

*На правах рукописи*

**КОРСУКОВ АЛЕКСАНДР СЕРГЕЕВИЧ**

**Инструментальный комплекс для разработки и применения  
гетерогенных распределенных вычислительных сред**

05.13.11 – Математическое и программное обеспечение  
вычислительных машин, комплексов и компьютерных сетей

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата технических наук

Иркутск – 2009

Работа выполнена в Учреждении Российской академии наук Институте динамики систем и теории управления Сибирского отделения Российской академии наук (ИДСТУ СО РАН).

Научный руководитель: кандидат технических наук, доцент  
**Феоктистов Александр Геннадьевич**

Официальные оппоненты: доктор физико-математических наук,  
профессор  
**Корольков Юрий Дмитриевич**

кандидат технических наук, доцент  
**Абасов Николай Викторович**

Ведущая организация: **Институт вычислительного  
моделирования СО РАН**  
(г. Красноярск)

Защита состоится 26 февраля 2009 г. в 14 часов на заседании диссертационного совета Д 003.021.01 при Институте динамики систем и теории управления СО РАН по адресу: 664033, г. Иркутск, ул. Лермонтова, 134.

С диссертацией можно ознакомиться в библиотеке ИДСТУ СО РАН.

Автореферат разослан 24 января 2009 г.

Ученый секретарь  
диссертационного совета  
д.ф.-м.н.

А.А. Щеглова

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** В последнее время для решения ресурсоемких фундаментальных и прикладных задач создаются различные распределенные и параллельные вычислительные системы и среды (см., например, работы В.С. Бурцева, В.В. Воеводина, В.Н. Коваленко, В.В. Корнеева, Д.А. Корягина, А.О. Лациса, В.В. Топоркова и др.). В частности, активно применяется технология создания специальной вычислительной сети, получившей название Grid<sup>1</sup>. Эта технология базируется на интеграции географически распределенных информационно-вычислительных и коммуникационных ресурсов и их совместном использовании в процессе вычислений. Инфраструктура Grid включает объединенные телекоммуникационной средой средства вычислений и обработки информации (суперкомпьютеры, вычислительные кластеры, отдельные персональные компьютеры, системы хранения данных и др.) и системное программное обеспечение (ПО), предназначенное для управления процессом выполнения заданий пользователей в этой среде.

Grid, организованная на базе кластеров, в большинстве случаев ее использования обеспечивает возможность удаленного доступа к ресурсам (узлам) вычислительной сети и позволяет определить вычислительные возможности конкретного узла (число процессоров, объем оперативной памяти и т.п.) и степень его работоспособности, а также выполнить на этом узле некоторое независимое задание или обработать один из блоков данных при решении «большой» задачи, допускающей распараллеливание по данным<sup>2</sup>.

Однако существуют типы научно-исследовательских задач, для решения которых необходимы дополнительные возможности:

- обеспечение вычислительных услуг нетиражируемых программных комплексов, размещенных в узлах Grid;
- выполнение ряда взаимозависимых заданий, составляющих процесс решения одной общей задачи и требующих интеграции распределенных вычислительных ресурсов на основе автоматического планирования последовательности их использования.

Анализ систем управления заданиями (СУПЗ) для кластеров показывает, что как правило в этих системах реализовано централизованное управление автономными заданиями. Между тем, решение задач с большим числом подзадач и информационно-логических связей между ними в значительной степени увеличивает нагрузку на управляющий компьютер кластера и может привести к снижению работоспособности всей вычислительной системы в целом<sup>3</sup>. Кроме того, формы ведения вычислительных работ в Grid, обусловленные используемым для ее организации сложным системным ПО, во многом ориентированы на

---

<sup>1</sup> Foster I. The Anatomy of the Grid: Enabling Scalable Virtual Organizations / I. Foster, C. Kesselman, S. Tuecke // Intern. J. of High Performance Computing Applications. – 2001. – Vol. 15, № 3. – P. 200-222.

<sup>2</sup> См., например, работу: Воеводин В.В. Решение больших задач в распределенных вычислительных средах / В.В. Воеводин // Автоматика и телемеханика. – 2007. – № 5. – С. 32-45.

<sup>3</sup> Каляев И.А. Распределенные системы планирования действий коллективов роботов / И.А. Каляев, А.Р. Гайдук, С.Г. Капустян. – М.: Янус-К, 2002. – 292 с.

специалистов с достаточно высоким уровнем квалификации в области системного программирования. Эти обстоятельства сдерживают широкое применение Grid специалистами-прикладниками и актуализируют вопросы создания «дружественных» средств доступа к ресурсам Grid и децентрализованных способов управления вычислительным процессом<sup>4</sup>.

**Цель работы** состоит в разработке способа создания гетерогенных распределенных вычислительных сред<sup>5</sup> (РВС), предоставляющих широкий выбор средств для решения ресурсоемких научно-исследовательских задач различных типов и обеспечивающих возможность интеграции с другими РВС. Достижение указанной цели осуществляется путем реализации программного комплекса, предназначенного для инструментальной поддержки основных этапов построения и применения РВС.

**Объектом исследования** является организация вычислительного процесса решения ресурсоемких фундаментальных и прикладных научных задач в РВС.

**Предмет исследования** – построение Grid на базе вычислительных кластеров. Важным требованием к такой Grid является обеспечение возможности включения в ее вычислительную инфраструктуру кластеров, организованных на разных программно-аппаратных платформах.

**Методы исследования.** При решении поставленных задач использовались методы системного программирования и искусственного интеллекта для организации, планирования и применения распределенных вычислительных ресурсов, а также методы разработки веб-приложений.

**На защиту выносятся следующие основные результаты:**

1. *Способ децентрализованного управления потоками заданий*, позволяющий (в отличие от известных) распределять задания в вычислительной среде в зависимости от типов решаемых задач. Для каждого типа задач разработан специализированный планировщик заданий.

2. *Средства организации унифицированного веб-ориентированного доступа* к распределенным вычислительным ресурсам. В том числе язык заданий и командный язык, обеспечивающие их настройку на особенности функционирования различных систем управления заданиями, используемых в узлах распределенной вычислительной среды.

3. *Системная архитектура и алгоритмы функционирования инструментального комплекса*, обеспечивающего поддержку основных этапов построения и применения гетерогенных распределенных вычислительных сред для решения ресурсоемких научно-исследовательских задач.

**Научная новизна.** Разработан способ организации распределенных вычислений, отличающийся от известных обеспечением таких дополнительных возможностей, как формирование потоков заданий в зависимости от типов решаемых задач и децентрализованное управление этими потоками в процессе

---

<sup>4</sup> См., например, работу: Durfee E.H. Distributed problem solving and planning / E.H. Durfee // Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence / Ed. by G. Weiss. – Cambridge: MIT Press, 1999. – P. 121-164.

<sup>5</sup> Под РВС следует понимать как отдельные вычислительные кластеры, так и Grid.

