

Система управления суперкомпьютером SCMS

А.Л. Головинский¹, А.Л. Маленко¹, Л.Ф. Белоус², С.В. Баранник³

Институт кибернетики им. В.М. Глушкова НАН Украины¹,
Физико-технический институт им. Б.И. Веркина НАН Украины²,
Институт сцинтилляционных материалов НАН Украины³

В работе представлена система управления суперкомпьютером, соединяющая в себе простой интерфейс и широкие возможности. Система предназначена как для обычных пользователей суперкомпьютера, так и для администраторов. Поддерживается большое количество аппаратных решений.

1. Введение

Проблема понятного и доступного пользователю программного интерфейса является актуальной и важной для всех видов программного обеспечения. Особенно это касается сферы высокопроизводительных вычислений, где традиционные консольные интерфейсы для доступа пользователей очень специфичны и требуют дополнительных технических знаний. Кроме особенностей своей научной области, для работы с суперкомпьютером пользователь должен понимать работу операционной системы кластера, процесс запуска задач, работу с компиляторами и т.д.

Развитие грид-технологий не улучшило текущее состояние вещей. Ведь работа в гриде – это еще один дополнительный уровень сложности, который требует знания грид-инструментов командной строки, нового синтаксиса запуска задач, межкластерной совместимости программных окружений и т.д.

С другой стороны, задача администрирования кластерных систем остается сложной и трудоемкой, требует высокого уровня квалификации администратора, при этом отнимает много времени.

Система управления суперкомпьютером SCMS 4.2 – это попытка предложить комплексное решение как для доступа пользователей, так и для администрирования кластеров. Она представляет собой веб-приложение, на котором пользователи могут легко запускать и контролировать свои задачи без необходимости изучать многочисленные детали работы суперкомпьютера и операционного окружения грида.

Для администрирования кластера в системе предусмотрены удобные средства мониторинга состояния оборудования суперкомпьютера, управления ресурсами кластера, очередями задач и пользователями.

Проект SCMS является продолжением работ, посвященных системам управления, интерфейсам суперкомпьютера и работе в гриде [1–3].

Система управления суперкомпьютером SCMS 4.2 является готовым продуктом, который успешно используется на грид-кластерах Института кибернетики им. В.М.Глушкова НАН Украины, г. Киев, Института сцинтилляционных материалов НАНУ, в Физико-техническом институте низких температур им. Б.И. Веркина НАН Украины, г. Харьков и Институте механики сплошных сред РАН, г. Пермь.

2. Интерфейс пользователя

В рамках проекта решена задача построения эффективного интерфейса как специализированного инструмента, имеющего, с одной стороны, гибкие возможности, а с другой стороны, достаточного простого и интуитивно понятного для современного специалиста.

Исследуя типовые способы работы пользователей на украинских суперкомпьютерах, мы отметили три главных направления: работа ученого, грид-пользователя и программиста.

ста. Для каждого из этих типов работы предложены соответствующие средства, обеспечивающие удобную и эффективную работу.

Большинство ученых, использующих кластер, работают с пакетами готового программного обеспечения. Для них разработана простая и удобная среда для редактирования исходных файлов, запуска параллельных программ и он-лайн просмотра результатов задач.

Прикладные программисты, в свою очередь, используют кластер в качестве инструмента разработки и тестирования параллельных программ. Для этого создано окружение для компиляции с поддержкой популярных компиляторов и библиотек, а также редактор исходного кода с подсветкой синтаксиса.

Ключевым элементом интерфейса пользователя является форма запуска вычислительных задач (рис. 1), позволяющая запускать как программы собственной разработки, так и большие программные пакеты, установленные на кластере. Форма оборудована системой профилей, облегчающих повторные запуски задач и предустановленных пакетов ПО.

