

ВЕБ-ПОРТАЛ СИСТЕМЫ УПРАВЛЕНИЯ СУПЕРКОМПЬЮТЕРОМ

А.Л. Головинский, Л.Ф. Белоус

Введение

На сегодняшний день доступ к вычислительным ресурсам предполагает достаточно высокие требования к квалификации пользователей. Им необходимо знать массу дополнительных деталей, по сути чисто технического характера, которые необходимы специалисту в своей области помимо знания свойств применяемых пакетов программ и их специфического языка. Не упростило ситуацию и появление грид-доступа, скорее наоборот, добавился еще один уровень сложности, обычно предполагающий умение работать в режиме командной строки в операционной среде Unix. Отметим также, что проблема администрирования вычислительного кластера остается непростой и трудоемкой задачей.

Рассматриваемая работа является попыткой комплексного решения обеих проблем. Предлагается интегрированное решение в виде веб-портала для системы управления кластером, которое позволяет пользователям управлять прохождением своих задач, не обязуя их изучать многочисленные детали операционного окружения суперкомпьютера. Что касается администрирования кластера, то для этой цели разработан удобный сервис для управления кластером, включающий, конечно, и режим консоли.

Данный проект внедрен и успешно используется на грид-кластерах Института кибернетики им. В.М. Глушкова НАН Украины [1], г. Киев (<http://icybcluster.org.ua>), в Физико-техническом институте низких температур им. Б.И. Веркина НАН Украины, г. Харьков (<http://cluster.ilt.kharkov.ua>), а также в ряде других академических учреждений, входящих в состав Украинского Академического Грида (УАГ) (<http://lcg.bitp.kiev.ua>).

Концепция графического интерфейса пользователя кластера

Задача интерфейса - обеспечить выполнение всех возможных операций пользователя только средствами интерфейса. Он должен как можно полнее удовлетворять потребностям пользователя, учитывать его интересы, привычки, специфику его задачи. Большинство ученых работают с готовыми программными пакетами. Поэтому им необходима удобная, не перегруженная дополнительными функциями, среда для редактирования файлов входных данных, запуска параллельных программ, оперативного просмотра в режиме on-line выходных файлов. Прикладные программисты используют кластер как инструмент настройки параллельных программ, и кроме вышеуказанных средств им дополнительно необходима среда для компиляции с поддержкой популярных компиляторов и прикладных библиотек, редактор для исходных текстов программ.

Основными операциями, которые выполняют пользователи комплекса, являются:

- файловые операции;
- выполнение задач на ресурсах кластера;
- слежение за процессом решения и просмотр результатов выполнения задач;
- взаимодействие между пользователями и администраторами.

Рассмотрим их более детально.

Файловые операции. Файлы пользователей находятся в персональных каталогах, защищенных от доступа посторонних лиц. Интерфейс доступа к файлам обеспечивает все традиционные операции с файлами: создание, редактирование, удаление, и т.д. Дополнительно предусмотрены сортировки по полям "имя", "размер", "время создания", что необходимо для удобства работы с каталогами при большом количестве файлов.

Большой объем файловой системы кластера обуславливает необходимость направленного поиска по файловой системе. В графическом интерфейсе предусмотрен поиск по регулярным выражениям, по дате создания файла.

Обмен данными между рабочей станцией и кластером осуществляется путем загрузки и отправки файлов. Для эффективной работы предусмотрен также обмен каталогами, который обеспечивается их предварительной архивацией.

При работе с текстовыми файлами используется синтаксически ориентированное редактирование.

Выполнение задач на ресурсах кластера. Постановка задачи в очередь осуществляется через соответствующий интерфейс, который позволяет задать все необходимые параметры вычислительной задачи менеджеру ресурсов кластера.

Для исходных текстов программ предусмотрена интеллектуальная система компиляции, которая анализирует текст программы, выбирает соответствующий язык программирования и сценарий компиляции.

В интерфейсе предусмотрен режим работы с общесистемными программными пакетами. В этом режиме некоторые параметры запуска задаются автоматически со значениями по умолчанию, что значительно упрощает работу с ними.

Просмотр хода выполнения процесса и результатов выполнения задач. Главным средством контроля процесса выполнения задачи является просмотр в реальном времени журналов задач, в котором отражается состояние выполнения и обнаруженные ошибки. Для этого поддерживается отдельный режим просмотра файлов - слежение за файлом. При этом все изменения в файле журнала сразу же отображаются в соответствующем окне. Дополнительными средствами контроля выступают просмотр уровня загрузки процессоров, объема занятой задачей оперативной памяти на узлах.

