

УДК 519.687

ТЕХНОЛОГИЯ ПОДДЕРЖКИ УЧЕТА РЕСУРСОВ В ВИРТУАЛЬНЫХ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНЫХ ЛАБОРАТОРИЯХ

Федосин М.Е.

*ФГБУН «Институт системного программирования Российской академии наук»,
Москва, e-mail: fedosin@ispras.ru*

В статье рассматривается концепция виртуальных информационно-вычислительных лабораторий, которая появилась в результате поиска решения задачи по упрощению доступа к программному обеспечению для высокопроизводительных вычислений. В результате анализа этой концепции автор приходит к выводу, что она обладает недостатком, связанным с отсутствием контроля над аппаратными и программными ресурсами, что ограничивает класс приложений, которые можно установить в лабораторию, бесплатными, нетребовательными к ресурсам программными продуктами. Исходя из этого автор предлагает модифицированную концепцию виртуальных информационно-вычислительных лабораторий, которая содержит в себе дополнительный элемент – систему учета ресурсов. В статье подробно описывается работа лаборатории в новой концепции и приводятся отличительные особенности запуска приложений от уже существующей концепции. В статье также приводится обзор средств поддержки исследовательских проектов и способов их администрирования.

Ключевые слова: аппаратно-программные комплексы, облачные вычисления, SaaS, суперкомпьютеры, высокопроизводительные вычисления, учет ресурсов

THE TECHNOLOGY OF SUPPORTING OF RESOURCE ACCOUNTING IN VIRTUAL INFORMATION-COMPUTATIONAL LABORATORIES

Fedosin M.E.

*Federal state budgetary institution of science «Institute for system programming
of Russian academy of sciences», Moscow, e-mail: fedosin@ispras.ru*

The article discusses the conception of virtual information-computational laboratories that was the result of finding a solution to the problem to facilitate access to the software for high-performance computing. After the analysis of the conception, author concludes it has a disadvantage, linked with its lack of control over the hardware and software resources that limits the class of applications that can be installed in the lab, to free, resource-sparing software products. Based on this, the author proposes a modified conception of virtual information-computational laboratories, which contains an additional element - resource accounting system. The paper gives a detailed description of laboratory functioning in the new conception and discusses the distinctive features of running applications of already existing conception. The article also provides an overview of techniques to support research projects and ways to administer them.

Keywords: computer appliance, cloud computing, SaaS, supercomputers, high performance computing, resource accounting

Для проведения сложных расчетов, связанных с математическим моделированием, зачастую недостаточно обычных персональных компьютеров ввиду того, что они не обладают необходимой вычислительной мощностью. Поэтому для подобных задач применяются так называемые «суперкомпьютеры», которые позволяют получать результаты на несколько порядков быстрее.

К основным недостаткам использования суперкомпьютеров относятся их высокая стоимость и большие затраты, связанные с их эксплуатацией. Кроме того, отсутствие для ряда отраслей промышленности открытых программных пакетов (open source) для моделирования вынуждает пользователей покупать их коммерческие аналоги. В некоторых случаях стоимость программных решений и их дальнейшая поддержка сопоставима со стоимостью покупки суперкомпьютера, а иногда и значительно превышает ее.

Известно решение вышеназванной проблемы, описанное, в частности, в работе [1]. В ней в качестве упрощения доступа к

суперкомпьютерам описывается концепция виртуальных информационно-вычислительных лабораторий.

Одним из основных недостатков такой концепции является отсутствие контроля над аппаратными мощностями лаборатории и установленными программными продуктами, что не позволяет использовать лаборатории для предоставления клиентам облачного сервиса на коммерческих условиях.

Решением этой задачи может стать модификация существующей концепции и внедрение в нее ЭВМ с установленной системой учета ресурсов, которая бы обеспечивала контроль над использованием аппаратных и программных возможностей лаборатории, ограничивала доступ к вычислительным ресурсам в случае, если пользователь, запрашивающий доступ, не прошел проверку, осуществляла хранение результатов использования лаборатории и поддерживала формирование отчетов о проделанной работе. Подобное улучшение концепции расширяет класс программных

продуктов, которые можно устанавливать в инфраструктуру лаборатории, и создает возможность поддержки исследовательских проектов.

Модифицированная концепция виртуальных информационно-вычислительных лабораторий

Основная задача лабораторий – предоставление клиентам облачного сервиса (Softwareasa Service, SaaS), обеспечивающего удаленный интерактивный доступ к прикладному программному обеспечению без его модификации. Лаборатория представляет собой платформу, на которую можно установить программный продукт, предназначенный для работы на суперкомпьютерах, и обеспечить доступ к нему и к вычислительным ресурсам через интернет [2].

