

УДК 004.9; 004.382.2

СОЗДАНИЕ БАЗОВОЙ КИБЕРИНФРАСТРУКТУРЫ СОЮЗНОГО ГОСУДАРСТВА

С.М. Абрамов¹, В.В. Анищенко², А.М. Криштофик², А.А. Московский¹

¹Институт программных систем им. А.К. Айламазяна РАН, Переславль-Залесский;
Объединенный институт проблем информатики НАН Беларуси, Минск

Рассматриваются вопросы создания и развития информационно-вычислительного пространства Союзного государства, а также этапы создания базовой киберинфраструктуры.

Введение

Будущие технологии и открытия напрямую зависят от используемых вычислительных ресурсов. Так, заместитель министра энергетики США Рэймонд Орбах отметил, что суперкомпьютер «позволит ученым моделировать физические процессы, об изучении которых прежде не было и речи», а также назвал «развитие суперкомпьютерной техники, наряду с проведением экспериментов и разработкой теорий, одним из ключевых факторов для научных открытий» [1].

Вопросам создания суперкомпьютеров и развитию новых вычислительных технологий было посвящено заседание Совета Безопасности Российской Федерации [2], состоявшееся 28 июля 2009 г. в Москве, на котором рассмотрены:

- недостатки в освоении суперкомпьютерных технологий, недопонимание этого вопроса федеральными органами, представителями науки, промышленности, бизнеса;
- необходимость создания требуемого мощного вычислительного суперкомпьютерного ресурса на основе грид-технологий и телекоммуникационной инфраструктуры;
- финансирование производства суперкомпьютеров;
- стимулирование их востребованности.

1. Общая характеристика киберинфраструктуры

Информационно-вычислительное высокопроизводительное пространство (киберинфраструктура) Союзного государства - это совокупность суперкомпьютерных центров, систем хранения данных, систем связи и технологий, обеспечивающих предоставление ресурсов пользователям и приложениям и их эффективное использование, а также реализующих функции глобальных вычислений. Для создания киберинфраструктуры используются технологии метакомпьютинга. Метакомпьютинг определяется как «использование мощных вычислительных ресурсов, прозрачно доступных посредством телекоммуникационной среды». В дополнении к условию прозрачности применимы также такие характеристики, как бесшовность, масштабируемость и глобальность. Таким образом, в новой парадигме метакомпьютинга наличие телекоммуникаций полностью скрыто, а подключенные к сети компьютеры используются как единый объединенный вычислительный ресурс. Основной акцент в метакомпьютинге делается на то, что потенциальный пользователь может получить практически неограниченные ресурсы для вычислений и хранения данных. Распределенные ресурсы спрятаны в архитектуру метакомпью-

тера с использованием грид-технологий. Грид является согласованной, открытой и стандартизированной средой, которая обеспечивает гибкое, безопасное, скоординированное объединение ресурсов и предоставление широкому кругу пользователей, как индивидуальных, так и в рамках виртуальных организаций, т. е. динамически формирующейся совокупности независимых пользователей, учреждений и ресурсов.

Отличительной особенностью является возможность создания мощного вычислительного ресурса. В качестве процессорных ресурсов рассматриваются, например, рабочие станции и ПК. На самом деле, основные вычислительные мощности сосредоточены вовсе не в суперкомпьютерном парке. Если организация располагает, скажем, тремя тысячами рабочих мест на базе рабочих станций, то за время их регулярного простоя потерянные циклы составят существенную долю даже терафлопной производительности. Мощные ресурсы - суперкомпьютеры, кластеры, SMP-системы - остаются важнейшей составляющей ресурса. Кроме того, новая трактовка применима к разнообразным типам ресурсов: телекоммуникациям, системам массовой памяти, хранилищам данных, а также измерительным и научным инструментам, например радиотелескопам.

Выход за рамки высокопроизводительных систем и приложений выявляет реальное содержание технологии грид - это инфраструктура для поддержки любой глобально распределенной вычислительной деятельности. Из инфраструктуры грид могут извлечь пользу всевозможные приложения: электронный бизнес, кооперативное проектирование, исследование данных, системы обработки высокой пропускной способности (High Throughput Computing - HTC) и, конечно, распределенный суперкомпьютинг (метакомпьютинг). Для многих из этих приложений, в том числе и с большим объемом вычислений, но с «хорошими» свойствами (грубо гранулированных, конвейеризуемых), не требуются высокопроизводительные телекоммуникации, как для метакомпьютинга. Тогда в качестве телекоммуникационной составляющей инфраструктуры грид может выступать обычный Интернет - неограниченно масштабируемый, всеобъемлющий и повсеместный уже сейчас.