выполнения заданий. В основе этого способа лежат специализированные языковые и инструментальные средства, предназначенные для быстрой и гибкой настройки веб-ориентированного доступа к ресурсам гетерогенной РВС.

**Практическая значимость.** Разработанное ПО предоставляет пользователям возможность самостоятельно (без участия высококвалифицированных системных программистов) описывать исследуемую предметную область и программно-аппаратную часть РВС, а также формировать задания для решения своих задач и проводить вычислительные эксперименты в РВС. Тем самым сокращаются сроки и повышается эффективность проведения научно-исследовательских работ в целом.

Разработанные в рамках выполнения диссертационной работы инструментальные средства использованы при создании Grid ИДСТУ СО РАН. Эти программные средства зарегистрированы в Федеральной службе по интеллектуальной собственности, патентам и товарным знакам [3, 4] и применяются для проведения экспериментальных расчетов по плановым НИР в ИДСТУ СО РАН, а также в учебном процессе в Иркутском государственном университете: Институте математики, экономики и информатики и Международном институте экономики и лингвистики.

Исследование, разработка и применение программных средств, представленных в диссертации, выполнялись в рамках проекта СО РАН № 6 «Планирование и оптимизация схем решения задач в распределенной мультиагентной вычислительной среде» программы фундаментальных исследований Президиума РАН № 21 «Разработка фундаментальных основ создания научной распределенной информационно-вычислительной среды на основе технологий GRID» (2004 г.); интеграционного проекта №3 СО РАН «Методы, технологии и инструментальные средства создания вычислительной инфраструктуры в Internet» (2003–2005 гг.); проекта СО РАН 3.2.6 «Интегрированные информационно-вычислительные и коммуникационные ресурсы: интеллектуальные методы организации, автоматизации разработки и применения» (2004–2006 гг., блок 1 «Распределенная вычислительная САТУРН-среда»); проекта СО РАН «Разработка научных основ распределенной информационно-аналитической системы на основе ГИС и Веб-технологий для междисциплинарных исследований» междисциплинарной программы 4.5.2 (2007–2009 гг., блок 2 «Интеллектуальные методы и инструментальные средства разработки и комплексирования распределенных информационно-вычислительных ресурсов»); проекта РФФИ № 08-07-00163 «Технологии интеллектуального анализа данных и высокопроизводительных информационно-вычислительных ресурсов для поддержки междисциплинарных фундаментальных исследований в области геоэкологии и природопользования» (2008–2010 гг.).

**Достоверность результатов диссертации.** Основные результаты диссертации базируются на исследованиях отечественных и зарубежных ученых и подтверждаются корректным использованием современных теоретических и экспериментальных методов их обоснования, а также успешным применением программного комплекса для решения практических задач.

**Апробация работы.** Основные результаты работы представлены на III Всероссийской молодежной конференции «Под знаком  $\Sigma$ » (Омск, 2005 г.), на IV Всероссийской конференции «Математика, информатика, управление» (Иркутск, 2005 г.), на VI–VIII Школах-семинарах молодых ученых «Математическое моделирование, управление и информационные технологии» (Иркутск, 2005 г., 2005 г., 2006 г.), на VI Международной научно-практической конференции «Компьютерные технологии в науке, производстве, социальных и экономических процессах» (Новочеркасск, 2005 г.), на Всероссийской конференции «Инфокоммуникационные и вычислительные технологии и системы» (Энхалук, 2006 г.), на XI Байкальской Международной конференции «Информационные и математические технологии в научных исследованиях» (Иркутск, 2006 г.), на II Международной конференции «Параллельные вычисления и задачи управления» (Москва, 2006 г.), на V Школе-семинаре «Распределенные и кластерные вычисления» (Красноярск, 2006 г.), на Международной научной конференции РАУТ'2007 «Параллельные вычислительные технологии» (Челябинск, 2007 г.), на XII Байкальской Всероссийской конференции с международным участием «Информационные и математические технологии в науке и управлении» (Иркутск, 2007 г.), на VI Сибирской научной школе-семинаре с международным участием «Компьютерная безопасность и криптография» (Горно-Алтайск, 2007 г.), на Международной научно-практической конференции «Информационные технологии в образовании» (Иркутск, 2007 г.), на Международной конференции «Геоинформатика: технологии, научные проекты» (Иркутск, 2008 г.), на Школе-семинаре молодых ученых «Информационные технологии и моделирование социальных эколого-экономических систем» (Иркутск, 2008 г.), а также неоднократно на семинарах ИДСТУ СО РАН.

**Публикации и личный вклад автора.** Результаты диссертации отражены в 15 научных работах (в том числе 1 статья [1] в журнале, рекомендованном ВАК для опубликования научных результатов диссертации на соискание ученой степени доктора или кандидата наук). В публикациях все результаты, связанные с алгоритмизацией, программной реализацией и вычислительным экспериментом на ЭВМ, получены автором лично. Результаты, относящиеся к разработке моделей и методов организации распределенных вычислений, получены совместно с А.Г. Феоктистовым и являются неделимыми. Из совместных работ с А.А. Александровым, В.И. Дмитриевым и А.В. Лариной в диссертацию включены результаты, принадлежащие лично автору.

**Структура работы.** Диссертация состоит из введения, трех глав, заключения, библиографии из 79 наименований, списка сокращений и 6 приложений. Общий объем работы – 147 страниц, из которых 120 страниц основного текста, включающего 28 рисунков и 12 таблиц.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обоснована актуальность темы диссертации, сформулирована цель исследования, показана научная новизна и практическая значимость полученных результатов, приведена структура работы.

**Первая глава** диссертации посвящена анализу современных способов и программных средств организации и применения распределенных вычислений.

В первом разделе главы приведены основные понятия, разъяснены назначение и сущность таких вычислительных структур, как кластер, Grid и PBC. В общем случае эти структуры предназначены для интеграции и использования распределенных вычислительных ресурсов.

Во втором разделе главы рассмотрены технические аспекты создания кластеров и основные вопросы, связанные с реализацией пользовательского веб-интерфейса к ним. В большинстве случаев использование вычислительных ресурсов кластеров не рассчитано на рядовых пользователей и требует привлечения системных программистов с высоким уровнем квалификации. Таким образом, создание простого и гибкого веб-интерфейса существенно облегчает процесс запуска и управления заданиями на вычислительных ресурсах кластера и значительно сокращает время, затрачиваемое неопытными пользователями на освоение кластерных систем.

Третий раздел главы посвящен анализу существующих СУПЗ и формулировке основных требований к их функционированию.

В четвертом разделе представлена архитектура СУПЗ Condor и обоснован выбор этой системы для организации кластеров невыделенных рабочих станций, вычислительные узлы которых представляются обычными персональными компьютерами.

В пятом разделе главы исследованы вопросы организации Grid, проведен анализ существующего ПО Grid и выполнен обзор ряда Grid-проектов. На основе проведенного анализа конкретизирован способ построения Grid, сформулированы возникающие при этом проблемы и определены пути их решения (см. рис. 1).

