

Копия текста публикации со страницы <http://www.ferra.ru/online/integral/101004/>

12.07.2010

Отчет о Дне суперкомпьютерных технологий, часть первая

Олег Горобец

В Новосибирске под эгидой корпорации состоялся Intel День Суперкомпьютерных технологий. Видные ученые и специалисты в области высокопроизводительных вычислений собрались, чтобы рассказать общественности, насколько важной и нужной является эта отрасль научно-технического развития России.

Немного истории вместо предисловия

Некогда, в 80е годы, и в начале 90х, слово «Суперкомпьютер» было у нас, тогда еще в Советском Союзе, очень популярно. Политики ставили его рядом с такими громкими терминами, как «программа «Звездных Войн» и «стратегическое экспортное ограничение». Физики, астрономы и простые инженеры с благоговением произносили в курилках слово «Крэй». Любители кинофантастики, посмотревшие в видеосалонах «Одиссею-2001» и «Звездный Путь», рассуждали в кулуарах о реальности создания искусственного интеллекта.



Наконец триумфальное шествие персональных компьютеров добралось и до российских просторов. Мегагерцы и мегабайты, еще недавно бывшие прерогативой больших машин, поселились прямо на рабочих столах. Громкое же слово, еще недавно владевшее умами,

отступило в тень, сохранив свою важность лишь для узкого круга посвященных. В условиях экономики переходного периода, когда одна страна прекратила свое существование, а другая еще только формировалась на обломках, о каких-то крупных проектах, связанных с массивными объемами вычислений, не могло быть и речи.

Время шло, и страна все чаще стала поднимать голову с больничной койки, на которую ее уложило колесо истории. Поднимать, обозревать мир вокруг — и осознать, что за это время он не только не стоял на месте, но успел ускакать далеко вперед. Высокопроизводительные вычисления (HPC, High Performance Computing) — важнейшая часть многих научных процессов, от них зависят производства, формирование рынка. Обороноспособность большинства развитых стран опирается на них как на один из краеугольных камней. Так чего же мы стоим?

Начало 2000х годов можно считать периодом, когда HPC в России начали оживать. Первым значимым событием российской эпохи HPC-возрождения можно назвать открытие в 1999г. Межведомственного Суперкомпьютерного Центра (МСЦ) РАН. Сердцем Центра стал комплекс МВС-1000, 1 TFLOPS система, собранная на 768 процессорах Compaq Alpha.



Вскоре после этого один за другим на свет стали появляться проекты новых высокопроизводительных систем. Не все они были реализованы, но те, что все же были претворены в жизнь, успешно доказали свою нужность. В их числе — проект перевооружения Росгидромета и создание 0,7 TFLOPS комплекса для расчета авиадвигателей в НПО «Сатурн».

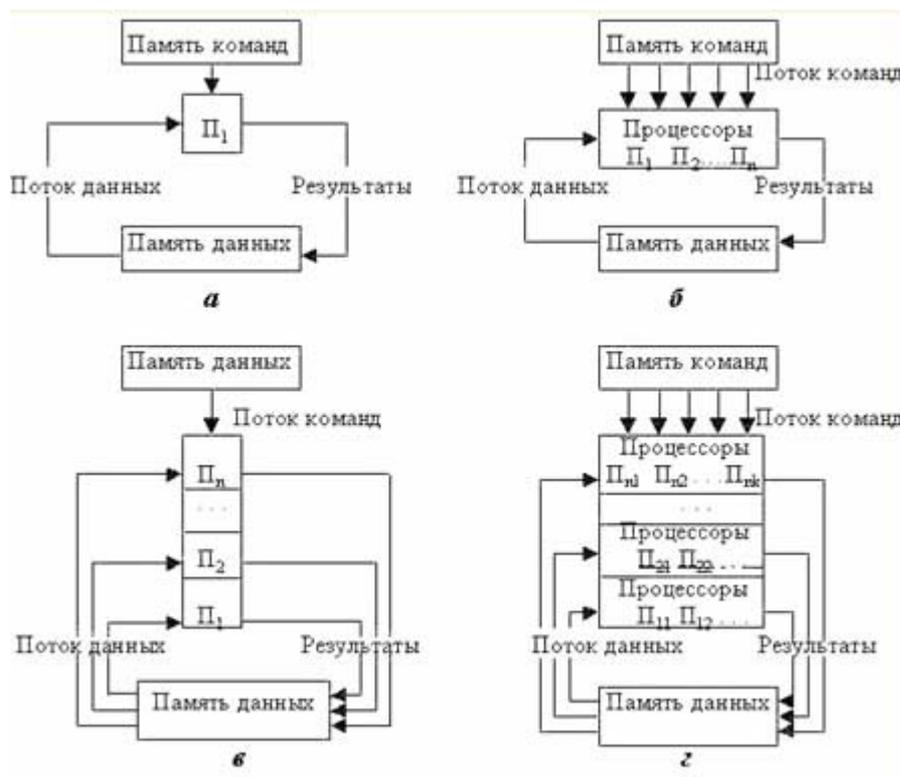
Однако отставание от мирового уровня в области HPC преодолеть не так просто. Для этого нужна не только поддержка госструктур, но и глубокое понимание задач, которые могут быть решены с помощью больших машин. Да и вообще, что это есть такое — High Performance Computing?

