

На правах рукописи

Кореньков Владимир Васильевич

Методология развития научного
информационно – вычислительного комплекса в составе
глобальной грид-инфраструктуры

05.13.01 – Системный анализ, управление и обработка информации
(информационные, телекоммуникационные и инновационные технологии)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
доктора технических наук

Дубна 2012

Работа выполнена в Объединенном институте ядерных исследований

Научный консультант: доктор физико-математических наук, профессор, академик РАН Кадышевский Владимир Георгиевич, научный руководитель ОИЯИ

Официальные оппоненты: Воеводин Владимир Валентинович, доктор физико-математических наук, профессор, член-корреспондент РАН, заместитель директора Научно-исследовательского вычислительного центра Московского государственного университета им. М.В. Ломоносова

Оныкий Борис Николаевич, доктор технических наук, профессор, президент Научно-исследовательского ядерного университета «МИФИ»

Сахаров Юрий Серафимович, доктор технических наук, профессор, заведующий кафедрой ПР7 «Персональная электроника» Московского государственного университета приборостроения и информатики

Ведущая организация: Институт системного анализа Российской академии наук, г. Москва

Защита состоится «___» _____ 2012 г. в ___ часов на заседании диссертационного совета Д 800.017.02 на базе Международного университета природы, общества и человека «Дубна», по адресу: Московская обл., г. Дубна, ул. Университетская, д. 19, аудитория _____.

С диссертацией можно ознакомиться в библиотеке Международного университета природы, общества и человека «Дубна», по адресу: Московская обл., г. Дубна, ул. Университетская, д. 19.

Автореферат разослан «___» _____ 2012 г.

Ученый секретарь
диссертационного совета
кандидат физико-математических наук _____

Токарева
Надежда Александровна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

Развитие исследований в физике высоких энергий, астрофизике, биологии, науках о Земле и других научных отраслях требует совместной работы многих организаций по обработке большого объема данных в относительно короткие сроки. Для этого необходимы географически распределенные вычислительные системы, способные передавать и принимать данные порядка сотен терабайт в сутки, одновременно обрабатывать сотни тысяч задач и долговременно хранить сотни петабайт данных.

Грид – географически распределенная инфраструктура, объединяющая множество ресурсов разных типов (процессоры, долговременная и оперативная память, хранилища и базы данных, сети), доступ к которым пользователь может получить из любой точки, независимо от места их расположения. Грид предполагает коллективный разделяемый режим доступа к ресурсам и к связанным с ними услугам в рамках глобально распределенных виртуальных организаций, состоящих из предприятий и отдельных специалистов, совместно использующих общие ресурсы. В каждой виртуальной организации имеется своя собственная политика поведения ее участников, которые должны соблюдать установленные правила.

Современные грид-инфраструктуры обеспечивают интеграцию аппаратных и программных ресурсов, находящихся в разных организациях в масштабах стран, регионов, континентов в единую вычислительную среду, позволяющую решать задачи по обработке сверхбольших объемов данных, чего в настоящее время невозможно достичь в локальных вычислительных центрах. Наиболее впечатляющие результаты по организации глобальной инфраструктуры распределенных вычислений получены в проекте WLCG (Worldwide LHC Computing Grid или Всемирный грид для Большого адронного коллайдера) в Европейской организации ядерных исследований (ЦЕРН) при обработке данных с экспериментов на LHC (Large Hadron Collider) или БАК (Большой адронный коллайдер). На семинаре 4 июля 2012 года, посвященном наблюдению частицы, похожей на бозон Хиггса, директор ЦЕРН Р.Хойер дал высокую оценку грид-технологиям и их значимости для мировой науки. Грид-инфраструктура на LHC позволила обрабатывать и хранить колоссальный объем данных, поступающих от экспериментов на коллайдере, и, следовательно, совершать научные открытия. В настоящее время ни один крупный научный проект не осуществим без использования распределенной инфраструктуры для обработки и хранения данных.

В 2003 году российские институты и ОИЯИ включились в работу по проекту WLCG. В 2004 году начался масштабный европейский проект

создания глобальной научной грид-инфраструктуры – EGEE (Enabling Grids for E-science in Europe). Для обеспечения полномасштабного участия России в этом проекте был образован консорциум РДИГ (Российский Грид для интенсивных операций с данными – Russian Data Intensive Grid, RDIG), который организовал выполнение работ по этому проекту и способствовал развитию российского сегмента грид-инфраструктуры EGEE.

Большую роль в этих проектах играют информационно-вычислительные комплексы институтов и университетов разных стран, которые являются ресурсными центрами (грид-сайтами) глобальной системы распределенных вычислений.

Ресурсный центр Объединенного института ядерных исследований (ОИЯИ) является крупнейшим в России (РДИГ) центром в составе глобальной инфраструктуры WLCG для обработки, хранения и анализа данных для ЛНС.

Кроме ЛНС сотрудники ОИЯИ участвуют в экспериментах на базовых установках ОИЯИ и в других научных центрах мира. Полным ходом идет подготовка проекта создания ускорительного комплекса НИКА.

Все вышеперечисленные направления исследований требуют организации компьютеринга для полноценного участия специалистов института и других стран в физических экспериментах ОИЯИ и международных коллаборациях. Под компьютерингом понимается применение средств вычислительной техники и телекоммуникаций для целей научных исследований.

Функциональные требования к вычислительному комплексу отличаются для разных экспериментов и групп пользователей в силу различия моделей компьютеринга, потребностей в ресурсах, специфики решаемых задач, специализации программного обеспечения и т.д.

Алгоритмы планирования и управления потоком задач и данных для разных групп пользователей отличаются. Методы оценки эффективности и надежности функционирования ресурсных центров в составе глобальной грид-инфраструктуры пока недостаточно проработаны. В существующей модели компьютеринга ЛНС постоянно растут потоки передаваемых данных и количество заданий пользователей, что приводит к замедлению анализа данных. Чтобы решить эти проблемы, необходимо совершенствовать модель компьютеринга в соответствии с требованиями со стороны пользователей ЛНС. Для новых крупных научных проектов необходимо совершенствовать модели компьютеринга при проектировании распределенной компьютерной инфраструктуры (например, для экспериментов на ускорительном комплексе НИКА).

Таким образом, при развитии ресурсных центров глобальной грид-инфраструктуры возникает актуальная проблема, связанная с необходимостью повышения эффективности их функционирования при постоянном увеличении потоков задач и данных.

Цель диссертационного исследования

Целью диссертационной работы является разработка методологии развития научного информационно – вычислительного комплекса в составе глобальной грид-инфраструктуры для повышения эффективности и надежности его функционирования с учетом требований проектов, связанных с обработкой и хранением больших и сверхбольших объемов информации.

Основными **задачами**, решаемыми в работе, являются:

1. Обоснование концепции организации научно-исследовательского информационно-вычислительного комплекса в составе глобальной грид-инфраструктуры.
2. Разработка архитектуры, структурной схемы построения крупного информационно-вычислительного комплекса, являющегося ресурсным центром глобальной грид-инфраструктуры.
3. Развитие модели компьютинга крупных научных проектов, связанных с хранением и обработкой больших объемов информации.
4. Разработка интегральной оценки эффективности функционирования ресурсного центра в составе глобальной грид-инфраструктуры с учетом основных показателей качества работы.
5. Разработка универсальной архитектуры системы мониторинга и статистического учета ресурсов, сервисов, задач, виртуальных организаций, пользователей для повышения эффективности и надежности функционирования глобальной грид-инфраструктуры и ресурсных центров.

Научная новизна работы

1. Разработана концепция, архитектура, структурная схема научно-исследовательского высокопроизводительного вычислительного комплекса, который является ресурсным центром глобальной грид-инфраструктуры и сочетает новые возможности для локальных пользователей и пользователей виртуальных организаций различных грид-сред.
2. Разработана новая модель компьютинга для Большого адронного коллайдера, включающая инфраструктурный слой Tier3, которая расширяет возможности для анализа данных пользователями крупных коллабораций.
3. Создана базовая модель компьютинга экспериментов на коллайдере НИКА, позволяющая осуществлять различные стратегии управления потоками данных и задач.
4. Разработана интегральная оценка эффективности функционирования ресурсного центра в составе глобальной грид-инфраструктуры.
5. Впервые разработана многоуровневая система грид-мониторинга ресурсов, сервисов, задач, виртуальных организаций, пользователей, представляющая инструментальные средства для повышения эффективности, качества и надежности распределенных систем.

Защищаемые положения

1. Разработанные концепция, архитектура, структурная схема научного вычислительного комплекса снимают противоречия в требованиях разных групп пользователей, улучшают функциональные характеристики комплекса.

2. Развитая модель компьютеринга Большого адронного коллайдера с введением нового инфраструктурного слоя Tier3 повышает эффективность и качество анализа данных экспериментов

3. Базовая модель компьютеринга проекта НИКА позволяет выполнить анализ различных сценариев и выбрать наиболее эффективное решение для построения распределенной системы обработки и хранения информации экспериментов на коллайдере НИКА.

4. Разработанная интегрированная оценка эффективности функционирования ресурсного центра в составе глобальной грид-инфраструктуры позволяет повысить объективность принятия решений по устранению недостатков в работе ресурсного центра и его развитию.

5. Разработанная архитектура и созданная на ее основе комплексная система грид-мониторинга (ресурсов, сервисов, задач, виртуальных организаций, пользователей) повышают эффективность и надежность функционирования грид-инфраструктур и служит основой для прогнозирования их развития.

Практическая значимость

1. На основе разработанной в диссертации концепции, архитектуры, структурной схемы научно-исследовательского высокопроизводительного вычислительного комплекса создан ресурсный центр в составе грид-инфраструктуры активно востребованный пользователями грид: только за 8 месяцев 2012 года в ОИЯИ было выполнено более 5 миллионов задач, а потребление процессорного времени при этом составило более 100 миллионов часов в единицах NEPSpec06.

2. Развитая модель компьютеринга ЛНС использована в ОИЯИ при создании программно-аппаратной среды для пользователей экспериментов ATLAS, CMS и ALICE, что обеспечило участие специалистов в работах по массовому моделированию физических событий, обработке и анализу реальных данных с действующих установок ЛНС.

3. На базе разработанной архитектуры создан комплекс систем грид-мониторинга и статистического учета (ресурсов, сервисов, задач, виртуальных организаций, пользователей), включающий:

- мониторинг и учет российской грид-инфраструктуры РДИГ;
- мониторинг сервиса передачи файлов (FTS);
- мониторинг вычислительного комплекса ОИЯИ;

- мониторинг функционирования глобальной инфраструктуры WLCG в реальном масштабе времени с применением интерфейса Google Earth;
- мониторинг и учет национальной нанотехнологической сети ГридННС;
- мониторинг ресурсов проекта «СКИФ-ГРИД»;
- мониторинг центров анализа информации уровня Tier3 в модели компьютеринга ЛНС.