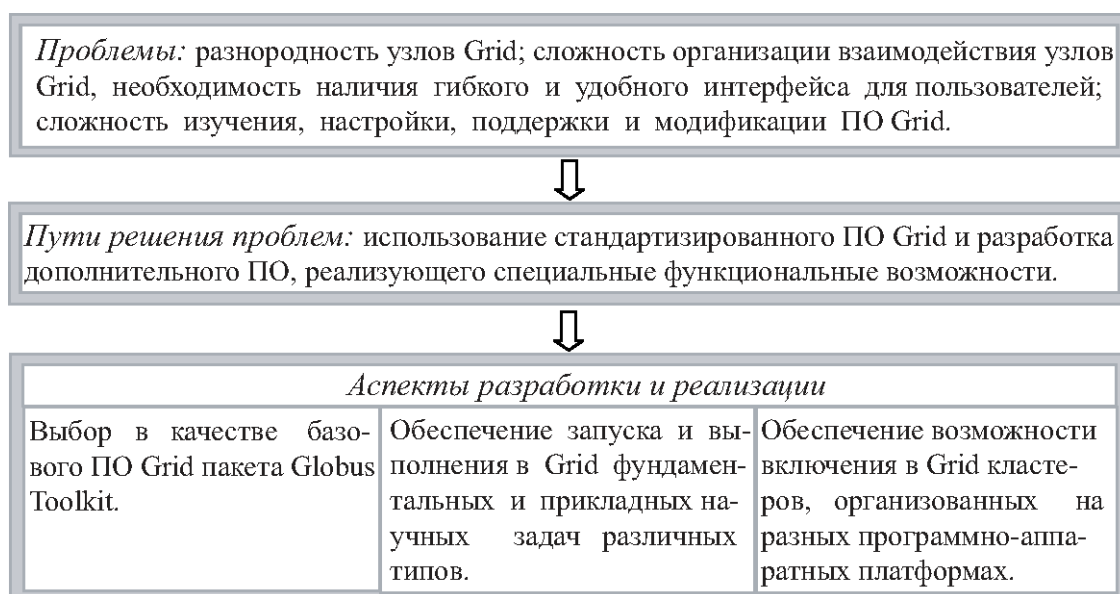


Рис. 1. Способ построения Grid

В последнем разделе главы проведен анализ технологий создания веб-приложений и обоснован выбор средств для реализации веб-ориентированного

доступа к ресурсам РВС.

**Во второй главе** представлены разработанные автором способы и средства построения гетерогенных РВС.

В первом разделе главы предложена модель РВС, обеспечивающая накопление знаний о вычислительных ресурсах, а также администрирование, планирование и применение этих ресурсов при решении пользовательских задач. В рассматриваемой модели выделены следующие основные уровни:

- пользовательский уровень (*UL*), включающий в себя: способы и средства доступа к вычислительным узлам РВС; множество пользователей РВС, их классификацию по категориям и правам доступа к РВС; множество заданий, запускаемых пользователями в РВС;
- уровень ПО (*SWL*), отражающий свойства и характеристики программных приложений, запуск которых требуется при выполнении заданий;
- аппаратный уровень (*HWL*), определяющий характеристики и комплектующие вычислительных узлов и коммутирующих устройств РВС;
- уровень планирования вычислений и загрузки ресурсов (*PLL*), обеспечивающий выбор необходимых ресурсов РВС для запуска заданий;
- вычислительный уровень (*CL*), отображающий процесс выполнения заданий на конкретных вычислительных узлах РВС.

Таким образом, модель РВС можно представить в виде структуры  $M = \langle UL, SWL, HWL, PLL, CL \rangle$ . Разные уровни модели РВС содержат свои наборы объектов. Для каждого объекта определены его атрибуты.

Для обозначения характеристик задач, важных в плане формирования и выполнения заданий в РВС, будем использовать следующую структуру:  $(a / b / c / d)$ , где символы  $a$ ,  $b$ ,  $c$  и  $d$  имеют следующее назначения:

- $a$  – характеристика алгоритма решения задачи с точки зрения наличия подзадач; фиксированными значениями для  $a$  являются 1 (отсутствие подзадач) и  $K$  (наличие подзадач);
- $b$  – характеристика алгоритма решения задач с точки зрения наличия в нем параллелизма; фиксированными значениями для  $b$  являются  $L$  (крупноблочный параллелизм алгоритма),  $F$  (мелкозернистый параллелизм алгоритма) и  $S$  (последовательный алгоритм); недопустимыми комбинациями значений  $a$  и  $b$  являются  $(a = 1 / b = L / \dots)$  и  $(a = K / b = F / \dots)$ ;
- $c$  – характеристика процесса решения задачи с точки зрения необходимости выполнения многовариантных расчетов; фиксированными значениями для  $c$  являются 1 (один единственный вариант данных) и  $N$  (наличие множества вариантов данных);
- $d$  – характеристика процесса решения задачи с точки зрения места размещения выполняемого программного приложения; фиксированными значениями для  $d$  являются  $G$  (выполнение приложения, размещенного в узлах РВС) и  $R$  (выполнение удаленного приложения).

Определим основные типы задач для РВС:

- $(a = 1 / b \neq L / c / d = R)$  – стандартные задачи (скомпилированные программы пользователей), для выполнения которых используется штатный



набор функций СУПЗ;

- $(a / b / c = N / d)$  – многовариантные задачи, для решения которых необходимо выполнение программы с различными наборами данных;
- $(a / b / c / d = G)$  – задачи, в которых требуется выполнение приложений, размещенных в узлах РВС;
- $(a = K / b = L / c / d)$  – взаимосвязанные задачи, требующие выполнения набора взаимосвязанных заданий, включенных в процесс решения одной общей задачи.

Тип решаемой задачи определяет соответственно тип задания, т.е. задает формат паспорта этого задания и способ его обработки (режим запуска) в РВС.

Во втором разделе приведена схема взаимодействия пользователя с узлами РВС, представленными вычислительными кластерами (рис. 2). В рамках приведенной схемы в качестве интеллектуальной управляющей надстройки к СУПЗ, установленных в узлах РВС, используется разработанная автором диссертационной работы программная система Web-Interface Manager (WIM).