Среди основных направлений использования грид-архитектуры на данный момент можно выделить:

- организацию эффективного использования ресурсов для небольших задач, с утилизацией временно простаивающих компьютерных ресурсов;
- распределенные супервычисления, решение очень крупных задач, требующих огромных процессорных ресурсов, памяти и т. д.;
- вычисления с привлечением больших объемов географически распределенных данных, например в метеорологии, астрономии и физике высоких энергий;
- коллективные вычисления, в которых одновременно принимают участие пользователи из различных организаций.

Киберинфраструктура Союзного государства создается последовательно в четыре этапа:

- создание опытного участка грид-сети;
- создание базовой киберинфраструктуры, разработка технологий ее использования;
- развитие киберинфраструктуры и ее использование.

1. Создание опытного участка грид-сети

Создание опытного участка грид-сети выполняется в рамках выполнения программы Союзного государства «СКИФ-ГРИД». При этом работы выполняются в два этапа. На первом этапе создан СКИФ-Полигон как конфедерация суперкомпьютерных центров научных организаций и вузов Союзного государства. Объединенные вычислительные ресурсы используются участниками программы для решения задач отработки

технологий, разработки программного обеспечения и развертывания приложений.

На втором этапе идет расширение ресурсной составляющей за счет подключения суперкомпьютерных центров и их объединение в грид-инфраструктуру (рис. 1).

В созданное общее вычислительное пространство подключается российский консорциум РДИГ (Российский грид для интенсивных операций с данными - Russian Data Intensive Grid, RDIG). Грид-инфраструктура консорциума создана для интенсивных операций с научными данными. Такая инфраструктура необходима для участия российских и белорусских ученых в экспериментах по физике высоких энергий, химфизике и биологии, в науках о земле и т. д. Консорциум РДИГ входит в структуру EGEE в качестве региональной федерации для обеспечения полномасштабного участия России в этом глобальном грид-проекте.

Важной задачей второго этапа является разработка приложений в грид-среде. Основными направлениями разработки приложений являются:

- физика высоких энергий, включая атомную энергетику;
- машиностроение (моделирование и виртуальные испытания);
- проектирование цифровых устройств;
- материаловедение и взрывная обработка материалов;
- управление тепловыми и энергетическими установками;
- фармацевтика, медицина и биоинформатика и др.

Характерной особенностью данного этапа является освоение грид-технологий и технологий построения суперкомпьютерных систем нового поколения. Это дает основание для создания базовой киберинфраструктуры.

2. Создание базовой киберинфраструктуры, разработка технологий ее использования

Создание базовой киберинфраструктуры подразумевает выполнение следующих видов работ (рис. 2):

- наращивание вычислительных мощностей;
- развитие телекоммуникационной инфраструктуры и в первую очередь научно-образовательной сети;
- объединение в общую грид-инфраструктуру;
- разработка технологий эффективного использования созданной инфраструктуры.

Наращивание вычислительных мощностей выполняется путем создания дополнительных суперкомпьютерных ресурсных центров на основе разработанных в рамках выполнения программы «СКИФ-ГРИД» технологий. Это технологии, созданные в ОИПИ НАН Беларуси по созданию гибридных архитектур кластерных вычислительных систем, технологии построения суперкомпьютерных систем высокой и сверхвысокой производительности на основе серверных решений с жидкостным охлаждением и технологии интерконнета на FPGA.

Дополнительные суперкомпьютерные ресурсы создаются на региональном уровне (рис. 3). При этом предусматривается создание одного-двух суперкомпьютерных центров петафлопсного диапазона. В результате общая вычислительная мощность будет рекордной для Союзного государства.

Развитие телекоммуникационной инфраструктуры и в первую очередь научно-образовательной сети проводится путем ее модернизации, использования технологий уплотнения информации и повышения пропускной способности с достижением значений не менее одного Гбит/с. При этом предусматривается использование космических каналов связи.

