

ЧТОБЫ ПОБЕДИТЬ В КОНКУРЕНЦИИ, НУЖНО ПОБЕДИТЬ В ВЫЧИСЛЕНИЯХ

Рассказывает директор Института программных систем РАН Сергей Михайлович Абрамов



За почти 16-летнюю историю списка ведущих суперкомпьютеров TOP500 в него попадали только шесть машин отечественной разработки. Первая из них – МВС1000М (2002 год, НИИ "Квант"). Остальные пять компьютеров – семейства СКИФ: СКИФ К-500 (2003 г.), СКИФ К-1000 (2004 г.), СКИФ Cyberia (2007 г.), СКИФ Урал и СКИФ МГУ (обе – 2008 г.). Известно, что создавались они в рамках программ СКИФ и СКИФ-ГРИД – совместного Российско-Белорусского проекта. О программе СКИФ и ее перспективах мы попросили рассказать директора Института программных систем РАН – главного исполнителя с российской стороны.

Сергей Михайлович, коротко расскажите об Институте программных систем.

Институт программных систем РАН (ИПС РАН) был организован 25 лет назад, в апреле 1984 года в Переславле-Залесском. Сначала – как Филиал Института проблем кибернетики АН СССР, в 1986 году он был преобразован в Институт программных систем АН СССР.

У истоков института стоял академик Е.П.Велихов, он готовил соответствующее постановление Правительства СССР. ИПС РАН и сегодня входит в отделение нанотехнологий и информационных технологий РАН, которое возглавляет Е.П.Велихов. Первым директором института был Альфред Карлович Айламазян. Кстати, недавно вышло постановление о присвоении ИПС РАН его имени. В институте изначально была очень сильная кадровая политика, подкрепленная строящимся жильем. Благодаря этому удалось из самых разных городов и организаций СССР собрать очень сильных специалистов.

Институт создавался под эгидой известной в те годы концепции "асимметричного эффективного ответа" на американскую программу СОИ ("Звездные войны"). Практически одновременно было организовано около десятка институтов в разных регионах СССР и в разных областях фундаментальных исследований. Нашему институту были предписаны три направления – высокопроизводительные вычисления (суперкомпьютеры), искусственный интеллект и информационные технологии (операционные системы, языки программирования,

базы данных). Сегодня ИПС РАН, сохранив эти три направления, продолжает развиваться. Особо отмечу, мы не занимались выживанием, все годы шло планомерное развитие.

Чем занимается ИПС РАН сегодня?

Институт состоит из пяти научных подразделений – исследовательских центров (ИЦ). ИЦ мультипроцессорных систем, помимо суперкомпьютинга, занимается телекоммуникациями, региональными компьютерными сетями, специализированными под условия России. В исследовательском центре медицинской информатики развиваются комплексные средства автоматизации лечебно-профилактических учреждений, заслуженное признание получила наша технология поддержки лечебно-диагностического процесса ИНТЕРИН. Это лучшая российская разработка в данной области, она используется в медицинских учреждениях ЦБ РФ и РАО РЖД, в Чазовской клинике, в клинике и поликлинике аппарата Президента РФ и др.

Также действуют исследовательские центры искусственного интеллекта, системного анализа и процессов управления, они занимаются соответствующими направлениями теории и прикладных аспектов информатики и математики. Еще одно наше подразделение – Международный детский компьютерный центр имени А.К.Айламазяна. Он уже 20 лет работает со школьниками.

Рассказывая об ИПС, я просто обязан упомянуть о формально сторонней организации – Негосударственном образовательном учреждении "Институт программных систем – Университет города Переславля имени А.К.Айламазяна". Он был создан в 1993 году по инициативе академика Е.П.Велихова, профессора А.К.Айламазяна и научных сотрудников ИПС РАН при поддержке ведущих предприятий и администрации города. Ежегодно он принимает 70–95 абитуриентов, выпускает в среднем 50–65 специалистов в области информационных технологий (прикладная математика, информационные технологии и системы, информационные технологии в экономике).

Университет Переславля – это фактически городской университет, 85% абитуриентов составляют жители Переславля. Остальные 15% – это приезжие, в основном – из



бывших республик СССР. Но есть студенты и из российских городов, в том числе – из Москвы и Подмосковья. Среди них – очень сильные студенты, которые легко могли бы поступить в любой вуз на бюджетное место. Среди учредителей Университета, помимо ИПС РАН, – и городская власть, и городской бизнес. Ведь основная задача Университета – позволить детям не уезжать за образованием. И действительно, у нас в городе остается 55% выпускников, в том числе – многие приезжие. Это очень важно для развития города, вот почему в нем заинтересованы все.