High Performance Computing

У обычного человека с этим термином — если он его вообще слышал — ассоциация простая: большие машины, занимающие иногда несколько комнат, и аббревиатура FLOPS

(количество операций с плавающей запятой в секунду) с приставкой «Гига» или «Тера». Специалист, использующий в работе НРС, думает о задачах, которые способен выполнить компьютер, будь то расшифровка структуры ДНК, расчет траектории полета на Марс или создание математической модели какого-то физического процесса.

Но в чем же принципиальное отличие компьютеров с приставкой «супер» от привычных всем персональных рабочих станций — ну, кроме того что они большие? Ответ заключается в одном слове: параллельность.



Научно-технический прогресс не стоит на месте, и линейная производительность систем постоянно растет — но ведь многие фрагменты вычислительного процесса могли бы выполняться одновременно!

Наиболее известные сейчас типы параллелизма — это:

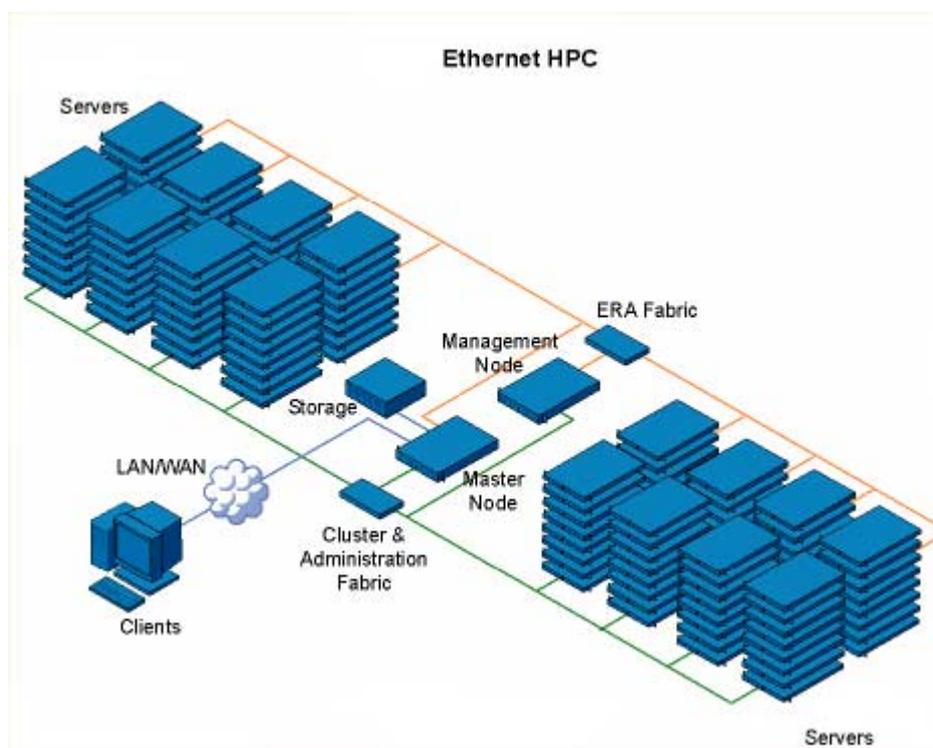
- *Параллелизм на уровне данных.* В этом случае одна операция выполняется сразу над несколькими элементами массива данных. Задача программиста в данном случае сводится к настройке и оптимизации процесса параллельной компиляции, или же к использованию специальных языков для параллельных вычислений.
- *Параллелизм на уровне задач.* Уже представляет собой особый подход к программированию, предполагающий разбиение задачи на несколько подзадач, которыми можно одновременно загружать процессоры системы.