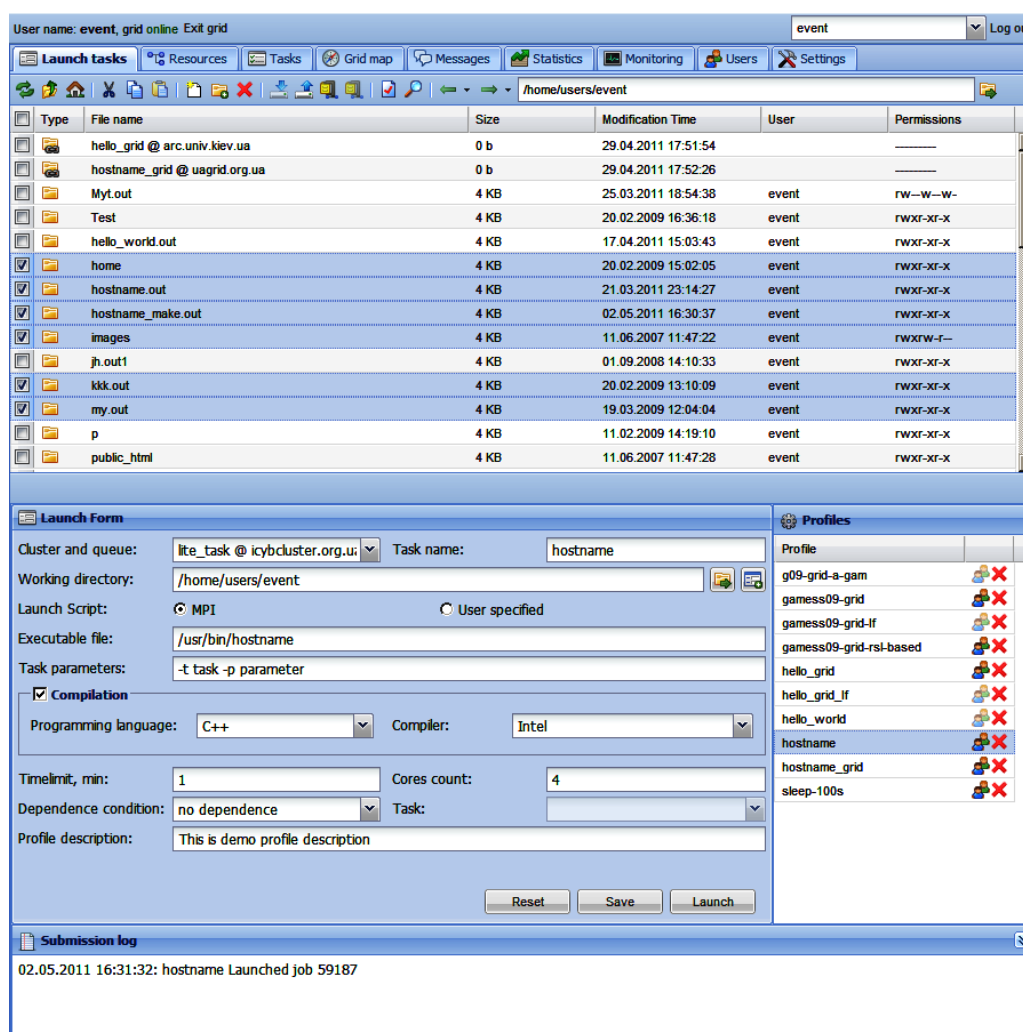


Рис. 1. Интерфейс пользователя SCMS

Для исходных файлов предусмотрена интеллектуальная система компиляции. Она определяет язык программирования и выбирает соответствующий сценарий компиляции. Поддерживаются сценарии языков программирования C/C++, Fortran 77, 90, 95 для компиляторов Intel и GNU. Для гибридного кластера на GPU-ускорителях доступно автоматическое построение сценария компиляции кода программы на языках CUDA и OpenCL.

Работа в гриде, в том числе запуск грид-задач, и на локальном кластере максимально

унифицированы. С точки зрения пользователя практически нет разницы, где запускается его задача: на локальном кластере или в гриде. Для полной поддержки грид-технологий пользователю достаточно иметь действительный грид-сертификат и грид-пароль.

3. Система контроля оборудования

Основной задачей суперкомпьютера является предоставление безотказного, круглосуточного, надежного и безопасного сервиса высокопроизводительных вычислений.

Основным способом решения этой задачи является создание единой интеллектуальной системы управления оборудованием вычислительного комплекса, которая может обеспечить непрерывный контроль всего аппаратного обеспечения (вычислительных узлов, хранилища, системы охлаждения, сетевого оборудования, каналов связи, блоков бесперебойного питания и т.д.) и базового программного обеспечения.

Было проведено исследование основных видов неисправностей и поломок больших вычислительных систем, приводящих к серьезным нарушениям их работы. Неисправности классифицированы по степени критичности и вероятности возникновения. Например:

- выход из строя системы кондиционирования, который приводит к перегреву всего кластера;
- выход из строя дисков системы хранения;
- выход из строя узлов;
- выход из строя различных программных сервисов.

Существующие системы мониторинга и диагностики, например, Ganglia и Nagios, успешно решают задачу получения подробной информации о текущем состоянии кластера по требованию администратора. Эти системы можно самостоятельно дополнить модулями извещения администратора по E-mail/SMS о критических событиях, например, перегреве узлов.

Однако, такие меры являются недостаточными, так как они предполагают наличие дежурного специалиста, который постоянно следит за состоянием кластера и в любой момент сможет решить возникшую проблему. На практике, такая методика имеет множество недостатков. Логично поручить заботу о "здоровье" суперкомпьютера самому суперкомпьютеру.

Авторами проанализирован опыт администраторов ряда суперкомпьютеров и выделены типовые сценарии поведения администратора во время выявления проблем аппаратного обеспечения и базового ПО кластера. Большинство из них можно и нужно автоматизировать.

Разработанная система SCMS позволяет оперативно решать проблемы кластера в автоматическом и полуавтоматическом режиме. Ее интеллектуальность заключается в возможности самостоятельного выявления опасной ситуации, ее оценки и принятия решения про выполнение необходимых действий.