Взаимодействие между пользователями и администраторами. Для общения между пользователями и администраторами предусмотрена система передачи сообщений и форум. Некоторым пользователям нужно совместно использовать ресурсы с другими пользователями, такая возможность тоже предусмотрена.

Графический интерфейс администратора кластера

Администратор выполняет функции организации вычислительного процесса суперкомпьютера.

Каждый администратор имеет учетную запись, как и обычный пользователь, но с расширенными возможностями. Специфические функции, которые выполняет администратор, следующие:

- вход в систему от имени произвольного пользователя;
- администрирование очередей задач;
- просмотр состояния оборудования кластера;
- управления ресурсами кластера;
- работа с базой учетных записей пользователей;
- запуск диагностических задач;
- просмотр системных журналов;

Опишем более детально эти функции.

Вход в систему от имени произвольного пользователя. Необходимость в таком средстве возникает при появлении у пользователя проблем с работой на кластере. Вход от имени конкретного пользователя позволяет достичь повторяемости ошибок, локализовать их в среде, где они возникают.

Администрирование очередей задач. Очереди задач требуют от администратора постоянного надзора. В интерфейсе предусмотрена возможность просмотра параметров задач в очереди, отмены их в случае возникновения ошибки или по каким либо иным причинам.

Просмотр состояния оборудования кластера. Состояние оборудования требует постоянного внимания со стороны администратора. Своевременное информирование об авариях является одной из основных задач интерфейса. Средства мониторинга включают просмотр состояния узлов, системы хранения данных, температуры в кластерном зале и т.д.

Управления ресурсами кластера. Основным ресурсом кластера являются вычислительные узлы. Администратор имеет возможность динамически менять общее количество узлов, доступных для назначения задачам. Предусмотрены возможности отключения, включения, блокировки назначения узлов конкретным очередям, возможность блокировки отдельной очереди или всех очередей задач. Также предусмотрена возможность приостановки очередей задач.

Работа с базой учетных записей пользователей. Работа с учетными записями предусматривает просмотр всей базы, изменение полей записей, регистрацию и удаление пользователей.

Запуск диагностических задач. Диагностические задачи – это особый класс задач. Они позволяют определить характеристики системы, проверить надежность кластера в целом. Запуск таких задач может быть осуществлен как по расписанию, так и по требованию. Предусмотрен интеллектуальный анализ журнала диагностики с определением компонент с пониженными характеристиками.

Просмотр системных журналов. Просмотр журналов компонентов кластера позволяет выявить неизвестные ранее проблемы, детектирование которых не предусмотрено в соответствующих разделах мониторинга. Предусмотрена возможность фильтрации журналов по определенным ключевым словам, что упрощает анализ больших объемов текста.

Веб-портал кластерных вычислений

Графический интерфейс должен быть доступным с произвольной рабочей станции как для пользователя, так и для администратора. Поэтому он должен быть кроссплатформенным и не требовать установки дополнительного программного обеспечения. Именно по этой причине в качестве технологической основы для реализации данного проекта избраны веб-сервисы. Они предоставляют наилучшую кроссплатформенность на сегодняшний день и их средств достаточно для обеспечения выполнения заданного круга задач. Общий вид интерфейса показан на Рис. 1. Дизайн портала и его функциональность в значительной степени могут быть произвольными, и определяется на стадии проектирования по согласованию с заказчиком.

