

# ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА ОРГАНИЗАЦИИ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ В ИНТЕГРИРОВАННЫХ КЛАСТЕРНЫХ СИСТЕМАХ

И.В. Бычков

*Институт динамики систем и теории управления СО РАН*

e-mail: idstu@icc.ru

Г.А. Опарин

e-mail: oparin@icc.ru

А.П. Новопашин

e-mail: apn@icc.ru

А.Г. Феоктистов

e-mail: agf@icc.ru

## Аннотация

Обсуждаются вопросы, связанные с организацией распределенных вычислений. В частности, рассматриваются архитектура и принципы функционирования инструментальных средств для автоматизации разработки и применения в интегрированных кластерных системах сложных прикладных программных комплексов.

**1. Введение.** Стремительный прогресс в области разработки компьютерной техники и средств коммуникации способствовал появлению и развитию высокопроизводительных вычислительных сред (кластеров, метакомпьютеров, сетей Grid). Тем не менее, в условиях быстро развивающейся компьютерной базы параллелизм, реализованный практически на всех уровнях вычислительной аппаратуры, используется далеко не полностью. Отсутствуют развитые языки, технологии и инструментальные средства, ориентированные на прикладных специалистов и обеспечивающие адекватное отображение задач на конкретную программно-аппаратную платформу. Крайне мало внимания уделяется построению прикладных программных систем на основе повторного использования программных блоков, особенно – большого количества таких блоков. Эффективность создаваемых приложений до сих пор во многом зависит от степени владения низкоуровневым системным программным обеспечением (MPI, OpenMP, Globus Toolkit и т.д.), что является уделом узкого круга программистов высокой квалификации. Многие распределенные прикладные программные комплексы разрабатываются без соблюдения требований к переносимости, совместимости и адаптируемости к динамически изменяющейся аппаратной среде. Зачастую отсутствуют характеристики их функционирования в соответствии с требуемыми показателями качества.

Анализ мировых тенденций в области автоматизации решения прикладных задач в параллельных и распределенных вычислительных средах позволяет утверждать, что решение многих проблем непосредственно связано с интеллектуализацией, так называемого, промежуточного программного обеспечения (ПО), позволяющего динамически интегрировать распределенные гетерогенные ресурсы в виртуальную исполнительную среду и предоставляющего возможности для прозрачного использования этой среды.

Наши исследования (см., например, [1-4]) показывают, что комплексное применение методов концептуального крупноблочного программирования, мультиагентных технологий,

логических методов синтеза и конструирования программ, технологий баз знаний и других средств искусственного интеллекта, является ключевым фактором в создании удобных информационно-вычислительных сред для прикладных специалистов, стремящихся эффективно использовать высокопроизводительные ресурсы, не вдаваясь в особенности низкоуровневого параллельного и/или распределенного программирования решаемой задачи. В таких средах построение параллельной (распределенной) крупноблочной программы на основе библиотеки специфицированных прикладных модулей выполняется автоматически по целевому содержательному запросу.

В докладе представлены разработанные в Институте динамики систем и теории управления СО РАН (ИДСТУ СО РАН) высокоуровневые средства автоматизации параллельного и распределенного программирования прикладных задач, а также инструментальные средства и технологии организации интегрированных кластерных систем.

**2. Интегрированная кластерная система.** Интегрированная кластерная система (рис. 1) – это распределенная вычислительная среда (РВС), предназначенная для решения фундаментальных и прикладных вычислительных задач и характеризующаяся следующими особенностями:

- в качестве узлов системы выступают вычислительные кластеры;
- на разных уровнях интеграции системы существуют различные категории пользователей, в их числе пользователи, нуждающиеся в высокоуровневых средствах организации вычислительного процесса решения задачи;
- задания пользователей относятся к задачам разных классов и требуют специализированных средств спецификации вычислительных процессов;
- множество заданий пользователей рассматривается с точки зрения теории очередей и представляется в виде совокупности потоков заданий;
- в рамках системы могут функционировать специализированные распределенные проблемно-ориентированные программные комплексы;
- в системе нет единой политики администрирования вычислительных кластеров, на кластерах могут применяться различные принципы и механизмы обработки потоков заданий различных типов;
- средства объединения вычислительных кластеров в единую систему должны обеспечивать прозрачность, открытость и масштабируемость интегрируемых ресурсов.

