

ИНФОРМАЦИОННАЯ ИНФРАСТРУКТУРА РЕСПУБЛИКИ ТАТАРСТАН: ОТ «SENet-Tatarstan» К «SEGrid-Tatarstan»

INFORMATION INFRASTRUCTURE OF TATARSTAN: FROM «SENet-Tatarstan» TO «SEGrid-Tatarstan»

М.Р. Биктимиров², Е.В. Биряльцев³, Д.Е. Демидов^{1,3}, А.М. Елизаров^{1,3}, Д.В. Чачков¹

M.R. Biktimirov², E.V. Biryaltsev³, D.E. Demidov^{1,3}, A.V. Elizarov^{1,3}, D.V. Chachkov¹

¹ Казанский филиал Межведомственного суперкомпьютерного центра РАН
Kazan Branch of Joint Supercomputer Center of the Russian Academy of Sciences

² Национальная ассоциация исследовательских и научно-образовательных электронных инфраструктур e-АРЕНА
National Association of Research and Educational Electronic Infrastructures E-ARENA

³ Казанский (Приволжский) федеральный университет

Kazan Federal University

Аннотация. В докладе описаны этапы формирования в Республике Татарстан телекоммуникационных, информационных и суперкомпьютерных ресурсов организаций науки, образования и государственного управления. Представлены также результаты в области высокопроизводительных вычислений, полученные к настоящему времени.

Summary. The report describes the formation stages of the Republic of Tatarstan telecommunication, information and supercomputer resources for institutions of science, education and public administration. We also present current scientific results obtained in the field of high performance computing.

Ключевые слова: информационная инфраструктура, информационно-коммуникационные технологии, грид-системы, высокопроизводительные вычисления, квантово-химические расчеты, синтез информационно-вычислительных систем, математическое моделирование, гидродинамика, геофизика, пакеты прикладных программ

Keywords: information infrastructure, information and communication technology, grid system, high performance computing, quantum-chemical calculations, synthesis of information systems, mathematical modeling, fluid dynamics, geophysics, application packages

1. Эволюция информационной инфраструктуры Республики Татарстан. История формирования информационной инфраструктуры Республики Татарстан (РТ) отражает наблюдаемое в мире нарастающее проникновение информационно-коммуникационных технологий (ИКТ) и компьютерных систем в научные исследования, образовательную деятельность и государственное управление. Общность целей и задач внедрения ИКТ позволила организовать в РТ многостороннее сотрудничество научных организаций, вузов и органов государственной власти, определяющим фактором которого стали государственная поддержка на федеральном и республиканском уровнях и реализация совместных проектов учреждений бюджетной сферы с республиканскими органами государственной власти. Результатом стала успешная реализация межведомственных проектов по созданию и последующей реконструкции телекоммуникационной инфраструктуры, охватившей большинство вузов Казани и институтов Казанского научного центра (КазНЦ) РАН и Академии наук Республики Татарстан (АН РТ), а также целый ряд других научных организаций, государственных и бюджетных учреждений. Единая сетевая инфраструктура стала базой совместных проектов в различных сферах, обеспечила внедрение прикладных информационных систем. Положительный опыт выполнения совместных проектов закреплён в принципах деятельности компьютерной сети науки и высшего образования Республики Татарстан «SENet-Tatarstan», которая обеспечивает полнофункциональное взаимодействие с российскими научно-образовательными сетями и международной сетью GEANT и соответствует всем организационным

принципам и техническим требованиям этих сетей.

Как известно, для России временем активных работ по созданию развитой инфраструктуры, основанной на телекоммуникационных технологиях, были 1990-е годы. На территории РТ в это время функционировал ряд телекоммуникационных сетей (см. [1]). Первые коммерческие сети были организованы в 1993 – 1994 гг., а некоммерческие сети, ориентированные на обслуживание организаций науки, образования, государственного управления, здравоохранения и культуры, были представлены несколькими ведомственными сетями, а также Гражданской сетью (ГС) РТ и сетью «SENet-Tatarstan».

Гражданская сеть РТ, созданная в 1994 году на базе Казанского государственного университета (КГУ, ныне – Казанский федеральный университет, КФУ) совместно с КазНЦ РАН и рядом вузов Казани, была крупной телекоммуникационной системой регионального масштаба, объединявшей более 130 организаций. Организационной и финансовой основой ГС был проект федеральной программы [2], реализованный на средства Министерства науки РФ, Российского фонда фундаментальных исследований (РФФИ) при финансовой поддержке Правительства РТ. Фактически ГС объединила на общей технической базе компьютерные сети практически всех вузов Казани и Набережных Челнов, институтов КазНЦ РАН и АН РТ, республиканских больниц и клиник, организаций государственного управления, культуры и многих других некоммерческих организаций. До появления ГС в Республике практически не было школ, библиотек, организаций здравоохранения, имевших доступ в глобальные сети.

Развитие научных телекоммуникаций и информационной инфраструктуры в РТ на рубеже веков характеризовалось мощными интеграционными тенденциями, наиболее ярким выражением которых стало создание в 2000 году сети «SENet-Tatarstan», которая объединила институты КазНЦ РАН, АН РТ, вузы Казани и была построена КазНЦ РАН на средства РАН и фонда НИОКР РТ. При финансовой поддержке РАН, РФФИ и Федеральной целевой программы (ФЦП) «Интеграция» в 2001 г. был реализован проект перевода коммуникационной основы SENet на оптоволоконные соединения. Функционирование ГС и SENet было скоординировано.

Созданные в рамках ГС и SENet информационные ресурсы позволили РТ войти со своим вкладом в мировое информационное пространство. Ярким достижением здесь стало создание в 2000 – 2001 гг. официального сервера РТ (<http://www.tatar.ru>), содержащего весьма представительные материалы о Татарстане на русском, татарском и английском языках и ставшего самым популярным информационным ресурсом Татарстана.