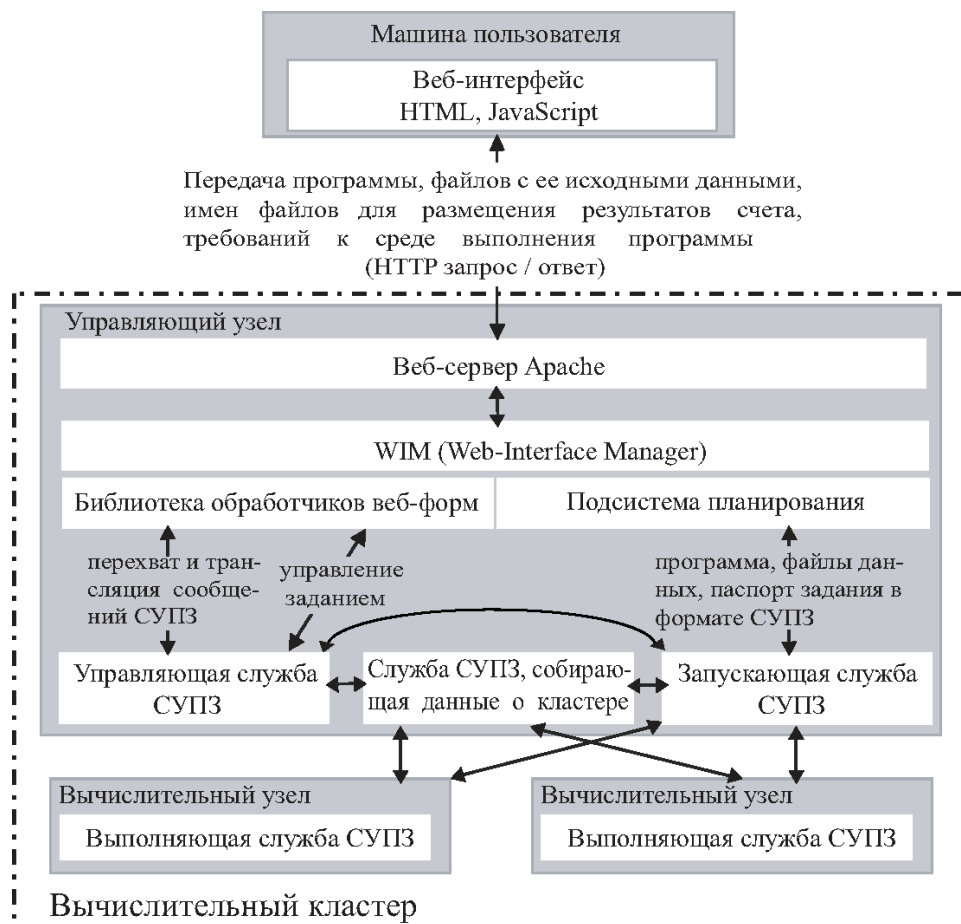


Рис. 2. Схема взаимодействия пользователя с ресурсами РВС на примере кластера

В третьем разделе второй главы представлен инструментальный комплекс (ИК) Distributed Computing Environment Toolkit (DISCENT), предназначенный для организации и применения РВС. ИК DISCENT состоит из трех основных компонентов (рис. 3): конструктора, базы данных и системы WIM.

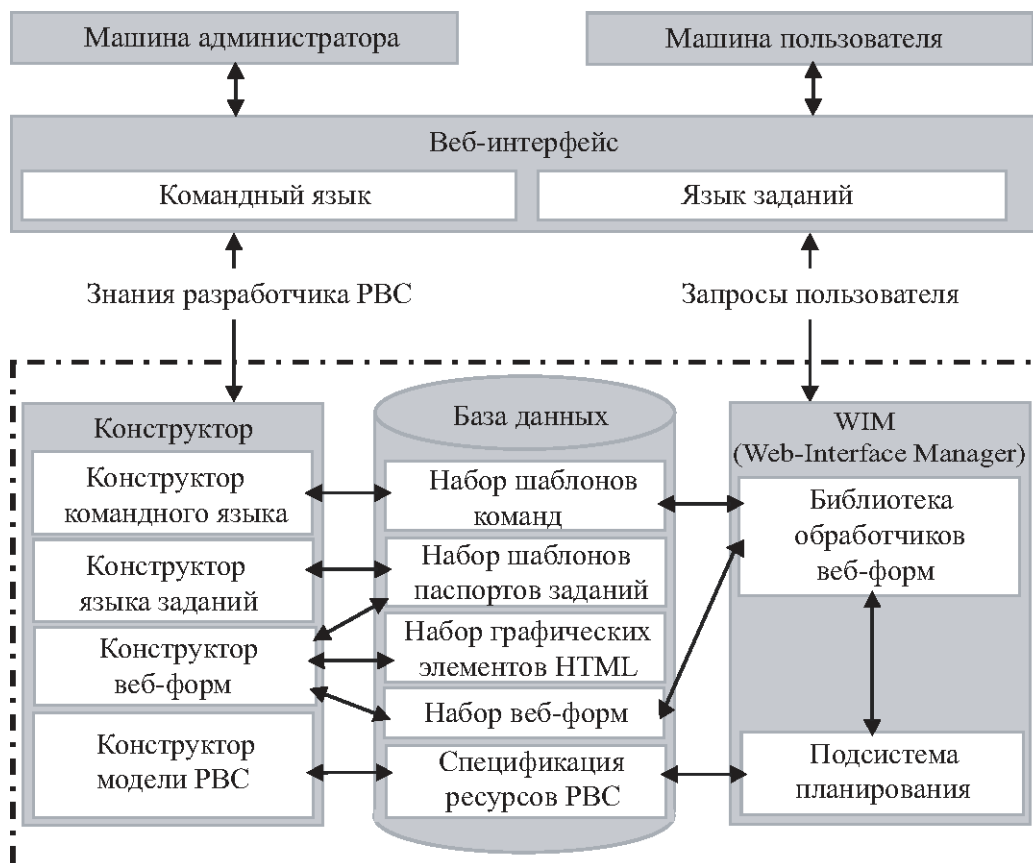


Рис. 3. Архитектура ИК DISCENT

В четвертом разделе главы рассмотрены основные функции конструктора ИК DISCENT, назначение языка заданий и командного языка. *Конструктор* включает четыре подсистемы.

1. *Конструктор командного языка* служит для создания, редактирования или удаления шаблонов команд для различных СУПЗ.

2. *Конструктор языка заданий* используется для построения и модификации шаблонов паспортов заданий для различных СУПЗ.

3. *Конструктор веб-форм* предназначен для создания шаблонов веб-форм, используемых для заполнения паспортов заданий. Создание веб-формы паспорта задания выполняется путем размещения на форме и задания свойств графических элементов HTML, соответствующих параметрам формируемого паспорта задания.

4. *Конструктор модели РВС* применяется для описания и модификации данных о вычислительных ресурсах РВС. Такая информация необходима подсистеме планирования для эффективного распределения различных типов задач по вычислительным узлам РВС.

*Язык заданий* определяет взаимодействие пользовательских задач с системой WIM и предназначен для формирования паспорта задания, включающего в себя уникальное имя задания, тип решаемой задачи, исполняемые программы, исходные данные, минимальные и/или желаемые требования к вычислительным ресурсам РВС и др.

*Командный язык* представляет собой совокупность команд, предназна-

ченных для управления пользовательскими заданиями в узлах РВС. С помощью командного языка система WIM может быть легко и гибко настроена для взаимодействия с используемой СУПЗ за счет применения унифицированных шаблонов описания команд СУПЗ.