Именно последний тип притягивает к себе сегодня наибольшее внимание — и порождает наибольшее количество проблем. Наибольшая эффективность достигается, когда авторы программы мыслят категориями параллельного программирования, начиная с этапа формирования задачи.

К сожалению, это направление долгое время оставалось уделом узкой группы специалистов, причем не только в нашей стране, но и во всем мире. И лишь в последние

несколько лет, благодаря распространению многоядерных процессоров, разработчики стали все больше внимания этому вопросу. Ведь именно благодаря параллельным процессам производительность вычислительных систем может быть резко увеличена. Тем более важен такой подход для НРС, где возможно распараллеливание на нескольких уровнях — на уровне процессорных ядер, процессоров и узлов.

Наиболее распространенным и доступным типом НРС-системы в настоящее время является кластер.



Кластер — мультисерверная система, состоящая из вычислительных узлов (nodes) и одного (реже нескольких) управляющего узла, объединенных между собой частной высокоскоростной сетью (интерконнектом). Управляет кластером специальное программное обеспечение, чаще всего на базе *NIX.

Несмотря на то, что собрать простейший кластер несложно, настоящая высокопроизводительная система требует значительных вложений сил и средств на всех подготовительных этапах, равно как и в процессе эксплуатации. Оправдать эти расходы могут решаемые задачи — а их, на самом деле, существует великое множество. Прикладная наука, ресурсодобыча, гидрометеорология, системы раннего оповещения о стихийных бедствиях и, конечно, оборонный комплекс — все эти отрасли способны загрузить расчетами ресурсы во много раз большие, чем есть сейчас у нас в стране. Проблема в том, что заказчики часто просто не знают о возможностях НРС в применении к своим проблемам, а уважаемые чиновники не осознают всей важности развития суперкомпьютерных технологий по всей стране.

Хотите примеров? Так давайте отправимся в Сибирь вместе с компанией Intel, одним из традиционных знаменосцев высокопроизводительных и параллельных вычислений в России.

Сибирский наукоград

Новосибирск не без основания претендует на звание научной и академической столицы Сибири. На территории Академгородка, что в 20 км от центра города, расположились десятки научно-исследовательских институтов, президиум Сибирского отделения РАН, Новосибирский Государственный университет, Физико-математическая школа. Основанный в 1957 году по инициативе академика Лаврентьева, Академгородок стремительно вырос в настоящий большой наукоград. Среди прочих дисциплин, важное место с самого начала заняли прикладная математика и информатика. В 1964 году был основан Вычислительный Центр (ныне Институт Вычислительной Математики и Вычислительной Геофизики РАН), на базе которого 9 лет назад был создан Сибирский суперкомпьютерный центр коллективного пользования. Сотни интереснейших научных работ явились на свет в этом уникальном научном инкубаторе.



Для корпорации Intel инновационность и наукоемкость — синонимы успеха, и открытие в 2004 году отделения компании в новосибирском Академгородке было более чем логичным. Разумеется, высокопроизводительные вычисления заняли в его работе одно из важнейших мест.

Трудно придумать лучшее место для мероприятия, пропагандирующего развитие ИТ. Именно поэтому в конце июня 2010 год и состоялся День Суперкомпьютерных Технологий в Новосибирске.

