

РЕШЕНИЕ
четвёртого Национального суперкомпьютерного форума

24–27 ноября 2015 г.

г. Переславль-Залесский

Четвёртый Национальный суперкомпьютерный форум прошел успешно, подтвердив свой статус крупнейшего мероприятия суперкомпьютерной отрасли России, комплексной и полной, нейтральной и равнодоступной, универсальной и многоаспектной, многогранной общеотраслевой площадки.

По итогам Форума принято настоящее Решение:

1. Обратиться к профильным ФОИВам и Институтам Развития с предложением назначить своих представителей для работы в оргкомитете и программном комитете Форума;
2. Констатировать роль Форума, как инструмента доступа к мнению научного и экспертного сообщества отрасли, открытого круглый год для использования ФОИВами и другими организациями. Отметить первый положительный опыт ФАНО России в этом направлении. Обратить внимание остальных ФОИВ и институтов развития на данную возможность.
3. По вопросам, поставленным перед экспертным сообществом Форума ФАНО России, отметить:
 - 3.1. России жизненно необходима суперкомпьютерная инфраструктура (киберинфраструктура), без которой все разговоры о новой экономике останутся разговорами. Эта стратегическая задача государства. Отставание России от передовых экономик мира становится критическим. Это тормозит внедрение и развитие передовых производственных технологий, стратегических направлений научных исследований и, в конечном счёте, угрожает обороноспособности страны.
 - 3.2. Киберинфраструктура должна строиться по иерархическому принципу: на первом (национальном) уровне 3–4 суперкомпьютера класса первой мировой десятки; на втором (региональный и отраслевой) — 15–20 суперкомпьютеров уровня первой мировой сотни; третий уровень — корпоративный (Top500).
 - 3.3. Национальные центры киберинфраструктуры должны создаваться на базе научных организаций имеющих наибольшие компетенции в данной области (к таким организациям, безусловно, можно отнести ИПМ им. М.В. Келдыша РАН)
 - 3.4. Полный текст меморандума по вопросам, поставленным ФАНО России, приведён в приложении.
4. Провести пятый Национальный суперкомпьютерный форум в период ориентировочно 29.11.2016–02.12.2016.

По поручению участников Форума

Председатель оргкомитета Форума,
чл.-корр. РАН



С.М. Абрамов

После обсуждения на круглом столе Форума вопросов поставленных ФАНО России, сформулировано следующее мнение сообщества:

1. Выработка подходов к оценке важности решения сверхмасштабных исследовательских вычислительных задач, их классификации и категоризации, как по вычислительным масштабам, так и по особенностям архитектуры требуемых вычислительных систем;

Оценка важности:

Суперкомпьютерная отрасль — дело государственное. Главная и самая важная задача государства — обеспечить развитие суперкомпьютерной отрасли.

Развивать отрасль государство должно не путём «раздачи всем сёстрам по серьгам», а путём поддержки реально существующих точек роста, т.е. тех коллективов, которые уже имеют заделы и результаты и поддерживать их не вливанием денег, но размещением госзаказов и стимулирующими мерами для бизнеса. Размещать заказы именно у отечественных компаний реализующих отечественные разработки (как в части аппаратных средств, так и в части программного обеспечения и математических методов).

Что касается задач, которые должны решаться отраслью, то существует Указ Президента Российской Федерации от 07.07.2011 г. № 899 «Об утверждении приоритетных направлений развития науки, технологий и техники в Российской Федерации и перечня критических технологий Российской Федерации». Ни одна из перечисленных там технологий и ни одно из направлений не может развиваться на современном уровне без применения суперкомпьютерных решений. Суперкомпьютерные проекты именно в этих направлениях и по этим технологиям и должны поддерживаться в первую очередь. При этом важно учитывать, что:

— Высшие уровни суперкомпьютерной инфраструктуры (национальные, отраслевые и региональные суперкомпьютерные центры) во всем мире применяются не для сиюминутных индустриальных задач, а для проведения научных исследований дающих знания, необходимые для обеспечения будущего прорыва по приоритетным направлениям развития науки, технологий и техники. Речь идет о новых знаниях, которые лягут в основу новых технологий и изделий, обладающих конкурентным превосходством.