В пятом разделе описана база данных и особенности ее реализации. Основным назначением такой базы является накопление информации о вычислительных ресурсах РВС, средствах создания и обработки паспортов заданий. Эта информация используется планировщиками системы WIM при формировании потоков заданий. Содержимое базы данных включает наборы шаблонов для командного языка и языка заданий, наборы графических элементов HTML и веб-форм, спецификацию ресурсов РВС.

Последний раздел главы посвящен описанию WIM. Данная система обеспечивает выполнение заданий пользователя путем распределения указанных в заданиях приложений на узлы РВС, подходящие для их запуска. Основными составляющими системы являются библиотека обработчиков веб-форм и подсистема планирования.

*Библиотека обработчиков веб-форм* предназначена для выполнения следующих функций: вывода пользователям требуемых веб-форм; обработки поступивших из этих форм данных; формирования паспортов заданий; предоставления пользователю информации о состоянии заданий и результатах счета; добавления, удаления и приостановления заданий пользователей. Данная библиотека состоит из набора скриптов языка РНР (обработчиков веб-форм), каждый из которых реализует алгоритмы для работы с одной или несколькими веб-формами на стороне веб-сервера. Имя обработчика веб-формы содержится на ней в специальном скрытом поле. Обработчики веб-форм имеют доступ к шаблонам команд, хранящимся в базе данных. В шаблонах команд для каждой операции управления заданием, размещенной на веб-форме, определяется команда СУПЗ, с помощью которой данная операция будет выполнена в РВС.

*Подсистема планирования* служит для обработки, постановки в очередь и распределения заданий, соответствующих различным типам задач. Данная подсистема включает следующие планировщики:

1. *Планировщик стандартных заданий MST (Manager of Standard Tasks)* используется для распределения заданий типа  $(a = 1 / b \neq L / c / d = R)$ .
2. *Планировщик распределенных приложений MDA (Manager of Distributed Applications)* предназначен для отправки заданий на вычислительные узлы РВС, на которых установлены приложения, необходимые для выполнения заданий типа  $(a / b / c / d = G)$ . Планировщик MDA взаимодействует с планировщиком MST с целью закрепления за заданием выбранных ресурсов перед передачей этого задания обработчику веб-форм для запуска задания.
3. *Планировщик взаимосвязанных заданий MIT (Manager of Interrelated Tasks)* ориентирован на распределение заданий типа  $(a = K / b = L / c / d)$  и контроль процесса их выполнения. Планировщик MIT, используя информационно-логические связи между объектами модели РВС, выполняет

частичное упорядочение подзадач общего задания, находит необходимые для выполнения задания вычислительные ресурсы, взаимодействует с планировщиком MST с целью закрепления за заданием выбранных ресурсов, передает задание обработчику веб-форм для его запуска и осуществляет дальнейший контроль выполнения этого задания.

При распределении заданий по узлам РВС планировщики системы WIM используют спецификацию ресурсов РВС из базы данных. Затем приводятся разработанные алгоритмы планирования ресурсов в системе WIM. Алгоритм A0 осуществляет распределение заданий по планировщикам WIM. Алгоритмы A1, A2 и A3 представляют схему работы планировщиков MST, MDA и MIT. В рамках алгоритма A3 реализованы вспомогательные алгоритмы A4, A5, A6 и A7. Алгоритмы A4 и A5 служат для построения плана решения задачи, относящейся к типу  $(a = K / b = L / c / d)$ , и базируются на использовании принципов «прямой и обратной волны». Алгоритмы A6 и A7 производят соответственно резервирование свободных вычислительных ресурсов для решения на них подзадач и распределение этих подзадач по зарезервированным ресурсам.

**В третьей главе** рассматриваются вопросы, связанные с построением гетерогенных РВС.

В первом разделе главы приведена общая схема организации и применения РВС. Данная схема включает следующие основные этапы:

- формирование программно-аппаратной части РВС:
  - определение вычислительных узлов РВС;
  - установка и настройка ИК DISCENT и системного ПО;
- настройка веб-интерфейса:
  - построение модели РВС;
  - параметризация и настройка шаблонов паспортов заданий и шаблонов команд для каждой СУПЗ, используемой в РВС;
  - создание веб-форм в графическом интерфейсе конструктора ИК DISCENT, которые будут выполнять запуск определенных типов заданий;
  - отладка и тестирование режимов запуска заданий;
- реализация процесса решения пользовательских задач:
  - регистрация пользователей в системе WIM;
  - формирование паспортов и запуск заданий;
  - мониторинг узлов РВС и распределение заданий в узлах РВС планировщиками системы WIM;
  - обмен данными между пользователями и СУПЗ.

Второй раздел главы посвящен созданию Grid. Архитектура экспериментальной Grid (организованной автором диссертации) состоит из трех основных частей: серверной, клиентской и исполнительной. Grid построена на базе кластеров с различными ОС и СУПЗ (рис. 4).

*Клиентская часть* предоставляет средства доступа пользователей к Grid. Сформировать и запустить задание в Grid можно одним из двух способов: либо путем запуска заданий (оформленных в специальном формате языка RSL –

Resource Specification Language) из другой Grid или с отдельной рабочей станции (машины с выделенным IP-адресом в Интернет), на которой установлен пакет Globus Toolkit, либо через веб-интерфейс с помощью языка заданий.

*Серверная часть* предназначена для организации соединения пользователей с ресурсами Grid. Структура серверной части является распределенной и включает Grid-шлюз и веб-сервер.

