра кол-

---

\*Институт вычислительной математики и математической геофизики СО РАН, Новосибирск, <http://www2.sccc.ru>.

лективного пользования (ССКЦ). На начальной стадии создания ССКЦ (первые четыре года) важную организационную роль сыграл доктор физ.-мат. наук Г. Н. Ерохин, собравший квалифицированный коллектив инженеров и программистов-системщиков, трудом которых был создан данный Центр коллективного пользования ННЦ. В 2000 г. Президиум СО РАН выделил финансирование на приобретение высокопроизводительной ЭВМ МВС-1000М, штатная конфигурация которой предусматривала существование 32 процессоров очень хорошего качества с внутренней конвейерно-векторной структурой. К концу 2003 года ЭВМ МВС-1000М достигла проектной конфигурации.

Формально ССКЦ создан в 2001 г. на основании Постановления Президиума СО РАН № 100 от 06.06.2001 г. “О создании Сибирского Суперкомпьютерного центра коллективного пользования СО РАН” на базе Института вычислительной математики и математической геофизики (ИВМиМГ) СО РАН и является нештатной структурой, объединяющей Отдел вычислительных систем и сетей ИВМиМГ, а также временный научно-технический коллектив “Параллель” из специалистов ИВМиМГ и других Институты СО РАН. Сайт: <http://www2.sccc.ru>. Научный руководитель – академик А. С. Алексеев.

С 2001 г. ССКЦ устойчиво работает в круглосуточном и круглогодичном режиме. Выработано более 100 тысяч часов процессорного времени. Услугами ССКЦ воспользовались 14 институтов СО РАН. В настоящее время зарегистрировано 22 организации как пользователей ССКЦ. За истекший период было решено более 20-ти больших задач, свыше 50-ти проходят стадию отладки. Важным моментом данного периода является распространение технологии параллельного программирования. В 2002 и 2003 гг. проведены весенние и осенние курсы лекций по параллельным алгоритмам и программам. Выпущено второе издание пособия по MPI. Проведена 7-я Международная конференция по параллельным вычислительным технологиям [8].

На сегодняшний день основными задачами ССКЦ являются:

- обеспечение работ Институтов СО РАН и НГУ по математическому моделированию в фундаментальных и прикладных исследованиях в механике, физике, химии, геологии, биологии и в других дисциплинах высокопроизводительной информационно-вычислительной технологией, системно-техническими средствами и квалифицированным обслуживанием;
- организация обучения специалистов СО РАН и студентов НГУ методам параллельных вычислений на суперкомпьютерах;
- сетевое взаимодействие с другими Суперкомпьютерными центрами СО РАН, Москвы и других городов России, а также зарубежных стран, совместная разработка технологий распределенных вычислений.

## 1. Информационно-вычислительные ресурсы и оборудование ССКЦ

Многопроцессорная вычислительная система МВС-1000/50Н производства 2001 г. ГУП НИИ “Квант” (Россия) с пиковой производительностью более 50 Гфлоп, в состав которой входят:

- 1) 16 вычислительных модулей (2×DEC Alpha 21264/823 МГц/4 Мб SLC, 2 Гб RAM, 9 Гб HDD SCSI);
- 2) управляющий компьютер (DEC Alpha 21264/633 МГц/4 Мб SLC, 2 Гб RAM, 36 Гб HDD SCSI-2);
- 3) 2 коммутатора: Intel EtherExpress E510T (24 порта FastEthernet и 1 порт GigabitEthernet) и Myricom M2M-SW16 (16 портов со скоростью 1.2 Гб/с);
- 4) операционная система LINUX RedHat 7.2, библиотека параллельного программирования MPICH-1.2.5, параллельный отладчик TotalView, компиляторы C, C++, Fortran (Compaq), подсистема коллективного доступа ИПМ им. Келдыша РАН.