В течение последнего десятилетия в РТ выполнены крупные телекоммуникационные и информационные проекты, финансируемые РФФИ, РАН и государственными программами РТ: по созданию корпоративной библиотечной сети Казани и электронных библиотек; построения магистральной АТМ-сети КазНЦ РАН; корпоративной сети передачи данных органов государственной власти и управления РТ; создания информационного Центра коллективного пользования КазНЦ РАН и действующего на его базе «зеркала» крупнейшей в России научной электронной библиотеки РФФИ; проекты республиканских межведомственных программ «Создание единой информационной научно-образовательной инфраструктуры Республики Татарстан», «Электронный Татарстан» и другие.

2. Развитие высокопроизводительных вычислений. Создание в 2000 году в КазНЦ РАН в рамках ФЦП «Интеграция» первого 10-процессорного вычислительного кластера на базе рабочих станций Compaq Alpha положило начало внедрению высокопроизводительных вычислений в научные и прикладные исследования, проводимые институтами КазНЦ РАН и вузами Казани. Тестовые квантово-химические расчеты, выполненные нами на базе общедоступных бесплатных программ (GAMESS, NWChem и др.), подтвердили большие возможности использования высокопроизводительных кла-

стеров для решения широкого круга научно-технических задач, в первую очередь, за счет сокращения времени вычислений. Отметим, что и сейчас вычислительные кластеры наиболее востребованы в Казани специалистами в области квантовой химии. С целью создания наиболее благоприятных условий для проведения научных исследований с использованием современных методов квантовой химии и подготовки соответствующих специалистов в КГУ и Казанском государственном технологическом университете (ныне Национальном исследовательском технологическом университете, КНИТУ) в 2001 году были организованы вычислительные кластеры на базе учебных классов, что обеспечило возможность проведения расчетов на кластере КазНЦ РАН посредством удаленного доступа по сети «SENet-Tatarstan» с высокой скоростью обмена данными.

В 2007 году был создан Суперкомпьютерный центр коллективного пользования КазНЦ РАН, основным вычислительным ресурсом которого стал вычислительный кластер MBC-15000 с девятью базовыми блоками IBM BladeCenter и двухпроцессорными платами JS20, коммуникационной средой Myrinet и Gigabit Ethernet с пиковой производительностью 2,25 Тфлопс. Соответствующее системное и прикладное программное обеспечение (ПО) функционировало под управлением ОС Linux и обеспечивало полный технологический цикл: организацию и управление проведением массовых расчетов посредством системы очередей в многозадачном и многопользовательском режимах; организацию распределенных вычислений с использованием грид-технологий на базе Globus ToolKit и ССРВ; обеспечение защищенного удаленного доступа пользователей к вычислительным ресурсам по сетям общего пользования; сбор данных об использовании вычислительных ресурсов; управление и мониторинг технического состояния компьютерного оборудования и инженерной инфраструктуры; информирование пользователей о подключении к кластеру и установленном ПО, технологиях запуска расчетных задач и др.. В течение 2007 – 2009 гг. ежегодно выполнялось около 10 тыс. расчетных заданий общей продолжительностью 16 – 19 тыс. процессорных дней. Основной объем расчетов приходился на квантовохимические расчеты [3], в которых использовалось свободно распространяемое и лицензионное ПО, в том числе, пакеты Priroda, GAMESS, Gaussian.

Совместное использование вычислительных ресурсов и обмен опытом по организации распределенных вычислений лежат в основе тесного сотрудничества КазНЦ РАН и вузов Казани. В рамках SENet действует Соглашение «О сотрудничестве в области обеспечения высокопроизводительными вычислительными ресурсами научных исследований и учебного процесса», работает городской семинар по высокопроизводительным вычислениям на базе КГУ. В 2009 году к этой работе присоединился Казанский государственный технический университет (ныне Национальный исследовательский технический университет, КНИТУ-КАИ), запустивший вычислительный кластер пиковой производительностью до 2 Тфлопс.

В деятельности по внедрению высокопроизводительных расчетов в научные исследования и развитию аппаратного комплекса КазНЦ РАН тесно сотрудничает с ведущей научной организацией РАН в области суперкомпьютинга – Межведомственным суперкомпьютерным центром (МСЦ) РАН. В 2009 году на базе КазНЦ РАН был создан Казанский Филиал (КазФ) МСЦ РАН, который продолжил начатые работы, значительно расширил тематику суперкомпьютерных приложений, а также вычислительную базу научных исследований в Татарстане: в 2010 году в IT-парке РТ установлен суперкомпьютер MBC-100КФ пиковой производительностью 10 терафлопс, получивший имя «Сююмбике» и ориентированный на функционирование в составе распределенной инфраструктуры суперкомпьютерных приложений РФ; в этом же году КазФ МСЦ РАН запустил в эксплуатацию вычислительный кластер на базе Microsoft Windows Server 2008 R2 edition, предназначенный для обеспечения вычислительными ресурсами разрабо-

ток в среде Microsoft Windows. К настоящему моменту суммарная пиковая производительность вычислительных ресурсов КазФ МСЦ РАН достигла 20 терафлопс, расширились их функциональность и, соответственно, спектр прикладных задач. Все это позволяет успешно решать следующие задачи: квантово-химическое изучение механизмов химических реакций; исследование стабильности и функционирования молекулярных наноустройств на основе ультрананокристаллического алмаза; разработка параллельных алгоритмов и программ решения задач в области аэрогидродинамики, теории упругости; применение вычислительных систем с массовым параллелизмом при анализа естественного микросейсмического поля и разведке залежей природных углеводородов; решение задач механики набухающих систем в медицинской физике и биомеханике.