ПО промежуточного уровня, созданное к настоящему времени, позволяет решить многие вопросы, связанные с построением сложноорганизованных РВС. Тем не менее, важными проблемами остаются автоматизация процесса решения разнотипных задач прикладных специалистов, а так же согласование политик администрирования вычислительных кластеров. Последняя проблема является темой отдельного исследования. Перспективным подходом к ее решению является использование технологий многоагентных систем [5].



Рис. 1. Интегрированная кластерная система

**3. Концептуальная модель РВС.** К настоящему времени разработаны различные подходы и средства моделирования РВС, базирующиеся на использовании аналитических, имитационных или полунатурных моделей. Однако, учитывая сложность исследуемых систем, применение моделей одного определенного вида позволяет зачастую получить лишь наиболее общие, ориентировочные результаты. Таким образом, возникает необходимость построения согласованного семейства моделей разных видов, которое бы позволило осуществлять всесторонний анализ сложных вычислительных систем на разных уровнях детализации.

Авторами доклада разработана оригинальная концептуальная модель разнородной РВС [6]. Модель имеет сложную многоуровневую структуру и позволяет описывать проблемно-ориентированный, программно-аппаратный, имитационный и вычислительный слои знаний о распределенной вычислительной среде, а так же определять и совместно использовать различные модели параллельных и распределенных вычислений, такие как модели программирования приложений, модели планирования вычислительных процессов и загрузки ресурсов.

**4. Инструментальный комплекс СИРИУС.** Для построения концептуальной модели РВС модели в ИДСТУ СО РАН разработан высокоуровневый инструментальный комплекс (ИК) СИРИУС [7], предназначенный для описания и анализа эффективности функционирования проектируемых РВС на основе средств и методов визуального программирования, распределенного имитационного моделирования, динамического анализа программ и автоматизации процессов построения моделей исследуемых систем и проведения по этим моделям широкомасштабных ресурсоемких вычислительных экспериментов.

ИК СИРИУС (рис. 2) включает следующие основные подсистемы: графическую инструментальную среду, генератор распределенной модели, исполнительную подсистему и модуль анализа результатов моделирования.



Рис. 2. Архитектура инструментального комплекса СИРИУС

Данная инструментальная среда обеспечивает выявление комплексного влияния характеристик как прикладного программного обеспечения, так и аппаратных средств на вероятностно-временные показатели решения прикладных задач в исследуемой системе.

**5. Интеллектуальные технологии и инструментальные средства автоматизации крупноблочного параллельного и распределенного программирования.** Известно, что проблема автоматизации составления программ является одной из ключевых в параллельном и распределенном программировании.

В ИДСТУ СО РАН разработаны высокоуровневые инструментальные средства, позволяющие в проблемно-ориентированных терминах как конструировать крупноблочные параллельные и распределенные вычислительные процессы, так и автоматически их синтезировать на основе непроцедурных постановок задач на вычислительной модели предметной области.

*5.1. Синтез параллельных планов решения вычислительных задач.* Предложен новый подход [8] к построению параллельных асинхронных планов требуемой длины для решения вычислительных задач в мультипроцессорной среде. Условия задачи планирования сформулированы в виде системы булевых уравнений (ограничений), решения которой определяют возможные планы активизации программных модулей. Учитываются ограничения на число используемых узлов мультипроцессорной вычислительной среды,

временные задержки, возникающие при выполнении программных модулей, а также множественность установки модулей в узлах.