Вычислительный комплекс из двух серверов RM600-E30 фирмы Siemens (Германия) с общей пиковой производительностью более 5 Гфлоп производства 1999 г. SMP-серверы архитектуры cc-NUMA с общим полем оперативной памяти (8 Гб – максимум и 4 Гб) на базе процессоров R10000/250 МГц/4 Мб SLC фирмы MIPS (8 процессоров – максимум и 5 процессоров) соединены между собой с помощью адаптеров SCI-2 (1 Гб/с). Операционная система Reliant UNIX 5.44, компиляторы C, C++, Fortran 77, EPC – Fortran, библиотеки MPI и система ADAPTOR версии 7.0, а также DQS 3.2.7.

Графический сервер Power Challenge M фирмы Silicon Graphics Inc. (США) производства 1996 г. (R7000/75 МГц/2 Мб SLC фирмы MIPS, RAM 4 Гб, HDD 16 Гб). Операционная система IRIX, СУБД Oracle 7.1, ГИС Arcinfo 7.0 и Grass и система визуализации научной графики IRIS Explorer 2.0.

Файловый сервер производства 2002 г. (Athlon MP / 1 ГГц, 2 Гб RAM, контроллер RAID-5/IDE, HDD 7×120 Гб=840 Гб). Операционная система Windows 2000.

WWW-Сервер производства 2001 г. (2 × Intel PentiumIII / 1000 МГц, 2 Гб RAM, HDD 36 Гб/SCSI). Операционная система Windows 2000.

Дисковый массив MA8000 фирмы Compaq Computer Corporation (США) с физическим дисковым объемом 648 Гб (максимум 3 Тб), RAID-контроллером и коммутатором FC на 8 оптических портов (2 Гб/с).

Графическая рабочая станция Blade2000 фирмы Sun Microsystem (США) производства 2003 г. (2×UltraSPARC III/1,2 ГГц/8 Мб SLC, RAM 8 Гб (максимум), HDD 2×73 Гб/10000 RPM/SCSI-2, графический акселератор SUN XVR1000/320 Мб, LAN 1000Base-SX). Два монитора: LCD 24У (1920×1200) и CRT 21У. Операционная система Solaris 9, пакет SUN Star Office (текстовый редактор, редактор векторных изображений и средства для создания презентаций), пакет Forte Compiler (C, C++ и Fortran), библиотеки Open GL и QT, а также среда разработки SUN One Studio (Java).

Коммутационное оборудование состоит из коммутаторов фирмы Cisco Systems (США): 3 коммутатора Catalyst 2950 (24 порта 100Base-T) производства 2002 г. и Catalyst 3550-12T (10 портов 1000Base-T и 2 порта 1000Base-LX) производства 2003 г.

Система бесперебойного электропитания включает модульный ИБП PW-9170 фирмы Power Wave Inc.(США) производства 2002 г. (18 кВт), 4 ИБП

Smart-UPS 1500 фирмы APC (США) производства 2001 г., 2 ИБП Masterguard фирмы Siemens (Германия) производства 1999 г. (3 и 6 кВт).

Система кондиционирования: 2 кондиционера FTY600GV1B фирмы DAIKIN (Япония) производства 2000 г. (6 кВт по холоду суммарно) и кондиционер OSA-51C-H фирмы “Tespair” (Италия) производства 2003 года (15 кВт по холоду, парогенератор).

Согласно утвержденному Положению серверами ССКЦ могут пользоваться бесплатно организации и институты СО РАН, в том числе и иногородние, на основании гарантийного письма руководителя (директора). Доступ к серверам осуществляется через корпоративную сеть ССКЦ или через сеть Интернет. Пользователь заполняет регистрационную форму, расположенную на www-сервере по адресу [http://www2.sssc.ru/Information/Information\\_ru.htm](http://www2.sssc.ru/Information/Information_ru.htm), и получает от администратора сервера имя и пароль доступа. Результаты расчетов пользователи получают самостоятельно через сеть ССКЦ или по сети Интернет.