Перспективным направлением развития высокопроизводительных вычислений в РТ является присоединение республиканских суперкомпьютерных ресурсов к научно-образовательным грид-инфраструктурам, создаваемым в РФ, что позволяет повысить эффективность использования компьютерного оборудования, увеличить на порядки объем вычислительных ресурсов и расширить доступ ученых к уникальному прикладному ПО в рамках совместных исследований с ведущими исследовательскими центрами России. Одним из первых проектов в этом направлении, реализованных КазНЦ РАН в 2008 году, стало подключение имеющегося вычислительного комплекса к распределенной инфраструктуре для суперкомпьютерных приложений РАН [4], работы по созданию которой осуществлялись МСЦ РАН. Сегодня эта инфраструктура объединяет ряд вычислительных систем РАН в городах Москва, Казань и Санкт-Петербург. В результате пользователи получили доступ по привычному им интерфейсу к распределенным вычислительным ресурсам РАН, суммарная производительность которых превышает 100 Терафлопс.

КазФ МСЦ РАН принял участие в проекте ГридННС [5], реализуемом в рамках ФЦП «Развитие инфраструктуры наноиндустрии в Российской Федерации на 2008 – 2010 годы». Цель проекта – обеспечение географически распределенных научных и инженерных коллективов-участников Национальной нанотехнологической сети (ННС) возможностью эффективного удаленного использования информационной, коммуникационной и вычислительной инфраструктуры. В рамках этой инфраструктуры КазФ МСЦ РАН совместно с ВЦ ДВО РАН провел ряд мероприятий: создание виртуальной организации для расчетов на базе популярного квантово-химического программного пакета AVINIT; синхронизация локальных систем управления очередями расчетных заданий между собой и грид-сервисами; согласование систем предоставления доступа пользователей к вычислительным системам; использование ГридННС в интересах территориально распределенных исследовательских групп. Дополнительным аргументом к созданию распределенных инфраструктур является возможность проведения компьютерного моделирования сложных объектов с высокой точностью на суперкомпьютерах, уникальных по своей архитектуре. Практическим примером могут служить исследования структуры пленок ультрананокристаллического алмаза (УНКА) в КазФ МСЦ РАН (см., например, [6]). Использование программного пакета квантово-химических расчетов CPMD (Car-Parinello Molecular Dynamics) фирмы IBM и суперкомпьютера факультета ВМК МГУ позволило проанализировать стабильность УНКА, содержащих более 1500 атомов, с точностью, беспрецедентной для систем такого размера. При расчетах структур УНКА на универсальных вычислительных кластерах, составляющих основную часть грид-систем, время расчетов увеличивалось в несколько раз.

Другим ярким примером является изучение механизмов химических реакций, т. е. всех элементарных стадий от исходного соединения до продуктов реакции [7, 8]. Без применения высокопроизводительных вычислительных кластеров данное исследование могло растянуться на несколько лет, а за счет использования вычислительных кластеров

(при достаточной квалификации исследователя) заняло лишь несколько недель.

Еще одним направлением применения высокопроизводительных компьютеров является их использование в интересах повышения нефтедобычи в Республике Татарстан, а также в других нефтедобывающих регионах РФ. Сегодня это обусловлено, в первую очередь, тем, что большие месторождения РТ уже длительное время находятся в эксплуатации, добыча нефти из них с использованием только традиционных технологий падает, а расходы на амортизацию инфраструктуры растут. Применение новых технологий сопровождается (особенно в условиях отраслевого и бюрократического лоббирования) большой долей неопределенности и, следовательно, риска при использовании новых методов непосредственно на месторождении, а лабораторная экспертиза перспективных инноваций зачастую не может быть расценена как достоверная из-за непереносимости результатов лабораторных экспериментов на реальный геологический объект. Существенную роль здесь может сыграть так называемое 3D-моделирование процессов нефтеизвлечения, опирающееся на созданные теоретические модели многофазной многокомпонентной фильтрации и реализуемое на современных высокопроизводительных компьютерных системах. Масштабное его использование позволяет сократить число дорогостоящих натуральных испытаний технологических новшеств и тем самым поднять рентабельность отрасли в целом.

Другим направлением применения суперкомпьютеров в нефтяной отрасли РТ является их широкое использование в интерпретации данных 3D-сейсмоки, а также других методов волнового исследования структуры земных недр. Детальное моделирование расположения и структуры продуктивных пластов с использованием натуральных данных позволит минимизировать затраты на разбуривание мелких месторождений, которые ранее относились к неперспективным, и ввести их в разряд эксплуатируемых.

Таким образом, развитие сложившейся в Татарстане инфраструктуры сети «SENet-Tatarstan», востребованность ее участниками новых типов сетевых сервисов, растущая заинтересованность научных учреждений и вузов РТ в использовании суперкомпьютерных технологий стали определяющим фактором в переходе на новый этап региональной кооперации по созданию современной ИКТ-инфраструктуры. Основной целью данного этапа является построение в РТ электронной грид-инфраструктуры, интегрированной в общероссийскую грид-сеть, для решения в распределенной вычислительной среде прикладных заданий в различных научных, инженерных и технологических областях, которые связаны с ресурсоемкими высокопроизводительными расчетами, а также обмена, обработки и доступа к информации. Поставленные задачи тесно связаны с организационно-технической грид-структурой, расширением сотрудничества организаций по формированию кадрового потенциала, разработкой прикладного и системного ПО, интеграцией региональных высокопроизводительных систем с общероссийскими инфраструктурами. Нацеленным на решение этих задач является проект создания на базе инфраструктуры сети «SENet-Tatrstan» грид-сети «SEGrid-Tatrstan». С учетом сложности поставленных задач и практического задела, имеющегося у участников, в проекте SEGrid определены центры компетенции по различным направлениям деятельности, в частности, ведущие университеты ориентированы: КФУ – на обучение в области суперкомпьютерных технологий, КНИТУ-КАИ – на разработку прикладного ПО, КНИТУ и КазФ МСЦ РАН – на внедрение грид-сервисов и взаимодействие с общероссийскими инфраструктурами. Участники SEGrid получают возможность удаленного выполнения высокопроизводительных расчетов и обработки данных на широком спектре географически распределенных вычислительных ресурсов, от уникальных суперкомпьютеров до средних и малых суперкомпьютерных кластеров и отдельных серверов, за счет эффективного, прозрачного и безопасного доступа к ресурсам грид-сети.