Для ИПС он также чрезвычайно важен. Не было бы сегодня нашего института, если бы 15 лет назад не был создан этот вуз. Ведь 40% штата ИПС составляют его выпускники и студенты. А всего в ИПС РАН трудятся 500 человек, два года назад их было 300. Причем сегодня 300 сотрудников занимаются наукой, остальные работают во вспомогательных подразделениях.

За счет чего все эти годы развивался ИПС?

Наш первый директор, Альфред Карлович Айламазян, пришел из прикладной науки. И он всегда говорил, что должна быть цепочка: фундаментальные исследования – инженерная разработка – внедрение – эксплуатация и получение обратной связи. Поэтому мы еще с советских времен работали в различных прикладных направлениях, у нас было много заказчиков в совершенно разных областях. Это были и государственные программы, связанные со "звездными войнами", и разработки отдела искусственного интеллекта для прогнозирования рыбных запасов и разумного вычисления квот на рыбную ловлю в интересах рыбной промышленности. Были работы в области экологии (система контроля качества воды) и медицины. И этот подход сохранен до сих пор. Например, на базе разработок в области компьютерных сетей была создана отдельная фирма-провайдер интернет-доступа. Сегодня она – фактически монополист в Переславле, несмотря на присутствие других операторов. Отмечу, все прикладные работы, эксплуатация – это немалые деньги, которые успешно осваивает наш институт. Следование идеологии "фундаментальные исследования – инженерная разработка – эксплуатация" и позволяет ИПС РАН постоянно развиваться.

Известно, что ИПС РАН – головная организация в суперкомпьютерных программах СКИФ и СКИФ-ГРИД. Расскажите об этом направлении.

В 1998 году мы установили контакты с белорусскими коллегами и по их инициативе, с учетом наших наработок, стали готовить решение о совместной программе в области высокопроизводительных вычислений (High-Performance Computing – HPC). По решению Национальной академии наук Белоруссии (НАН Беларуси) при поддержке Президента Белоруссии было решено сформировать научную програм-

Сергей Михайлович Абрамов,

член-корреспондент РАН, родился в 1957 году. После окончания МГУ и аспирантуры, с 1983 года работал в НИЦЭВТ. Курировал разработки спецпроцессоров ЕС ЭВМ (серии ЕС 27xx). С 1986 года работает в Институте программных систем (ИПС) АН СССР. С 1991 года является директором Исследовательского центра мультипроцессорных систем ИПС РАН. В 2003 г. возглавил ИПС РАН. Научный руководитель от России программ СКИФ и СКИФ-ГРИД. С 2003 года С.М.Абрамов – ректор Университета города Переславля им. А.К.Айламазяна, где он заведует кафедрой вычислительной техники и сетевых технологий. Лауреат премии Правительства РФ в области науки и техники. Специалист в области системного программирования и информационных технологий (суперкомпьютерные системы, телекоммуникационные технологии, теория конструктивных метасистем и метавычислений). Автор и соавтор более 100 научных работ и двух монографий.



му Союзного государства. Белоруссией были выделены деньги, организован временный научный коллектив по написанию программы, куда вошли и специалисты ИПС РАН. В 1998 году программу написали, еще год ушел на согласование с российской стороной (в Белоруссии на это потребовался 1–1,5 месяца). И в 2000 году стартовала союзная суперкомпьютерная программа СКИФ. Ее полное название – разработка и освоение в серийном производстве семейства высокопроизводительных вычислительных систем (суперкомпьютеров) и прикладных систем на их основе. Название очень точно отражает суть. Это была комплексная программа, включавшая даже разработку СБИС для вычислительных установок специального назначения (ускорители, спецвычислители). Мы разрабатывали конструктивы и схемотехнику для кластеров и гибридных машин (т.е. со спецвычислителями). Было создано программное обеспечение (ПО) для этих установок. ПО компоновали частично из свободного ПО (Open Source), но и очень много написано своего. ПО охватывает все системные средства (ядро операционной системы (ОС), параллельная файловая система, библиотеки поддержки параллельных вычислений, разные языковые средства поддержки параллельного программирования, системы очередей, системы мониторинга и т.п.), а также инструментальные средства для прикладных систем и 20 различных прикладных систем. Мы поддерживали два ядра ОС – наш

вариант ядра Linux, усиленный с точки зрения безопасности, и стандартное ядро.

В рамках программы СКИФ были вспомогательные мероприятия – например, создание единого информационного пространства проекта. Это и система поддержки коллективной разработки, и сервера хранения и обмена информацией, и каналы связи. Была организована подготовка кадров – ведь мало создать технику, нужны еще и специалисты, умеющие на ней работать. А к тому моменту рынка средств НРС в России практически не было.