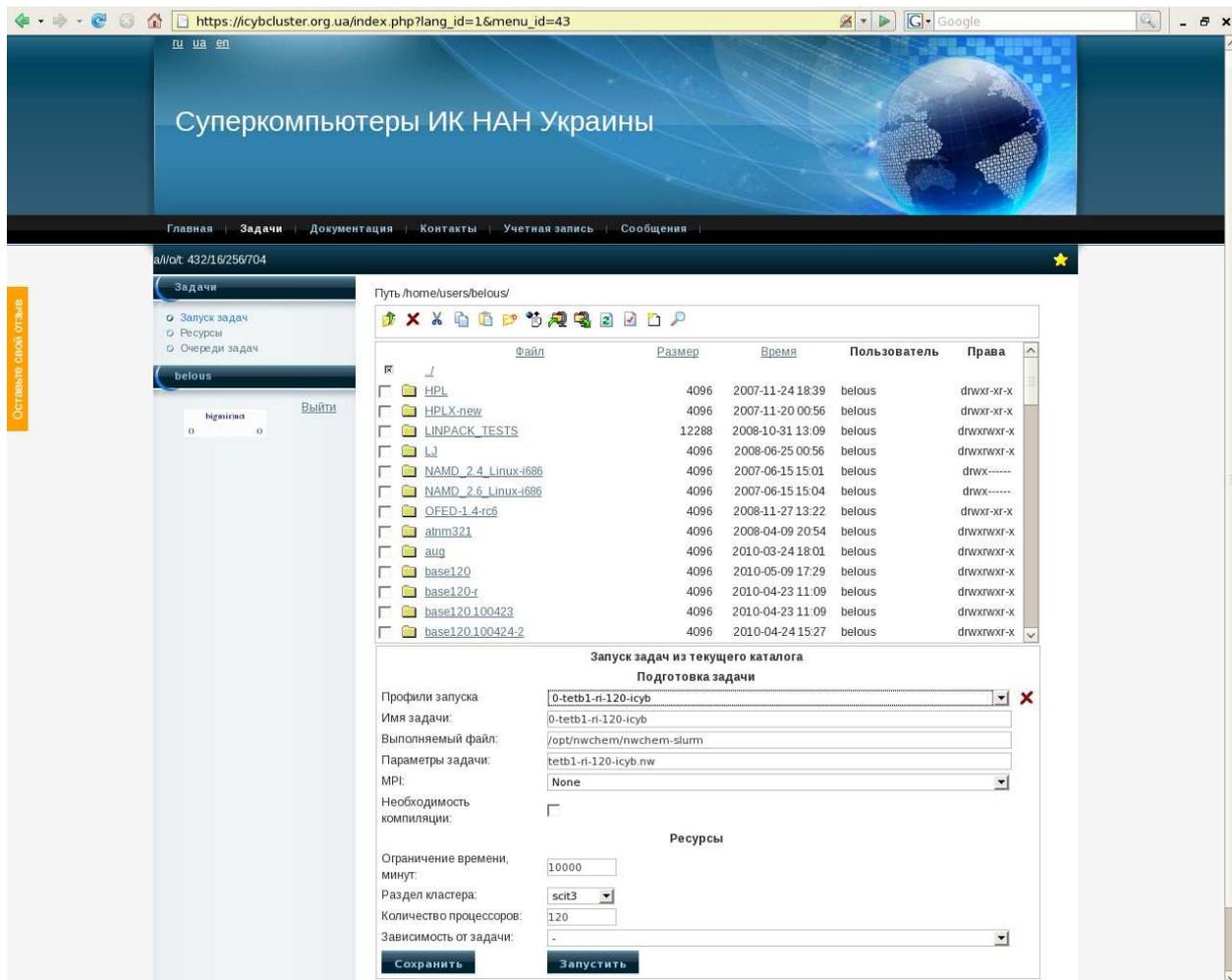


Рис. 1. Графический интерфейс управления суперкомпьютером

Структура интерфейса

Интерфейс имеет модульное строение и состоит из двух слоев программного обеспечения: веб-части, которая отвечает за диалог с пользователем, и промежуточного уровня взаимодействия с системным программным обеспечением суперкомпьютера (Рис. 2).