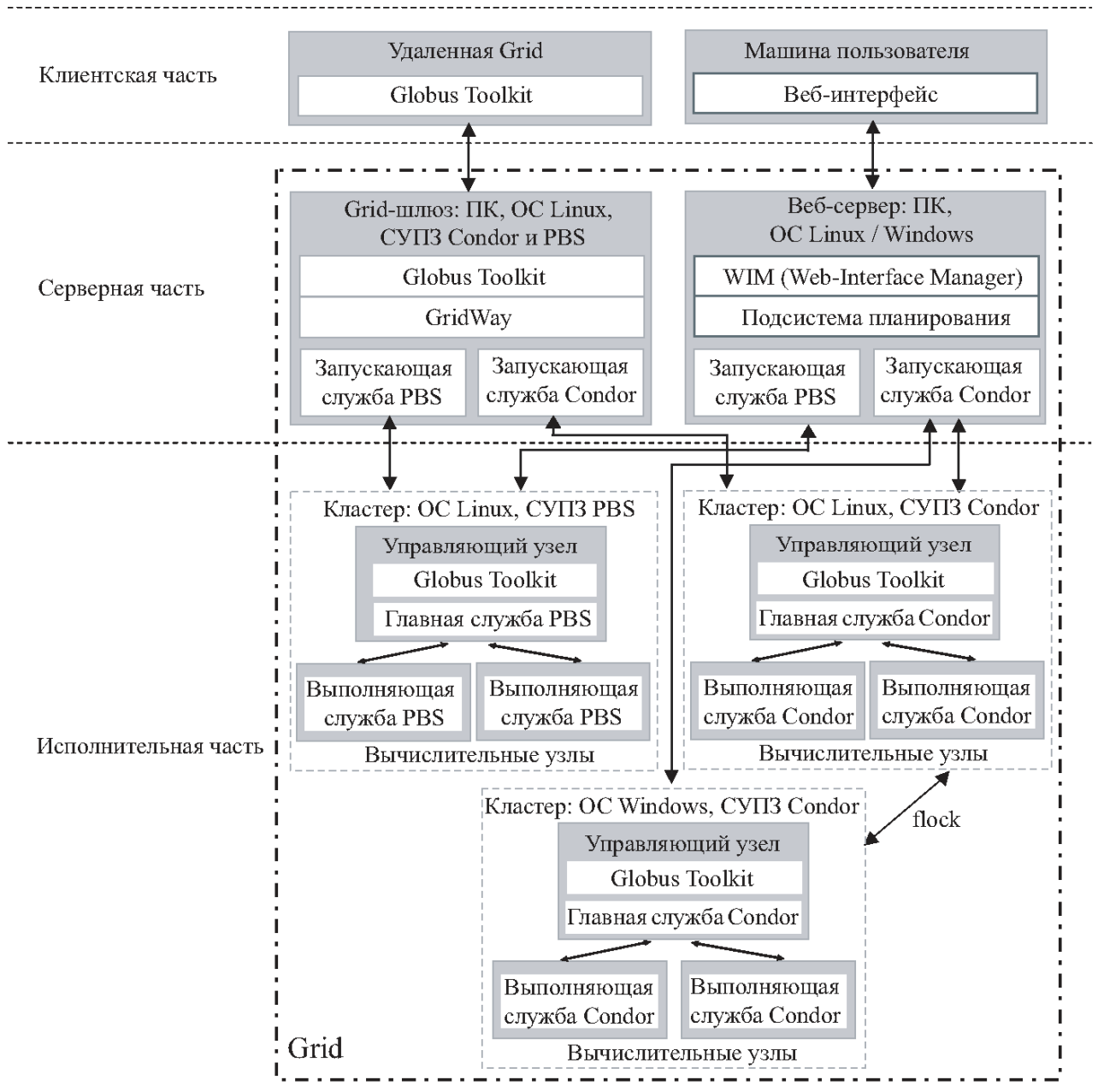


Рис. 4. Схема функционирования Grid

Grid-шлюз представляет собой рабочую станцию с установленными на ней пакетами Globus Toolkit (включающим утилиту GridWay) и СУПЗ, используемыми для управления кластерами Grid. Grid-шлюз является точкой входа пользователей из другой Grid или с рабочей станции. Grid-шлюз обеспечивает

следующие стандартные для Grid функции<sup>6</sup>: подключение пользователей, их сертификацию и авторизацию, получение заданий пользователей, выбор необходимых ресурсов для выполнения этих заданий и получение результатов вычислений. Планирование ресурсов для выполнения поступившего задания осуществляется с помощью утилиты GridWay. Остальные функции реализуются средствами Globus Toolkit.

На веб-сервере Grid размещается система WIM, позволяющая пользователю отправлять его задание в Grid и передавать это задание на тот или иной кластер в зависимости от заданных пользователем системных требований. Спецификация и запуск задания производится с помощью командного языка.

*Исполнительная часть* обеспечивает выполнение заданий в Grid на вычислительных узлах, которые представляют собой кластеры под управлением ОС Windows или Linux и СУПЗ PBS или Condor. В качестве СУПЗ кластера в Grid можно использовать и другие СУПЗ, совместимые с пакетом Globus Toolkit (например, Load Shared Facility).

При выборе механизма обслуживания очереди заданий на выполнение в Grid учитываются следующие факторы: выполнение задания в пакетном режиме; отсутствие априорной информации о времени выполнения задания; степень важности заданий в зависимости от их приоритетов; типы задач, к которым относятся задания с точки зрения планирования процесса их выполнения. В Grid дисциплина обслуживания очереди заданий базируется на комбинированном применении следующих принципов: FIFO (первый пришедший обслуживается первым); переключения выполнения; учета приоритетов заданий. Приоритеты для заданий разных типов устанавливаются администратором Grid.

Поступающие в Grid задания разделены на два потока (рис. 5). *Первый поток* формируется из заданий типа ( $a = 1 / b \neq L / c / d = R$ ), поступивших из других Grid. Поступившее на Grid-шлюз задание ставится в очередь и ждет назначения на один из свободных вычислительных ресурсов средствами пакета Globus Toolkit. При наличии свободных ресурсов утилита GridWay направляет находящееся на данный момент в голове очереди задание на выполнение СУПЗ освободившегося кластера. СУПЗ формирует локальную очередь заданий на кластере и передает задание на выполнение его вычислительным узлам. *Во второй поток* поступают задания, отправленные через веб-интерфейс. Подход к планированию работ WIM основан на следующих принципах. Во-первых, планировщики системы WIM используют модель Grid, что обеспечивает быстрый поиск информации о характеристиках и состоянии ресурсов. Во-вторых, для управления кластерами использованы СУПЗ, совместимые с инструментарием пакета Globus Toolkit и обладающие возможностью резервирования вычислительных ресурсов. В-третьих, в WIM реализована смешанная стратегия планирования, включающая как способ мониторинга состояния ресурсов, так и способ их резервирования. С помощью системы WIM, помимо заданий типа ( $a = 1 / b \neq L / c / d = R$ ), которые распределяются по узлам Grid с помощью планировщика *MST*, можно запускать так же следующие типы заданий:

---

<sup>6</sup> С точки зрения пакета Globus Toolkit.

( $a/b/c/d = G$ ) и ( $a = K/b = L/c/d$ ). Как правило, резервирование ресурсов увеличивает время ожидания заданий в очереди<sup>7</sup>. Однако без применения такого способа время ожидания возрастает в значительно большей степени для заданий типов ( $a/b/c/d = G$ ) и ( $a = K/b = L/c/d$ ).