Реализация результатов работы

Результаты диссертации были получены под руководством и личном участии соискателя в следующих международных проектах: EDG (проект европейской DataGrid-инфраструктуры), WLCG: проект грид для Большого адронного коллайдера (Worldwide LHC Computing Grid), EGEE: развертывание грид-систем для научных исследований (The Enabling Grids for E-sciencE), SKIF-Grid (программа развития высокопроизводительных вычислений союзного государства Россия – Беларусь), EGI-InSPARE (проект развития европейской грид-инфраструктуры).

Автор диссертации внес определяющий вклад при выполнении ряда национальных российских проектов, из которых как наиболее значительные можно выделить: «Разработка компьютеринговой системы для развития Грид-комплекса RuTier2/РДИГ для проведения российскими институтами распределенного анализа данных экспериментов на Большом адронном коллайдере ЦЕРН в составе глобальной Грид-системы WLCG/EGEE» (НИР), «Создание прототипа центра базовых Грид-сервисов нового поколения для интенсивных операций с распределенными данными в федеральном масштабе» (ОКР), «ГридННС - инфраструктура региональных центров программы развития нанотехнологий и nanoиндустрии» (ОКР).

Основные направления работ по созданию и развитию информационно-вычислительного комплекса ОИЯИ и глобальной системы распределенных вычислений были поддержаны в период с 1995 года по 2012 год девятнадцатью грантами РФФИ. В настоящее время автор является руководителем двух международных проектов РФФИ: «Глобальная система мониторинга передачи данных в инфраструктуре проекта WLCG» и «Обработка данных в T2-центрах ЛИТ ОИЯИ и ННЦ ХФТИ грид-инфраструктуры эксперимента CMS в условиях быстрого увеличения светимости Большого адронного коллайдера».

Базовая модель компьютеринга НИКА реализуется в проекте «Модель распределенной системы коллективного пользования для сбора, передачи и обработки сверхбольших объемов информации на основе технологии Грид для ускорительного комплекса НИКА», а развитие модели компьютеринга для экспериментов на ЛНС – в проекте «Создание автоматизированной системы обработки данных экспериментов на Большом адронном коллайдере (ЛНС) уровня Tier1 и обеспечения грид-сервисов для распределенного анализа этих

данных». Оба проекта реализуются в рамках федеральной целевой программы (ФЦП) Министерства образования и науки РФ "Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы", причем автор диссертации является руководителем в первом проекте и координатором от ОИЯИ во втором.

Результаты диссертации используются при реализации совместных проектов ОИЯИ с организациями стран-участниц.

Апробация диссертации:

Результаты работы являются итогом более чем 20-летней научной и организационной деятельности соискателя. Основные результаты диссертации докладывались и обсуждались на научных семинарах ЛВТА (ЛИТ) и других подразделений ОИЯИ, университета «Дубна», на Международных совещаниях и конференциях в научных центрах и университетах России, Армении, Беларуси, Болгарии, Венгрии, Испании, Италии, Казахстана, Молдовы, Монголии, Польши, Румынии, Словакии, Украины, Узбекистана, Франции, ФРГ, Чехии, Швейцарии, в том числе:

- Международные конференции «Computing in High Energy Physics» (СНЕР): 1997 (Берлин, ФРГ), 2004 (Интерлакен, Швейцария), 2009 (Прага, Чехия)
- Международные конференции «Математика. Компьютер. Образование», Дубна: 2000, 2002, 2004, 2006, 2008, 2010, 2012
- Международные конференции «Modern Trends in Computational Physics», Дубна (2000, 2009), Словакия (2011)
- Всероссийская конференция "Научный сервис в сети ИНТЕРНЕТ", Новороссийск, 2000, 2002
- Всероссийская конференция "Высокопроизводительные вычисления и их приложения", Черногоровка, 2000
- Международная конференция "Решения по управлению данными в научных исследованиях", Дубна, 2001
- Международный конгресс по математическому моделированию, Дубна, 2002
- Международная конференция «Advanced computing and analysis techniques in physics research» (АСАТ-2002), Москва, 2002
- Всероссийские конференции "Электронные библиотеки: перспективные методы и технологии, электронные коллекции" (RCDL), Дубна: 2002, 2008, 2012
- Сессия отделений информатики и математики РАН, Москва, 2003
- Международные конференции RDMS CMS: Дубна (2003), Минск, Беларусь (2004), Варна, Болгария (2006), Дубна (2009), Варна, Болгария (2010)
- Конференция «Телематика-2006», Санкт-Петербург, 2006

- Международная конференция «GRID Activities within Large Scale International Collaborations», Синая, Румыния, 2006
- Международная конференция Open Grid Forum и EGEE User Forum, Манчестер, Англия, 2007
- Международное совещание “Physics and Computing at ATLAS”, Дубна, 2008
- Международная конференция "Distributed Computing before and during LHC data-taking", Москва, 2008
- Международная конференция «Украинский академический грид» (УАГ-2009), Киев, 2009
- Международная конференция «Annual Meeting of the Romanian LCG Federation», Бухарест, Румыния: 2009, 2010
- Международная конференция "Telecommunications, Electronics and Informatics“ (ICTEI 2010), Кишинев, Молдова, 2010
- Международное совещание по компьютерингу коллаборации эксперимента ATLAS, Дубна, 2011
- Международное совещание "Perspectives on Physics on CMS at Very High Luminosity, HL-LHC“, Алушта, 2012

Соискатель являлся сопредседателем оргкомитета шести международных симпозиумов по ядерной электронике и компьютерингу (Варна, Болгария) в 2001, 2003, 2005, 2007, 2009 и 2011 году, зам. председателя пяти международных конференций «Распределенные вычисления и ГРИД-технологии в науке и образовании» (Дубна) в 2004, 2006, 2008, 2010 и 2012 году, на которых также представлялись результаты настоящей диссертации.

Публикации и личный вклад автора

Изложенные в диссертации результаты получены соискателем в результате его многолетней научной и организационной работы по развитию и совершенствованию информационно-вычислительного комплекса ОИЯИ и внедрению современных архитектурных инфраструктурных решений в масштабе России и стран-участниц ОИЯИ.

Все исследовательские работы и разработки по теме диссертации – от постановки задачи и выбора методики до получения результатов – были выполнены под непосредственным руководством соискателя, и его вклад является определяющим.

Немаловажным при принятии решений о развитии информационно-вычислительного комплекса ОИЯИ как базового сегмента глобальной системы распределенных вычислений являлось активное участие соискателя в работе ряда международных и российских комитетов и рабочих групп, занимающихся выработкой концептуальных решений в сфере вычислительной техники и сетевых коммуникаций.

По теме диссертации автором опубликовано свыше 80 печатных работ, в том числе по основным результатам – 41 работа (из них 13 работ в изданиях

из перечня ведущих рецензируемых научных изданий, рекомендованного ВАК РФ для публикации основных результатов диссертаций). Также результаты работы опубликованы в отчетах по руководимым автором инфраструктурным проектам в рамках Федеральных целевых программ и проектам, поддержанных РФФИ. Список основных публикаций приведен в конце автореферата

Структура и объем диссертации

Диссертация состоит из введения, пяти глав, заключения, списка литературы из 202 наименований; полный объем работы составляет 260 страниц.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **Введении** обоснована актуальность темы, сформулированы цель и основные задачи работы, описаны результаты, показана их научная новизна и практическая значимость, приведены положения, выносимые на защиту.

В **первой главе** диссертации разработана концепция, архитектура, структурная схема научного вычислительного комплекса, учитывающие требования разных групп пользователей.

Приведен краткий обзор этапов развития распределенного вычислительного комплекса ОИЯИ. Для каждого этапа выполнялся анализ требований пользователей и классов решаемых задач с учетом тенденций развития вычислительной техники, сетевого оборудования, архитектурных решений.

На основе этого анализа были сформулированы основные требования к современному высокопроизводительному научному вычислительному комплексу:

- многофункциональность,
- высокая производительность,
- развитая система хранения данных,
- высокая надежность и доступность,
- информационная безопасность,
- масштабируемость,
- развитая программная среда для различных групп пользователей.

В разные периоды развития компьютерной инфраструктуры проводился анализ классов задач пользователей, чтобы максимально соответствовать их требованиям. Так, например, в середине девяностых годов было следующее распределение задач, решаемых на вычислительном комплексе ОИЯИ:

- задачи моделирования, обработки и анализа информации в физике высоких энергий составляли около 70%;
- задачи теоретической и математической физики, требующие больших вычислительных ресурсов и использующие параллельные вычисления (квантовая хромодинамика на решетках, расчет магнитных полей, молекулярная динамика и др.) – 15%;

- задачи ядерной физики, низких и промежуточных энергий, не использующие параллельных вычислений – 8%;
- другие задачи (аналитические преобразования, разработка численных методов, пакетов и т.д.) – 7%;

Со временем это распределение по классам задач меняется в сторону увеличения доли задач в физике высоких энергий. Эта тенденция развивалась в связи с активным участием физиков ОИЯИ в экспериментах на LHC. Таким образом, при разработке концепции и архитектуры компьютерного комплекса необходимо учитывать в первую очередь требования задач физики высоких энергий, которая является пионером в области применения компьютеринга в экспериментальных исследованиях.

Особенность задач в физике высоких энергий состоит в следующем:

- исследования проводятся в больших международных коллективах (коллаборации с числом участников более 1000);
- огромные объемы экспериментальных и смоделированных данных, которые один раз формируются и много раз используются для обработки и анализа;
- использование сложного ПО для обработки и анализа данных, которое разрабатывается большими коллективами и постоянно совершенствуется;
- необходимость повторной обработки данных после уточнения условий работы детекторов и ускорителя и/или проведения калибровки каналов считывания;
- моделирование условий работы современных установок и физических процессов, одновременно с набором и обработкой «реальных» данных;
- на основных этапах обработки и анализа нет необходимости применения параллельных алгоритмов, так как можно обрабатывать параллельно независимые данные (события).

Следовательно, для решения задач в области физики высоких энергий можно использовать любые вычислительные ресурсы, которые имеют эффективный доступ к системам хранения экспериментальных и смоделированных данных. Особенность этого класса задач делает неэффективным применение классических суперкомпьютеров, так как выполняется огромный поток независимых задач, каждая из которых независимо обрабатывает события, расположенные в системах распределенного хранения. При этом высокоскоростная коммуникационная среда суперкомпьютеров, связывающая вычислительные узлы практически не используется. Поэтому наиболее эффективным решением для этого класса задач является объединение слабосвязанных кластеров с высокопроизводительными распределенными системами хранения больших и сверхбольших объемов данных.