— Нельзя механически, с порога отменить направления и технологии не вошедшие в президентский указ. Какую-то часть ресурсов необходимо стараться выделять и на совершенно новые направления. Кто знает, что будет главным приоритетом завтра? Но здесь нужен более осторожный поход. Если актуальность и высокая важность направлений из президентского указа не вызывает сомнений, то другие технологии и подходы необходимо тщательно анализировать с привлечение широкого экспертного сообщества. Механизмы такой широкой экспертизы существуют и ими необходимо пользоваться. Одним из таких механизмов реализован и круглогодично доступен — Национальный Суперкомпьютерный Форум.

Классификация задач:

Можно предложить много вариантов классификации и все они имеют право на существование. Различные группы специалистов (разработчики, пользователи, администраторы и т.п.) пользуются различными классификациями, наиболее адекватными их видению вопроса.

Классификация по математическим методам решения задачи основана на том, что многие, на первый взгляд совершенно различные задачи решаются при помощи одних и тех же математических методов (например, любые процессы, связанные с обтеканием тел потоками жидкостей или газов: движение корабля, автомобиля, дозвукового летательного аппарата, атмосферные явления, морские течения, работа ветряных и гидравлических турбин и многое другое описывается одной и той же системой уравнений Навье-Стокса).

Преимущество такого метода классификации в том, что он сразу же определяет и желаемую архитектуру вычислительной установки, поскольку применяемые математические методы диктуют те или иные особенности архитектуры: для каких-то методов. Например, методы Монте-Карло не требуют одновременного доступа ко всей памяти установки и можно использовать более дешёвую массово-параллельную архитектуру. Для других методов (например, вычисления на неравномерных сетках) наоборот требуется прямая адресация огромных массивов памяти и здесь более уместно решение на базе симметрично-мультипроцессорной (SMP) архитектуры. Наконец, есть математические методы, требующие специальных вычислительных возможностей для отдельных частей задачи, здесь оптимальным является асимметричный параллелизм (ASMP архитектура).

Недостатком этой классификации является то, что это классификация по математическим методам, а не по прикладной сути задачи. Неспециалист в математических методах не сможет ничего отклассифицировать, да и специалисты часто будут подолгу спорить.

Другой поход к классификации – по прикладному назначению. Это полная противоположность предыдущему подходу и их достоинства и недостатки симметричны. Такая классификация понятна прикладному специалисту и он сразу понимает о чём речь. С другой стороны, при такой классификации, например, задачи распознавания фронта (или спада) сигнала в электронике и задача определения точек разворота биржевых курсов окажутся в «разных корзинах», хотя с точки зрения математики это одна и та же задача.

Категоризация по масштабам и особенностям архитектуры

Также можно классифицировать задачи по вычислительной ёмкости (по необходимым объёмам вычислений). Вычислительная ёмкость, определяется не только используемыми математическими методами, но и природой задачи.

Существуют две полярные точки в такой классификации.

Для некоторых задач (например, для определения геометрических размеров строительных конструкций) сверхвысокая точность не нужна, всё равно никто и никогда не изготовит кирпич с точностью до нанометра просто потому, что это не нужно. Т.е. это такие задачи, которые можно решить с достаточной точностью (более точное решение просто не нужно), если иметь достаточные (но вполне «конечные») вычислительные ресурсы.

На другом полюсе стоят задачи с потенциально бесконечной (по крайней мере с сегодняшних позиций) вычислительной сложностью. Например, при моделировании атмосферных явлений (прогноз погоды, прогноз последствий применения оружия массового поражения и т.п.) или, скажем, при моделировании ядерных реакций, или процессов разрушения материалов, увеличение вычислительных ресурсов будет давать увеличение точности (или объёма, или временных отрезков), и это будет разумное, нужное с прикладной точки зрения увеличение. Никакими современными и «видимыми из

сегодняшнего дня» вычислительными ресурсами нельзя добиться такой точности решения этих задач, при которой можно было бы сказать: «точнее уже не требуется».