Проект SEGrid и его инновационная составляющая получили поддержку со сто-

роны Правительства Татарстана: в крупнейшем в Поволжье IT-парке размещен суперкомпьютер «Сююмбике» и предоставлены площади для размещения узла общероссийской инфраструктуры суперкомпьютерных приложений, работы по созданию которой осуществляются в рамках деятельности Национальной ассоциации исследовательских и научно-образовательных электронных инфраструктур «е-Арена».

3. Современные подходы к синтезу сложных информационно-вычислительных систем. При рассмотрении архитектуры информационно-вычислительных систем (ИВС) массового пользования становится очевидной тенденция перемещения клиентских или пользовательских данных в области удаленного хранения. В первую очередь это связано с виртуализацией вычислений и общей направленностью на разделение программ и данных ИВС. Такая организация работы с данными позволяет не только снизить затраты на их хранение, но и обеспечить существенное повышение надежности, а также снизить угрозы нарушения конфиденциальности.

Отметим, что значительное внимание в новых технических решениях и патентах уделяется поддержанию работоспособности ИВС, а также обеспечению их управляемости, осуществлению аудита и оптимизации информационных процессов. В связи с этим в научной и методической литературе последних лет наряду с терминами «безопасность» все чаще встречаются термины «надежность» и «управляемость», а также «доверие» и «доверенные системы» [9]. Использование термина «доверие» в применении к ИВС означает переход от обеспечения безопасности компонент системы к методологии обеспечения выполнимости целевой функции всей системы на протяжении ее жизненного цикла (ЖЦ). Таким образом, свойство доверенности является проекцией понятия защищенности на более широкий класс объектов и понятий. С этой точки зрения ИВС должна четко соответствовать той целевой функции, для реализации которой она создается. В этом случае естественным является рассмотрение всего ее жизненного цикла, начиная с формулирования корректной, непротиворечивой целевой функции и рассмотрения архитектуры как целостной совокупности компонент ИВС и связей между ними, описанной в терминах системного анализа.

Кроме типовых угроз безопасности, характерных для всех автоматизированных систем, включающих нарушения свойств доступности, конфиденциальности и целостности [10], для доверенных систем возникает ряд специфических угроз, определяемых тем, что потенциально нарушаются свойства защищенности на определенных этапах их ЖЦ, причем нарушение свойств доверия на одном из этапов ЖЦ может приводить к нарушению указанных свойств на следующем этапе. Поэтому этапы ЖЦ ИВС также должны обладать атрибутом доверенности: доверенные проектирование, разработка, реализация, эксплуатация, модернизация, развитие и вывод из эксплуатации. Определив доверенную среду как область существования и функционирования компонент ИВС, обладающих свойством доверенности, выделим следующие основные компоненты: доверенное коммуникационное оборудование и доверенная среда передачи данных; доверенный вычислительный комплекс на основе доверенной операционной среды; доверенная система хранения, доверенная система управления.

Для синтеза защищенных ИВС необходимо сформулировать требования к каждому доверенному компоненту и этапу жизненного цикла таких систем, а также обеспечить их соблюдение. Можно сформулировать группы требований к современной ИВС: доверенность на всех или большинстве этапов жизненного цикла, нормативно-методическая обеспеченность процессов ЖЦ; высокая техническая надежность и доступность ресурсов ИВС, резервирование и катастрофоустойчивость, удаленное надежное хранение данных; расширяемость и масштабируемость с сохранением свойств доверия; стабильность предоставления сервисов, изоляция пользователей от данных, замкнутая индивидуальная среда; аудит и подконтрольность ресурсов ИВС и вычисли-

тельного процесса в целом; управляемость системы, переход к управлению безопасностью через управление пользовательскими процессами (задачами).

Эволюционный переход к созданию доверенной среды ИВС с использованием изложенных подходов дает следующие стратегические преимущества:

- существенно снижаются системные риски для ИВС большого масштаба;
- поэтапно повышаются текущие уровни надежности и доступности системы, для которой усиливаются свойства доверия;
- создаются высокие потенциалы для собственной технической поддержки и развития, снижаются технологические риски;
- снижается совокупная стоимость владения компонентами ИВС за счет уменьшения затрат на техническую поддержку и модернизацию;
- оптимизируется архитектура ИВС за счет консолидации инфраструктуры вокруг доверенных компонент среды;
- минимизируются затраты на реализацию подсистем обеспечения информационной безопасности за счет использования штатных средств доверенных компонент и снижения затрат на сертификацию, аттестацию и проверку корректности встраивания;
- минимизируются затраты на систему управления и повышается управляемость и мобильность ИВС.

4. Использование и развитие технологии NVIDIA CUDA

4.1. Модификация алгебраического многосеточного метода для эффективного решения нестационарных задач подземной гидродинамики с применением технологии NVIDIA CUDA

Многосеточные методы (ММ) считаются оптимальными для решения эллиптических дифференциальных уравнений и базируются на использовании сеточной иерархии и операторов перехода от одной сетки к другой (см., например, [11]). Основная идея заключается в ускорении сходимости итерационного процесса, лежащего в основе ММ, за счет коррекции решения, полученного на точной сетке, с помощью решения уравнения на грубой сетке.