Среди задач 14-ти Институтов СО РАН, решавшихся в ССКЦ в 2003 г., был ряд крупных задач: моделирование трехмерного высокоскоростного обтекания аэрокосмических аппаратов, эволюции протопланетного пылевого облака, эволюции земной коры, распространения сейсмических волн, переноса частиц в физике, фильтрации нефти в упругопористых средах и т. д.

## 2. План развития вычислительных ресурсов на 2004–2005 гг.

Анализ больших вычислительных задач СО РАН и прогноз потребности вычислительных ресурсов показывает следующую закономерность: требуется увеличивать число процессоров и объемы оперативной памяти в 2–3 раза каждые два года. Например, задача моделирования предбиологических исследований с учетом химических реакций между газами требует на сегодняшний день не менее 100 процессоров, – В.Н. Снытников (ИК СО РАН), В.А. Вшивков, В.Э. Мальшкин (ИВМиМГ СО РАН).

С учетом рекомендаций исполнителей больших задач Бюро Межинститутского Совета по супервычислениям СО РАН и руководством ССКЦ разработана следующая программа развития вычислительных ресурсов:

**2004 г.** Монтаж и запуск многопроцессорной вычислительной системы МВС-1000М из Межведомственного суперкомпьютерного центра г. Москва в составе 128 процессоров DEC Alpha 21264/633 МГц/4 Мб SLC (64 двухпроцессорных вычислительных модуля с 2 Гб оперативной памяти и дисковой памятью 9 Гб) и коммутаторами FastEthernet и Myrinet на базе операционной системы LINUX RedHat 7.2.

**2005 г.** Закупка, монтаж и запуск многопроцессорной вычислительной системы в составе 128 процессоров Opteron 246/2 ГГц (64 двухпроцессорных вычислительных модуля с 4 Гб оперативной памяти) и коммутаторов GigabitEthernet и Myrinet на базе операционной системы LINUX.

Для проверки вычислительной мощности новых образцов серверов на реальных задачах наших пользователей на свободно распространяемом программном обеспечении и на коммерческих продуктах ССКЦ налаживает контакты с фирмами-производителями. Так были получены во временное пользование сервера hp Integrity rx5670 с 4 процессорами Intel Itanium 2, 1.3 ГГц, 3 Мб кэш, 4 Гб оперативной памяти (может наращиваться до 96 Гб) и 2-процессорные серверы Celestica A2210 с микропроцессорами AMD Opteron 2 ГГц, 1 Мб кэш, 4 Гб оперативной памяти. Оценка производительности этих серверов на реальных задачах наших пользователей показала, что в настоящее время решения на базе AMD Opteron по критерию отношение производительность/стоимость являются более эффективными. Поэтому они и были выбраны в качестве базовых процессорных элементов для 128 процессорной системы.

Опыт эксплуатации вычислительных мощностей ССКЦ для решения больших задач показывает необходимость увеличения объема дискового пространства, доступного для оперативного использования. Как было указано выше, в распоряжении ССКЦ имеется 144 (16×9) Гб в виде локальной дисковой памяти узлов и 36 Гб на хост-узле системы МВС-1000, 117 Гб дисковой памяти системы RM-600, дискового массива 840 Гб файлового сервера и 648 Гб дискового массива MA8000, что в сумме составляет 1785 Гб (~1.7 Тб). Однако реально доступный в настоящее время объем дискового пространства для решения больших задач составляет существенно меньший объем и ограничен как особенностями коммуникаций между различными частями общего дискового пространства, так и порядком его выделения пользователям ССКЦ. Для дальнейшего повышения эффективности работы ССКЦ потребуется пересмотр обоих этих факторов. Основным направлением развития системы дисковой памяти ССКЦ является создание высокопроизводительного файлового сервера, управляющего имеющимся дисковым пространством и позволяющего наращивать его объем в дальнейшем. Этот файловый сервер должен быть функционально узкоспециализированным и обеспечивающим максимальную скорость доступа при средней эффективной нагрузке, характерной для большинства вычислительных задач. Высокая производительность файлового сервера необходима для поддержания эффективной работы общей файловой системы, объединяющей все доступное дисковое пространство [4–6].