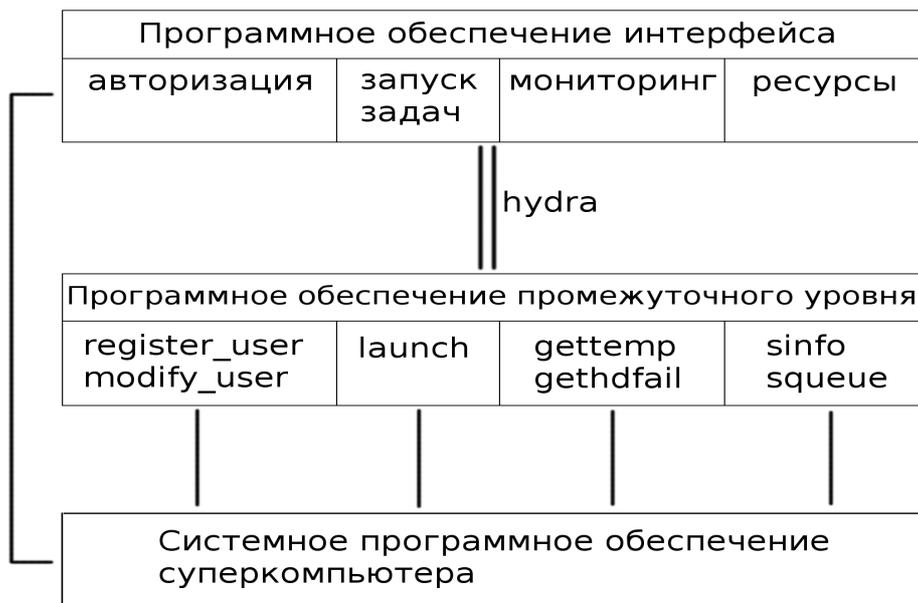


Рис. 2. Структура ПО интерфейса

Среди модулей можно выделить: подсистему авторизации и работы с базой пользователей, модуль взаимодействия с менеджером ресурсов кластера и модули диагностики разных подсистем суперкомпьютера.

Коротко об их основных функциях.

Модуль авторизации. Поддерживается авторизация с помощью систем PAM и LDAP [2]. Авторизация осуществляется непосредственно, без промежуточного программного обеспечения. Работа с базой пользователей включает операции авторизации пользователя, добавление и удаление пользователей, редактирование их учетных записей. Для работы с LDAP в файле конфигурации `hydra-adm.conf` указывается адрес `ldap`-сервиса, суффикс базы, учетные данные администратора базы. Для работы с PAM используется модуль `php-auth-pam`, который настраивается отдельно.

Модули диагностики. Разработанные тесты анализируют основные параметры работы суперкомпьютера: состояние жестких дисков, сетевых соединений, сети Infiniband, температуры узлов, данные сенсоров IPMI. Скрипты диагностики и мониторинга находятся на промежуточном уровне программного обеспечения. Они выполняются от имени суперпользователя на шлюзе кластера. Выполнение осуществляет системный планировщик `sgon` в соответствии с расписанием. Расписание устанавливается в соответствии с потребностью администраторов в предоставляемой информации.

Выполнение административных операций. Часть административных операций выполняется от имени суперпользователя. К таким операциям относятся: редактирование базы пользователей, удаление произвольных задач из очереди менеджера ресурсов. Для выполнения таких задач в интерфейсе предусмотрен следующий подход. В суперкомпьютере создается служебный пользователь `portaladmin`, который через `sudo` может выполнять заданные команды из `/opt/hydra/lib` от имени `root`. Таким образом, графический интерфейс инициирует выполнение операции от имени `portaladmin`, после чего происходит переключение на суперпользователя и выполнение команды.

Дистанцирование веб-сервера от суперкомпьютера

Высокопродуктивные файловые системы часто недостаточно надежны. Аварии в файловых системах Lustre [3], NFS-RDMA [4] чаще всего являются причинами остановок в работе суперкомпьютеров, поскольку ведут к фатальным сбоям в работе компонент: серверов, узлов, в частности веб-сервера.

Для того, чтобы решить проблему надежности работы веб-сервера, разработано специализированное клиент-серверное ПО, которое позволяет скриптам веб-интерфейса инициировать выполнение команд на шлюзе кластера от имени и пользователей, и администраторов, а также обмениваться данными.

Клиентская часть `hctl` находится на веб-сервере. Серверная часть `hydrad` выполняется как системная служба на шлюзе суперкомпьютера. Данное ПО работает в следующих режимах:

- выполнение команд пользователя и предоставление результата выполнения;
- чтение данных из файла;
- запись данных, полученных веб-сервером, в файл файловой системы кластера (реверсный режим передачи данных);

- выполнение команд администратора от имени суперпользователя осуществляется через служебного пользователя rootadmin, который переключается на суперпользователя и выполняет команду из списка разрешенных.

Таким образом, даже в случае значительной аварии веб-сервер будет работать, и пользователи смогут получить информацию о причине проблемы, сроке ее устранения и т.д.

Запуск задач с помощью менеджера ресурсов кластера

Запуск задачи на выполнение осуществляется через соответствующую форму интерфейса. На показанном ранее Рис. 1 выбрано меню запуска задач с этой формой. Пользователь задает параметры задачи и название файла с программой, параметры командной строки, количество процессоров, время выполнения, выбирает MPI-среду. При необходимости указывается режим компиляции программы из исходных текстов. В таком случае автоматически создается сценарий компиляции программы из ее исходных текстов, автоматически запускается компиляция, а после ее успешного завершения начинается выполнение программы.