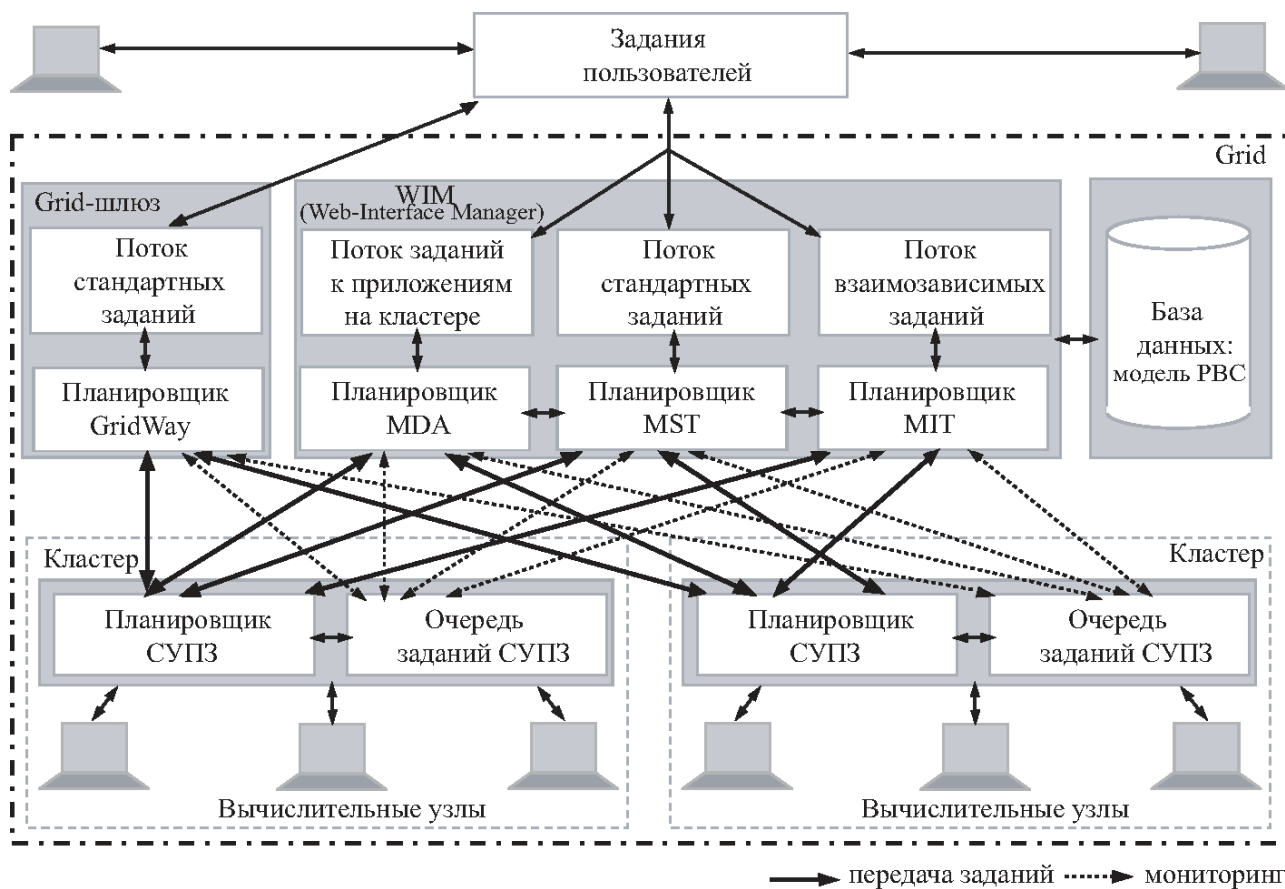


Рис. 5. Управление потоками заданий

В третьем разделе главы описан проведенный вычислительный эксперимент по моделированию процессов формирования и обработки заданий в Grid. Grid представляет собой стохастическую динамическую вычислительную сеть, состояние которой в некоторый момент времени определяется распределением в ней потоков заданий<sup>8</sup>. Для получения оценки эффективности различных способов распределения потоков заданий ( $E$ , %) генерируется поток заданий, осуществляется их распределение и выполнение в Grid двумя способами (с помощью GridWay и GridWay + WIM) и сравниваются следующие показатели эффективности функционирования вычислительных ресурсов Grid: число заданий с нулевым временем ожидания ( $T_0$ , сек.), среднее время ожидания задания в очереди ( $T_q$ , сек.), среднее число заданий в очереди ( $C_q$ , ед.), среднее время пребывания задания в Grid ( $T_g$ , сек.), коэффициент полезного использования ресурсов Grid ( $K_g$ , %), количество заданий в потоке ( $C_j$ , ед.), общее время решения заданий потока ( $T_s$ , сек.).

<sup>7</sup> Топорков В.В. Модели распределенных вычислений / В.В. Топорков. – М.: ФИЗМАТЛИТ, 2004. – 320 с.

<sup>8</sup> Попков Ю.С. Макросистемы и GRID-технологии: моделирование динамических стохастических сетей / Ю.С. Попков // Проблемы управления. – 2003. – № 8. – С. 10-20.

Поток заданий в Grid характеризуется следующими свойствами: неоднородностью (задания соответствуют разным типам задач и отличаются друг от друга по своей специфике); отсутствием обратной связи (число заданий, поступивших за один промежуток времени, не зависит от числа заданий, поступивших за другой промежуток времени); неординарностью (возможно поступление двух и более заданий в один и тот же момент времени); стационарностью (число событий, поступивших за определенный промежуток времени, зависит от длины этого промежутка и не зависит от момента его начала).

Для формирования потока заданий и отправки их на выполнение в Grid ИДСТУ СО РАН разработан специальный генератор. В качестве приложений в заданиях используются программы-имитаторы, выполняющие в Grid реальную загрузку вычислительных ресурсов и обмен заданными объемами данных. Генератор заданий был запущен на удаленной машине в Институте вычислительного моделирования СО РАН, которая играла роль сторонней Grid, а также на независимых рабочих станциях в Интернет, представляющих машины пользователей веб-интерфейса к Grid ИДСТУ СО РАН. Результаты вычислительных экспериментов (см. таб. 1) показывают устойчивое преимущество второго способа распределения заданий (GridWay+WIM) по всем показателям эффективности функционирования ресурсов Grid.

Таблица 1. Результаты вычислительных экспериментов

Grid	2 кластера с ОС Linux (168 выч. ядер)			2 кластера с ОС Linux, кластер с ОС Windows (184 выч. ядер)			2 кластера с ОС Linux, кластер с ОС Windows (184 выч. ядер)		
Типы Задач	Все типы			Все типы			$(a = 1 / b \neq L / c = N / d = R)$ и $(a = K / b = L / c / d = R)$		
	GridWay	GridWay+ WIM	$E$ (%)	GridWay	GridWay+ WIM	$E$ (%)	GridWay	GridWay+ WIM	$E$ (%)
$T_0$	49060	52620	7,26%	24340	26680	9,61%	17980	19920	10,79%
$T_a$	87260	80100	-8,21%	44020	39740	-9,72%	28240	25540	-9,56%
$C_a$	7360	7260	-1,36%	3760	3620	-3,72%	2380	2280	-4,20%
$T_g$	123920	118440	-4,42%	62580	58460	-6,58%	39220	35420	-9,69%
$K_g$	93,34%	94,13%	0,85%	92,41%	95,12%	2,93%	93,74%	94,97%	1,31%
$C_s$	18688	18688	-	9344	9344	-	157	157	-
$T_s$	4917080	4716820	-4,07%	2486420	2336520	-6,03%	1585160	1459720	-7,91%

В четвертом разделе главы рассматриваются примеры решения практических задач в Grid ИДСТУ СО РАН. В частности, в Grid был решен ряд задач складской логистики, представляющих все основные типы задач. Проведенные расчеты подтвердили эффективность децентрализованного управления потоками заданий в Grid.