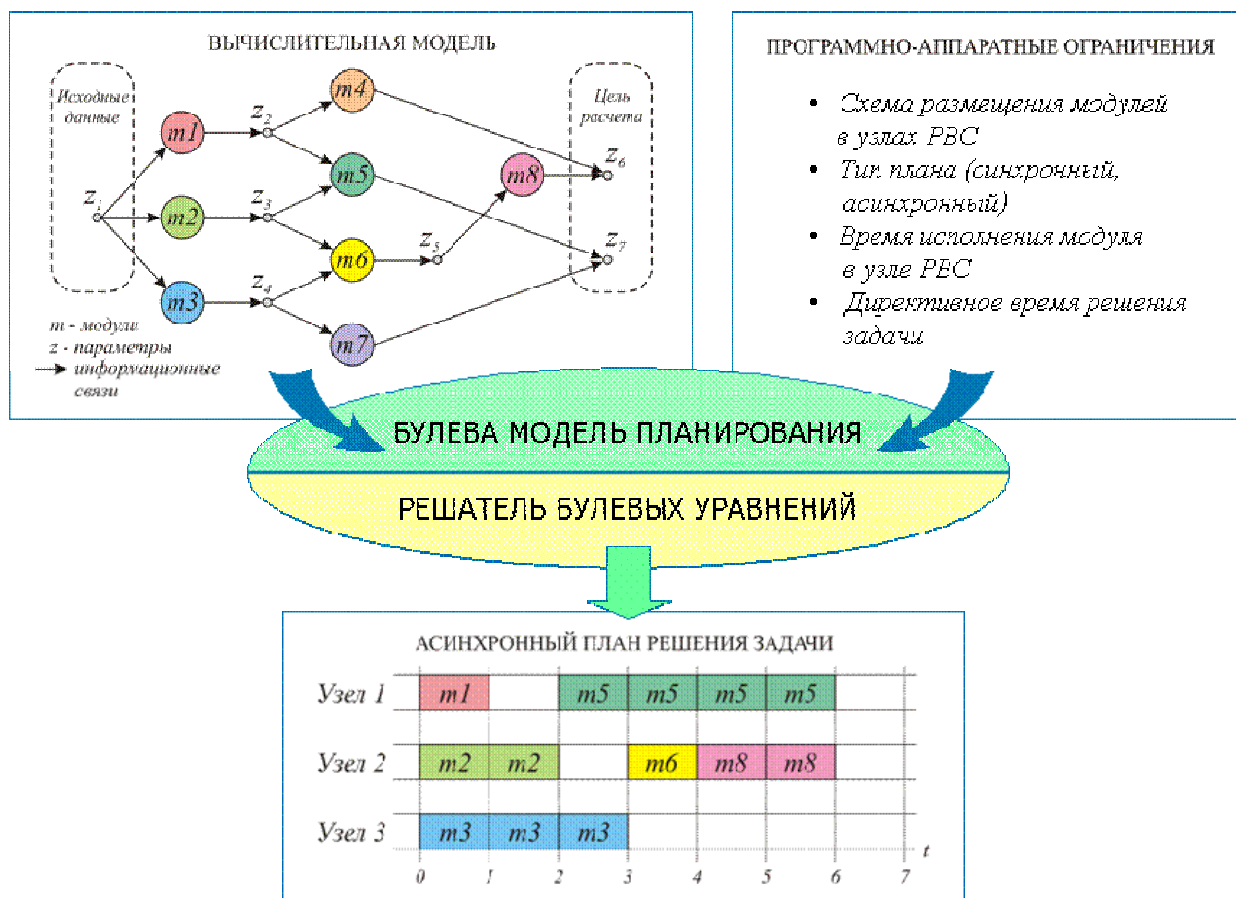


Рис. 3. Схема построения параллельных планов решения вычислительных задач

5.2. *Среда непроцедурного программирования СиКрус* [9]. На основе рассмотренного выше подхода к построению параллельных асинхронных планов решения задач требуемой длины разработана среда непроцедурного программирования СиКрус для синтеза параллельных программ на языке Fortran-DVM с учетом ресурсных ограничений используемой вычислительной системы и времени исполнения предметных модулей. Выбор языка Fortran-DVM в качестве выходного языка генератора управляющей программы транслятора-синтезатора обеспечивает простоту генерации результирующего кода и удобство его использования прикладным специалистом в плане его читаемости, отладки и ручной модификации (в случае такой необходимости).

5.3. *Инструментальный комплекс Orlando Tools* [10]. Данный инструментальный комплекс предназначен для конструирования асинхронных параллельных программ с использованием коммуникационной библиотеки PVM. Комплекс содержит модели, алгоритмы и программные средства, обеспечивающие возможность реализации в создаваемых параллельных программах как параллелизма по управлению, свойственного системам крупноблочного программирования, так и параллелизма по данным, характерного для задач многовариантного анализа и обработки больших массивов данных, допускающих естественную декомпозицию на независимые сегменты.

ИК Orlando Tools (рис. 4) включает следующие основные компоненты: многооконный текстовый редактор, транслятор, подсистему компиляции, подсистему запуска, базы расчетных данных. Все эти компоненты устанавливаются и запускаются на рабочей станции пользователя.