В середине 90-х годов под руководством автора диссертации была разработана концепция и проект развития сетевой и информационно-

вычислительной инфраструктуры ОИЯИ (проект CONET), который сыграл важную роль в системном разделении работ по следующим направлениям:

- системы телекоммуникаций, включая развитие внешних каналов связи и высокоскоростной опорной сети ОИЯИ (Backbone);
- системы мощных вычислений и массовой обработки данных;
- системы оперативного хранения данных и архивной массовой памяти;
- системы программного, информационного обеспечения и визуализации.

Высокоскоростная, надежная, хорошо защищенная сетевая инфраструктура создает прочный фундамент для развития научного вычислительного комплекса, который является ресурсным центром системы распределенных вычислений.

Одной из значимых работ было развертывание в ОИЯИ системы распределенных вычислений на основе пакета Condor, в которой процессорное время и другие ресурсы серверов и рабочих станций, расположенных по всему миру, предоставляются всем участникам в те моменты, когда эти ресурсы локально не используются или используются незначительно. При этом администратор каждого из компьютеров может сформулировать свое собственное понимание «загруженности». В результате участники этой динамически развивающейся системы приобретают возможность резко ускорить обработку своих заданий без расширения своих локальных ресурсов.

В 1994 году пул из трех SUN SPARC-станций ОИЯИ был включен в европейский пул ресурсов с центром администрирования в Амстердаме (Голландия). Этот пул, в свою очередь, был составной частью объединенного пула с центром администрирования в университете штата Висконсин (США). В результате создания небольшого собственного пула и включения его в глобальную систему пользователи ОИЯИ получили доступ к 250 рабочим станциям в США и Европе, которые использовались для расчетов в физике высоких энергий. Это был первый в России опыт участия в крупном международном проекте, который продемонстрировал новый архитектурный подход к реализации глобальной системы распределенных вычислений на невыделенных ресурсах.

В 1998 году была разработана и реализована архитектура комплекса, состоящего из массивно-параллельной вычислительной системы S-класса фирмы Hewlett Packard (HP) SPP2000 с 8 процессорами (общая пиковая производительность 6,4 GFlops), файлового сервера D-класса и системы массовой памяти на основе автоматизированной ленточной библиотеки ATL-2640 емкостью до 10 Терабайт. Система массовой памяти допускала две технологические схемы использования: с помощью программных средств HP OmniBack (автоматическое резервное копирование институтских компьютеров) и HP OmniStorage (управление мигрирующей файловой системой). Для интеграции этих элементов и оптимизации работы комплекса

было проведено исследование с целью определения надежности функционирования и выбора конфигурации с максимальной скоростью выполнения резервного копирования. Была выполнена интеграция различных элементов компьютерного комплекса и кластеров в единую систему массовой памяти на базе HP OmniStorage.

Начиная с 2000 года, развитие вычислительного комплекса ОИЯИ продолжилось на базе кластеров и компьютерных ферм общего и специального назначения. Так как требования различных групп пользователей отличались, была разработана архитектура распределенного кластера, состоящего из четырех отдельных взаимосвязанных компонент различной аппаратной структуры и функционального назначения. Он включал в себя комплекс интерактивных серверов, на которых было установлено базовое и специализированное ПО, необходимое для физических экспериментов. В состав кластера входили 3 специализированные вычислительные фермы: ферма общего назначения, ферма для экспериментов ЛНС и ферма параллельных вычислений. Все эти фермы состояли из 2-процессорных персональных компьютеров (ПК), объединенных на базе сетевой технологии Ethernet, а узлы в ферме параллельных вычислений были соединены коммуникационной средой Myrinet-2000.

В 2004 году была выполнена работа по настройке грид-сервисов и включению вычислительного комплекса ОИЯИ в состав глобальной грид-инфраструктуры. Чтобы соответствовать требованиям глобальных грид-проектов, требовалось обеспечить необходимый уровень всех составляющих инфраструктуры, включая широкополосные каналы связи, надежную высокоскоростную локальную сеть, развитую систему управляющих серверов, мощную систему обработки и хранения данных, установку программного обеспечения грид-среды и поддерживаемых виртуальных организаций.

На этом этапе грид-сайт ОИЯИ функционировал только для пользователей виртуальных организаций, а локальные пользователи и пользователи параллельных вычислений работали специальных вычислительных фермах. При этом решении каждая ферма обслуживала разные группы пользователей (локальные, параллельные, грид) с различными алгоритмами планирования задач и распределения ресурсов.

Почему для разных групп пользователей используются различные алгоритмы планирования заданий? Для локальных пользователей наиболее приемлемая стратегия: использование приоритетного планирования (чем проще и короче задача, тем выше приоритет) для минимизации среднего времени выполнения заданий. Для параллельных задач существует много стратегий распределения ресурсов и планирования заданий, которые направлены на нахождение компромисса между средним временем выполнения заданий и эффективной загрузкой вычислительных ресурсов. При большом потоке параллельных заданий некоторые ресурсоемкие задания

могут стоять в очереди на выполнение непредсказуемо долго. Поэтому существуют алгоритмы, ограничивающие максимальное время нахождения задания в очереди, понижая загрузку вычислительных ресурсов. Для задач, запущенных пользователями виртуальных организаций в рамках грид-проектов, обычно применяются другие стратегии планирования. Как правило, между виртуальными организациями и ресурсными центрами подписываются соглашения об уровне обслуживания (Service Level Agreement, SLA), в котором указаны предоставляемые ресурсы, в том числе вычислительные. Можно поступить просто – закрепить вычислительные ресурсы за определенными виртуальными организациями, которые только они могут использовать. Если в какой-то период времени от какой-то виртуальной организации нет заданий, то выделенные для нее ресурсы простаивают. Как только задания этой виртуальной организации появляются в очереди, они сразу без задержки начинают выполняться. Эту стратегию легко реализовать, она устраивает виртуальные организации, но может привести к неэффективному использованию вычислительных ресурсов. Наиболее разумная стратегия состоит в том, чтобы закрепить за виртуальной организацией процент вычислительных ресурсов, указанных в соглашении и обеспечить этот процент за определенный срок. Эта стратегия не предполагает закрепления ресурсов, поэтому они могут использоваться заданиями разных виртуальных организаций, поэтому вычислительные ресурсы используются эффективно.

В течение нескольких лет этот комплекс, состоящий из трех вычислительных ферм, развивался, но были очевидные недостатки этой архитектуры, которые выражались в неравномерности загрузки разных вычислительных ферм комплекса, что приводило к простоям одних и перегрузкам других ферм.

Были проведены исследования с использованием пакета имитационного моделирования для оценки параметров кластера при объединении вычислительных ферм. Рассматривались три варианта:

- 1) Специализированные фермы, т.е. для каждой фермы свой поток заданий.
- 2) Универсальные фермы, т.е. задания из общего потока могут выполняться на любой ферме (объединение вычислительных ферм).
- 3) Компромиссный вариант, когда поток параллельных заданий может выполняться только на вычислительной ферме для параллельных вычислений, а остальные потоки заданий могут выполняться на любой вычислительной ферме.

В работе приведены параметры и распределения трех потоков заданий для разных дисциплин планирования с заданными характеристиками трех вычислительных ферм. Были получены результаты для каждого из трех вариантов: общее время прохождения заданий, размер очереди заданий, загрузка кластеров. Результаты моделирования показали, что при заданных параметрах потоков заданий общее время выполнения заданий уменьшается

на 30%, а средняя загрузка комплекса увеличивается на 20% в третьем варианте по сравнению с первым вариантом.

В результате была разработана архитектура вычислительного комплекса, который организован как единый вычислительный ресурс для всех групп пользователей и виртуальных организаций с определенной специализацией внутри кластера для параллельных вычислений. Разработана структурная схема, в рамках которой вычислительные ресурсы и ресурсы хранения данных управляются базовым программным обеспечением (БПО), позволяющим использовать ресурсы как в международных проектах распределенных вычислений, предоставляя доступ виртуальным организациям, так и локальным пользователям.

Для обоснования концепции вычислительного комплекса и его архитектуры были исследованы распределения потоков задач и данных, поступающих на обработку. Имеются следующие потоки заданий:

- Задания локальных пользователей, которые характеризуется большим разнообразием по времени выполнения, по запрашиваемым ресурсам;
- Задания, которые запрашивают более 1 процессора для своего выполнения (параллельные задачи);
- Задания, запускаемые пользователями виртуальных организаций через грид-инфраструктуру. Эти задания являются основными и среди них львиную долю занимают задачи моделирования, обработки и анализа данных экспериментов ЛНС. Требования у заданий этого класса очень высокие: высокоскоростные каналы связи для передачи данных, быстрый доступ к распределенным системам хранения данных и др.

Для решения задач моделирования, обработки и анализа данных экспериментов ЛНС необходимо обеспечить следующие потоки данных:

- передача файлов из центров хранения и их запись в локальную систему хранения;
- передача файлов из локальной системы хранения в центры хранения информации;
- обработка файлов, находящихся в локальной системе хранения.

В разрабатываемой архитектуре и структурной схеме необходимо было обеспечить компромиссную стратегию планирования заданий, чтобы учесть противоречивые требования разных групп пользователей в сочетании с высокой загрузкой вычислительных ресурсов и поддержкой наиболее универсальных и защищенных методов доступа к хранилищам данных. Распределение и учет вычислительных ресурсов реализован на основе системы пакетной обработки torque и планировщика ресурсов maui. Алгоритм планирования заданий настраивается в соответствии с обязательствами перед виртуальными организациями, с одной стороны, и обеспечения высокого уровня показателя загрузки вычислительного комплекса с другой стороны.

Для надежной и эффективной работы научного вычислительного комплекса необходимо:

- обеспечить надежную инженерную инфраструктуру (система бесперебойного питания, климат-контроль и т.д.) и сетевую инфраструктуру;
- проводить тщательный выбор приобретаемого оборудования. Были разработаны стенды и методика тестирования вычислительных серверов, управляющих серверов и систем хранения. Проведение предварительного тестирования особенно важно, когда модернизация и развитие комплекса происходит постоянно и новое оборудование должно быть эффективно включено в работающую систему без остановки ее функционирования.
- выполнять обоснованный выбор программного обеспечения и автоматизировать процесс перехода на новые версии в работающей системе;
- развивать многоуровневую систему мониторинга, которая выполняет функции сбора, хранения, анализа информации об инженерной и сетевой инфраструктуре, аппаратных и программных ресурсах, сервисах и пользователях, виртуальных организациях и приложениях с реализацией оперативной обратной связи.
- постоянно проводить анализ потоков задач и данных для учета требований разных групп пользователей и адаптации к ним архитектуры, алгоритма планирования заданий и схемы функционирования вычислительного комплекса.