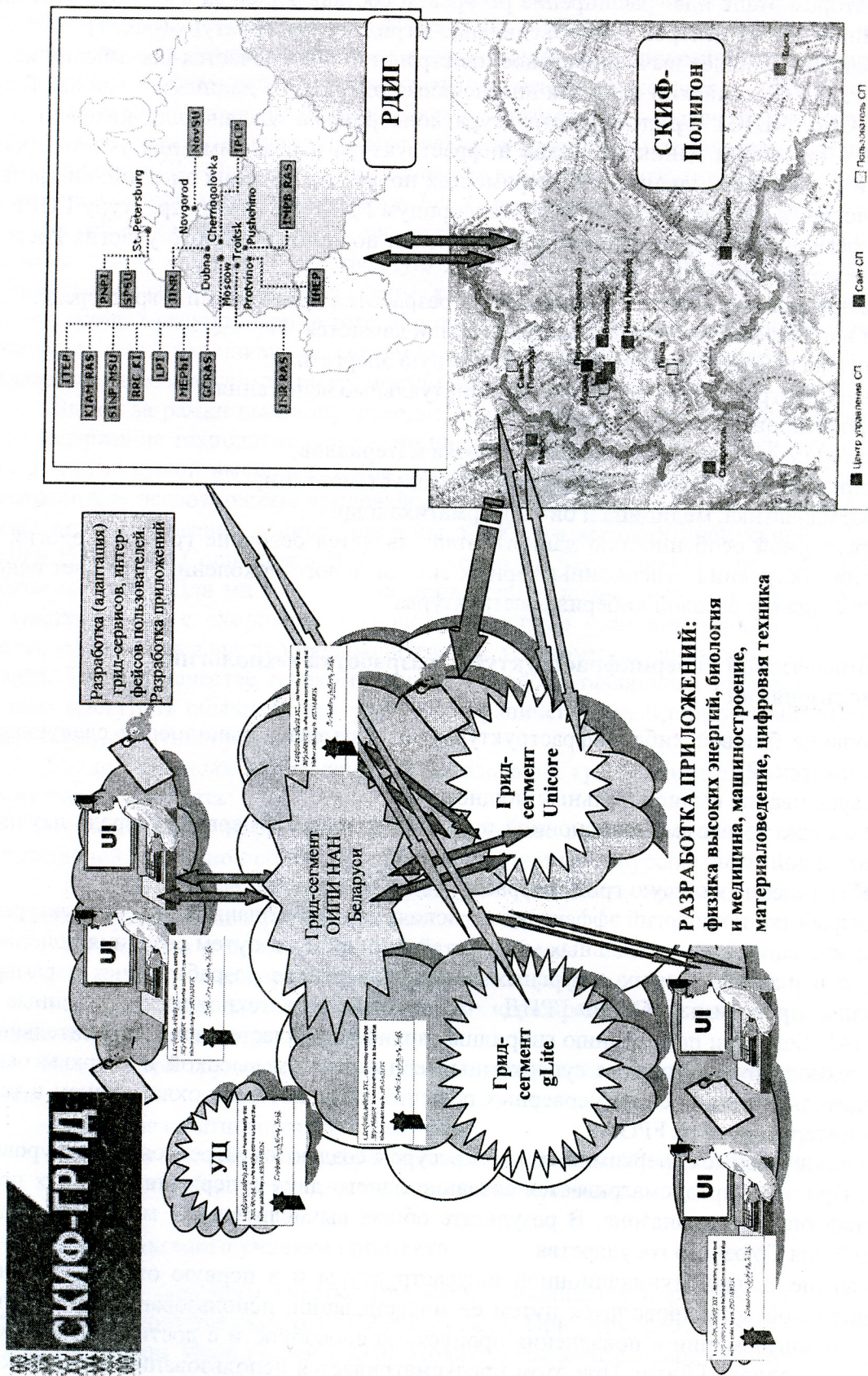


Рис. 1. Опытный участок научно-образовательной грид-сети Союзного государства



Рис. 2. Направления создания базовой киберинфраструктуры

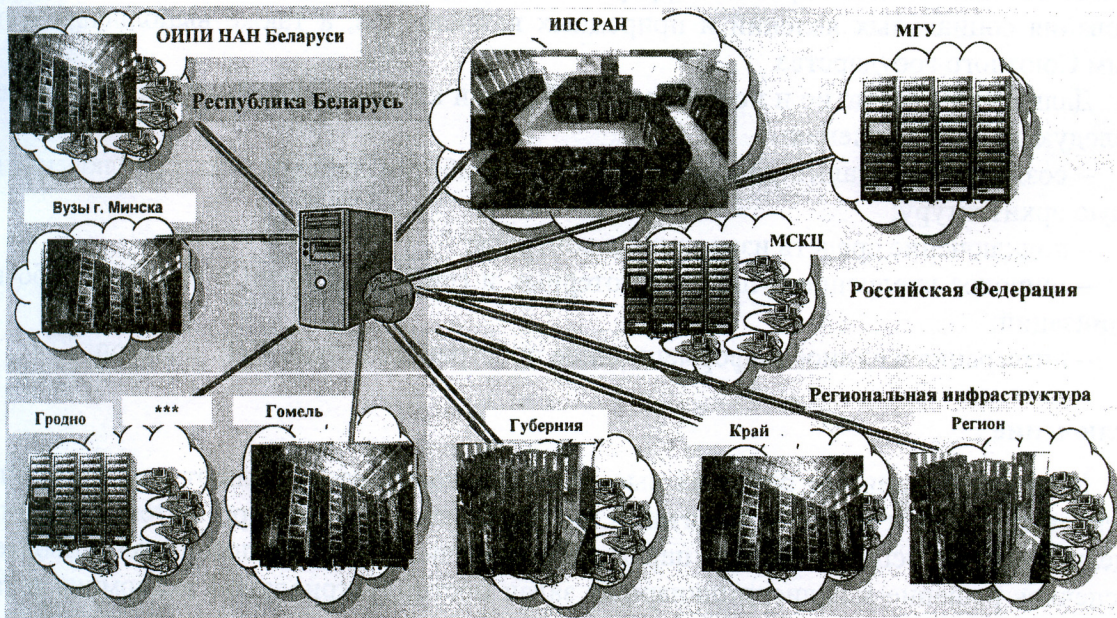


Рис. 3. Наращивание суперкомпьютерных мощностей Союзного государства

На основе созданных ресурсных центров и телекоммуникационной инфраструктуры создается единая вычислительная среда с использованием грид-технологий. При этом предусматривается дальнейшее расширение ее функциональности и разработка сервисов для приложений, однако рост степени интеграции ресурсов для решения сложных наукоемких задач приводит к возникновению ряда отрицательных последствий. Основными из них являются:

- увеличение потребления электроэнергии и, следовательно, увеличение затрат на содержание вычислительных ресурсов;

- проблема их эффективного использования из-за увеличения времени простоя ресурсов;

- усложнение программного обеспечения.

Поэтому немаловажным фактором использования планируемого к созданию единого информационно-вычислительного пространства является разработка технологий ее эффективного использования. Мощности современных компьютеров используются только на 10-15 % (по информации компаний HP и Intel). Это приводит к простоям оборудования организаций и компаний, имеющих свои вычислительные ресурсы.

В настоящий момент в мировой практике рассматриваются два пути решения проблемы: виртуализация и консолидация, что приводит к решениям с меньшим числом более крупных систем.

Эффективное использование такой инфраструктуры может быть реализовано посредством массивного целенаправленного развития такой стратегической области, как обеспечение единой вертикали при построении высокопроизводительных систем для прикладных сервисов, т. е. обеспечение вычислительными сервисными услугами научных и производственных организаций, учреждений социальной сферы. Это в первую очередь сервисные технологии в широком их понимании с позиций стандарта OASIS soa-gm-csru [4], технологий cloud computing.