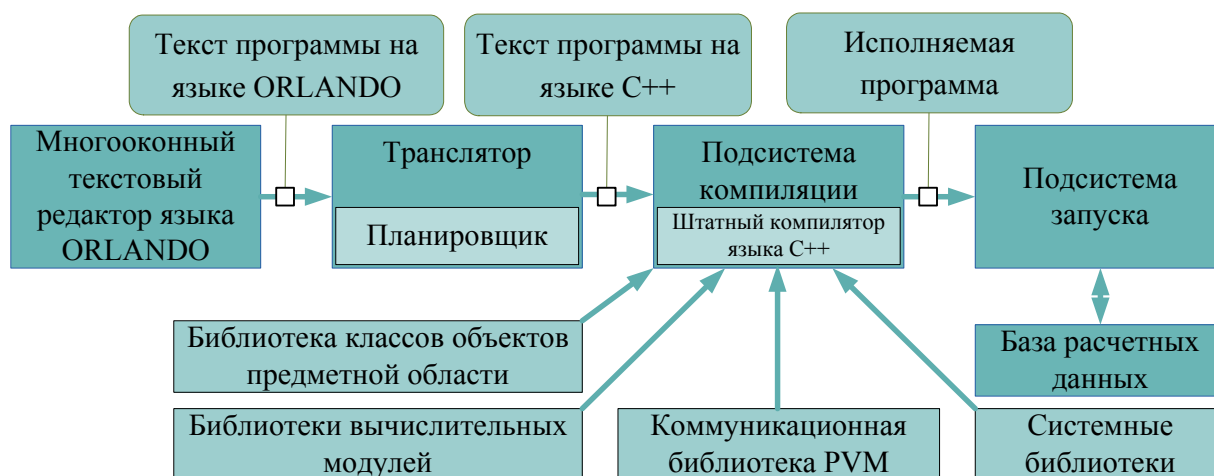


Рис. 4. Архитектура инструментального комплекса Orlando Tools

5.4. *Инструментальный комплекс DISCOMP* [11]. ИК DISCOMP предназначен для организации распределенных пакетов прикладных программ в разнородных вычислительных средах. Функциональное наполнение таких пакетов может включать модули, реализованные на различных языках программирования, нетиражируемые программные комплексы, а также унаследованное ПО. ИК DISCOMP обеспечивает возможность создания как автономных распределенных пакетов прикладных программ, так и интеграцию данного инструментария с другими программными комплексами для автоматизации распределенных вычислений при решении исследовательских задач.

ИК DISCOMP включает следующие основные компоненты (рис. 5): систему управления РВС (СУРВС), набор вычислительных клиентов РВС, систему хранения данных и средства доступа пользователей к распределенному пакету прикладных программ.

Эффективность решения задач в распределенных пакетах прикладных программ обеспечивается следующими особенностями работы ИК DISCOMP: возможностью максимального использования разнородных ресурсов РВС; поддержкой высокопроизводительного взаимодействия между распределенными компонентами ИК DISCOMP; гибкой диспетчеризацией очередями задач; наличием средств оптимизации объемов данных, передаваемых между модулями в процессе решения задач в РВС; возможностью динамического управления процессами решения задач.