Эффективность предложенных методов и подходов к разработке архитектуры вычислительного комплекса, выбора программно-аппаратных средств можно проанализировать на примере вычислительного комплекса ОИЯИ.

Для выбора системы хранения больших объемов информации был составлен список требований и критериев, среди которых отметим следующие:

- Широкие функциональные возможности, гибкость, масштабируемость.
- Возможность работы с протоколами dcap, SRM2, gsiFTP, xrootd.
- Интероперабельность с другими системами хранения.

Проведен анализ различных систем и обоснован выбор системы хранения dCache, которая наиболее полно удовлетворяет сформулированным критериям.

Важнейшей задачей является обеспечение надежной и эффективной работы системы дисковых массивов с учетом балансировки нагрузки на дисковые пулы. Для этой цели созданы средства мониторинга и анализа

нагрузок на дисковые пулы, которые позволяют решать задачу равномерного распределения нагрузок на систему дисковых массивов.

Большую роль в эффективности функционирования вычислительного комплекса играет правильно организованная сетевая инфраструктура, соединяющая управляющие серверы, вычислительную систему и систему хранения информации. В ОИЯИ центральный маршрутизатор сети ресурсного центра соединен с основным граничным маршрутизатором сети ОИЯИ на скорости 10 Гигабит Ethernet. Для обеспечения высокой пропускной способности локальной сети и минимального времени доступа к данным и файлам применяется агрегирование нескольких соединений 1 Гигабит Ethernet в единый виртуальный канал (TRUNK) с увеличенной пропускной способностью от 4 до 8 Гигабит Ethernet.

На рис. 1 приведена архитектура вычислительного комплекса ОИЯИ и общая схема взаимодействия уровней, ресурсов, сервисов.

В настоящее время вычислительный комплекс ОИЯИ состоит из 2582 64-х битных процессоров и системы хранения данных общей емкостью более 1800 Тбайт.

В качестве практического результата можно представить следующие аргументы:

Показатели надежности и доступности ресурсного центра ОИЯИ составляют в 2012 году 99%, а за 8 месяцев 2012 года было выполнено более 5 миллионов заданий пользователей. Более подробный анализ проведен в следующих главах.

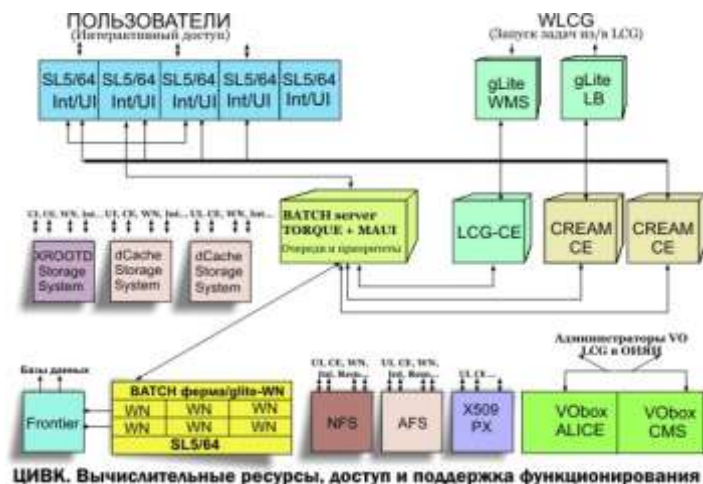


Рис. 1. Структурная схема вычислительного комплекса ОИЯИ.

Результаты первой главы подтверждают *первое защищаемое положение*:

Разработанные концепция, архитектура, структурная схема научного вычислительного комплекса снимают противоречия в требованиях разных групп пользователей, улучшают функциональные характеристики комплекса.

Вторая глава диссертации посвящена участию в работах по развитию компьютерной инфраструктуры для обработки, хранения, анализа данных экспериментов на Большом адронном коллайдере.

Для моделирования распределенной вычислительной инфраструктуры центров для LHC в 1998 году был создан специальный международный проект MONARC (**M**odels of **N**etworked **A**nalysis at **R**egional **C**entres for LHC **E**xperiments). Задача организации компьютеринга была очень сложной, поскольку требовалось:

- обеспечить быстрый доступ к массивам данных колоссального объема;
- обеспечить прозрачный доступ к географически распределенным ресурсам;
- создать протяженную надежную сетевую инфраструктуру в гетерогенной среде.

В рамках проекта MONARC была разработана базовая модель компьютеринга для экспериментов LHC как иерархическая централизованная структура региональных центров, включающая в себя центры нескольких уровней. Суть распределенной модели архитектуры компьютерной системы состоит в том, что весь объем информации с детекторов LHC после обработки в реальном времени и первичной реконструкции должен направляться для дальнейшей обработки и анализа в региональные центры.

Методом имитационного моделирования был обоснован иерархический принцип организации информационно-вычислительной системы LHC, предполагающей создание центров разных уровней или ярусов (Tier's): Tier0 (CERN) → Tier1 → Tier2 → Tier3 → компьютеры пользователей.

Уровни различаются по масштабу ресурсов (сетевые, вычислительные, дисковые, архивные) и по выполняемым функциям:

Tier0 (ЦЕРН)	первичная реконструкция событий, калибровка, хранение копий полных баз данных
Tier1	полная реконструкция событий, хранение актуальных баз данных по событиям, создание и хранение наборов анализируемых событий, моделирование, анализ
Tier2	тиражирование и хранение наборов анализируемых событий, моделирование, анализ
Tier3	кластеры отдельных исследовательских групп

В рамках этого проекта были проработаны требования к ресурсам и функции региональных центров уровней Tier0, Tier1, Tier2. Разработанная модель была реализована и успешно функционирует с момента запуска Большого адронного коллайдера в 2009 году. Ежегодно собираются и обрабатываются данные объемом в десятки и даже сотни петабайт. Эта централизованная модель компьютеринга хорошо зарекомендовала себя в первые годы работы экспериментов на LHC, однако наблюдается тенденция к

децентрализации сервисов и переносу акцентов с массовой обработки данных на обеспечение индивидуального анализа. Происходит эволюция инфраструктуры и модели компьютеринга экспериментов на LHC:

- переход от иерархической структуры к сетевой, а в идеале к полностью связанной, где возможны связи между центрами всех уровней;
- развитие средств распределенного управления данными, поддержка очень высоких скоростей передачи огромных массивов данных;
- создание инфраструктуры центров уровня Tier3 для индивидуального анализа данных;
- развитие и применение средств виртуализации и облачных вычислений (проект «Helix Nebula – научное облако»).

Под руководством автора диссертации работы по организации компьютеринга для экспериментов на LHC в ОИЯИ начались еще в 1996 году с создания информационного WWW-сервера для коллаборации CMS. В 1997 году в ОИЯИ был создан кластер из трех SunSparc станций. Программное окружение кластера полностью соответствовало программному окружению на ЦЕРН-кластере cms.cern.ch. На кластере в ОИЯИ проводилась обработка данных с прототипов детекторов CMS, осуществлялась постановка актуальных версий специализированного программного пакета CMS, велись работы по моделированию физических процессов и подсистем детектора. В 2000 году в ОИЯИ была создана вычислительная ферма для экспериментов на LHC, которая входила в состав вычислительного комплекса ОИЯИ. Эти ресурсы использовались для текущей работы специалистов ОИЯИ и проведения Монте-Карло моделирования физических событий в рамках сеансов массового моделирования, осуществляемого коллаборацией CMS. В 2004 году был организован ресурсный центр уровня Tier2, интегрированный в глобальную грид-инфраструктуру проекта WLCG. Tier2 центр ОИЯИ имеет полный набор необходимых грид-сервисов и отвечает всем требованиям по техническому и программному обеспечению, необходимым для успешного функционирования виртуальных организаций CMS, ATLAS, ALICE. За период своего функционирования с 2004 года по настоящее время ресурсный центр уровня Tier2 в ОИЯИ по своим показателям надежности и доступности работы является одним из лучших среди всех ресурсных центров уровня Tier2.

После запуска в эксплуатацию LHC в эксперименте ATLAS возникла серьезная проблема с обеспечением целостности распределенного хранения информации. Это выразалось в том, что большой процент удаляемых файлов из глобального каталога оставался в локальных каталогах систем хранения данных ресурсных центров. Чтобы решить эту серьезную проблему, под руководством автора диссертации была разработана архитектура и спроектирована подсистема обеспечения целостности распределенного хранения информации для эксперимента ATLAS. В рамках этой подсистемы был реализован один из центральных сервисов системы управления данными

эксперимента ATLAS - сервиса удаления данных (Deletion Service), который предназначен для централизованного удаления ненужных данных эксперимента ATLAS на сайтах Tier0, Tier1 и Tier2.

Основные функции сервиса удаления данных:

- обслуживает запросы на удаление,
- организует балансировку нагрузки,
- обеспечивает масштабируемость и отказоустойчивость системы,
- выполняет корректную обработку исключений, возникающих в процессе работы,
- реализует стратегию повторения операций в случае возникновения отказов.

Данные эксперимента ATLAS распределены более, чем на 100 грид-сайтах с общим объемом дискового пространства более 150 петабайт, в котором хранятся сотни миллионов файлов. Недельный объем удаляемых данных составляет около 2 Петабайтов (20 млн. файлов).

Созданный сервис обеспечивает целостность хранения информации в географически распределенной среде.

Опыт выполнения анализа данных с использованием центров уровня Tier2 показал, что эти центры не всегда оперативно решают задачи анализа, так как загружены многими другими важными задачами, в том числе моделированием событий, и часто поддерживают несколько виртуальных организаций. Поэтому задачи анализа данных LHC могут достаточно долго находиться в очереди на выполнение. Вследствие вышеперечисленных обстоятельств, многие группы пользователей в разных странах для анализа экспериментальных данных стали использовать дополнительные разнообразные вычислительные ресурсы (серверы, кластеры, суперкомпьютеры), которые находятся вне централизованного управления и планирования, на которые не распространяются какие-либо единые требования, касающиеся технических решений. Для этих центров характерно большое разнообразие систем анализа данных, систем хранения данных и пакетной обработки задач (PBS, OGE, CONDOR, Lustre, xROOTd, PROOF и др.). Практически, эти центры составили новый инфраструктурный уровень (Tier3) в модели компьютеринга экспериментов LHC, который вносит существенный вклад в физический анализ экспериментальных данных.