С точки зрения пользователя лаборатория представляет собой веб-сайт, с помощью которого осуществляется доступ к программному обеспечению для высокопроизводительных вычислений [3]. Для получения доступа к программному обеспечению лаборатории пользователю необходимо авторизоваться на сайте, выбрать

нужное ему приложение из предложенного списка и запустить его.

Запуск приложения приводит к созданию сессии, которая выполняется на вычислительных узлах лаборатории и отображается в браузере пользователя. При этом работа с приложением происходит в интерактивном графическом режиме и не отличается от работы с локальным настольным приложением.

Аппаратная часть лаборатории в существующей концепции содержит (рис. 1) [4]:

- Группу ЭВМ, обеспечивающих функционирование веб-ядра (1). Эта группа предоставляет интерфейс для работы с лабораторией и отвечает за авторизацию пользователей.
- Систему хранения данных для бинарных файлов приложений и домашних каталогов пользователей (2).
- Группу ЭВМ для хранения программных сессий (3). Группа хранит сеансы запущенных в лаборатории приложений (сессии), удаленный доступ к которым пользователь получает через веб-браузер.
- Высокопроизводительные вычислительные ресурсы (суперкомпьютеры) (4). Оборудование, на котором происходит основной счет приложений.

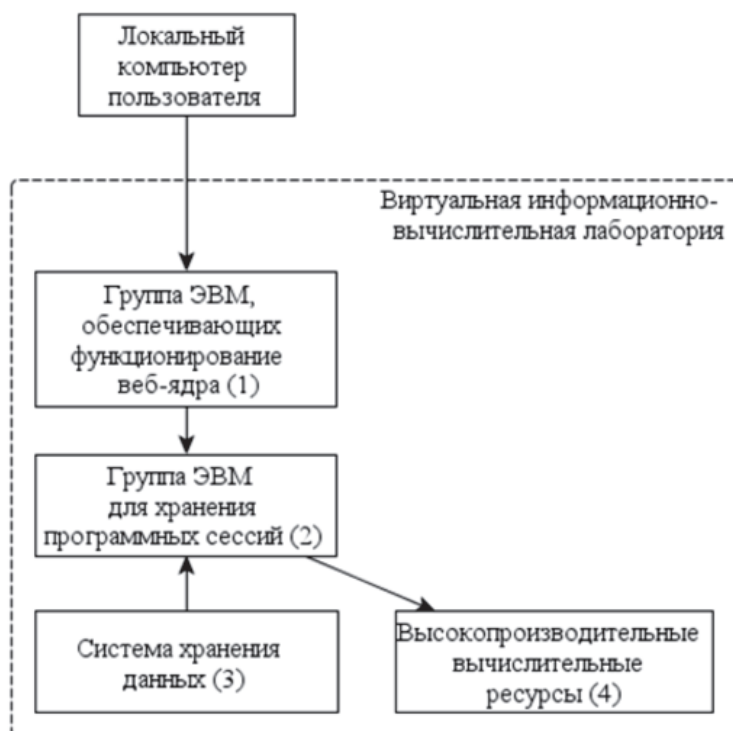


Рис. 1. Существующая концепция лабораторий

Ввиду отсутствия в существующей концепции контроля над ресурсами пользователь может запускать приложения на

неограниченный срок без ограничений на используемые вычислительные ресурсы. Из-за этого в лабораториях устанавливаются

ся только бесплатные, нетребовательные к ресурсам приложения.

Модифицированная концепция предполагает ввод дополнительной ЭВМ с работающим на ней сервером системы учета ресурсов (5) (рис. 2), в задачи которого входит:

1. Контроль над использованием ресурсов лаборатории.

2. Осуществление проверки на возможность пользователя получить доступ к приложениям лаборатории.

3. Ограничение доступа в случае, если проверка не пройдена.

4. Хранение результатов запусков.

5. Поддержка формирования отчетов по проектам.