Геометрический многосеточный метод (ГММ) основан на предопределенной сеточной иерархии. При решении сеточного уравнения $Au=f$ на каждом уровне иерархии задаются сетка Ω_l , оператор задачи A_l , оператор продолжения $P_l: u_{l+1} \rightarrow u_l$ и оператор сужения $R_l: u_l \rightarrow u_{l+1}$. На каждом уровне l иерархии один шаг (V-цикл) геометрического многосеточного метода выполняет следующие действия:

- текущее решение сглаживается несколькими шагами релаксации;
- вычисляется невязка $r_l=f_l-A_l u_l$, которая затем сужается в правую часть более грубого уровня: $f_{l+1}=R_l r_l$;
- если уровень $l+1$ – последний в иерархии, то уравнение $A_{l+1} u_{l+1}=f_{l+1}$ решается прямым методом; иначе на уровне $l+1$ рекурсивно выполняется один шаг многосеточного метода с нулевым начальным приближением $u_{l+1}=0$;
- решение на уровне l корректируется: $u_l \rightarrow u_l + P_l u_{l+1}$;
- выполняется несколько шагов релаксации.

Алгебраический многосеточный метод (АММ) [12] является одним из наиболее эффективных методов решения разреженных неструктурированных систем линейных уравнений большой размерности. В отличие от ГММ он не требует постановки задачи на сетке, а работает непосредственно с разреженной системой линейных алгебраических уравнений $Au=f$. Алгоритм АММ состоит из двух основных этапов: настройка метода и решение. Настройка заключается в конструировании проблемно-зависимой иерархии и включает выбор вложенных подмножеств $\Omega_{l+1} \subset \Omega_l$ исходных переменных Ω_0 ,

построение операторов продолжения $P_l: \Omega_{l+1} \rightarrow \Omega_l$ и операторов на грубых сетках $A_{l+1} = P_l^T A_l P_l$. Эта часть алгоритма АММ основана только на алгебраической информации относительно матрицы A и не зависит от правой части уравнения. Второй этап алгоритма состоит в решении уравнения $Au = f$ для заданной правой части. Если это уравнение нужно решить для нескольких правых частей, то настройку достаточно выполнить один раз. Алгоритм этапа решения АММ практически совпадает с алгоритмом ГММ.

Этап настройки в классическом варианте АММ довольно трудно распараллелить, поэтому в предлагаемом подходе он полностью выполняется на CPU. Этап решения достаточно просто реализовать на параллельной архитектуре, причем использование технологии NVIDIA CUDA на этом этапе позволяет достичь почти 10-кратного ускорения [13, 14]. Так как последовательная часть алгоритма (этап настройки) при реализации на последовательной вычислительной архитектуре занимает примерно 30 % общего времени решения, то в соответствии с законом Амдала метод в целом не может быть ускорен более чем в три раза. Это ограничение можно смягчить при решении нестационарных задач с неизменной матрицей, так как настройку при этом достаточно выполнить один раз. Однако на практике часто встречаются задачи, в которых матрица системы изменяется на каждом шаге по времени. Например, при моделировании разработки нефтяного месторождения на каждом временном шаге изменяются режимы работы и число активных скважин. В этом случае метод приходится перенастраивать на каждом временном шаге, что приводит к резкому увеличению общего времени работы.

Предложенная модификация. На практике АММ чаще всего используется не сам по себе, а в качестве предобуславливателя для одного из методов для подпространства Крылова (например, метода сопряженных градиентов) [15]. Это позволяет повысить робастность метода и уменьшить число итераций, необходимых для сходимости. В применении к проблеме, рассматриваемой нами, это также позволяет сэкономить на настройке многосеточного метода.

Допустим, что на определенном временном шаге для системы $Au = f$ получен настроенный многосеточный метод. На следующем временном шаге коэффициенты задачи изменяются: требуется решить систему $A'u = f'$. Можно надеяться, что многосеточный метод, настроенный для матрицы A , будет являться хорошим предобуславливателем и для матрицы A' (напомним, что этап настройки не зависит от правой части уравнения). На этом предположении и основан предложенный нами алгоритм решения нестационарной задачи:

- на первом временном шаге выполняются этапы настройки и решения;
- на каждом последующем временном шаге: 1) если этап решения на предыдущем шаге потребовал числа V -циклов, большего, чем заданное пороговое значение N_{cr} , то выполняется перестройка многосеточного метода; иначе используется уже настроенный метод; 2) выполняется этап решения АММ.

Значение N_{cr} в общем случае должно зависеть от скорости сходимости конкретной задачи. Нами использовалось значение $N_{cr} = 20$.

При использовании описанного алгоритма этап настройки будет выполнен в худшем случае на каждом временном шаге. Более вероятно, что в большинстве случаев многосеточный метод, настроенный на одном из предыдущих шагов, будет работать удовлетворительно. Так как при использовании графического процессора на этапе решения этап настройки может занимать до 90 % общего времени решения, можно ожидать значительной экономии расчетного времени.

Сравнение с классическим вариантом. Эффективность предложенного метода была проверена на примере четырех реальных нефтяных месторождений. В таблице 1 приведены их характеристики, для каждого указаны число активных ячеек расчетной

области, участвующих в расчете (т. е. число неизвестных в алгебраической системе уравнений), число скважин, моделируемый (расчетный) период разработки в месяцах и число временных слоев. Расчеты проводились по схеме IMPES. В таблице 2 указано время, затраченное на получение поля давления, при использовании центрального и графического процессоров, а также в случае полного перестроения АММ на каждом временном слое. В расчетах использовались CPU Intel Core i7 920 и GPU NVIDIA Tesla C2070. Все расчеты проводились в двойной точности.