Американские коллеги выделяют четыре группы задач, разделённых по характеру вычислительной сложности. В основе такого деления лежит предположение о том, что удастся сделать в задаче, если вдруг увеличить производительность вычислительной установки на 2-3 порядка. Группы такие:

- I. задачи, которые решались бы также как сегодня, но с большим количеством точек (т.е. с возможностью обчислять большие площади/объёмы). Это обычно «слабо-масштабируемые» задачи, например, глобальное (по всей планете) моделирование погоды;
- II. задачи, которые можно решить с той же точностью, что и сейчас, но значительно быстрее. Например, гидродинамическая задача стабилизации крыла самолёта могла бы решаться в реальном времени за секунды, а не за десятки часов как сейчас;
- III. задачи, которые можно решать с той же точностью, что и сейчас, но при этом моделировать ситуацию на большие временные отрезки. Например, долгосрочный прогноз погоды;
- IV. задачи, которые можно решать со значительно большей точностью за счёт учёта большего количества факторов. Например, моделирование процессов в мировом океане или приливных течений с учётом преломления волн и турбулентности.

2. Формулирование предложений по вопросу создания суперкомпьютерной инфраструктуры Российской Федерации с учётом необходимости развития её региональной составляющей;

Киберинфраструктура сегодня — это просто инфраструктура современной экономики. В своё время современная для того времени экономика была невозможна без железнодорожной сети или единой энергосистемы. Сегодня современная экономика невозможна без киберинфраструктуры.

Киберинфраструктура России на наш взгляд должна состоять из трёх уровней. Первые два уровня обеспечиваются государством, третий — корпоративный, его мы здесь рассматривать не будем.

Национальный уровень: 3—4 установки уровня первой мировой десятки. Устанавливаются в крупнейших государственных центрах коллективного пользования.

Региональный и отраслевой уровень: 15—20 установок уровня первой мировой сотни. Устанавливаются в отраслевых и региональных центрах коллективного пользования.

Все ЦКП должны быть соединены высокоскоростными каналами связи.

ЦКП должны предоставлять предприятиям и научным коллективам не только (и не столько) машинное время, но и сервисы по математическому моделированию, программированию и т.п. Т.е. предприятие может обратиться в такой центр с описанием своей технической проблемы, а специалисты центров должны уметь построить модели, написать программы и решить проблему.

Создание и эксплуатация ЦКП национального уровня осуществляется за счёт федерального бюджета. Речь идет о полном покрытии всех оперативных расходов ЦКП из средств федерального бюджета: электроэнергии, текущий ремонт, зарплата всех специалистов и сотрудников, задействованных в обеспечении разностороннего сервиса (математическое моделирование, программирование, обеспечение счета, поддержка эксплуатации и безопасности и т.д.). ЦКП регионального и отраслевого уровня создаются и содержатся за счёт совместно федерального бюджета и бюджетов регионов.

Суперкомпьютерная киберинфраструктура это инфраструктура страны для нового уклада экономики — экономики, основанной на знаниях. Поэтому, ЦКП равнодоступны и бесплатны для всех

предприятий и научных центров на конкурсной основе. Речь идет не о конкурсе денег: кто больше заплатит за сервис. Нет, суперкомпьютерный сервис от ЦКП высшего уровня для всех бесплатный (оплачен из бюджета страны). Речь идет о конкурсе идей: чьи заявки на суперкомпьютерный сервис могут привести к созданию в дальнейшем новых технологий и изделий с максимальным эффектом для страны. Окупаемость этих ЦКП должна быть не прямой, но через поступление налогов от создания и выпуска на рынок конкурентоспособной (и «конкурентопревосходящей») продукции.

Приблизительно так работает программа INCITE в США и программа PRACE в Европе. Очевидно, что и в России, если нам нужна конкурентоспособная экономика, необходимо нечто подобное.