Опишем для примера модуль автоматического отключения оборудования кластера при перегреве.

При отсутствии средств автоматического выключения выход из строя одного кондиционера системы охлаждения провоцирует эффект "домино": остальные кондиционеры работают на грани своих возможностей и один за другим выходят из строя, в течение короткого промежутка времени (10-60 минут) наступает полный коллапс системы охлаждения. Температура в кластерном зале возрастает до 60–70°C и выше, вследствие чего начинают выходить из строя узлы.

Такого сценария можно избежать, если кластер контролирует ситуацию. После отключения части охладителей термодинамический баланс оказывается разбалансированным, возникают зоны перегрева, при этом другая часть кластера может работать в штатном режиме.

Модуль защиты находит такие зоны и отключает узлы, оказавшиеся под угрозой, уменьшая тепловыделение. Последовательное выключение узлов кластера продолжается до наступления нового термодинамического равновесия. Такой подход рациональнее полного выключения кластера.

Эффективность системы продемонстрировала ее успешная эксплуатация на ряде украинских кластеров с 2009 года. Благодаря этой системе процент аварий сведен практически к нулю, что позволило повысить надежность работы суперкомпьютеров, уменьшить количество простоев, сэкономить на устранении аварий.

4. Обзор возможностей администратора

Интерфейс рабочего места администратора обеспечивает удобное выполнение основных действий по управлению вычислительным комплексом:

- управление узлами и разделами кластера;
- управление потоком задач;
- контроль оборудования кластера;
- управление учетными записями пользователей;
- выполнение диагностических задач;
- анализ статистики использования ресурсов;
- контроль качества предоставляемых сервисов.

Администратор создает разделы кластера из его узлов, указывает образы операционных систем узлов, включает, выключает и перезагружает узлы.

Управление менеджером ресурсов позволяет выбрать оптимальный алгоритм планировщика и приоритеты пользователей, а также просматривать очереди задач и отменять их.

Интерфейс администратора позволяет визуальнo контролировать основные характеристики работы кластера. В состав системы входит модуль оповещения о критических событиях с помощью электронной почты или SMS.

Администратор имеет полный комплекс средств для управления учетными записями пользователей: прием запроса регистрации пользователя, редактирование данных учетной записи, удаление пользователей.

Модуль учета системы SCMS постоянно собирает информацию об использовании ресурсов кластера и состоянии оборудования, позволяет создавать отчеты по пользователям, организациям за любой период времени. Статистика отображается в системе в табличном и графическом виде. Ее можно экспортировать в CSV или Excel формат.

Авторами разработан метод оценки качества предоставления сервиса [4]. На основании собранной статистики система автоматически вычисляет индекс ожидания ресурсов. Визуальное отображение индекса по неделям или месяцам дает администраторам четкое понимание того, насколько быстро обслуживаются запросы пользователей, позволяет оценить зависимость качества предоставленных ресурсов от нагрузки и структуры потребления.

На основе индекса ожидания администратор принимает решение об изменении параметров счетного поля и алгоритма планировщика задач, а также выставляет приоритеты и квоты пользователям.

5. Выводы

Описанная в данной статье система удовлетворяет основные потребности пользователей и администраторов суперкомпьютера. Разработанные средства позволяют работать на кластере неспециалистам, сократить время обучения пользователей работе на суперкомпьютере.

Таким образом, система способствует более широкому применению отечественных многопроцессорных вычислительных систем в науке и промышленности, поскольку упрощает их использование учеными и программистами.

Система SCMS позволяет разгрузить администраторов кластеров от рутинной работы, переключить их внимание на более важные глобальные задачи.

Разработанная система доступна в сети Интернет для публичного ознакомления с ее функционалом [6]. Параметры тестовой учетной записи: логин – *demo*, пароль – *cluster*.

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. 97–105.
2. Головинский А.Л., Белоус Л.Ф., Маленко А.Л. Веб-портал системы управления суперкомпьютером // Вычислительные методы и программирование. 2010. 11/7.
3. Головинський А.Л., Маленко А.Л., Черепинець В.В. Система керування суперкомп'ютером з підтримкою роботи у ґріді SCMS 4.0 // Матеріали міжнародного наукового конгресу з розвитку інформаційно-комунікаційних технологій та розбудови інформаційного суспільства в Україні. — 2011. — С. 28–29.
4. Головинський А.Л., Маленко А.Л. Аналіз ефективності планувальників черги задач для суперкомп'ютера з кластерною архітектурою // Матеріали конференції Високопродуктивні обчислення НРС-UA 2011. — 2011. — С. 64–69. Перевод статьи на русский язык: <http://melkon.com.ua/files/article-scheduling-rus.pdf>
5. Документация системы SCMS: <http://melkon.com.ua/>
6. <http://scms.melkon.com.ua/>