В рамках развития модели компьютеринга экспериментов на LHC в диссертации проведены исследования центров уровня Tier3, их систематизации и интеграции с центрами Tier2. Было проведено анкетирование существующих центров и их классификация, выбраны наиболее распространённые конфигурации, проанализированы различные схемы взаимодействия между уровнями. В работе систематизированы различные конфигурации используемых систем и выделены наиболее распространённые из них.

Для исследования центров уровня Tier3 и систематизации используемых средств обеспечения интеграции с центрами уровня Tier2 в

ОИЯИ была разработана архитектура тестовой инфраструктуры, которая позволяет создавать прототипы различных конфигураций центров уровня Tier3. С применением технологий виртуализации такая инфраструктура была реализована на базе виртуальных кластеров, что позволило разработать для каждого варианта конфигурации документацию, настроить или разработать средства локального мониторинга, выработать полные рекомендации по системе сбора информации для глобального мониторинга центров уровня Tier3.

Разработанная система глобального мониторинга центров уровня Tier3 предоставляет обобщенную информацию о ресурсах, популярности анализируемых данных, что имеет большое значение для координации работ в рамках виртуальной организации, так как обеспечивается всесторонний взгляд на вклад Tier3 сайтов в вычислительный процесс.

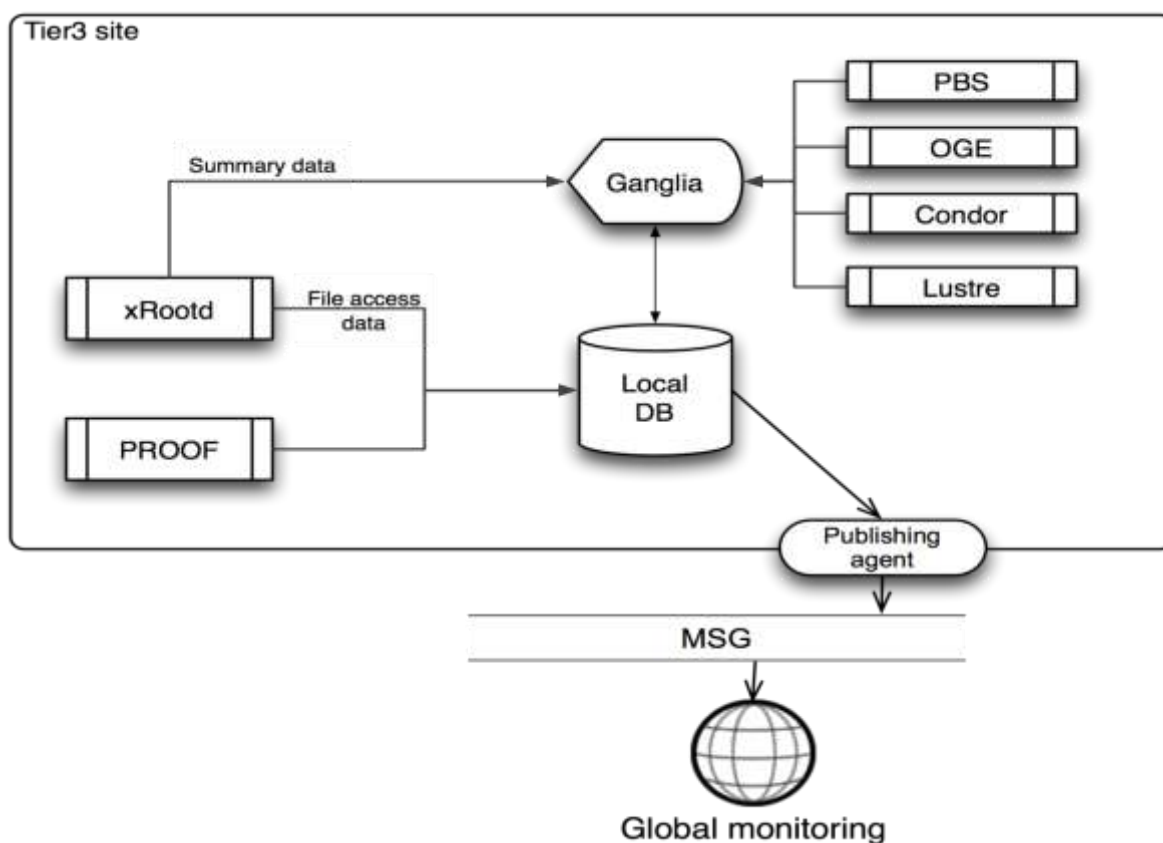


Рис. 2. Схема функционирования основных вариантов Tier3 центров и их взаимодействие с системой глобального мониторинга.

В диссертации представлена модель глобального мониторинга Tier3 центров (T3mon), который выполняет задачи сбора информации от систем локального мониторинга Tier3 центров, организации хранения этой информации для предоставления нужной информации администраторам Tier3 сайтов, пользователям виртуальной организации. Схема функционирования

основных вариантов Tier3 центров и их взаимодействие с системой глобального мониторинга представлены на рис. 2.

Результаты второй главы подтверждают *второе защищаемое положение*:

Развитая модель компьютеринга Большого адронного коллайдера с введением нового инфраструктурного слоя Tier3 расширяет возможности для анализа данных пользователями крупных коллабораций

В третьей главе обсуждаются вопросы создания имитационной модели для исследования планируемой архитектуры распределенного комплекса хранения, обработки и анализа данных экспериментов на проектируемом ускорительном комплексе НИКА.

Моделирование системы off-line обработки и хранения данных экспериментов должно дать ответ на следующие вопросы:

- какие вычислительные ресурсы требуются для обработки данных, чтобы получить результат в заданное время;
- как должны быть связаны между собой центры обработки;
- какое должно быть разделение функций между центрами;
- как организовать работу пользователей;
- какая стратегия запуска заданий должна применяться;
- какой объем дисковых массивов необходимо выделить для обеспечения надёжного хранения информации.

Создаваемая имитационная модель должна использоваться длительное время на разных этапах создания ускорительного комплекса: от технического проекта до создания системы обработки и хранения данных.

Модель должна включать:

- интерфейс пользователя,
- ядро моделирования,
- систему визуализации результатов моделирования.

Ядро моделирования обрабатывает входные данные (сценарий обработки, структура и параметры комплекса, характеристики потока заданий и др.) и определяет параметры прохождения заданий. Необходимо также смоделировать основные функции и сервисы грид, которые наиболее существенно влияют на прохождение заданий и передачу данных.

Для решения этой задачи необходимо выбрать пакет имитационного моделирования, исходя из сформулированных требований.

В результате сравнительного анализа за основу для моделирования был взят пакет GridSim, который имеет следующие преимущества:

- моделирование гетерогенных типов ресурсов;
- возможность моделирования приложений с различными стратегиями параллельных вычислений;
- отсутствие ограничений на количество заданий, которые могут быть отправлены на обработку в ресурсный центр;

- возможность задания пропускной способности сети между ресурсами;
- поддержка моделирования статических и динамических планировщиков заданий;
- гибкие возможности реализации различных стратегий управления данными;
- возможность регистрации статистики всех или выбранных операций.

Модель на основе пакета GridSim строится из базовых блоков, которые реализуют следующие классы моделируемых объектов: Пользователь, Ресурс, Информационный сервис грид, Ввод/вывод, Сеть и др. (Рис. 3.). Логически их можно объединить в две группы: описание структуры системы и описание потока заданий.

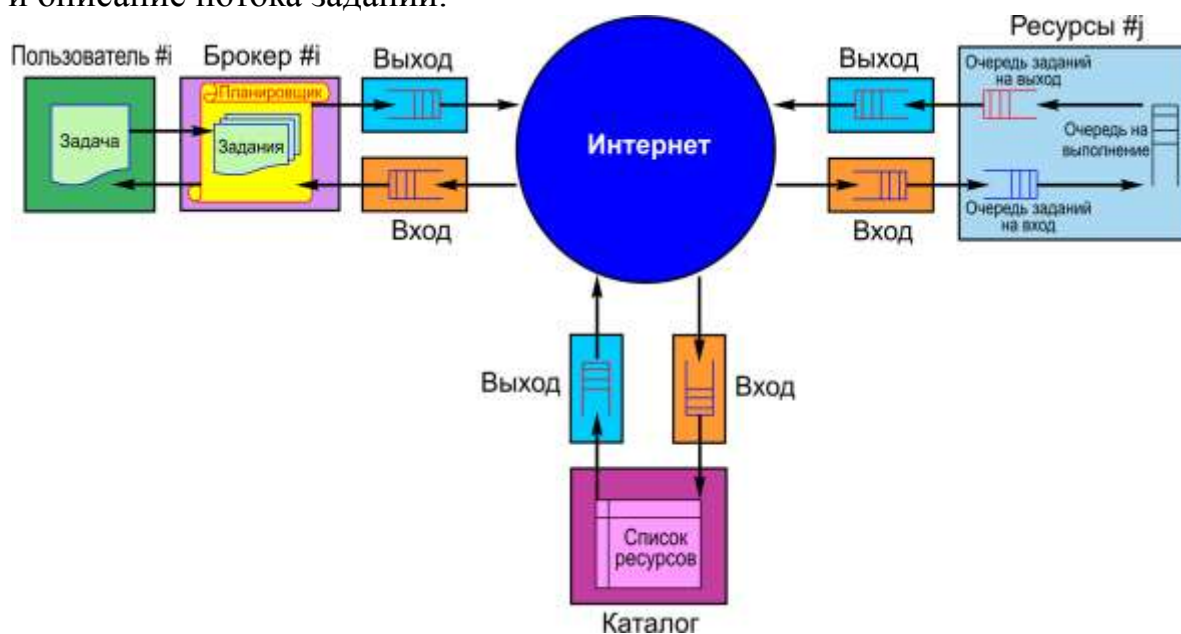


Рис. 3. Диаграмма основных блоков базовой модели, потоков данных и задач в GridSim.

Дается описание и характеристики более сложных объектов модели: книга сценариев, таблица ресурсов, смесь задач сценария, дисковый пул, ленточная подсистема, задача, пользователь, каталог файлов и реплик и другие объекты.

Модель рассчитывает 8 параметров, определяемых как процент, или абсолютное значение:

- Средняя загрузка сети по дням [%]
- Количество активных /ожидающих заданий
- Количество запрошенных и используемых процессоров
- Использование грид-структуры по часам [%] в день
- Использование ресурсов хранения данных [%]
- Процент отказавших процессоров по дням [%]
- Объем переданных данных в час
- Загрузка кластеров [%]

Данный набор параметров достаточен для оценки эффективности топологии распределенного вычислительного комплекса, оценки технических характеристик ресурсных центров, эффективности алгоритмов распределения заданий и передачи данных между ресурсными центрами.