Таблица 1. Модели месторождений, использовавшиеся в численных экспериментах

Модель	Число ячеек	Число скважин	Расчетный период (мес)	Число временных шагов
MD	922 298	451	291	99
MI	649 316	105	168	57
MK	591 080	479	384	386
MM	70 806	1 534	336	114

Как видно, для всех моделей перестройка многосеточного метода потребовалась не более чем в 50 % случаев. Заметим, что при расчете на CPU предложенная методика приводит к увеличению общего времени расчета. Это объясняется тем, что среднее число V-циклов, необходимых для решения задачи на каждом временном слое, возрастает (т. к. в качестве предобуславливателя используется многосеточный метод, настроенный на одном из предыдущих шагов и потому неоптимальный).

Применение технологии NVIDIA CUDA с перестроением метода на каждом временном слое позволяет ускорить расчет примерно вдвое. Как видно, этап настройки при этом начинает занимать до 75 % времени. Применение предложенной нами методики расчета позволяет сократить долю настройки до 40 %, что приводит к трехкратному ускорению по сравнению с расчетом на CPU.

Заметим, что описанные задачи решались в самом «критичном» (с точки зрения изменения матрицы) случае – каждый шаг расчета поля давления соответствовал ежемесячной смене данных по технологическим характеристикам скважин, в частности, изменению интервалов перфорации. При таких изменениях значительно изменяются (как количественно, так и структурно) уравнения системы, соответствующие скважинам. Это часто приводит к заметному изменению структуры связей в матрице и необходимости перенастройки АММ. Особенно сильно это заметно на месторождении MM, которое имеет наибольшее число скважин по отношению к размеру расчетной сетки. Структура матрицы при этом изменяется настолько сильно, что предложенная методика приводит к значительному увеличению числа V-циклов и замедлению счета. Однако при решении большого количества задач подземной гидродинамики возникает необходимость пересчета поля давления при неизменной структуре связей скважина – пласт. При этом изменяются только количественные характеристики таких связей. В этом случае перенастройка АММ должна требоваться гораздо реже, и предложенная методика должна привести к гораздо большим выигрышам по времени.

Таблица 2. Расчетное время для классического варианта АММ (с) и модифицированного алгоритма (m)

Алгоритм, архитектура	Число шагов	Ср. число V-циклов	Настройка (с)	Решение (с)	Сборка матрицы (с)	Общее время (с)
Модель MD						
CPU (с)	99 (100%)	16	83.1 (19%)	352.9 (80%)	6.3 (1%)	442.6
CPU (m)	43 (43%)	30	34.2 (6%)	529.5 (93%)	5.5 (1%)	569.6

Алгоритм, архитектура	Число шагов	Ср. число V-циклов	Настройка (с)	Решение (с)	Сборка матрицы (с)	Общее время (с)
GPU (с)	99 (100%)	16	114.6 (74%)	35.4 (23%)	5.4 (3%)	155.7
GPU (m)	44 (44%)	30	50.6 (40%)	69.3 (55%)	5.2 (4%)	125.5
Модель МI						
CPU (с)	57 (100%)	14	30.1 (24%)	90.4 (74%)	2.3 (2%)	122.9
CPU (m)	21 (37%)	19	10.6 (8%)	120.9 (90%)	2.4 (2%)	134.0
GPU (с)	57 (100%)	15	47.5 (74%)	14.3 (22%)	2.4 (4%)	64.4
GPU (m)	21 (37%)	19	18.7 (44%)	21.4 (50%)	2.7 (6%)	42.9
Модель МК						
CPU (с)	386 (100%)	16	210.6 (25%)	621.7 (73%)	13.4 (2%)	846.7
CPU (m)	164 (42%)	24	88.5 (8%)	981.5 (90%)	13.4 (1%)	1084.6
GPU (с)	386 (100%)	16	308.8 (73%)	99.8 (24%)	14.2 (3%)	423.8
GPU (m)	167 (43%)	24	133.6 (42%)	168.0 (53%)	13.7 (4%)	316.4
Модель ММ						
CPU (с)	114 (100%)	12	7.2 (28%)	18.0 (69%)	0.6 (2%)	25.9
CPU (m)	97 (85%)	90	5.0 (4%)	118.3 (95%)	0.5 (0%)	123.9
GPU (с)	114 (100%)	12	11.0 (54%)	8.7 (43%)	0.5 (2%)	20.2
GPU (m)	97 (85%)	90	9.5 (13%)	65.3 (87%)	0.5 (1%)	75.3

Отметим также, что предложенная модификация решения нестационарных задач алгебраическим многосеточным методом с применением технологии NVIDIA CUDA позволила существенно снизить время расчета за счет уменьшения числа перестроений многосеточного метода. Модификация проста в реализации и позволяет сэкономить до 50 % расчетного времени по сравнению с классической версией АММ. Ожидается, что для многих других задач механики, в которых изменения матрицы носят монотонный характер без структурных изменений, применение данной методики сможет привести к еще большему выигрышу по времени. Далее, описанная модификация не эффективна при решении только на центральном процессоре или при решении задач с быстро меняющимися коэффициентами. В первом случае V-циклы намного дороже перестройки метода; во втором случае резко возрастающее число V-циклов перевешивает экономию, достигнутую на настройке многосеточного метода.

4.2. Моделирование распространения микросейсмических волн на процессорах NVIDIA. Как известно, стоимость разведки и добычи углеводородного сырья неуклонно возрастает, а число месторождений, простых в геологическом строении, следовательно, легко обнаруживаемых традиционными методами геологоразведки, становится все меньше. Сейчас поисковыми объектами являются небольшие месторождения – спутники ранее открытых крупных месторождений и месторождения в труднодоступных районах планеты. Такая ситуация заставляет искать новые методы геологоразведки, эффективные и одновременно малозатратные.