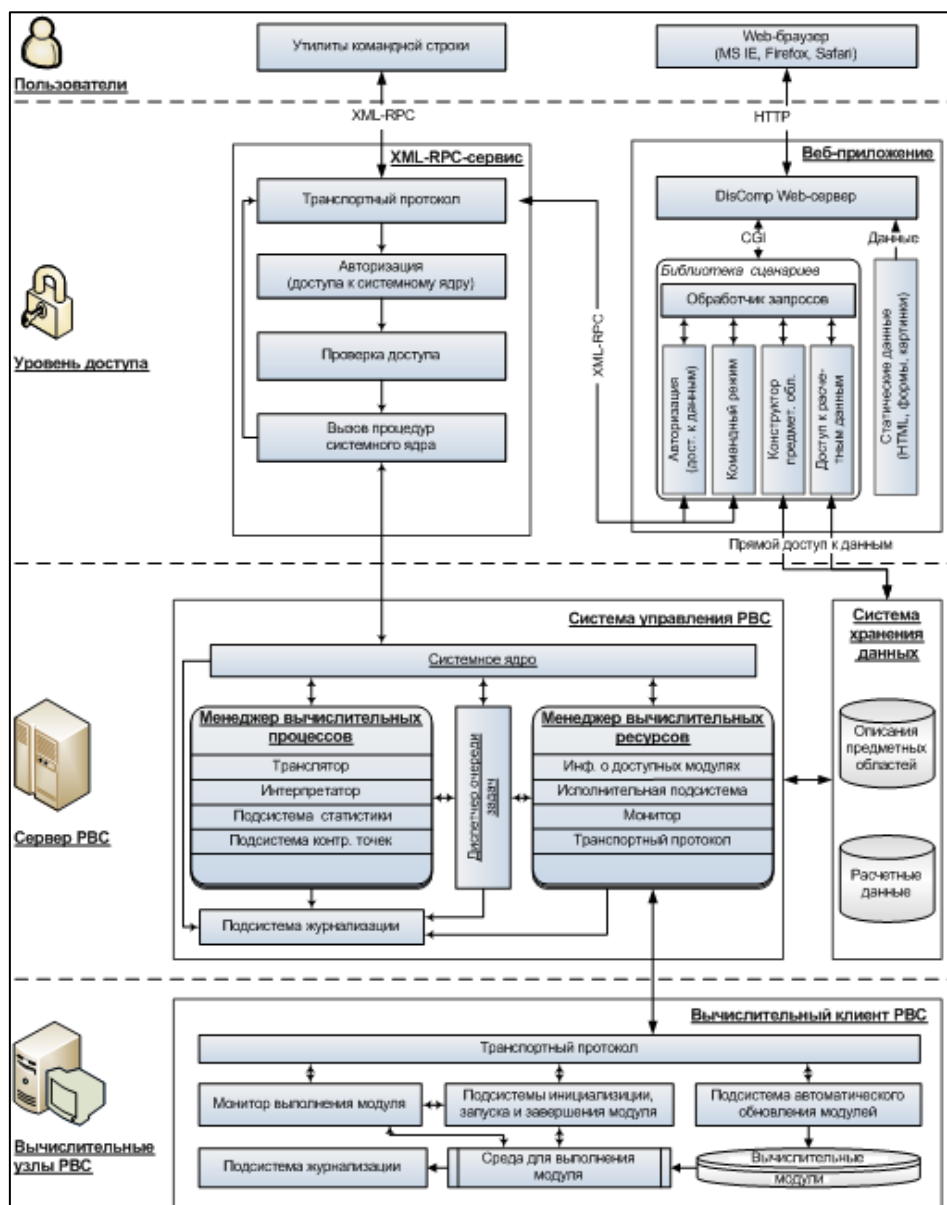


Рис. 5. Архитектура инструментального комплекса DISCOMP

**6. Организация интегрированных кластерных систем.** Для организации интегрированных кластерных систем в ИДСТУ СО РАН разработан инструментальный комплекс DISCENT [12]. В состав комплекса включены средства, обеспечивающие:

- организацию интегрированных кластерных систем с использованием стандартизированного базового программного обеспечения: пакета Globus Toolkit и согласованных с ним по управлению и форматам данных систем управления прохождением заданий (СУПЗ), таких, например, как PBS, SGE или Condor;
- разработку программных настроек для решения сложных задач, требующих дополнительных возможностей организации вычислений и децентрализованного управления ими;
- эффективное управление потоками разнотипных заданий в интегрированной кластерной системе на основе интеллектуального планирования вычислений;

- прохождение задач, поступающих с помощью различных средств доступа к ресурсам интегрированной системы (в их числе задачи локальных пользователей вычислительных кластеров), исключительно через кластерные СУПЗ, что позволяет децентрализованным службам управления вычислительными процессами поддерживать целостность информационной картины, складывающейся в процессе вычислений;
- полунатурное моделирование вычислительной инфраструктуры интегрированной кластерной системы с целью оценки эффективности ее функционирования;
- объединение нескольких интегрированных кластерных систем в единую глобальную систему с многоуровневой (по управлению потоками заданий) архитектурой путем взаимной сертификации их вычислительных ресурсов средствами пакета Globus Toolkit через стандартизированные шлюзы;
- унифицированный прозрачный доступ пользователей ко всем ресурсам глобальной интегрированной кластерной системы через специализированные веб-серверы.

ИК DISCENT состоит из трех основных компонентов (рис. 6): конструктора, базы данных и подсистемы планирования.