Разработанная модель обобщает опыт компьютеринга на ЛНС, который показал, что для выбора архитектуры распределенной системы обработки и хранения информации необходимо разработать базовую модель, управление параметрами которой позволяет исследовать различные архитектуры системы, сценарии обработки и хранения данных, в зависимости от интенсивности потоков задач и данных и других факторов.

Планируется длительное использование созданной модели в процессе разработки технического проекта, последующего создания системы обработки и хранения данных и всего жизненного цикла функционирования системы. Для этой цели создана база данных сценариев, исходных данных, результатов моделирования, которые можно использовать на всех этапах проекта.

Сформулированы результаты по созданию базовой модели обработки экспериментальных данных ускорительного комплекса НИКА:

- Была разработана имитационная модель;
- Выделены ключевые параметры оценки модели;
- Созданы средства визуализации результатов;
- Имитационная модель прошла отладку и верификацию;
- Результаты моделирования позволяют сформулировать требования к архитектуре системы обработки.
- Имеется возможность создавать различные сценарии обработки информации, структур и параметров системы обработки, для выбора архитектуры и технических решений в соответствии с критериями, которыми руководствуется проектировщик.

Результаты третьей главы подтверждают *третье защищаемое положение*:

Базовая модель компьютеринга проекта НИКА позволяет выполнить анализ различных сценариев и выбрать наиболее эффективное решение для построения распределенной системы обработки и хранения информации экспериментов на коллайдере НИКА.

В четвертой главе разработана интегральная оценки эффективности грид-сайтов в составе распределенной инфраструктуры. Отметим, что общепринятой оценки эффективности грид-сайтов на данный момент не существует.

Интегральная оценка разрабатывалась на примере проекта WLCG. Этот проект имеет наиболее развитую инфраструктуру с большим количеством ресурсных центров разного уровня, а также средства мониторинга и учета ресурсов, сервисов и пользователей виртуальных организаций. В настоящий

момент каждый месяц в проекте WLCG публикуется информация о доступности/надежности грид-сайтов в процентах, что является важным показателем функционирования ресурсных центров.

Эти показатели вычисляются по результатам периодических тестов и следующим формулам:

$$\text{Доступность} = T_p / (T_o - T_{сн})$$

$$\text{Надежность} = T_p / (T_o - T_{зп} - T_{сн}),$$

где T_p - время работы, T_o - общее время, $T_{сн}$ - время, в течение которого статус был неизвестен, $T_{зп}$ - время запланированного простоя.

Время работы (T_p) определяется следующим образом: каждый час на грид-сайт автоматически запускаются тесты, чтобы проверить работоспособность всех сервисов. Если хотя бы один тест завершился неудачно, считается, что грид-сайт этот час был неработоспособен ($T_i=0$). Если все тесты выполнены успешно, то грид-сайт этот час был работоспособен ($T_i=1$). T_p определяется как сумма T_i за период T_o . На рис. 4 приведена средняя надежность по годам с 2009 по 2011 грид-сайта ОИЯИ и российских грид-сайтов РДИГ.



Рис 4. Показатели надежности российских грид-сайтов РДИГ и ОИЯИ с 2009 по 2011.

Важная информация для оценки работы грид-сайтов сосредоточена на портале статистического учета работы грид-сайтов (EGI Accounting Portal; <http://accounting.egi.eu/egi.php>):

- количество выполненных задач;
- количество часов, выданных пользователям, в различных единицах измерения (процессорное время, астрономическое время, нормализованное время в единицах NEPSpec06);
- средняя эффективность использования процессоров (отношение процессорного времени к астрономическому времени).

Главные показатели работы грид-сайтов, принятые в проекте WLCG, следующие:

- надежность грид-сайтов в процентах,
- количество нормализованных часов процессорного времени в единицах NEPSpec06, выданных пользователям виртуальных организаций.

Каждый из этих показателей очень важен, но всегда возникает вопрос, какой из грид-сайтов более эффективен, а для этого нужно иметь интегрированную оценку эффективности грид-сайтов.

Автором диссертации предложен базовый вариант интегрированной оценки эффективности грид-сайтов инфраструктуры WLCG, в которую входят основные и общедоступные показатели.

Базовая оценка эффективности (БОЭ) грид-сайтов:

$$\text{БОЭ} = \text{КН} * \text{СЭП} * \text{ПВ} / \text{КЯ} * (\text{ПВ} / \text{ОПВ})$$

КН – коэффициент надежности за указанный период времени,

СЭП - средняя эффективность использования процессоров, как отношение процессорного времени к астрономическому времени,

ПВ – количество нормализованных часов процессорного времени, выданных пользователям виртуальных организаций LHC (WLCG) на грид-сайте за указанный период,

ОПВ – количество нормализованных часов процессорного времени, выданных пользователям всех виртуальных организаций на грид-сайте за указанный период,

КЯ – количество вычислительных ядер, входящих в состав грид-сайта.

Например, для грид-сайта ОИЯИ за июль-август было выдано пользователям 30,371,000 часов в NEPSpec06 (из них пользователям LHC - 30,046,224).

Надежность за этот период составила 100%, поэтому КН=1.

СЭП=0.894.

КЯ - 2582 (количество процессорных ядер).

Таким образом, базовая оценка эффективности грид-сайта ОИЯИ за июль-август 2012 года составляет 10.52.

Был проведен анализ наиболее производительных грид-сайтов уровня Tier2 инфраструктуры WLCG, т.е. тех центров, которые за этот период предоставили пользователям виртуальных организаций LHC наибольшее количество часов процессорного времени в единицах NEPSpec06. Полученные результаты представлены в таблице 1.

Таблица 1.

Базовые оценки эффективности грид-сайтов уровня Tier2 инфраструктуры WLCG.

Грид-сайт	ПВ	ОПВ	КЯ	СЭП	КН	БОЭ
MWT2 (США)	59312824	59312824	4504	.881	0.995	11.54
GRIF (Франция)	58766864	65639928	6289	.829	1.0	8.65
DESY-НН (ФРГ)	46775568	61261172	6355	.908	0.99	8.67
WT2 (США)	45210192	45210192	7292	.883	0.985	5.39
IN2P3-CC-T2 (Франция)	37374520	37375044	3244	.822	0.99	9.38
AGLT2 (США)	33825440	33825440	4452	.822	0.975	6.09

INFN-LNL-2 (Италия)	30830072	30965604	2416	.833	1.0	10.68
JINR-LCG2	30046224	30371000	2582	.894	1.0	10.52
CYFRONET (Польша)	28059368	29248312	10656	.895	0.975	2.40
UKI-LT2-IC-HEP (Великобритания)	27701236	30477856	3084	.86	0.995	8.46
UFlorida (США)	23980912	23980912	2496	.877	1.0	8.43
SiGNET (Словения)	22969032	23033152	4416	.915	0.985	4.70
INFN-BARI (Италия)	22524940	23726584	4626	.758	0.84	3.27
prague1cg2 (Чехия)	22390428	23462516	3860	.757	0.925	4.26
UKI- SOUTHGRID- RALPP (Великобритания)	21032132	22906928	2872	.899	0.955	6.85
UKI- NORTHGRID- MAN-HEP (В)	19890948	23624760	2592	.89	0.9	7.30
TOKYO-LCG2 (Япония)	17981292	17981372	3712	.88	0.99	4.22
UKI-LT2-QMUL (Великобритания)	16764776	16929052	3496	.844	0.965	3.94
UKI-SCOTGRID- GLASGOW (Великобритания)	15913660	16076612	3880	.838	0.95	3.30
SFU-LCG2 (Канада)	15661156	15661212	4448	.907	0.99	3.16

Исходя из базовой оценки эффективности, можно сделать вывод, что грид-сайт ОИЯИ имеет высокий показатель и входит в число мировых лидеров по эффективности. Эта оценка могла быть еще выше, если учитывать тот факт, что часть процессорных ресурсов используется на решение задач локальных пользователей, в том числе параллельных приложений.

Предложенная оценка учитывает основные показатели эффективности грид-сайтов, но существуют много показателей, которые также оказывают влияние на эффективность.

Например, имеет смысл ввести коэффициент сложности грид-сайта, который зависит от количества ядер, количества поддерживаемых виртуальных организаций, количества задач и т.д.

Чем больше ядер, тем сложнее обеспечить их эффективную загрузку.

Чем больше задач, тем больше накладных расходов на их поддержку.

Чем большее разнообразие классов задач и поддерживаемых виртуальных организаций, тем ниже эффективность.

Этот коэффициент сложности является функцией нескольких переменных.

Кроме системы вычислений в каждом ресурсном центре есть система хранения, которая тоже влияет на общую эффективность грид-сайта. Для оценки эффективности системы хранения (ЭСХ) предложено учитывать следующие показатели:

$$ЭСХ = НСХ + СПФ$$

НСХ - нагрузка на систему хранения (среднее количество обменов в единицу времени)

СПФ – среднее количество переданных с/на грид-сайт файлов или данных в единицу времени (средний темп передачи данных).

В настоящий момент в полной мере информация о параметрах и нагрузке на систему хранения грид-сайтов пока не публикуется, кроме информации о передаче файлов с/на грид-сайт. В настоящее время ведется работа по проекту, в котором недостающая информация о нагрузке на систему хранения и передачи данных будет публиковаться. Как только информация начнет публиковаться, то в формуле эффективности грид-сайта будут учтены показатели эффективности системы хранения.

Анализ интегрированной оценки эффективности работы ресурсного центра за различные периоды (а также сравнение с аналогичными оценками для других грид-сайтов) дает возможность выделить показатели, по которым можно определить причины неэффективной работы грид-сайта.

Результаты этой главы подтверждают *четвертое защищаемое положение*:

Разработанная интегрированная оценка эффективности функционирования ресурсного центра в составе глобальной грид-инфраструктуры позволяет повысить объективность принятия решений по устранению недостатков в работе ресурсного центра и его развитию.

В пятой главе разработаны архитектурные принципы, методы и средства для реализации мониторинга распределенных вычислительных систем.

Ключевым элементом в обеспечении бесперебойной и надежной работы глобальных систем распределенных вычислений и входящих в их состав ресурсных центров являются качественные системы мониторинга, своевременно оповещающие о сбоях, позволяющие проводить комплексный анализ работы системы, дающие подробную картину функционирования и производительности отдельных элементов – объектов мониторинга.

Мониторинг предоставляет как общую информацию о функционировании инфраструктуры, систем, сервисов, пользователей, так и информацию об ошибках, нестандартных ситуациях и «узких местах».