Одной из важнейших задач на текущий период является создание в рамках ССКЦ высокопроизводительной вычислительной системы на разнородных вычислительных кластерах. В частности, кластеров на основе микропроцессоров DEC Alpha 21264 и Opteron 246. Реализация этого направления потребует решение следующих задач:

- выбор архитектуры технических средств и программного обеспечения (коммуникационное оборудование, операционной системы, подсистемы управления операционными ресурсами, трансляторов и библиотек программ);
- создания библиотеки прикладных программ, ориентированной на параллельные вычисления с использованием разнородных кластеров, в частности, выбор пакетов прикладных программ с учетом запросов пользователей (AMBER, Gaussian 03, GROMACS и др.);

- отработка организации и проведения крупномасштабных вычислительных экспериментов на разнородных кластерах с использованием технологии Grid.

Как показывает опыт эксплуатации всех высокопроизводительных вычислительных систем, всегда существуют задачи, для которых мощностей одной вычислительной системы не достаточно. Потребность в более эффективном использовании имеющихся разнородных вычислительных систем приводит к необходимости объединения вычислительных и информационных возможностей как ССКЦ, так и ННЦ и других городов (в частности, ЮНИИ ИТ г. Ханты-Мансийск). Основным препятствием, мешающим в настоящее время широкому распространению Grid-технологий, является большая сложность и ограниченные функции свободно доступных разработок в этой области. Для повышения эффективности использования разнородных вычислительных ресурсов предполагается осуществить следующие шаги: разработать концепцию и прототип интерфейса пользователя для доступа к объединенным вычислительным ресурсам. Предпочтительной основой для этого является веб-интерфейс; создать систему распределенной диспетчеризации ресурсов на основе специализированных серверов, позволяющую эффективно распределять различные типы задач на вычислительные ресурсы, контролировать процесс выполнения задач и взаимодействовать с пользователями через указанный выше интерфейс; разработать концепцию и создать прототип сервиса межкластерных коммуникаций, коммутирующего передачу сообщений библиотеки MPI между кластерами для одного параллельного приложения.

### 3. Некоторые тенденции развития высокопроизводительных вычислений

Опыт работы ВЦН, ВЦК и ВЦИ в режиме коллективного пользования подтверждает ряд общих тенденций развития математического моделирования в различных науках. Одну из них, по-видимому главную, четко сформулировал известный американский физик Кен Вильсон: “Научные вычисления стали третьей компонентой научных методов, равнозначной эксперименту и теории” (Computer Physics Communications. – 2001. – Vol. 1. – P. 1–6, Elsevier Science).

Эта тенденция как результат развития прикладной математики была осознана еще в пятидесятые годы прошлого века многими математиками, в том числе основателями Сибирского Отделения академиком: М. А. Лаврентьевым, С. Л. Соболевым, С. А. Христиановичем, Г. И. Марчуком. Поэтому развитие математического моделирования вычислительной математики и центров вычислительной техники в доперестроечный период в СО АН имели высокий приоритет, широкую востребованность и быстрый рост.

Вторая важная тенденция состоит в осознании научным сообществом существенной зависимости продуктивности своей работы от развития количественных методов исследования [7]. Отсюда проистекают две формы самоорганизации “незримых” научных коллективов:

- образование крупных междисциплинарных команд, включающих кроме специалистов в конкретных областях, смежников физического, химиче-

ского профилей, а также прикладных математиков и вычислителей для формулировки количественных постановок задач (это является традицией в СО РАН, которая усилена в последние годы поддержкой интеграционных проектов со стороны Президиума СО РАН);

- формирование оперативных информационно-вычислительных взаимосвязей интерактивного типа между теорией, экспериментом и научными вычислениями. (Эта тенденция соответствует изречению Р. Декарта: “Если не все ясно – давайте посчитаем”.)