В результате на втором этапе будет создана технологическая основа для развития стратегических отраслей: нано- и биотехнологии, космических исследований, авиа-, машино- и судостроения, создания образцов военной техники нового поколения, обеспечения безопасности и борьбе с террористической угрозой, прогнозирования и моделирования социальных явлений и природных катаклизмов, а также выполнения программ Союзного государства.

Дальнейшее развитие и использование базовой киберинфраструктуры проводится по следующим направлениям:

- создание специализированных центров обработки данных и их подключение в общую архитектуру;

- подключение средств измерения и контроля;

- широкое внедрение суперкомпьютерных и грид-технологий в практику работы организаций;

- интеграция в единое информационное пространство стран СНГ и Европы.

Заключение

Мировые тенденции развития высоких технологий не оставляют сомнений, что суперкомпьютеры и грид-технологии - это наиболее мощное ядро математического моделирования с использованием параллельных вычислений. Их потенциал позволяет обеспечить решение междисциплинарных задач постиндустриальной экономики и таких фундаментальных проблем, как низкие производительность и эффективность производства. В целях повышения эффективности развития экономики Союзного государства необходимо создавать технологическую основу для развития и использования наукоемких технологий.

Список литературы

1. Govt Supercomputer Upgraded to 54 Teraflops [Electronic resource]. - 2006. - Mode of access : http://www.spacemart.com/reports/Jaguar_Supercomputer. - Date of access: 14.04.2010.

2. Russian Data Intensive Grid [Electronic resource]. - 2006. - Mode of access : www.egee-rdig.ru. - Date of access : 2.04.2010.

3. Reference Model for Service Oriented Architecture [Electronic resource]. - 2006. - Mode of access : <http://soa.skatin.ru/soa-rm-csru.pdf>. - Date of access : 2.04.2010.

4. Начало совещания с членами Совета Безопасности по вопросам создания и применения суперкомпьютеров [Электронный ресурс]. - 2009. - Режим доступа : www.kremlin.ru/text/appears/2009/07/220213.shtml. -Дата доступа : 1.03.2010.