Основная цель мониторинга заключается в слежении за состоянием и работоспособностью системы в целом и всех ее элементов в реальном времени, а также накопление и представление статистических данных для анализа функционирования системы.

Наряду с понятием «мониторинг» повсеместно используется термин «аккаунтинг» (Accounting), т.е. статистический учет использования ресурсов. Основной целью учета ресурсов в распределенных вычислительных системах является отслеживание и сбор статистики использования ресурсов, таких как дисковая память и процессорное время, а также данных о количестве задач, выполненных на этих ресурсах отдельными пользователями и целыми виртуальными организациями.

Система мониторинга должна быть масштабируемой, расширяемой, переносимой, доступной, независимой от модели доставки данных, иметь развитую клиентскую часть и обеспечивать анализ собранной информации. Исходя из этих требований, разработана архитектура и основные составные части системы мониторинга (рис.5):

- Подсистема сбора данных – осуществляет с заданной регулярностью опрос объектов, подлежащих мониторингу, для получения исследуемых значений параметров системы. Может также включать в себя анализ полученных данных с целью реагирования на критические ситуации.
- Подсистема оповещения – отвечает за уведомление лиц, ответственных за функционирование проверяемых объектов о нештатных ситуациях и иных значимых событиях, возникающих в системе.
- Подсистема хранения – отвечает за накопление, хранение, архивирование данных о результатах проверок. Может включать в себя, например, компоненты для работы с базой данных (БД), программные средства для усреднения значений за некоторый отрезок времени для уменьшения объема хранимой информации и т.п.
- Подсистема анализа данных – включает компоненты, производящие исследования данных, получаемых системой, поиск закономерностей, сбор статистики и тому подобные операции.
- Подсистема вывода – отвечает за представление информации о работе системы и результатов проверок в виде, удобном для восприятия пользователем.
- Подсистема коррекции – в том случае, если предусмотрена возможность выполнения системой некоторых действий для устранения возникших нештатных ситуаций, данная подсистема будет включать компоненты для выбора и осуществления подходящих действий в соответствии с типом проблемы и другими параметрами.

Сформулированные требования и предложенная структура системы мониторинга распределенных вычислительных систем имеют универсальный характер и могут быть применимы к различным архитектурам (кластеры, суперкомпьютеры, грид-инфраструктуры, облачные вычисления и т.д.).

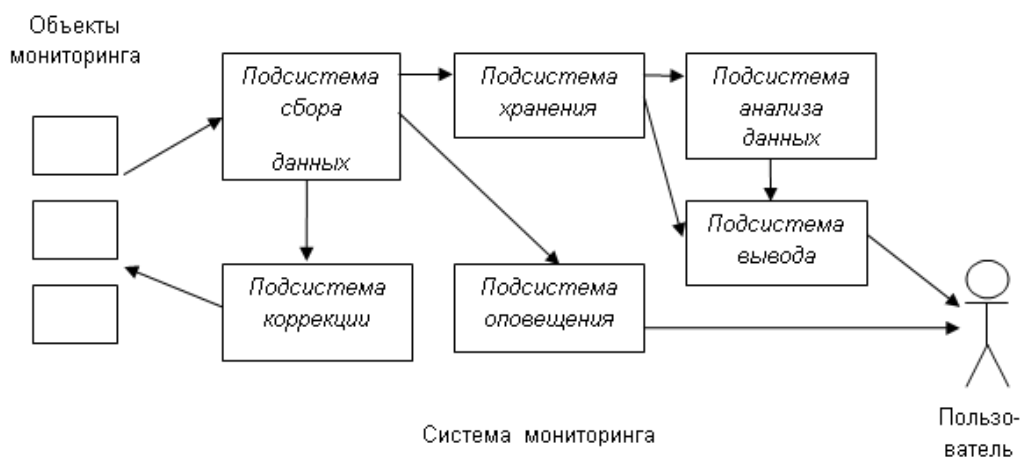


Рис 5. Структура системы мониторинга

Распределенные вычислительные системы – очень сложный объект, который состоит из большого числа компонент, разделенных на несколько уровней. Например, если рассматривать глобальную грид-инфраструктуру, осуществляющую скоординированное взаимодействие множества географически распределенных грид-сайтов (ресурсных центров), то мы можем выделить несколько уровней мониторинга:

- глобальный мониторинг с отображением состояний сервисов, взаимодействия ресурсных центров и активности пользователей в реальном масштабе времени;
- мониторинг виртуальной организации как динамического объединения пользователей, ресурсов и сервисов: контроль и информация о заданиях пользователей, их распределении между грид-сайтами, о взаимоотношениях отправителей и исполнителей, о распределении и использовании систем хранения данных, а также слежение и контроль передачи информации между грид-сайтами и т.д.;
- локальный мониторинг с отображением состояний инфраструктуры, ресурсов, сервисов, пользователей каждого грид-сайта (или ресурсного центра), входящего в грид-инфраструктуру.

В глобальной распределенной системе можно выделить следующие уровни мониторинга:

- система мониторинга ресурсных центров, или грид-сайтов;
- система мониторинга глобальных грид-сервисов;
- система мониторинга виртуальных организаций, приложений и пользователей (с возможностью мониторинга выполнения задач).

Базовым элементом грид-инфраструктуры является грид-сайт, или ресурсный центр, который представляет собой сложный программно-аппаратный комплекс, поэтому система мониторинга ресурсного центра имеет несколько уровней:

- Аппаратный, который выполняет мониторинг инфраструктурных инженерных систем (система обеспечения бесперебойного питания,

система климат – контроля и другие), аппаратного комплекса (серверы, вычислительные узлы, системы хранения информации, сетевое оборудование и т.д.).

- Сетевой уровень выполняет мониторинг работы компьютерной сети ресурсного центра, осуществляет постоянное наблюдение за сетью, сетевыми элементами, портами в поисках сбоев, нарушений работы или перегрузки сети или отдельных ее элементов для оперативного их устранения.
- Сервисный уровень выполняет мониторинг использования ресурсов грид-сайта (количество активных задач, длина очереди задач, выделенные и использованные дисковые ресурсы, количество обменов с дисками, загрузка процессоров и другие показатели), мониторинг и учет пользователей, виртуальных организаций, сервисов и приложений.

Все эти уровни мониторинга должны функционировать на общих принципах, использовать единую платформу, единый репозиторий для оперативного и эффективного контроля работоспособности ресурсного центра, оперативно реагировать на выявленные проблемы.

Автором диссертации разработаны методика и средства для проектирования многоуровневой системы мониторинга, удовлетворяющей перечисленным выше требованиям.

Большое значение имеет разработка соответствующих методов обнаружения и локализации сбоев отдельных сервисов, компонентов и других объектов распределенной системы, а также выявления взаимосвязей между возникающими проблемами для их прогнозирования и скорейшего устранения.

В этой главе диссертации автором систематизированы существующие и обоснованы методы и механизмы сбора данных об объектах мониторинга, способы организации хранения этих данных, алгоритмы и методы анализа полученных данных для прогнозирования поведения системы, а также средства визуализации результатов мониторинга.

Под руководством автора диссертации разработаны системы мониторинга на основе концепции, архитектуры, методов и средств, представленных в диссертации:

- Создана система мониторинга и сбора статистики для Российского грид-сегмента РДИГ глобальной инфраструктуры WLCG. Данная система отслеживает состояние грид-инфраструктуры РДИГ по множеству параметров, а также позволяет получать статистическую информацию с большой степенью детализации в наглядной форме.
- Разработана и функционирует система локального мониторинга грид-сайта ОИЯИ, которая обеспечивает корректную и надежную работу вычислительного комплекса и предоставляет актуальную оперативную информацию о его работе на более высокие уровни мониторинга. Данные, предоставляемые данной системой, имеют большое значение как для сетевых администраторов, ответственных за предоставление

оборудования и каналов, так и для разработчиков и пользователей грид-сервисов.

- Существенный вклад внесен в развитие системы мониторинга для виртуальных организаций ЛНС, разрабатываемой и поддерживаемой в отделении информационных технологий в ЦЕРН (система Dashboard) в рамках совместного проекта РФФИ-ЦЕРН “Грид мониторинг для эффективного функционирования виртуальных организаций”. Например, разработана система визуализации для динамического мониторинга в реальном времени функционирования грид-инфраструктуры на платформе Google Earth.
- Разработана система мониторинга сервиса FTS для контроля качества передачи файлов в грид-среде.
- Разработана система мониторинга грид-инфраструктуры с веб-интерфейсом для проекта СКИФ-Грид.
- Создана система мониторинга и учета ресурсов грид-инфраструктуры национальной нанотехнологической сети ГридННС.
- Разработана система статистического учета для анализа использования системы хранения информации dCache.
- Разработана система мониторинга нового инфраструктурного уровня Tier3-центров для анализа данных экспериментов ЛНС.
- Разработана система мониторинга и учета ресурсов Российской грид-сети (РГС).
- Разработан, в рамках совместного проекта РФФИ-ЦЕРН «Глобальная система мониторинга передачи данных в инфраструктуре проекта WLCG», прототип универсальной системы мониторинга, способной собирать подробную информацию о каждой передаче файлов, независимо от метода осуществления передачи, статуса ресурсного центра в многоярусной иерархии (Tier-0, Tier-1, Tier-2, Tier-3) и принадлежности данных определенной виртуальной организации.

Результаты этой главы подтверждают *пятое защищаемое положение*:

Разработанная архитектура и созданная на ее основе комплексная система грид-мониторинга (ресурсов, сервисов, задач, виртуальных организаций, пользователей) повышают эффективность и надежность функционирования грид-инфраструктур и служит основой для прогнозирования их развития.

В **Заключении** сформулированы основные выводы и результаты диссертации.

На основе разработанной в диссертации концепции, архитектуры, структурной схемы научно-исследовательского высокопроизводительного вычислительного комплекса создан ресурсный центр в составе грид-инфраструктуры, имеющий высокий мировой рейтинг по показателям производительности и надежности, активно востребованный пользователями грид: только за 8 месяцев 2012 года в ОИЯИ было выполнено более 5

миллионов задач, а потребление процессорного времени при этом составило более 100 миллионов часов в единицах NEPSpec06.

Основные результаты, выводы и рекомендации, полученные в диссертации, используются при реализации крупных национальных и международных проектов (создание российской национальной нанотехнологической грид-сети, создание автоматизированной системы обработки данных экспериментов на ЛНС уровня Tier1 и обеспечения грид-сервисов для распределенного анализа этих данных и др.).

СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ АВТОРА ПО ТЕМЕ ДИССЕРТАЦИИ

*Работы из перечня ведущих рецензируемых научных изданий,
рекомендованных ВАК РФ для публикации основных результатов
диссертаций:*

1. В.В.Кореньков, В.В.Мицын, П.В.Дмитриенко: Архитектура системы мониторинга центрального информационно-вычислительного комплекса ОИЯИ // Информационные технологии и вычислительные системы, 2012, №3, стр. 3-14
2. В.В.Кореньков, Н.А.Кутовский, Р.Н.Семёнов: Опыт адаптации прикладных программных пакетов для работы в грид-средах // Компьютерные исследования и моделирование, 2012, Том 4, № 2, стр. 339-344.
3. E. Dushanov, Kh. Kholmurodov, V.Korenkov, G. Aru, W. Smith, Y. Ohno, T. Narumi, G. Morimoto, M. Taiji, K. Yasuoka: JINR CICC in Computational Chemistry and Nanotechnology Problems: DL_POLY Performance for Different Communication Architectures // Particles & Nuclei, Letters, 2009, No. 3, p. 251–259.
4. В.В.Кореньков, Н.А.Кутовский: Инфраструктура обучения grid-технологиям // Открытые системы, 2009, №10, стр. 48-51.
5. В.В.Кореньков, А.В.Ужинский: На пути к адаптивности grid // Открытые системы, 2009, №9, стр.18-19.
6. В.В.Кореньков, А.В.Ужинский: Система мониторинга сервиса передачи данных (FTS) проекта EGEE/WLCG // Вычислительные методы и программирование: Новые вычислительные технологии, 2009, том 10, стр.96-100.
7. Gh.Adam, V.Korenkov et al: Consistent Performance Assessment of Multicore Computer Systems// Romanian Journal of Physics, 2008, Vol. 53, No. 9-10, p. 3-8.
8. Gh.Adam, V.Korenkov et al: Performance assessment of the SIMFAP parallel cluster at IFIN-HH Bucharest // Romanian Journal of Physics, 2008, Vol. 53, No. 5-6, p. 665-677.

9. В.В.Кореньков, А.В.Ужинский: Архитектура сервиса передачи данных в grid // Открытые системы», 2008, №2, с.52-56.
10. В.А.Ильин, В.В.Кореньков, А.А.Солдатов: Российский сегмент глобальной инфраструктуры LCG // Открытые системы, 2003, №1, с. 56-60.
11. В.В.Кореньков, Е.А.Тихоненко: Организация вычислений в научных отраслях // Открытые системы, 2001, №2, с.30-35.
12. В.В.Кореньков, Е.А.Тихоненко: Концепция GRID и компьютерные технологии в эру LHC // Физика элементарных частиц и атомного ядра, 2001, т. 32, вып.6, с.1458-1493.
13. В.В.Мицын, В.В.Кореньков, К.Ф.Окраинец, Р.Г.Позе, М.Ю.Попов: О развитии информационно-вычислительной инфраструктуры ОИЯИ // Краткие сообщения ОИЯИ, 1995, №2[70], с.5-10.

Публикации в трудах конференций:

14. V.Boirisovskij, A.Ershov, V.Korenkov, Yu.Stolyarsky, P.Sychev: JINR Information Infrastructure Development Concept // Proc. of Int. Conf. "Programming and Mathematical techniques in Physics", World Scientific, Singapore, 1994, p.129-137.
15. V.Korenkov, R.Pose: Status and trend of development of networking and computing at JINR // Proc. of XVII International Symposium on Nuclear Electronics, Varna, 1997, p.168-171.
16. V.Korenkov: Status and Perspectives of JINR Computing Centre // Proc. of Int. Conf. «High Performance Computing on Hewlett-Packard Systems», Zurich, Switzerland, 1998, p. 224-227.
17. В.В.Кореньков, Е.А.Тихоненко: О необходимости создания российского GRID-сегмента // Сборник трудов Всероссийской научной конференции «Интернет в научных исследованиях», изд. МГУ, 2000, с.86-90.
18. А.В.Жучков, В.А.Ильин, В.В.Кореньков: Некоторые аспекты создания глобальной системы распределенных вычислений в России // Труды Всероссийской научной конференции "Высокопроизводительные вычисления и их приложения", Черногловка, изд. МГУ, 2000, с.227-229.
19. В.А.Ильин, В.В.Кореньков: Создание Российского сегмента европейской инфраструктуры EU DataGrid // Труды Всероссийской конференции «Электронные библиотеки», Дубна, 2002, с.239-248.
20. V. Pyin, V. Korenkov: Development of Russian Grid Segment in the frames of EU DataGRID, LCG and EGEE projects // Proc. of the XIX Int. Symposium on Nuclear Electronics & Computing (NEC"2003, Varna, Bulgaria), Dubna, 2004, p.154-156.
21. В.А.Ильин, В.В.Кореньков: Участие российских центров в европейских проектах LCG (LHC Computing GRID) и EGEE (Enabling GRIDS for E-science in Europe) // Сборник 10-й конференции представителей региональных научно-образовательных сетей ("RELARN-2004"), С-Петербург, 2004, с.302-305.

22. Y.Bugaenko, V.Korenkov et al: Participation of Russian sites in the Data Challenge of ALICE experiment in 2004 // Proc. of Int. Conf. "Computing in high energy and nuclear physics (CHEP'2004)", Interlaken, Switzerland, CERN-2005-002-V-2, 2004, p. 802-805.
23. В.В. Кореньков, П.С. Березовский, В.В. Галактионов, А.П. Демичев, В.Е. Жильцов, В.А. Ильин, В.Н. Коваленко, Д.А. Корягин: Создание центра базовых ГРИД-сервисов нового поколения федерального уровня // Труды конференции «Телематика-2006», С-Петербург, 2006
24. V.Korenkov, A.Nechaevskiy, A. Uzhinsky: System for simulation of networks based on GRID – technologies // Труды II Международной конференции "Распределенные вычисления и ГРИД-технологии в науке и образовании", Дубна, 2006, с.99-102.
25. V.Ilyin, V.Korenkov, A.Soldatov: RDIG (Russian Data Intensive Grid) e-Infrastructure // Proc. of XXI Int. Symposium of Nuclear Electronics & Computing (NEC'2007, Varna, Bulgaria), Dubna, 2008, p. 233-238.
26. V.Ilyin, V.Korenkov, A.Kryukov, Yu.Ryabov, A.Soldatov: Russian Data intensive Grid (RDIG): current status and perspectives towards national Grid initiative // Proc. of Int. Conf. "Distributed computing and Grid technologies in science and education, GRID-2008", Dubna, 2008, p. 100-108.
27. A.Uzhinsky, V.Korenkov: Monitoring system of the file transfer service for EGEE/WLCG // Proc. of Int. Conf. "Distributed computing and Grid technologies in science and education, GRID-2008", Dubna, 2008, p. 197-201.
28. A.Ayriyan, Gh.Adam, S.Adam, V.Korenkov, A.Lutsenko, V.Mitsyn: CICC JINR Cluster 2008 Performance Improvement // Proc. of Science XII Advanced Computing and Analysis Techniques in Physics Research (ACAT'2008), Erice, Italy, 2008.
29. S.Belov, V.Korenkov: Experience in development of Grid monitoring and accounting systems in Russia // Proc. of XXII Int. Symposium on Nuclear Electronics & Computing (NEC'2009, Varna, Bulgaria), Dubna, 2010, p.75-80.
30. V. Ilyin, V.Korenkov, A.Soldatov: RDIG (Russian Data Intensive Grid) e-Infrastructure: status and plans // Proc. of XXII Int. Symposium on Nuclear Electronics & Computing (NEC'2009, Varna, Bulgaria), Dubna, 2010, p.150-153.
31. V.Korenkov: Grid activities at the Joint Institute for Nuclear Research // Proc. of the 4th Intern. Conf. «Distributed Computing and Grid-Technologies in Science and Education, GRID-2010», Dubna, 2010, p. 142-147.
32. V. Gavrilov, I. Golutvin, V. Korenkov, E. Tikhonenko, S. Shmatov, V.Zhiltsov, V. Ilyin, O. Kodolova, L.Levchuk: Status of the RDMS CMS Computing // Proc. of the 15th Annual RDMS CMS Collaboration Conference, Kharkov, KIPT, 2011, p.140-144.

Публикации в прочих научных изданиях:

33. И.А.Голутвин, В.В.Кореньков, А.А.Лаврентьев, Р.Г.Позе, Е.А.Тихоненко: Поддержка компьютинга CMS в ОИЯИ // Сообщение ОИЯИ, Д11-98-122 Дубна, 1998, с.10.
34. V.Korenkov, N.Zaikin, T.Strizh: Computing and networking at JINR // Czechoslov.J.Phys. , 2001, № 51, с. 355-362.
35. В.В.Кореньков, Е.Н. Черемисина, А.О. Лацис, Ю.А. Крюков: Современные тенденции развития информационно-вычислительных систем (GRID-технологии) // Избранные труды университета "Дубна", Дубна, 2004, с. 121-147.
36. В.Ильин, В.Кореньков: Компьютерная грид-инфраструктура коллаборации RDMS CMS / сб.: «В глубь материи: физика XXI века глазами создателей экспериментального комплекса на Большом адронном коллайдере в Женеве» – М: Этерна, 2009, с. 361-372.
37. В.В.Кореньков, А.В.Нечаевский: Пакеты моделирования DataGrid // Системный анализ в науке и образовании, Электронный научный журнал, 2009, Выпуск 1, URL:<http://sanse.ru>
38. О.О.Бунецкий, Е.С.Горбенко, В.В.Кореньков и др.: Подготовка Грид-инфраструктур ЛИТ ОИЯИ и НИЦ ХФТИ к анализу данных эксперимента CMS (ЦЕРН) // Сообщение ОИЯИ P11-2010-11, Дубна, 2010, с.9.
39. J.Andreeva, D.Benjamin, S.Campana, A.Klimentov, V. Korenkov, D.Oleynik, S.Panitkin, A.Petrosyan: Tier-3 Monitoring Software Suite (T3MON) proposal // CERN Note, ATLASOFT-PUB-2011-001, 2011, p.7.
40. S. Belov, I. Kadochnikov, V. Korenkov, M. Kutouski, D. Oleynik, A. Petrosyan on behalf of the ATLAS Collaboration: VM-based infrastructure for simulating different cluster and storage solutions used on ATLAS Tier-3 sites // CERN Note, ATLASOFT-PROC-2012-057, 2012, p.4.
41. А. Климентов, В. Кореньков: Распределенные вычислительные системы и их роль в открытии новой частицы // Суперкомпьютеры, 2012, №3 (11), стр. 7-11.