Конструктор включает четыре подсистемы. Конструктор командного языка служит для создания, редактирования или удаления шаблонов команд для различных СУПЗ. Конструктор языка заданий используется для построения и модификации шаблонов паспортов заданий для различных СУПЗ. Конструктор веб-форм предназначен для создания шаблонов веб-форм, используемых для заполнения паспортов заданий. Создание веб-формы паспорта задания выполняется путем размещения на форме графических элементов HTML, соответствующих параметрам формируемого паспорта задания и описания их свойств. 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 подсистемой планирования 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, предназначенных для управления пользовательскими заданиями в узлах РВС. С помощью командного языка подсистема планирования может быть легко настроена для взаимодействия с используемой СУПЗ за счет применения унифицированных шаблонов описания команд СУПЗ.

Основным назначением базы данных является накопление информации о вычислительных ресурсах РВС, средствах создания и обработки паспортов заданий. Эта информация используется подсистемой планирования 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.



Подсистема планирования обеспечивает выполнение заданий пользователя путем распределения указанных в заданиях приложений на узлы РВС, подходящие для их запуска.

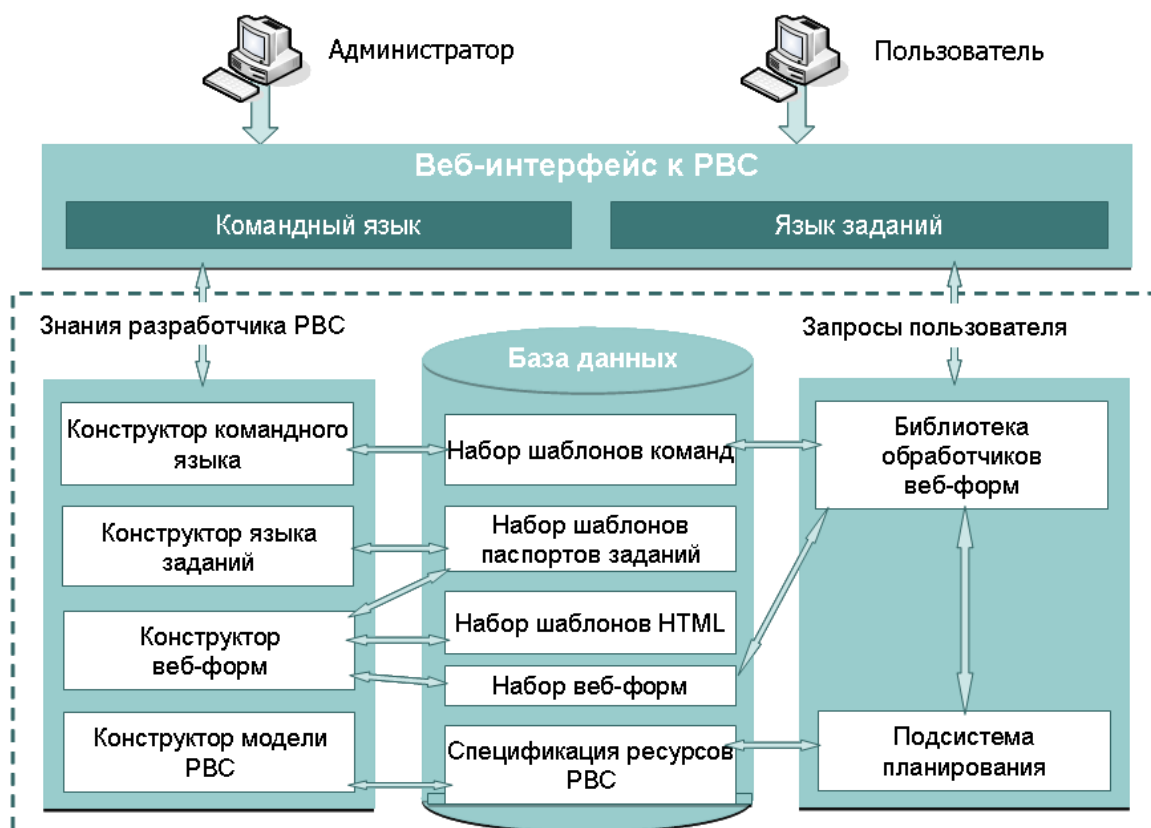


Рис. 6. Архитектура инструментального комплекса DISCENT

**Применение.** Перечисленные выше инструментальные средства используются в институтах ИНЦ СО РАН при создании пакетов прикладных программ для решения задач из различных предметных областей, в частности, исследования биоресурсов озера Байкал, решения систем булевых уравнений и задач оптимального управления, моделирования складского логистического комплекса и ряда других.