Единожды выполненные действия по заполнению формы запоминаются в виде профиля запуска с уникальным именем, так что при повторении этого действия достаточно только указать имя профиля (возможно, дополнительно отредактировав его), и нажать кнопку «запустить». Через некоторое время создается своего рода библиотека этих профилей, отражающая специфику повторяющихся действий пользователя. С другой стороны, для неподготовленных пользователей такая библиотека может быть заранее составлена системным администратором, что фактически позволит специалистам в своей области вообще не знакомиться с деталями запуска своего пакета, а сосредоточится только на специфике языка пакета и поведения задачи. Структура запуска вычислительной задачи показана на Рис. 3.

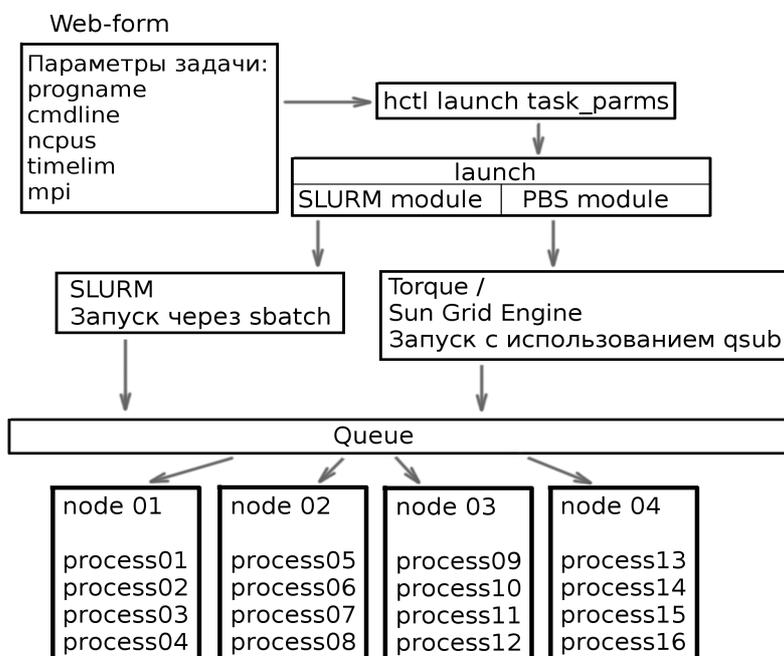


Рис. 3. Структура запуска задачи

Запуск производится через соответствующие модули графического интерфейса. Пользователь задает параметры задачи на странице "Запуск задачи". Все данные передаются модулю launch, который производит запуск через менеджер ресурсов кластера. Тип менеджера задается в файле конфигурации интерфейса. Далее задача ставится в очередь и после предоставления ей ресурсов начинается выполнение на вычислительных узлах.

Интерактивное взаимодействие с пользователем

Некоторые компоненты интерфейса отображают информацию, которая часто изменяется. К ним принадлежат модули ресурсов, очереди задач, просмотра файла вывода выполняющейся задачи. Интерфейс осуществляет обновление информации без необходимости перезагрузки страницы с помощью технологии AJAX [5]. Такой подход позволяет пользователям осуществлять интерактивную работу со своими вычислительными задачами на кластере, что очень важно для многих исследований в областях физики и химии с не вполне формализованными алгоритмами.

Заключение

В статье описана программная система, которая обеспечивает веб-интерфейс системы управления суперкомпьютером для пользователей и администраторов. С нашей точки зрения описанная среда способствует более широкому применению многопроцессорных вычислительных систем, поскольку сильно упрощает их использование учеными и программистами. Более подробно с системой и условиями ее распространения можно познакомиться на сайте разработчиков <http://www.melkon.com.ua/>.

ЛИТЕРАТУРА:

1. Якуба А.А., Головинский А.Л., Бандура А.Ю., Горенко С.А., Ефременюк Д.А. Портал кластерных вычислений для управления вычислительными процессами на суперкомпьютерном комплексе. Кибернетика и системный анализ, 2009, № 6, с. 97-105.
2. <http://ru.wikipedia.org/wiki/LDAP>
3. <http://www.lustre.org/>
4. <http://nfs-rdma.sourceforge.net/>
5. [http://en.wikipedia.org/wiki/Ajax_\(programming\)/](http://en.wikipedia.org/wiki/Ajax_(programming)/)