Но ведь до СКИФ в нашей стране были другие разработки. В чем особенность именно этой программы?

Разумеется, программа СКИФ не стала в России первой и единственной в этой области. Были разработки и других организаций. В 2002 году появилась система МВС-1000М, первый российский компьютер с пиковой производительностью в 1 TFlops (НИИ "Квант" в кооперации с рядом других учреждений). А до 1991 года в нашей стране была целая суперкомпьютерная индустрия, в основном сосредоточенная в Министерстве радиоприборостроения (МРП). Это и НИЦЭВТ (серия ЕС ЭВМ и спецпроцессоров ЕС 27xx) с огромной системой кооперации вокруг него, и Институт проблем управления АН СССР (суперкомпьютеры серии ПС), и ИТМиВТ ("Эльбрус"), и Институт проблем кибернетики АН СССР ("Электроника ССБИС") и многие другие. Тогда страна позволяла себе полтора десятка суперкомпьютерных проектов, абсолютно конкурирующих друг с другом. На совершенно разных платформах. И все они доводились до внедрения в различных прикладных областях.

Во времена СССР 50% мощностей МРП (производственных, интеллектуальных и др.) были в Белоруссии. Там выпускалось от 40 до 60% всех советских компьютеров. Поэтому СКИФ по сути явился программой восстановления утраченной кооперации, в ее рамках мы фактически воссоздали кусочки МРП СССР. Причем НИЦЭВТ выступил одним из ключевых участников программы СКИФ. Программа СКИФ была толчком для развития НРС-рынка России.

То есть с самого начала в программе речь шла о массовом производстве высокопроизводительных компьютеров?

Да, это отражено и в названии. Правда, деньги выделялись не на серийное производство, а на создание мощностей для него. Изначально речь шла о нормальной промышленной разработке семейства суперкомпьютеров СКИФ, со всей необходимой конструкторской и программной документацией по ЕСКД и ЕСПД, с нормоконтролем, с созданием опытных образцов, с прохождением предварительных и государственных испытаний и присвоением соответствующих этапам литер.

Программа СКИФ завершена?

Изначально программе отводилось четыре года, но из-за задержек с финансированием она длилась пять лет – с 2000 по 2004 год. Программа была абсолютно успешной. В ее рамках сделано 16 опытных образцов вычислительных установок, подготовлен весь комплект программной и конструкторской документации. Две машины вошли в TOP500 – СКИФ К-500 (2003 год, 407-е место) и СКИФ К-1000 (98-е место в 2004 году). Из 16 опытных образцов 10 были установлены в России, остальные – в Белоруссии, в том числе – две самых мощных. СКИФ К-1000 с пиковой производительностью 2,5 TFlops стала самым мощным суперкомпьютером на территории СНГ на тот период. Это были очень значимые результаты и для России, и для Белоруссии. В частности, в 2004 году Президент Белоруссии в своем новогоднем обращении к нации, отметив ряд трудностей, в первую очередь сказал, что белорусскому народу есть чем гордиться – "у нас есть суперкомпьютер". Вне программы, т.е. уже на деньги заказчиков, было порядка 60 инсталляций компьютеров СКИФ.

Руководство Союзного государства также сочло программу абсолютно успешной. Достаточно сказать, что в заключении о завершении программы вместо стандартного "Отчет принять" было написано: "Отчет одобрить, поручить Национальной Академии наук Белоруссии и Роснауке подготовить предложение о продолжении работ". Такого больше в истории Союзного государства не было. В 2006 году по ее результатам десять человек получили Правительственные премии (восемь из России, двое – из Белоруссии).

Каково было продолжение?

Еще в 2003 году мы написали план продолжения работ, поскольку понимали, что останавливаться нельзя. Ведь область суперкомпьютерных вычислений развивается очень интенсивно, по экспоненте. Каждые 18 месяцев производительность суперкомпьютеров удваивается. Стоять на месте – значит растерять все уже созданное. Поэтому мы заранее написали предложения на следующий этап, но несмотря на прямое поручение премьер-министра РФ М.Е.Фрадкова (он возглавлял тогда кабинет Союзного государства), программу согласовали только к 2007 году – т.е. мы потеряли два года.

