

## Научно-образовательный портал параллельных вычислений

*К. Е. Афанасьев, А. В. Демидов, С. Н. Трофимов*  
Кемеровский государственный университет, Кемерово

Высокопроизводительные вычисления на ЭВМ с параллельной архитектурой получили широкое распространение в науке и промышленности. Следуя тенденциям рынка, российские вузы все шире внедряют тематику параллельного программирования в свои учебные программы. Для выпуска качественных специалистов в данной области необходимо иметь как информационно-методические материалы для теоретической подготовки, так и многопроцессорные вычислительные системы для проведения практикумов.

Кафедра ЮНЕСКО по новым информационным технологиям КемГУ приступила к чтению спецкурсов по тематике параллельных вычислений с момента своего основания в 2000 г. За это время разработано несколько курсов и накоплен большой опыт по их ведению, а сотрудники кафедры используют высокопроизводительные ресурсы не только в учебных, но и в научно-исследовательских целях. Для систематизации имеющихся материалов и организации удаленного доступа к имеющимся кластерным системам, в КемГУ разрабатывается информационно-вычислительный портал (ИВП).

ИВП является единым пространством для организации учебной и научной деятельности вуза в сфере решения различных задач с использованием вычислительного эксперимента, организации научных конференций, хранения информационных ресурсов и работы с ними.

Создание типового информационно-вычислительного портала ведется с целью обеспечения:

- доступа к пополняемой и обновляемой электронной библиотеке информационных образовательных и научных ресурсов по тематике параллельных вычислений;
- проведения численного эксперимента на высокопроизводительном кластере в режиме удаленного доступа;
- представления информации о предстоящих и проходящих научных конференциях по тематике портала;

- представления информации о жизни университета в целом;
- организации виртуальных обсуждений (форумов) — новых видов коммуникации людей, способных конкурировать с традиционными СМИ;
- доступа к пополняемой и обновляемой коллекции аннотированных ссылок на ресурсы в Интернете.

В основе технологических решений Портала лежат следующие принципы.

**Открытость.** Для обеспечения перспективы развития Портала предусмотрена интеграция гетерогенных вычислительных компонент и различных приложений.

**Модульность.** Архитектура Портала является гибкой и допускает простое, без структурных изменений, развитие и наращивание функций и ресурсов системы в соответствии с расширением сфер и задач ее применения. Этот принцип также позволяет проводить отладку, удаление, добавление компонент системы без потери работоспособности.

**Масштабируемость.** Архитектура Портала позволяет наращивать производительность системы, объемы хранимой и обрабатываемой информации без длительной остановки работы Портала и модификации его программного кода.

**Индивидуализация.** Портал обеспечивает удобный индивидуальный доступ для всех потенциальных групп пользователей с функциональностью, соответствующей задачам, решаемым каждой группой пользователей.

**Функциональное соответствие.** Архитектура Портала отвечает текущим и перспективным целям и функциональным задачам создаваемой системы.

**Управляемость.** Предоставляются гибкие, полнофункциональные механизмы управления Порталом на всех уровнях архитектуры.

**Системность.** Архитектура Портала такова, что все взаимосвязанные подсистемы строятся по единой методологии и отвечают единым принципам взаимодействия, надежности и управления.

**Иерархическая структура.** Структура Портала основывается на трехуровневой модели:

- презентационный уровень, обеспечивающий представление информационных объектов и сервисов Портала пользователям;

- прикладной уровень — система серверов приложений, обеспечивающая выполнение преобразований информационных объектов в соответствии с заданными процедурами и правилами;
- уровень хранения данных, реализующий задачи контроля за целостностью и корректностью структуры информационных объектов.

Принципы интеграции данных и систем. Унификация обмена и представления данных на Портале основывается на:

- принятии формата XML в качестве основного стандарта для интеграции и представления данных в Интернете для всех систем общественного сектора: XML для хранения и транспортировки данных, XSLT для преобразования данных, XML-схемы для интеграции данных;
- принятии веб-браузеров в качестве единственной клиентской программы.