День Суперкомпьютерных Технологий в Новосибирске

В Большом зале Дома Ученых собралось множество людей. Именитые ученые, представители компаний, студенты и, разумеется, наша группа журналистов. Среди гостей были и приезжие из сопредельных государств — ведь тема волнует умы отнюдь не только в России.



После кратких приветственных речей видных деятелей науки, образования и бизнеса, докладчики приступили к освещению проблематики Дня НРС. Целью выступающих было рассказать о состоянии дел в области высокопроизводительных вычислений и помочь гостям увидеть многообразие российских задач, требующих использования суперкомпьютеров.

Открыл программу выступлений Сергей Михайлович Абрамов, директор знаменитого Института Программных Систем, что в Переславле-Залесском.

«Суперкомпьютерные технологии России: объективные потребности и реальные возможности» — так назывался его доклад, и приведенные в нем факты однозначно утверждают: почитать на лаврах России противопоказано.



Одна фраза, сказанная некогда Деборой Винс-Смит, Президентом совета по конкурентоспособности США, обрисовывает проблему предельно четко: «Страна, желающая победить в конкуренции, должна победить в вычислениях». Да, в нашей стране немало сделано за последнее время в области НРС, в эксплуатацию запущено несколько крупных вычислительных систем, таких как 60 TFLOPS «Чебышев» и 414 TFLOPS «Ломоносов»... но на общем фоне российская киберинфраструктура выглядит слабо. Производственных машин мало, и львиная доля финансирования на их создание поступает в центральные регионы страны — при том, что потребность в них от удаленности регионов зависит мало. По общей производительности отставание российской инфраструктуры от лидера — США — составляет не менее 7 лет, и это при

условии, что темпы ее развития всего лишь сохраняются. То есть, чтобы приблизиться к лидерам, нужно значительно ускорить собственное развитие.

Научно-техническая программа Союзного государства "СКИФ-ГРИД" 

Иерархия киберинфраструктуры 2010.06

	США	Единая Европа	Китай	Россия	Потребности России
Топ 1–20	11 шт. 7,096 Tflops	4 шт. 1,612 Tflops	3 шт.. 2,041 Tflops	1 шт. 350 Tflops	2–3 шт.
Топ 21–100	35 шт. 3,207 Tflops	26 шт. 2,626 Tflops	2 шт.. 283 Tflops	2 шт.. 209 Tflops	20–30 шт.
Топ 101–250	93 шт. 3,594 Tflops	32 шт. 1,372 Tflops	10 шт.. 395 Tflops	2 шт.. 85 Tflops	28–40 шт.
Топ 251–500	143 шт. 4,071 Tflops	72 шт. 2,056 Tflops	9 шт.. 272 Tflops	6 шт.. 170 Tflops	50–75 шт.
Топ 1–500: ВСЕГО	277 шт. 16,416 Tflops	134 шт. 7,667 Tflops	24 шт.. 2,993 Tflops	8 шт.. 646 Tflops	100–150 шт.

26–22 раз 12–9 раз 3.7–2.2 раза

22 июня 2010 г. СКИФ-ГРИД © 2009 Все права защищены Слайд 18

А ведь разработчики не сидят сложа руки! Так, в рамках российско-белорусской программы «СКИФ-ГРИД» с 2003 по 2009 год было построено 6 из 8 отечественных суперЭВМ, когда-либо входивших в мировой рейтинг Top500. Программа обладает интереснейшими наработками в области создания масштабируемых НРС-систем согласно потребностям заказчика. При этом суперкомпьютеры СКИФ «ряд 4» обладают уникальными показателями по энергопотреблению, массогабаритным характеристикам и акустическому комфорту. 1 PetaFLOPS? Почему нет, было бы финансирование. Модельная линия «СКИФ/П» имеет проектируемую мощность в 100 TFLOPS и энергопотребление в 0,146 МВт на 1 шкаф; вожденный 1 PFLOPS поместится всего лишь в пяти таких шкафах!