В 2007 году стартовала программа СКИФ-ГРИД. Она существенно расширена, в ней четыре направления. Первое – разработка грид-технологии. Основное в этом направлении – разработать ПО промежуточного слоя (middleware) для грид-систем. Второе направление – суперкомпьютеры семейства СКИФ, т.е. продолжение линии первой программы. Оно включает и аппаратные средства, и базовое ПО. Третье направление – это информационная безопасность. Если мы создаем грид-системы, причем на базе публичных каналов, то необходимо заниматься информационной безопасностью. Четвертое направление – пилотные демонстрационные сис-



темы. Ведь невозможно продемонстрировать достижения по первым трем направлениям, не показав их практически, при решении реальных задач.

Программа рассчитана на 2007–2010 годы. Ее бюджет вдвое больше (681 млн. руб.), чем у СКИФ. Объем работ также намного шире. Например, если в программе СКИФ со стороны РФ участвовало 10 организаций, то сегодня их 25. В структуре управления программой два заказчика-координатора – НАН Белоруссии и Роснаука. И два головных исполнителя – Объединенный институт проблем информатики НАН Белоруссии и ИПС РАН. Головные организации формируют всю кооперацию соисполнителей и распределяют между ними бюджет программы. Соотношение в бюджете между Россией и Белоруссией – 2:1, пропорционально вкладам в общий бюджет Союзного государства. Напомню, в первой программе был почти полный паритет, доля Белоруссии даже была чуть больше.

В программу СКИФ-ГРИД заложено много интересных работ. Чрезвычайно перспективно направление собственно грид-технологий. Оно может включать интеграцию вычислительных мощностей (объединение нескольких суперЭВМ в распределенную вычислительную систему), емкостей хранения (объединение дисковых пространств), источников данных (объединение территориально разнесенных разнородных БД) и канальных мощностей (одновременная передача данных по нескольким альтернативным каналам для интеграции их пропускной способности).

В этой области уже получены интересные результаты. Например – разработка распределенного хранилища научной информации для наук о Земле (ИКИ РАН), БД медицинской информации о результатах онкологических обследований в национальном масштабе. Интересна разработка для распределенных вычислений X-сом (НИВЦ МГЦ). Заслуживает внимания оригинальная отечественная система SKIF@HOME (ИПС РАН) интеграции неиспользуемых мощностей ПК для решения научных задач. В рамках направления демонстрационных проектов создаются очень интересные прикладные системы. Обо всем рассказать за один раз просто невозможно.

Тогда давайте сосредоточимся на направлении суперкомпьютеров СКИФ. В каком направлении развивается это семейство?

В развитии суперкомпьютеров СКИФ мы выделяем несколько поколений – рядов, все они программно совместимы снизу вверх. В рамках программы СКИФ создавались суперкомпьютеры ряда 1 и 2. Ряд 3 – в первой половине программы СКИФ-ГРИД, закончившейся в 2008 году. Ряд 4 включает проекты, над которыми мы работаем сейчас (в период 2009–2010 годов и далее). От ряда к ряду осваивались все более мощные средства суперкомпьютерных технологий.

Если говорить про пиковую производительность, ряд 1 включает системы с вычислительной мощностью порядка де-

сятков GFlops, ряд 2 – до 5 TFlops, ряд 3 – до 150 TFlops. Разумеется, это – возможности масштабирования, а не реально созданные системы. В 2009 году в ряде 4 мы планируем создать систему с пиковой производительностью 0,5 PFlops, в 2010 году – 1 PFlops, а в 2012 году – ~10 PFlops. И это – не фантастика, уже есть эскизная конструкторская документация, выпускаются опытные образцы модулей для таких систем.

Существенно, что производительность компьютеров СКИФ развивается по экспоненте, но быстрее, нежели зарубежные системы. Ведь мы стартовали с большим отставанием, догнать можно быстрее. Но эти темпы сохранились и после того, как компьютеры СКИФ вошли в TOP500. За пять лет вся мировая отрасль на тестах Linpack ускорила в 16 раз, а мы – в 185 раз, и готовы еще какое-то время поддерживать этот темп. Но, разумеется, производительность – это лишь интегральный показатель, компьютеры каждого ряда отличаются друг от друга конструктивными особенностями.

Во-первых, менялись семейства микропроцессоров. В этом нет нашей заслуги, но их нужно было освоить, поддержать на уровне ПО и т.д. Ряд 1 использовал 32-разрядные одноядерные процессоры, ряд 2 – 64-разрядные одноядерные, ряд 3 – двухъядерные 64-разрядные. В ряду 4 запланированы 64-разрядные процессоры с четырьмя и более ядрами. В принципе, приемлемы любые Linux-совместимые процессоры, но реально использовалась архитектура x86. Она открытая, и в рамках этой архитектуры есть конкуренция производителей – между компаниями Intel, AMD и другими, например VIA. Поэтому можно говорить о низких рисках импортозависимости.