В число технических и организационных задач, которые решаются в ходе разработки Портала, входят:

- разработка способов представления информации, механизмов доступа к ней с целью улучшения восприятия и упрощения освоения системы неподготовленными пользователями;
- стандартизация процедур и протоколов обмена информацией между различными информационными системами, взаимодействующими с Порталом;
- исключение дублирования информационных объектов и программных компонентов в различных пользовательских сервисах Портала;
- организация доступа к территориально распределенным вычислительным ресурсам параллельной архитектуры через единый пользовательский интерфейс.

Таким образом, Портал является основной технологической единицей общей инфраструктуры единой информационной научно-образовательной среды [1]. По сути, Портал — это точка с определенным регламентом доступа к научно-образовательной жизни вуза (информационным ресурсам, научным конференциям, вычислительным экспериментам) независимо от места расположения потенциального пользователя, учебного заведения или места подключения пользователя сети Интернет.

Доступ ко всем подсистемам Портала осуществляется через единую систему безопасности, что позволяет пользователю, зарегистрировавшись в системе, получить доступ ко всем его функциям в рамках имеющихся привилегий.

Целевой аудиторией портала являются студенты, преподаватели, научные сотрудники, организаторы и участники конференций, другие пользователи.

В качестве основных разделов Портала выделяются наиболее значимые, с точки зрения большинства пользователей, аспекты деятельности. По мере развития и реализации новых сервисов и информационных объектов Портала предусматривается возможность создания альтернативных (в т. ч. персонализированных) рубрикаторов, ориентированных на более узкие целевые группы пользователей.

### **Основные подсистемы Портала**

Система удаленного доступа и управления распределенными вычислительными ресурсами — единая точка доступа к набору удаленных кластерных установок [2].

Основные задачи системы:

- хранение пользовательских файлов с исходными и бинарными кодами программ, входными и выходными данными;
- хранение метаданных о пользовательских файлах, необходимой для организации запуска программ на кластерных установках;
- отслеживание состояния доступных вычислительных ресурсов;
- взаимодействие с вычислительными ресурсами для передачи файлов и команд между ресурсом и хранилищем данных.

Система построена по трехзвенной архитектуре «клиент-сервер-сервер», где клиентским приложением является web-браузер, промежуточным сервером выступает соответствующий компонент портала («Менеджер вычислительных ресурсов»), а конечным сервером — удаленный кластер, где производится запуск программ.

Через портал пользователь может размещать в базе данных свои расчетные программы и задавать для них условия запуска.

Менеджер вычислительных ресурсов просматривает базу данных и выбирает из нее новые задания, которые нужно отправить на удаленные кластеры. При обращении кластера за новым заданием, менеджер отдает ему файлы новой программы и забирает файлы результатов ранее отправленных программ. Для управления кластером на нем устанавливается удаленный агент, который осуществляет связь с менеджером и обрабатывает полученные от него команды по запуску приложений и передаче файлов.

Каталог ресурсов — является хранилищем описаний электронных ресурсов и их полнотекстовых экземпляров. Пользователь может размещать новые ресурсы, выполнять редактирование, удаление ресурсов, передавать данные на экспертную оценку. По результатам экспертной оценки ресурс может быть опубликован (размещен) в каталоге ресурсов для просмотра.

Основные сервисы каталога ресурсов:

**Поисковые механизмы** — для обеспечения поиска информации на Портале реализованы следующие поисковые механизмы:

- быстрый поиск — поиск по названию объекта;
- поиск по рубрикатору;
- многокритериальный поиск по информационным объектам определенного типа с возможностью задания значеный искомым метаданных.

**Управление просмотром** — выполняет действия по настройке параметров просмотра результатов поиска в каталоге ресурсов.

Результатом поиска является список объектов (описаний ресурсов), который содержит информацию об объекте в зависимости от выбранных параметров просмотра. В системе предусмотрены различные варианты просмотра:

- вывод обязательной информации о ресурсе;
- вывод дополнительной информации о ресурсе;
- вывод полной информации о ресурсе;
- вывод определенного набора полей описания ресурса.