Разработка и применение пакетов прикладных программ осуществляется в ИДСТУ СО РАН. Объединение ресурсов интегрированной кластерной системы и функциональных возможностей пакетов прикладных программ позволяет создавать проблемно-ориентированные кластерные вычислительные системы.

Схема технологического процесса построения проблемно-ориентированных кластерных вычислительных систем приведена на рис. 7. Данный процесс базируется на комплексном использовании методов концептуального крупноблочного программирования, мультиагентных технологий, логических методов синтеза и конструирования программ, булевых уравнений и методов их решения, технологий баз знаний и Grid-технологий. Для его реализации созданы оригинальные модели, методы, алгоритмы и высокоуровневые (ориентированные на предметных специалистов) инструментальные средства, обеспечивающие эффективное использование высокопроизводительных ресурсов без погружения в особенности низкоуровневого параллельного (распределенного) программирования решаемой задачи путем автоматического синтеза параллельной

(распределенной) крупноблочной программы по целевому содержательному запросу на вычислительной модели проблемно-ориентированного программно-аппаратного вычислительного комплекса.

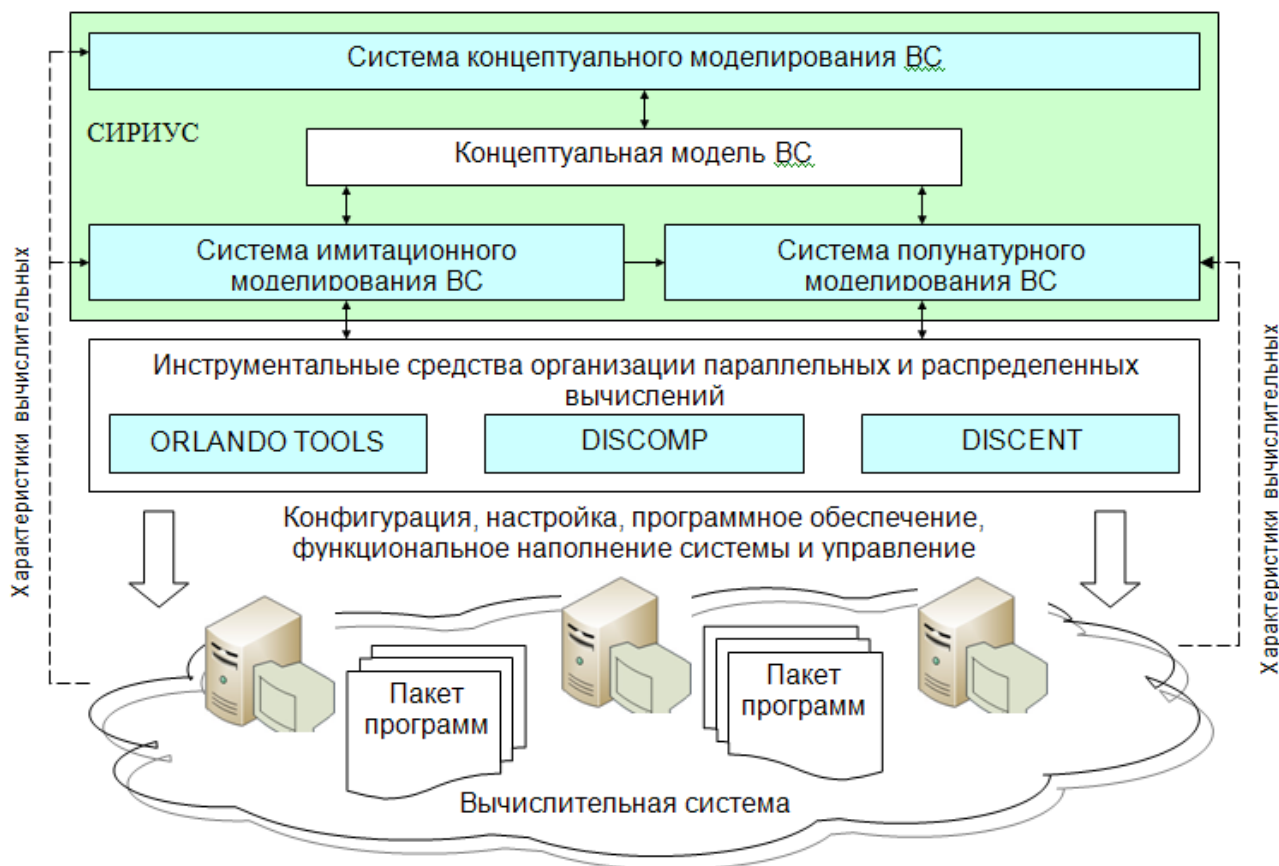


Рис. 7. Схема технологического процесса построения проблемно-ориентированных кластерных вычислительных систем

**Заключение.** Представленные в докладе инструментальные средства организации высокопроизводительных вычислений позволяют комплексно решать целый спектр задач, связанных с разработкой, эксплуатацией и модернизацией параллельных и распределенных проблемно-ориентированных сред, а также созданием кластерных Grid с учетом их возможной интеграции в многоуровневые структуры.

## ЛИТЕРАТУРА

- [1]. Опарин Г.А., Феоктистов А.Г., Новопашин А.П. Вопросы организации взаимодействия вычислительных модулей в распределенной интеллектуальной САТУРН-среде // Вычислительные технологии. Т. 8. Вестник КазНУ. № 3. Совместный выпуск. 2003. Ч. 2. С. 307-315.
- [2]. Опарин Г.А., Новопашин А.П. Булево моделирование планирования действий в распределенных вычислительных системах // Известия РАН. Теория и системы управления. 2004. № 5. С. 105-108.