Очень важная для суперкомпьютеров компонента – коммуникационная сеть. В ряду 1 использовались сети Myrinet и SCI (стандарт, продвигаемый в те годы компанией Dolphin) с топологией “двумерный тор”. В ряду 2 – уже трехмерный тор SCI и Infiniband, вспомогательная сеть – Fast Ethernet. В ряду 3 – Infiniband DDR, вспомогательная сеть – GB Ethernet. В ряду 4 мы сделаем отечественную коммуникационную сеть, по характеристикам превосходящую Infiniband QDR. Эта работа уже идет, есть опытные образцы плат, пишется встроенное ПО. В ней в рамках кооперации СКИФ участвуют и ИПМ им. Келдыша, и НИЦЭВТ, и ИПС РАН. К осени 2009 года такая системная сеть будет разработана. Работы по собственной коммуникационной сети особо важны, поскольку данную компоненту суперкомпьютеров легко поставить под экспортный контроль, наложить эмбарго на поставки. Ни с процессорами, ни с памятью, ни с системными платами этого сделать практически невозможно.

Еще один параметр, важный для суперкомпьютера – плотность упаковки аппаратуры. Чем она выше, тем короче линии связи, меньше задержки. Чем больше интеграция печатных плат, тем меньше кабелей. Соответственно, меньше разъемов и выше надежность. Последнее, в частности, означает, что можно строить систему из большего числа компонентов. Поэтому плотность упаковки – очень важный показатель уровня

суперкомпьютерных технологий. Но обратная сторона плотной упаковки – большее тепловыделение в единицу объема, необходимы специальные технологии отвода тепла. С этой точки зрения в ряде 1 было два процессора в модуле высотой 2U–4U. В ряде 2 – два процессора в модуле 1U. В ряде 3 появляются blade-системы (компания "Т-Платформы"), 10 узлов по два процессора в модуле 5U. Ряд 4 использует уже модули высотой 6U с 32 blade-серверами по два процессора в каждом (более 10 процессоров в пересчете на 1U). Соответственно, в конструкциях ряда 1 и 2 использовалась воздушная система охлаждения, в ряде 3 – трехзвенная система воздуха-фреон (вода на уровне шкафа), в ряде 4 – водяное охлаждение на уровне печатных плат.

Отмечу, что при переходе от ряда к ряду, от компьютера к компьютеру в огромной мере используется уже наработанная конструкторская документация. Так, запланированная в ряде 4 семейства СКИФ в 2010 году смена процессора и удвоение производительности системы потребует только смену BIOS'a. Следующий шаг развития в 2012 году повлечет переделку лишь 30% конструкторской документации.

В какой мере в разработках серии СКИФ используются отечественные аппаратные решения?

Можно ввести еще один параметр, характеризующий переход от ряда к ряду – уровень "отечественности". Он неуклонно возрастает. Если в ряде 1 собственными были шкафы, термодизайн и т.п., то в ряде 2 применялись уже модификации на уровне материнской платы. Их выполняли компании-производители, но по нашим заказам. Уже тогда появились отечественные средства построения коммуникационной сети (работа НИЦЭВТ), хотя они использовались и не во всех машинах. В ряде 3 задействованы blade-серверы собственной конструкции (компания "Т-Платформа"). В них хоть материнские платы и были импортные (Supermicro Atoka-2), но удалось придумать такую конструкцию и термодизайн, что на момент выпуска плотность упаковки оказалась на 18% выше по сравнению с любой другой. В рядах 2 и 3 была единственная отечественная печатная плата – контроллер сервисной сети ServNet (управление питанием, мониторинг, консоль управления и т.п.) разработки ИПС РАН.

ПО суперЭВМ СКИФ также полностью наше, пусть и на базе открытого ПО. С ряда 3 сборкой дистрибутива ПО СКИФ занимается российская компания "АльтЛинукс". Они на основе всех наших программных разработок формируют дистрибутив ОС (ALT Linux SKIF Cluster).

В работах по ряду 4 предусмотрено, что отечественным будет все, кроме микросхем. Как и в других наших проектах, при создании ряда 4 мы работаем совместно с западными партнерами. Но речь идет не о покупке лицензии, а именно о совместных разработках с паритетными условиями использования интеллектуальной собственности. Лицензионных отчислений либо нет, либо они зеркальны. Россия получит интеллектуальную собственность

и конструкторскую документацию на все конструктивы (вся "механика": шкафы, шасси, система водяного охлаждения и т.п.), на схемотехнику, на все печатные платы, включая объединительные (backplane) и материнские платы. У нас будет документация, права и возможность все выпускать в России. Партнеры будут использовать российские решения в области коммуникационных сетей, сервисной сети и т.п. А, скажем, наша сервисная сеть ServNet – это в своем классе действительно один из лучших продуктов в мире, рядом ее возможностей не обладает никто.