Результатом выполнения управления просмотром будут данные для просмотра, сформированные определенным образом.

Предоставление полнотекстовой, ознакомительной или полнотекстовой-постраничной версии электронного ресурса осуществляется в зависимости от прав доступа пользователя электронной библиотеки.

Сервер конференций — централизованная система для обслуживания и информационного сопровождения научных конференций, обладающая следующими функциями:

- создание и редактирование свойств конференции;
- публикация информационных сообщений (итогов конференции и другой информации);
- единая регистрация участников на сервере конференций;
- просмотр текущего состояния заявки;
- электронная рассылка информации участникам;
- подача заявки на текущую конференцию с возможностью указать тип участия и проживания в гостинице;
- публикация тезисов и статей.

Публикация информационных сообщений и других документов осуществляется посредством регистрации через веб-интерфейс самой конференции, а затем размещения необходимой информации.

### **Архитектура и технологии Портала**

Обмен данными проходит по стандартным сетевым протоколам в среде Intranet/Internet. Такая архитектура позволяет работать с системой из любой точки, имеющей выход в Intranet/Internet, а также обеспечивает независимость от платформы и минимальные требования к клиентам.

Клиент для системы выбирался максимально «тонким», поэтому в роли него выступает стандартный веб-браузер. Такой выбор позволяет минимизировать системные требования к клиентскому месту, а также снизить расходы за счет отсутствия потребности в установке дополнительного программного обеспечения, помимо уже имеющегося стандартного, и обучении пользователей работе с ним.

Посредством веб-браузера пользователь связывается с сервером Портала. Данные хранятся на сервере базы данных. Обмен данными между сервером приложений и сервером базы данных осуществляется по локальной сети (Intranet).

Такая архитектура позволяет работать с системой из любой точки, имеющей выход в Internet, а также обеспечивает независимость от платформы и минимальные требования к клиентам.

Подсистемы Портала располагаются на сервере приложений (Apache Tomcat), предназначенного для реализации технологии Java Servlets.

Вся логика приложений расположена на стороне сервера БД и реализована в виде набора пакетов. Для каждой таблицы, а также для функциональных блоков создаются пакеты для осуществления операций с данными. Тем самым обеспечивается возможность манипуляции необходимыми данными в каждой из подсистем путем вызова хранимых процедур.

Для реализации web-интерфейса используются технологии J2EE и XML:

- формирование выходных данных для отображения в браузере: XSLT;
- формирование списков и отчетов в формате PDF: XSL-FO;
- выходные данные XHTML, CSS.

На сервере приложений Apache Tomcat используется библиотека KemSUWeb, разработанная в центре новых информационных технологий (ЦНИТ) КемГУ [3]. Библиотека KemSUWeb располагается на сервере приложений и обеспечивает единую среду для создания приложений, основанных на трехуровневой архитектуре в среде Internet за счет адаптеров, которые удовлетворяют различные потребности разработчика: в операциях с БД, в защите информации, в управлении ходом приложения.

Реализация ИВП ведется в рамках выполнения проекта № РНП.3.2.3.4256 аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы (2006–2008 гг.)».

## Литература

1. **Гудов, А. М.** Система электронного документооборота [Текст] / А. М. Гудов, С. Ю. Завозкин // Тр. VI Всерос. науч.-практ. конф. «Системы автоматизации в образовании, науке и производстве». – Новокузнецк, 2007. – С. 278–281.

2. **Демидов, А. В.** Система удаленного доступа и управления вычислительными ресурсами [Текст] / А. В. Демидов, Е. В. Буслов, М. А. Пospelов // Гидродинамика больших скоростей

и численное моделирование : мат-лы III Международ. летней школы. — Кемерово : ИНТ, 2006. — С. 365—370.

3. **Гудов, А. М.** Об одной технологии создания web-интерфейса к программному обеспечению [Текст] / А. М. Гудов, Е. А. Ростовцев // Тр. III регион. науч.-практ. конф. «Информационные недра Кузбасса». — Кемерово, 2004. — С. 84—88.