3. Разработка предложений и концепций программ и мероприятий по повышению квалификации и обучению специалистов, способных эффективно использовать современную суперкомпьютерную технику для решения важнейших научно и научно-технических задач;

В области суперкомпьютерного образования необходимо:

I. Расширить перечень направлений подготовки высшего образования в рамках укрупненных групп направлений подготовки 02.00.00 Компьютерные и информационные науки и 09.00.00 Информатика и вычислительная техника. Перечень направлений подготовки в рамках этой группы должен обеспечить возможность подготовки квалифицированных специалистов на всех этапах разработки вычислительных комплексов, создания информационных технологий, обеспечения надлежащей инфраструктуры информационного пространства страны, средств информационного взаимодействия и информационного рынка. Такой перечень может включать направления подготовки: Конструкторско-технологическое обеспечение вычислительных комплексов и систем, Математическое моделирование технологических процессов и систем, Системное и прикладное программное обеспечение, Программное обеспечение высокопроизводительных вычислительных комплексов, Математическое моделирование социальных и экономических процессов, Экономика информационных и наукоемких производств и др.

II. В связи с необходимостью углубленного изучения математики, информатики по направлениям, связанным с разработкой математического и алгоритмического обеспечения для разработки современных информационных систем и технологий, включить соответствующие направления подготовки в перечень специальностей высшего образования — специалитета,— и разработать федеральные образовательные стандарты, обеспечивающие требуемый уровень математического образования по этим направлениям.

III. Разработать правовые механизмы, поддерживающие развитие взаимодействия между образовательными и научными учреждениями, в том числе в сетевой форме, что позволит использовать опыт высококвалифицированных специалистов в образовательном процессе.

IV. Из-за специфики суперкомпьютерных технологий, учебные планы подготовка специалистов по данным направлениям обязаны (1) проходить в формате специалитета; (2) не менее двух завершающих лет учебного плана проходить в формате стажировки (командирования) обучающегося на предприятиях, занятых в процессе исследований и разработок отечественных суперкомпьютерных технологий переднего мирового уровня (по некоторой отрасли технологий: аппаратные средства, программное обеспечение, математические методы, аспекты поддержания сервисов и эксплуатации ЦКП и т.п.) Необходимо законодательно поддержать данную особенность, в том числе и механизмами фи-

нансового обеспечения покрытия затрат как студента (проезд и проживание в месте командирования), так и так и наставников (оплата руководства ими данной НИОКР–практики стажера).

4. Прогноз необходимой производительности вычислительных систем с учётом сегодняшних и перспективных задач, в том числе по областям применения: ядерно-оборонный комплекс, высокотехнологические отрасли промышленности (авиа, судостроение, автомобилестроение, космическая отрасль), атомная энергетика, нефтегазовые отрасли, новые материалы на основе нанотехнологий, биотехнологий;

Оценка необходимой производительности, дело неблагодарное, особенно для «вычислительно бесконечных» задач (см. раздел 1).

Например, сегодня специалисты из «Росатома» заявляют, что для решения всех задач стоящих перед отраслью им достаточно иметь вычислительные мощности в 20 эксафлопс. Однако надо понимать, что их задачи как раз «бесконечны» и 20 эксафлопс им достаточно для сегодняшней постановки задач. Как только они получают эти 20 эксафлопс, они поставят задачи получения большей точности (больших временных отрезков и т.п.) и им снова потребуется в сотню раз больше.

Есть и другой подход. Сегодня в нефтегазовой отрасли создаётся специализированная вычислительная платформа для решения задач сейсмоки и гидродинамики с акцентом на скорость решения целевых задач, без всякой привязки к «флопсам». Т.е. не ставится задача «дать столько-то операций в секунду», а ставится задача обсчитать месторождение за приемлемое время. При этом никого не интересует, сколько там будет процессоров или флопс.

Наиболее разумным представляется подход Еврокомиссии – проект PRACE (Partnership for Advanced Computing in Europe). Там рассуждали примерно так: любые оценки в «количестве операций» временны, т.к. количество операций, которые выполняю суперкомпьютеры, возрастает в тысячу раз каждые 11 лет, но это не приводит к ситуации «все задачи решены», просто усложняются постановки задач. Поэтому, для обеспечения потребностей европейской науки и промышленности на современном (в каждый момент времени!) уровне следует иметь в доступности 5-6 вычислительных установок примерно уровня первой мировой десятки и два-три десятка установок уровня первой мировой сотни. При таком подходе как таковое количество операций не фиксируется. С ростом производительности суперкомпьютеров в мире, будет расти и производительность заявленного парка. Они посчитали, что такие вычислительные мощности будут достаточны для решения необходимых задач в их сегодняшней (в каждый момент времени!) постановке.