Низкочастотное сейсмическое зондирование (НСЗ), основанное на анализе поля естественных микросейсм, – одна из наиболее эффективных разработок в этой области. Описанный ниже метод – совместная разработка ЗАО «Градиент», НИИ математики и механики им. Н.Г. Чеботарева КФУ и КазФ МСЦ РАН. Он позволяет отказаться от применения опасных взрывотехнических работ или громоздких вибраторов и резко снизить стоимость полевых работ.

Одной из сложностей метода НСЗ является необходимость производить многовариантные ресурсоемкие расчеты. Для интерпретации данных нужно сопоставить заре-

гистрированные микросейсмические модели с модельными микросейсмическими, полученными численным моделированием для различного положения залежи углеводородов в геологической среде. При этом оптимальный объем вычислений при численном моделировании таков, что зачастую дешевле пробурить скважину на исследуемом участке, чем рассчитать необходимый набор вариантов на суперкомпьютере. Именно поэтому было принято решение перенести созданный ранее программный комплекс на GPU. Моделирование распространения микросейсмических волн на процессорах NVIDIA Tesla позволило существенно оптимизировать затраты на проведение исследований, а использование массивно-параллельной архитектуры GPU позволило резко снизить время и стоимость компьютерного моделирования и начать активно применять программный комплекс при решении поисковых задач.

При расчете типовой модели (около 8 млн. узлов, 100 тыс. шагов по времени) один графический процессор оказался сопоставим по вычислительной мощности с 48 процессорами суперкомпьютера МВС-100К (пиковая производительность 120 ТФлопс). Это позволило добиться уменьшения времени вычислений от 50 (1 GPU) до 250 раз (6 GPU), а также снизить стоимость вычислений до 1000 раз [16].

За последние 5 лет ЗАО «Градиент» провело поисково-разведочные работы методом НСЗ на более чем 100 площадях в Урало-Поволжье и Западной Сибири, на которых пробурено в общей сложности 95 поисково-разведочных и эксплуатационных скважин. Высокая эффективность метода НСЗ и оптимальная аппаратная поддержка процессов моделирования и интерпретации данных позволили достичь успешности прогноза более 80%.

4.3. VexCL: библиотека шаблонов C++ для разработки OpenCL-приложений. VexCL – это библиотека шаблонов векторных выражений с открытым исходным кодом, разработанная в КазФ МСЦ РАН [17]. Предназначение библиотеки – упрощение разработки C++-приложений, использующих технологию OpenCL для решения вычислительных задач с высокой степенью параллелизма. Библиотека поддерживает удобную и интуитивную нотацию для вычисления векторных выражений и позволяет прозрачно использовать несколько арифметических ускорителей [18, 19].

Научные результаты, представленные выше, получены при частичной финансовой поддержке РФФИ, проект 12-07-00007.

Литература

1. Электронная почта и доступ к Internet в Татарстане, 96. Справочно-информационный сборник материалов под редакцией Г.В. Васильева. – Казань: Изд-во «Книга Памяти» Республики Татарстан, 1996. – 96 с.
2. Межведомственная программа создания национальной сети компьютерных телекоммуникаций для науки и высшей школы. 1995-1998. Министерство науки и технической политики России, 1995. – 50 с.
3. Чачков Д.В., Астафьев М.Н., Биктимиров М.Р., Туриянский Е.А., Шамов Г.А., Шамов А.Г., Храпковский Г.М. Квантово-химические расчеты термического разложения нитросоединений на базе вычислительного кластера Казанского научного центра РАН // 8-я междунар. науч. конф. «Высокопроизводительные параллельные вычисления на кластерных системах» (НПС-2008). Сборник трудов. – Казань, 2008. – С. 341-346.
4. Савин Г.И., Шабанов Б.М., Корнеев В.В., Телегин П.Н., Семенов Д.В., Киселев А.В., Кузнецов А.В., Вдовикин О.И., Аладышев О.С., Овсянников А.П. Создание распределенной инфраструктуры для суперкомпьютерных приложений // Программные продукты и системы. – 2008. – №2. – С. 2-7.
5. <http://www.ngrid.ru>
6. Denis Tarasov, Ekaterina Izotova, Diana Alisheva, Natalia Akberova, Robert A. Freitas Jr. Structural stability of clean, passivated, and partially dehydrogenated cuboid and octahedral nanodiamonds up to 2 Nanometers in Size // J. Comput. Theor. Nanosci. – 2011/ – V. 8. – P. 147-167.