Приведенные обстоятельства проявляются в эволюции многодисциплинарного научного сообщества Сибирского отделения РАН. Раньше они отображались в существовании Координационных советов разных Министерств, многодисциплинарный состав которых обычно содержал математиков-прикладников и вычислителей.

В последние годы наблюдается стабилизация постановок “больших задач” типа: эволюция протопланетного вещества и планет; эволюция Земли; эволюция жизни (в количественной постановке); биоинформатика, модели климата и экологии и т. д. В связи с этими задачами, требующими интеграции знаний из разных областей, набирает силу процесс формирования своеобразных “клубов больших задач”. Эти межинститутские коллективы создают модели, численные методы, параллельные программы, технологии вычислительного эксперимента, его анализа и теоретического обобщения.

Потребности вычислительных мощностей и программного обеспечения, определяемые указанными коллективами должны удовлетворяться в приоритетном порядке руководством Центров супервычислений и финансирующих институции. При этом сами Суперцентры коллективного пользования оказываются не только организациями-производителями профессиональных информационно-вычислительных услуг, но и базовыми ячейками, вокруг которых формируются межинститутские целевые коллективы.

Если говорить об ориентирах количественного развития требуемых вычислительных ресурсов для фундаментальных научных исследований в областях: ядерной физики и ядерного оружия; молекулярной нанотехнологии; молекулярной химии, проектирования аэрокосмических аппаратов, климата и моделирования окружающей среды; прогноза штормов, цунами, землетрясений и т. д., то имеются оценки ряда авторитетных международных экспертных групп. Эти оценки интегрально характеризуются следующими цифрами. Производительность: десять в пятнадцатой степени операций с плавающей точкой в секунду (PetaFlops); оперативная память сотни гигабайт. Системы с такими параметрами смогут производить вычисления для современных реалистичных моделей в указанных выше проблемах за приемлемое время. Темпы роста производительности компьютеров, входящих в группу Top 500 показывают, что PetaFlops может быть достигнуто в 2007–2008 гг., скорее всего, в Японии или США.

### **Список литературы**

- [1] Алексеев А.С. Лаврентьев М.А. и информационные-вычислительные технологии // Наука в Сибири. – 2000. – № 44/45.

- [2] Алексеев А.С., Елепов Б.С., Котов В.Е., Метляев Ю.В. О программе работ по созданию сети информационно-вычислительных систем (Центров) в Сибирском отделении АН СССР. – Новосибирск, 1987. – 27 с. – (Препринт / РАН. Сиб отделение. ВЦ; 743).
- [3] Алексеев А.С., Кузнецов Е.П., Метляев Ю.В. и др. Академическая региональная сеть Сибири. – Новосибирск, 1983. – 39 с. – (Препринт / РАН. Сиб отделение. ВЦ; 467).
- [4] Foster I., Kesselman C. Globus: A Metacomputing Infrastructure Toolkit // Int. J. Supercomputer Applications. – 1997. – Vol. 11(2). – P. 115–128.
- [5] Foster I., Kesselman C., Nick J., Tuecke S. The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration // Open Grid Service Infrastructure WG, Global Grid Forum, June 22, 2002.
- [6] Selikhov A., Germain C. CMDE: a channel memory based dynamic environment for fault-tolerant message passing based on MPICH-V architecture // Proc. of 7-th Int. – Springer Verlag, 2003. – LNCS 2763. – P. 528–537.
- [7] Алексеев А.С. О Математике и информатике в современном мире // Наука на рубеже веков, взгляд из Сибири. – Новосибирск, 2000. – С. 5–13.
- [8] Proceedings of the International Conference on Parallel Computing Technologies (PaCT-2003) / Ed. V. Malyskin. – Nizhni Novgorod, Russia: Lecture Notes in Computer Science, 2003. – Vol. 2763.