Что конструктивно будут представлять из себя компьютеры СКИФ ряда 4?

Минимальная единица в системе ряда 4 (законченная машина) – это корзина, шасси высотой 6U с 32 blade-серверами по два процессора. Используется водяное охлаждение на уровне платы. Напряжение питания – 48 В постоянного тока, мощность потребления – 10 кВт. Пиковая производительность для моделей 2009 года – 3 TFlops. Система оснащена сетью Infiniband для файловых обменов по протоколу TCP/IP и отдельной отечественной системной коммутационной сетью с топологией трехмерного тора – для счета.

Второй уровень системы – шкаф. В него входит 8 корзин по 6U. Общая пиковая производительность одного шкафа для моделей 2009 года – 24 TFlops. Третий уровень масштабирования – система из нескольких шкафов. Так, система из 21 шкафа в 2009 году может иметь пиковую производительность 504 TFlops.

С точки зрения архитектуры, суперкомпьютер СКИФ – это кластер?

Компьютер семейства СКИФ – это система с распределенной памятью. Можно было бы сказать, что это – кластер. Но данное утверждение не совсем верно. Обычно под кластером понимают систему, объединяющую стандартные вычислительные узлы. Но в ряде 4 речь идет о суперкомпьютерах с технологией высшего уровня. И это – еще один параметр, который принципиально отличает суперкомпьютеры СКИФ уровня 4 от всех предыдущих. Впервые российский суперкомпьютер будет делаться по так называемым технологиям уровня N.

Что такое технологии уровня N?

Возьмем список TOP500. Он выглядит как пирамида. Его верхушка принципиально отличается тем, что в нее входят суперкомпьютеры, первые 10–15 машин, на которых создаются технологии. Эти пиковые технологии и называют "технологиями уровня N", их невозможно купить. А все остальные компьютеры используют уже отработанные, стандартные технологии, которые можно купить – технологии уровня N-1. Суперкомпьютеры на их основе – это наибольшая часть рынка HPC. Но новые системы строят по технологиям уровня N, а сами технологии уровня N создаются и отрабатываются в рамках этих проектов.



В системах ряда 4 мы используем именно такие технологии. Мы их не покупаем, а создаем сами. Это – первый шаг, когда мы перестаем использовать только готовые идеи и начинаем создавать решения, которых не было ни у кого. А в качестве аналога смотрим на технологии суперкомпьютеров из первой десятки TOP500. Это – принципиальный момент, поэтому про СКИФ ряда 4 не совсем верно говорить, что это – кластер.

Например, технология водяного охлаждения – это технология уровня N, поскольку ее нельзя купить. В машинах ряда 4 будет отдельная сеть синхронизации всех системных часов, отдельная сеть для реализации барьерной синхронизации между параллельными процессами (в некотором смысле – система глобальных прерываний). Часть операций библиотеки MPI будет реализована аппаратно, т.е. на уровне коммутационной сети. Можно продолжать и дальше. Такие подсистемы и решения можно увидеть только в пиковых суперкомпьютерах. В свободной продаже на рынке подобных технологий нет, но они будут определять облик НРС в ближайшие годы. И мы этим занимаемся.

Иными словами, ряд 4 в программе СКИФ-ГРИД – это не просто уровень петафлопсов. Это – принципиально новый шаг, новый по содержанию.

Каковы перспективы производства таких систем, есть ли на них покупатели?

В 2007 году мы сделали суперкомпьютер СКИФ МГУ с пиковой производительностью 60 TFlops. Бюджет программы предусматривал на это 100 млн. руб., еще около 140 млн. руб. вложил МГУ.

Вторая часть программы предусматривает создание только модулей и решений для ряда 4 и никаких серьезных больших суперЭВМ. В ИПС РАН для испытаний мы сформируем опытный образец из двух шкафов по две корзины в каждом и с двумя blade-серверами в каждой корзине. Он позволит проверить функциональность всех подсистем, отладить их. Но это – лишь опытный образец. Гигантскую установку только на средства программы СКИФ-ГРИД сделать нельзя. Конечно, мы работаем с различными инвесторами и прямыми заказчиками, интерес очень большой.

Большие, пиковые суперкомпьютеры стоят серьезных денег, это сотни миллионов долларов. Суперкомпьютер семейства СКИФ производительностью 0,5 PFlops в 2009 году можно оценить примерно в 2 млрд. руб. 1 PFlops в 2010 году – это около 3 млрд. руб. И 5–7 PFlops к весне 2012 года – около 5 млрд. руб.