7. Alexander G. Shamov, Grigori M. Khrapkovskii. A theoretical study of the gas-phase pyrolysis of nitroethylene// *Mendeleev's Communications*. – 2001. – No 4. – P. 163-164.
8. Шамов А.Г., Николаева Е.В., Храповский Г.М. Теоретическое изучение механизма газофазного распада нитроэтилена // *Журнал общей химии*. – 2004. – Т. 74, Вып. 8. – С. 1327-1342.
9. Лукацкий А. Как построить доверенную систему из недоверенных элементов. – http://lukatsky.blogspot.com/2011/12/blog-post_19.html.
10. Щербаков А.Ю. Современная компьютерная безопасность. Теоретические основы. Практические аспекты. Учебное пособие. – М.: Книжный мир, 2009. – С. 49-52.
11. Trottenberg U., Oosterlee C., Schuller A. Multigrid. – London: Academic Press, 2001. – 631 p.
12. Stuben K. Algebraic multigrid (AMG): an introduction with applications// GMD Report 70, GMD, Sankt Augustin, Germany, 1999.
13. Демидов Д.Е., Егоров А.Г., Нуриев А.Н. Решение задач вычислительной гидродинамики с применением технологии NVIDIA CUDA // *Ученые записки Казанского университета. Серия Физико-математические науки*. – 2010. – Т. 152, Кн. 1. – С. 142-154.
14. Демидов Д.Е., Егоров А.Г., Нуриев А.Н. Применение технологии NVIDIA CUDA для решения задач гидродинамики // *Основы работы с технологией CUDA / Под ред. А.В. Борескова, А.А. Харламова*. – М.: ДМК Пресс, 2010. – С. 194-205.
15. Barrett R., Berry M., Chan T.F., Demmel J., Donato J., Dongarra J., Eijkhout V., Pozo R., Romine C., der Vorst H. V. Templates for the solution of linear systems: building blocks for Iterative methods, 2nd Edition, SIAM, Philadelphia, PA, 1994.
16. Галимов М.Р., Биряльцев Е.В. Некоторые технологические аспекты применения высокопроизводительных вычислений на графических процессорах в прикладных программных системах // *Вычислительные методы и программирование*. – 2010. – Т. 11. – С. 77-93.
17. <https://github.com/ddemidov/vexcl>
18. Demidov D. VexCL: Vector expression template library for OpenCL. <http://www.codeproject.com/Articles/415058/VexCL-Vector-expression-template-library-for-OpenC>, July 2012.
19. Ahnert K., Demidov D. Solving ordinary differential equations with OpenCL in C++. <http://www.codeproject.com/Articles/429183/Solving-ordinary-differential-equations-with-OpenC>, July 2012.

Сведения об авторах

1. Биктимиров Марат Рамилевич, генеральный директор Национальной ассоциации исследовательских и научно-образовательных электронных инфраструктур "е-АРЕНА", окончил Казанский государственный авиационный институт в 1988 году, кандидат технических наук, автор более 20 печатных работ и монографий.

Область научных интересов: информационные технологии, распределенные вычислительные грид-системы, высокопроизводительные вычисления, информационная безопасность.

Адрес электронной почты и телефон: marat@e-arena.ru; +7(499) 1352598.

Biktimirov Marat R., CEO of the National Association of Research and Educational Electronic Infrastructures "E-ARENA", graduated from the Kazan Aviation Institute in 1988, Ph.D., author of 20 publications and monographs.

Research interests: information technology, GRID-systems, high performance computing, information security.

E-mail address and phone number: marat@e-arena.ru; +7(499) 1352598

2. Биряльцев Евгений Васильевич, зав. лаб. НИИ математики и механики им. Н.Г. Чеботарева Института математики и механики им. Н.И. Лобачевского Казанского (Приволжского) федерального университета, окончил Казанский государственный университет в 1982 году, кандидат технических наук, автор более 50 печатных работ и монографий.

Область научных интересов: системы интеллектуальной обработки информации и моделирование геофизических процессов.

Адрес электронной почты и телефон: IgenBir@yandex.ru, +7 (937)6113103

Biryaltsev Eugene V. Head Lab. of Lobachevskii Institute of Mathematics and Mechanics Kazan Federal University, graduated from the State University in 1952, Ph.D., author of 50 publications

and monographs.

Research interests: intelligent information processing systems and modeling of geophysical processes.

E-mail address and phone number: IgenBir@yandex.ru, +7 (937)6113103

3. **Демидов Денис Евгеньевич**, старший научный сотрудник Казанского Филиала МСЦ РАН, окончил Казанский государственный университет в 2002 году, кандидат физ.-мат. наук, автор 15 печатных работ и монографий.

Область научных интересов: вычислительная гидродинамика, высокопроизводительные вычисления и численные методы.

Адрес электронной почты и телефон: ddemidov@ksu.ru, +7(843) 2337707

Demidov Denis E., a senior fellow of the Kazan Branch of Joint Supercomputer Center of the Russian Academy of Sciences, graduated from the State University in 2002, Ph.D., author of 15 publications and monographs.

Research interests: computational fluid dynamics, high performance computing and numerical methods.

E-mail address and phone number: ddemidov@ksu.ru, +7(843) 2337707

4. **Елизаров Александр Михайлович**, директор Казанского Филиала МСЦ РАН, профессор Казанского (Приволжского) федерального университета, окончил Казанский государственный университет в 1976 году, доктор физ.-мат. наук, профессор, Заслуженный деятель науки Республики Татарстан, автор около 200 печатных работ и монографий.

Область научных интересов: математическое моделирование, аэрогидродинамика, информационные технологии, вычислительная математика, высокопроизводительные вычисления.

Адрес электронной почты и телефон: amelizarov@gmail.com, +7 (987) 2961781

Elizarov Alexander M., director of the Kazan Branch of Joint Supercomputer Center of the Russian Academy of Sciences, Professor of Kazan Federal University, graduated from the State University in 1976, Dr. Sci., Professor, Honored Scientist of the Republic of Tatarstan, the author of about 200 scientific publications and monographs.

Research interests: mathematical modeling, aerodynamics, information technology, computational mathematics, high performance computing.

E-mail address and telephone number: amelizarov@gmail.com, +7 (987) 2961781

5. **Чачков Денис Владимирович**, старший научный сотрудник Казанского Филиала МСЦ РАН, окончил Казанский государственный технологический университет в 2005 году, кандидат химических наук, автор более 150 научных публикаций.

Область научных интересов: квантовая химия, механизмы реакций, взрывчатые вещества.

Адрес электронной почты и телефон: chachkov@jscs.ru, +7(843) 2921608.

Chachkov Denis V., a senior fellow of the Kazan Branch of Joint Supercomputer Center of the Russian Academy of Sciences, Kazan State Technological University graduated in 2005, PhD Chemical Engineering science, author of 150 publications.

Research interests: quantum chemistry, mechanism of reactions, explosive materials.

E-mail address and phone number: chachkov@jscs.ru, +7(843) 2921608