- [3]. Бычков И.В., Корсуков А.С., Опарин Г.А., Феоктистов А.Г. Технология организации гетерогенных распределенных вычислительных сред // Научный сервис в сети Интернет:

масштабируемость, параллельность, эффективность: Труды Всерос. суперкомпьютерной конф. – М.: Изд-во МГУ, 2009. С. 353-359.

- [4]. Бычков И.В., Опарин Г.А., Новопашин А.П., Феоктистов А.Г., Корсуков А.С., Сидоров И.А. Высокопроизводительные вычислительные ресурсы ИДСТУ СО РАН: Текущее состояние, возможности и перспективы развития // Вычислительные технологии. 2010. Т. 15. № 3. С. 69-82.
- [5]. Бычков И.В., Опарин Г.А., Новопашин А.П., Феоктистов А.Г., Корсуков А.С. Управление потоками заданий в интегрированной кластерной системе // Параллельные вычисления и задачи управления: Пленарные доклады V Междунар. конф. – М.: Изд-во ИПУ РАН, 2010. С. 39-45.
- [6]. Опарин Г.А., Феоктистов А.Г. Модели и инструментальные средства организации распределенных вычислений // Параллельные вычисления и задачи управления: Материалы IV Междунар. науч. конф. – М.: Изд-во ИПУ РАН, 2008. С. 1126-1135.
- [7]. Опарин Г.А., Феоктистов А.Г., Александров А.А. Графическая инструментальная среда для описания модели распределенной вычислительной системы // Вестник ИрГТУ. 2006. Т. 3. № 2 (26). С. 35-40.
- [8]. Опарин Г.А., Новопашин А.П. Булевы модели и методы планирования параллельных абстрактных программ // Автоматика и телемеханика. 2008. № 8. С. 166-175.
- [9]. Опарин Г.А., Новопашин А.П. Технология синтеза модульных параллельных программ для DVM-системы // Интеллектуальные системы: Тр. VII Междунар. симпозиума. / Под ред. К.А. Пупкова. – М.: РУСАКИ, 2006. С. 468-471.
- [10]. Опарин Г.А., Феоктистов А.Г., Новопашин А.П., Горский С.А. Инструментальный комплекс ORLANDO TOOLS // Программные продукты и системы. 2007. № 4. С. 63-65.
- [11]. Сидоров И.А., Опарин Г.А., Феоктистов А.Г. Разработка и применение распределенных пакетов прикладных программ // Программные продукты и системы. 2010. № 2. С. 108-111.
- [12]. Бычков И.В., Корсуков А.С., Опарин Г.А., Феоктистов А.Г. Инструментальный комплекс для организации гетерогенных распределенных вычислительных сред // Информационные технологии и вычислительные системы. 2010. № 1. С. 45-54.