Есть ли перспективы продолжения программы, в рамках которого подобные средства изыщутся?

Программы СКИФ и СКИФ-ГРИД очень успешны. Результаты работы программы получили самую высокую оценку

на уровне Правительства России и Белоруссии (Правительственная премия 2007 года). От Союзного государства поступают предложения о продлении и расширении программы СКИФ-ГРИД. Сейчас мы готовим такой проект добавочных мероприятий, в которые заложено и создание больших моделей. В рамках этих идей прорабатывается построение суперкомпьютерного центра на базе ИПС РАН.

Почему именно в ИПС РАН?

Наш институт изначально строился как суперкомпьютерный центр. В его здании есть неиспользуемое помещение – машинный зал площадью 700 м². Он делался еще для "Эльбрусов" и "Электроники ССБИС". В составе комплекса ИПС РАН есть вся необходимая инженерная инфраструктура – градирни и бассейны для охлаждения, по территории института проходит ЛЭП, есть две электроподстанции с резервом мощности 3 МВт, и расширение резерва стоит недорого. Нами создана городская опорная оптоволоконная сеть с пропускной способностью 1 Гбит/с. Сеть имеет мощную систему внешних каналов связи – сегодня это более 100 Мбит/с, есть коммерческое предложение с разумным сроком реализации и приемлемой ценой об организации линии с пропускной способностью 1 Гбит/с непосредственно до известной точки обмена трафиком М9-IX.

Надо учесть, что у вычислительных установок петафлопсного уровня мощность потребления электроэнергии — на уровне нескольких мегаватт. В Москве энергетические мощности дефицитны, как и в любом мегаполисе. Москва — это редкий мегаполис, где создают суперкомпьютерные центры и data-центры. Ведь в подобных городах и площади, и энергоресурсы крайне дороги.

В большинстве стран мира мощные суперкомпьютерные центры организуют в национальных лабораториях, которые расположены вне крупных городов. Классический суперкомпьютерный центр LRZ в Германии находится в нескольких километрах от Мюнхена, в Гархинге, что не мешает работать с ним пользователям всего мира.

Таким образом, в ИПС РАН есть все — канал связи, помещение, энергетические мощности. Это уже сделанные огромные инвестиции, которые необходимо использовать. И со стороны властей такое понимание есть.

То есть можно надеяться, что мощные вычислительные системы СКИФ будут созданы?

Все зависит от политической воли руководства страны и от того, сколько времени займет согласование наших новых предложений. Конечно, мы работаем и с другими инвесторами. Например, уже год ведем переговоры с госкорпорацией РОСНАНО. Речь идет о суперкомпьютерных сервисах в интересах нанотехнологической отрасли — не просто о компьютерах, а о создании целого комплекса из аппаратных средств, специализированного ПО и сервисов для nanoиндустрии.

Рассматривается ли в рамках проекта возможность выпуска собственной элементной базы, например — микропроцессоров?

Исходя из общей логики, следующий шаг освоения — это именно элементная база. Очевидно, что переход на отечественную элементную базу должен быть постепенным. Там, где это выгодно и безопасно, разумно использовать импортные компоненты. Для СКИФов процессоры могут быть любыми, лишь бы они были Linux-совместимы — иначе придется с нуля писать все ПО, что совершенно нереально. И к отставанию в области микроэлектроники добавится отставание в области ПО.

В России есть несколько команд, которые создают современные микропроцессоры. Не желая никого обидеть, упомяну только две — НИИСИ РАН и МЦСТ. Если говорить о спецпроцессорах, то уже сегодня в рамках программы СКИФ-ГРИД работают многоядерные DSP-процессоры семейства Multicore компании ЭЛВИС. И это — далеко не все отечественные команды, способные работать в данной области. Сегодня мы пытаемся сформировать проект — небольшой, на три года, — предусматривающий создание аппаратуры вычислительной системы на базе отечественных микропроцессоров. Пока это проект, но по крайней мере мы знаем, с кем можно кооперироваться в этой области.

Чисто с экономической точки зрения, можно ли говорить о рынке суперкомпьютеров, могут ли они сформировать существенный заказ для полупроводниковых производств?