В последнем разделе данной главы приводится пример применения ИК DISCENT для организации вычислительного кластера на базе компьютерного класса вуза. Кластер использован в учебном процессе для проведения многовариантных расчетов при моделировании систем массового обслуживания.

**В заключении** перечисляются основные результаты, полученные в рамках диссертационной работы.



**Приложения** включают описание характеристик СУПЗ, графических элементов HTML, языка заданий, командного языка и объектов базы данных, а также другие дополнительные материалы.

## **ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИИ**

1. *Разработан способ децентрализованного управления потоками заданий*, позволяющий (в отличие от известных) распределять задания в вычислительной среде в зависимости от типов решаемых задач. Для каждого типа задач разработан специализированный планировщик заданий.

2. *Разработаны средства организации унифицированного веб-ориентированного доступа* к распределенным вычислительным ресурсам. В том числе язык заданий и командный язык, обеспечивающие их настройку на особенности функционирования различных систем управления заданиями, используемых в узлах распределенной вычислительной среды.

3. *Выполнена программная реализация системной архитектуры и алгоритмов функционирования инструментального комплекса DISCENT*, обеспечивающего поддержку основных этапов построения и применения гетерогенных распределенных вычислительных сред для решения ресурсоемких научно-исследовательских задач.

## **СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ**

1. Феоктистов А.Г. Разработка Grid-системы с децентрализованным управлением потоками заданий / А.Г. Феоктистов, А.С. Корсуков // Вестник НГУ. Серия: Информационные технологии. – 2008. – Т. 6, вып. 3. – С. 147-154.
2. Информационно-вычислительная среда для моделирования сложных систем в учебном процессе / А.Г. Феоктистов, В.И. Дмитриев, А.С. Корсуков, А.В. Ларина // Вестник ТГУ. Приложение. – 2007. – № 23. – С. 348-353.
3. Корсуков А.С., Феоктистов А.Г. Система управления заданиями для вычислительного кластера: Свидетельство об официальной регистрации программы для ЭВМ № 2007611624. – М.: Федеральная служба по интеллектуальной собственности, патентам и товарным знакам, 2007.
4. Корсуков А.С., Феоктистов А.Г., Дмитриев В.И. Инструментальный комплекс для организации Web-интерфейсов к вычислительным кластерам (WIDT): Свидетельство об официальной регистрации программы для ЭВМ № 2007613651. – М.: Федеральная служба по интеллектуальной собственности, патентам и товарным знакам, 2007.
5. Средства и методы организации учебной информационно-вычислительной среды / А.Г. Феоктистов, В.И. Дмитриев, А.С. Корсуков, А.В. Ларина // Вестник Московского городского педагогического университета. Серия: Информатика и информатизация образования. – 2007. – № 2(9). – С. 37-41.
6. Феоктистов А.Г. Ролевая модель системы безопасности распределенной вычислительной среды / А.Г. Феоктистов, А.С. Корсуков // Информационные и математические технологии в науке и управлении: Тр. XII Байкальской

- Всерос. конф.: в 2-х ч. – Иркутск: Изд-во ИСЭМ СО РАН, 2007. – Ч. 2. – С. 33-38.
7. Феоктистов А.Г. Организация распределенного имитационного моделирования в grid-системе / А.Г. Феоктистов, А.А. Александров, А.С. Корсуков // Параллельные вычислительные технологии: Тр. Междунар. науч. конф. – Челябинск: Изд-во ЮУрГУ, 2007. – Т. 2. – С. 28-36.
  8. Феоктистов А.Г. Средства выполнения распределенной GPSS-модели на вычислительном кластере / А.Г. Феоктистов, А.А. Александров, А.С. Корсуков // Инфокоммуникационные и вычислительные технологии и системы: Материалы II Всерос. конф. с междунар. участием. – Улан-Удэ: Изд-во Бурятского ун-а, 2006. – Т. 2. – С. 166-173.
  9. Феоктистов А.Г. Архитектура системы управления вычислительным кластером невыделенных рабочих станций / А.Г. Феоктистов, А.С. Корсуков // Параллельные вычисления и задачи управления: Тр. III Междунар. конф. – М.: ИПУ РАН, 2006. – С. 492-497.
  10. Феоктистов А.Г. Средства организации запуска прикладных задач на вычислительном кластере / А.Г. Феоктистов, А.С. Корсуков // Информационные и математические технологии в научных исследованиях: Тр. XI Байкальской междунар. конф.: в 2-х ч. – Иркутск: Изд-во ИСЭМ СО РАН, 2006. – Ч. 2. – С. 32-38.
  11. Корсуков А.С. Подсистема запуска задач на вычислительном кластере / А.С. Корсуков // Математическое моделирование и информационные технологии: Материалы VIII Школы-семинара молодых ученых. – Иркутск: ИДСТУ СО РАН, 2006. – С. 80-82.
  12. Корсуков А.С. Методы и средства организации вычислений на кластерах для класса многократно используемых программ / А.С. Корсуков, А.Г. Феоктистов // Компьютерные технологии в науке, производстве, социальных и экономических процессах: Материалы VI Междунар. науч.-практ. конф. – Новочеркасск: ЮРГТУ, 2005. – Ч. 1. – С. 46-50.
  13. Феоктистов А.Г. Удаленный доступ к кластерным системам / А.Г. Феоктистов, А.С. Корсуков // Математическое моделирование и информационные технологии: Материалы VI Школы-семинара молодых ученых. – Иркутск: ИДСТУ СО РАН, 2005. – С. 39-40.
  14. Корсуков А.С. Вопросы организации дистанционного доступа к кластерным системам / А.С. Корсуков // Под знаком  $\Sigma$ : Тез. докл. III Всерос. молодежной конф. – Омск: ОНЦ СО РАН, 2005. – С. 238.
  15. Корсуков А.С. Реализация удаленного доступа к кластеру лаборатории методов автоматизации исследований управляемых систем ИДСТУ СО РАН / А.С. Корсуков // Математическое моделирование и информационные технологии: Материалы VII Школы-семинара молодых ученых. – Иркутск: ИДСТУ СО РАН, 2005. – С. 17-18.

Редакционно-издательский отдел  
Института динамики систем и теории управления СО РАН  
664033, Иркутск, Лермонтова, 134  
Подписано к печати 21.01.2009 г.  
Формат бумаги 60×84 1/16, объем 1 п.л.  
Заказ №12. Тираж 100 экз.

---

Отпечатано в ИДСТУ СО РАН