Именно такой подход к построению киберинфраструктуры можно рекомендовать и для России (см. раздел 2 выше). Он требует постоянного обновления парка, но при этом гарантирует постоянное наличие вычислительных мощностей для решения актуальных задач в «сегодняшней» постановке.

5. Оценка потребности финансовой поддержки проектов увеличения производительности отечественной суперкомпьютерной инфраструктуры;

Для создания и содержания необходимой России инфраструктуры (см. раздел 2 выше) необходимо построить 3–4 суперкомпьютера уровня первой мировой десятки и 15–20 суперкомпьютеров уровня первой мировой сотни. Кроме того, необходимы современные каналы связи. Кроме того, для того чтобы постоянно оставаться на уровне «первой десятки» и «первой сотни» компьютеры должны обновляться примерно раз в три года.

В целом это потребует затрат (только из федерального бюджета) порядка 200-250 миллионов долларов в год. Также потребуются средства на оперативные расходы по содержанию этих центров (электроэнергия, инфраструктура, каналы связи, персонал). Эти расходы можно оценить, методики для этого есть.

б. Вопрос целесообразности проекта по созданию на базе Института прикладной математики им. М.В. Келдыша РАН суперкомпьютерного центра коллективного пользования. Оценка стоимости проектирования СЦКП, а также оценка полной стоимости эксплуатации системы.

ИПМ им. Келдыша РАН – признанный центр суперкомпьютерной отрасли России. Институт занимается как математическими проблемами эффективного использования суперЭВМ, так и разработкой системного программного обеспечения, а также разработкой архитектуры и собственно построением суперкомпьютеров.

С этой точки зрения создание ЦКП именно на базе этого института оправдано и разумно, т.к. они смогут не только обеспечить создание и эксплуатацию центра, но и смогут как сами загрузить его, так и оказать всю необходимую поддержку сторонним пользователям.

Судя по широте исследований и работ ведущихся в ИПМ им. Келдыша РАН, им необходима **по большей части универсальная машина (на универсальных процессорах)**, но с обязательной установкой специализированных ускорителей вычислений на некотором количестве узлов. Причём ускорителей всех типов: GPGPU от NVidia и от AMD, Intel Xeon Phi, FPGA и т.п. Такой гибридный компьютер с архитектурой ASMP должен покрыть все потребности как самого ИПМ им. Келдыша РАН, так и сторонних пользователей ЦКП.

Для такого компьютера (с архитектурой ASMP) и производительностью в 15 петафлопс, заявленная стоимость в 3,3 миллиарда рублей выглядит несколько заниженной. Если же предполагается строить компьютер на базе одного ускорителя (например, Intel Xeon Phi), то, такой компьютер не является универсальной вычислительной установкой, и не закроет ни всех нужд самого ИПМ им. Келдыша РАН, ни, тем более, сторонних пользователей.

Необходимо подчеркнуть, что во всем мире суперкомпьютерные центры высшего национального уровня создаются именно так: не просто в месте, где есть электроэнергия и можно обеспечить вспомогательный персонал для поддержания бессбойной работы 7×24, а в ведущих национальных исследовательских центрах, где сосредоточены максимальные компетенции в развитии новых математических методов, создании новых языков и систем программирования для суперЭВМ, разработки новых решений в части программного обеспечения и аппаратных средств суперкомпьютеров. Поскольку национальный уровень киберинфраструктуры призван не просто предоставлять «машинное время», а давать именно высокоуровневый суперкомпьютерный сервис. Сюда поступают самые сложные проблемы и задачи, для которых, скорее всего, сначала надо развивать математический аппарат, потом создавать расчетные методы и алгоритмы, потом уже — программную реализацию расчета, и только затем — выполнять суперкомпьютерный расчет и (по обратной цепочке) научную интерпретацию результатов расчета в терминах решаемой проблемы. Это все должно быть поддержано местным персоналом национального суперкомпьютерного центра. И это все в полной мере присутствует в ИПМ им. М.В. Келдыша РАН.