С точки зрения рынка, коммерческой привлекательности у рекордных НРС-систем нет. Нет нигде в мире, поэтому за крупными НРС-проектами стоят огромные государственные деньги. Сегодня практически никто не делает специальных процессоров для суперкомпьютеров. Для ведущих производителей микропроцессоров область НРС — это как "Формула-1" для автопроизводителей. Это полигон, где отрабатываются технологии для массовых применений — для ПК, серверов, игровых приставок и т.п. В подавляющем большинстве суперкомпьютеров используются процессоры с архитектурой x86. Ситуация с процессорами POWER от IBM аналогична — в этих процессорах используется ядро, которое также применяется в различных массовых устройствах, включая игровые консоли. Это же ядро интегрировано и в процессоры CELL. То есть технологии отрабатываются в области НРС, но окупаются они в массовой сфере. Поэтому рынок НРС, по крайней мере по отношению к топ-машинам, не коммерческий.

Но ведь не из любви к искусству, и не только на средства оборонных бюджетов во всем мире создаются сверхвысокопроизводительные системы, формируются государственные программы их развития?

Во всех развитых странах мира (пока кроме России) область высокопроизводительных вычислений понимают единственным образом — как ключевой инструмент обеспечения конкурентного превосходства. В любой отрасли экономики.

При президенте США была консультационная группа по информационным технологиям — PITAC (www.nitrd.gov/pitac/index.html). Они ввели термин "киберинфраструктура" — суперкомпьютерные центры, объединенные каналами связи, т.е. грид-системы. Очень интересны заголовки их отчетов: "Киберинфраструктура и здоровье нации", "Киберинфраструктура и новые подходы к образованию", "Киберинфраструктура и преодоление цифрового неравенства" и т. п. — перечислены практически все наши национальные проекты, только на пять лет раньше и с приставкой "киберинфраструктура". Но это писали представители IT-отрасли. Они — лица заинтересованные, им лишь бы бюджетом завладеть. Но вот что говорит Дебора Винс-Смит, председатель Совета по конкурентоспособности США, экономист: "Сегодня технологии, таланты и деньги доступны многим странам. Поэтому США стоят перед лицом беспрецедентной экономической конкуренции. Страна, желающая победить в конкуренции, обязана победить в вычислениях". Все сказано: НРС — необходимый, единственный инструмент победы в конкуренции.

В США постулируется, что НРС, как общественное благо, — это бизнес государства. Это — новая инфраструктура экономики, основанной на знаниях. Равно как в свое время желез-



ные дороги, автобаны, ЛЭП и нефтепроводы были элементами инфраструктуры экономик других типов. Экономика, основанная на знаниях, требует кибернетической инфраструктуры. Поэтому в США она создается на бюджетные деньги, порядка 4 млрд. долл. в год. В объединенной Европе этот показатель примерно такой же. За бюджетные деньги организуются и оснащаются суперкомпьютерные центры уровня N. Мало того, они и содержатся за счет бюджета. И что самое важное – этот ресурс свободно доступен всей экономике страны. В НРС-центры обращаются фирмы со своими задачами, из них выбираются наиболее важные и решаются за государственный счет. Интеллектуальная собственность остается у фирм-заказчиков. Они получают новые материалы, конструкции и т. п., создают на этой основе новые товары и услуги, превосходящие продукцию иностранных конкурентов, расширяют свои продажи, платят налоги и в этот момент рассчитываются с бюджетом. Такую схему называют не "коммерческая эффективность", а "бюджетная эффективность".

Так работают все ведущие суперкомпьютерные центры – они не продают компьютерное время, не берут денег с заказчиков. А вот когда мы согласуем свои программы, с нас требуют расчет экономической эффективности. Причем наши экономисты прекрасно понимают ситуацию, они знают термин "бюджетная эффективность". Но у них нет методики ее расче-

та, поэтому предложение напрямую и не проходит. В итоге все выливается в затягивание сроков согласования программ.

Опять же, в отчетах о развитии суперкомпьютерной отрасли в США прямо пишут, что эта отрасль – хрупкая, что задержка на год означает катастрофу, исчезают фирмы и "память фирм" – по-нашему, научная школа. Вывод – должен быть особый механизм формирования госзаказа на эти работы. Недопустимы годовые согласования. И в США, в Европе находят эти механизмы. Обратите внимание – тревогу бьют в США, стране – абсолютном лидере рынка НРС. Может быть, пора бить тревогу и нашим экономистам?

А строить мощные суперкомпьютеры России все равно придется. В мире стартуют программы создания систем на 10–20 PFlops. Все крупнейшие национальные лаборатории и США, и Европы говорят о транспетафлопсных системах. У нас есть реальные достижения, работающая команда, налаженная кооперация. Можно очень эффективно и быстро использовать этот задел. Но если мы продолжим на уровне государства тяжелые раздумья и долгие согласования, то у нас ничего хорошего не произойдет.

Будем надеяться на хорошее.

Спасибо за содержательный рассказ.

С С.М.Абрамовым беседовали П.П.Мальцев и И.В.Шахнович