Не стоит забывать и про вторую часть названия программы — «ГРИД». Она означает возможность объединения разнесенных в пространстве машин в единую вычислительную систему. И, кстати, для производства СКИФов импортировать нужно только микросхемы — все остальное, включая печатный монтаж и сборку, может осуществляться в России.

«СКИФ-ГРИД» — не единственный крупный НРС-проект в России (хотя определенно один из самых интересных). Но бытие таких проектов определяет их востребованность.

«Народу нужно что - решить задачу! Хоть синтезом, хоть лаптем — все равно!» С этой фразы началось выступление Виктора Владимировича Самофалова, менеджера компании Intel по НРС-решениям в России и СНГ.

Высокопроизводительные вычисления на сегодняшний день могут решать множество задач не только научного, но и прикладного, практического характера. Производительности в 100 TFLOPS хватит, чтобы полностью смоделировать поведение самолета в заданных условиях, предсказать поведение воздушных масс или собрать из отдельных образцов целый клеточный геном. Но не стоит обольщаться: полное моделирование процессов одной живой клетки потребует вычислительной мощности около 1 ExaFLOPS (10 в 18 операций с плавающей точкой в секунду).



По сути, чтобы узнать, как работает наш собственный организм на клеточном уровне, нужна производительность, во много раз превышающая ресурсы, доступные сейчас земной науке.

Разумеется, прогресс не стоит на месте. Каждый год Intel демонстрирует общественности новый технологический шаг, будь то более «тонкий» техпроцесс или новая система команд. Число ядер в процессоре возрастает, и разработчикам важно иметь возможность в полной мере задействовать эту мощь. Сознывая это, компания уделяет немалое внимание вопросам разработки алгоритмов, использующих преимущества параллельного исполнения. Специальные инструменты разработки, помогающие прог



intel
Software

Intel® Software Tools for Parallelism

- Architectural Analysis** Visualization of parallel (threaded or MPI) application execution and communication behavior – give valuable insights for application architects and programmers
- Introduce Parallelism** Highly optimized OpenMP® and MPI library and run-time system for scalable solutions. MKL, threaded and distributed mathematical library
- Confidence / Correctness** Detecting actual and potential Threading and MPI programming and API issues - to address challenges unique to parallel programming
- Optimize / Tune** Valuable insights for performance and scalability tuning of threaded and MPI applications

11

Copyright © 2010 Intel Corporation. All rights reserved.

Но уже сейчас необходимо смотреть в будущее; число процессорных ядер в НРС-системе требует многоуровневого распараллеливания и особого подхода к оптимизации кода. Количество ошибок при параллельном программировании существенно возрастает, и успешно бороться с ними можно только с помощью нового, более эффективного инструментария.

Чтобы простимулировать развитие новой области искусства программирования, корпорация Intel совместно с РОСНАНО, Суперкомпьютерным консорциумом университетов и программой СКИФ-ГРИД организовала конкурс проектов по применению высокопроизводительных вычислений. Фокус конкурса — практическая ценность проектов; многие творческие коллективы получают наконец возможность донести свои идеи до тех, кому они нужны с практической точки зрения — настоящих потребителей мощи НРС, ученых и промышленников.



«Невозможное стало возможным: реальные приложения для НРС»

Конкурс проектов по применению высокопроизводительных вычислений

2 апреля – 10 октября 2010 г.

Цели конкурса:

- Демонстрация возможностей практического применения высокопроизводительных вычислений для нужд различных отраслей науки и промышленности
- Поощрение творческих коллективов, работающих над проектами в области НРС

Призовой фонд:

- 950 000 рублей (6 денежных призов)
- 1 500 000 процессоро-часов работы на суперкомпьютерах

Подайте заявку на сайте: <http://www.intel.ru/hpc>
Или пришлите ее по e-mail: HPC.contest@intel.com

intel **РОСНАНО** **СУПЕРКОМПЬЮТЕРНЫЙ КОНСОРЦИУМ УНИВЕРСИТЕТОВ** **СКИФ-ГРИД**

Copyright © 2010 Intel Corporation. All rights reserved.