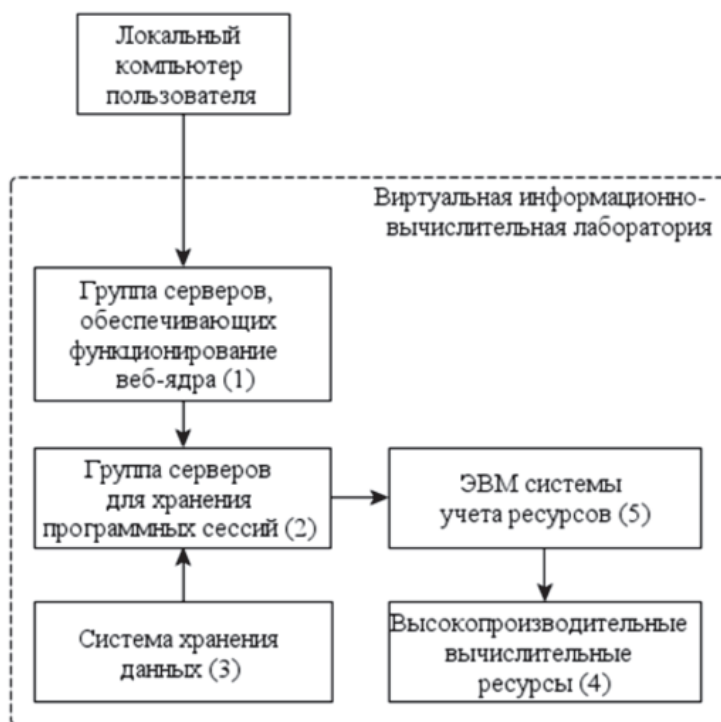


Рис. 2. Модифицированная концепция лабораторий

Ввод системы учета ресурсов в существующую концепцию позволяет расширить класс программных продуктов, которые можно устанавливать в инфраструктуру лаборатории, и привносит новые функциональные возможности, связанные с появлением средств поддержки проведения исследовательских проектов [5].

Работа системы учета ресурсов

Отличием работы лаборатории в модифицированной концепции является способ запуска приложений.

В существующей концепции после того, как пользователь получает удаленный доступ к приложению, им формируется запрос на счет, где указывается требуемое приложению количество ресурсов и время, в течение которого они будут использоваться. Затем запрос отправляется менеджеру ресурсов.

В свою очередь менеджер ресурсов обрабатывает запрос и загружает задание на вычислительные узлы. После завершения

счета менеджер возвращает результат приложению. На рис. 3 представлена UML-диаграмма активности, описывающая запуск приложений в существующей концепции.

Особенностью модифицированной концепции является то, что в отличие существующей концепции запрос на счет изначально пересылается приложением не менеджеру ресурсов, а системе учета ресурсов.

Система производит сверку запроса с данными в собственной базе данных, где хранится информация о доступном количестве ресурсов для каждого пользователя. Если запрашиваемые ресурсы меньше допустимых для пользователя, то система перенаправляет запрос менеджеру ресурсов, в противном случае возвращает сообщение о невозможности запуска.

Далее менеджер ресурсов, получив запрос, обрабатывает его и загружает задание на вычислительные узлы. После завершения счета менеджер возвращает результат, и система учета ресурсов заносит информа-

цию о нем в свою базу данных. Затем перенаправляет результаты приложению. На рис. 4 представлена UML-диаграмма актив-

ности, описывающая запуск приложений в модифицированной концепции.

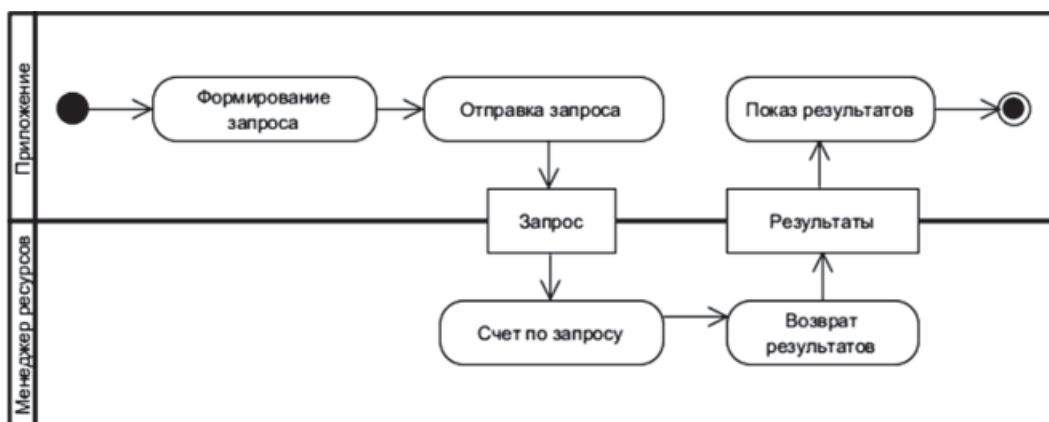


Рис. 3. Запуск приложений в существующей концепции

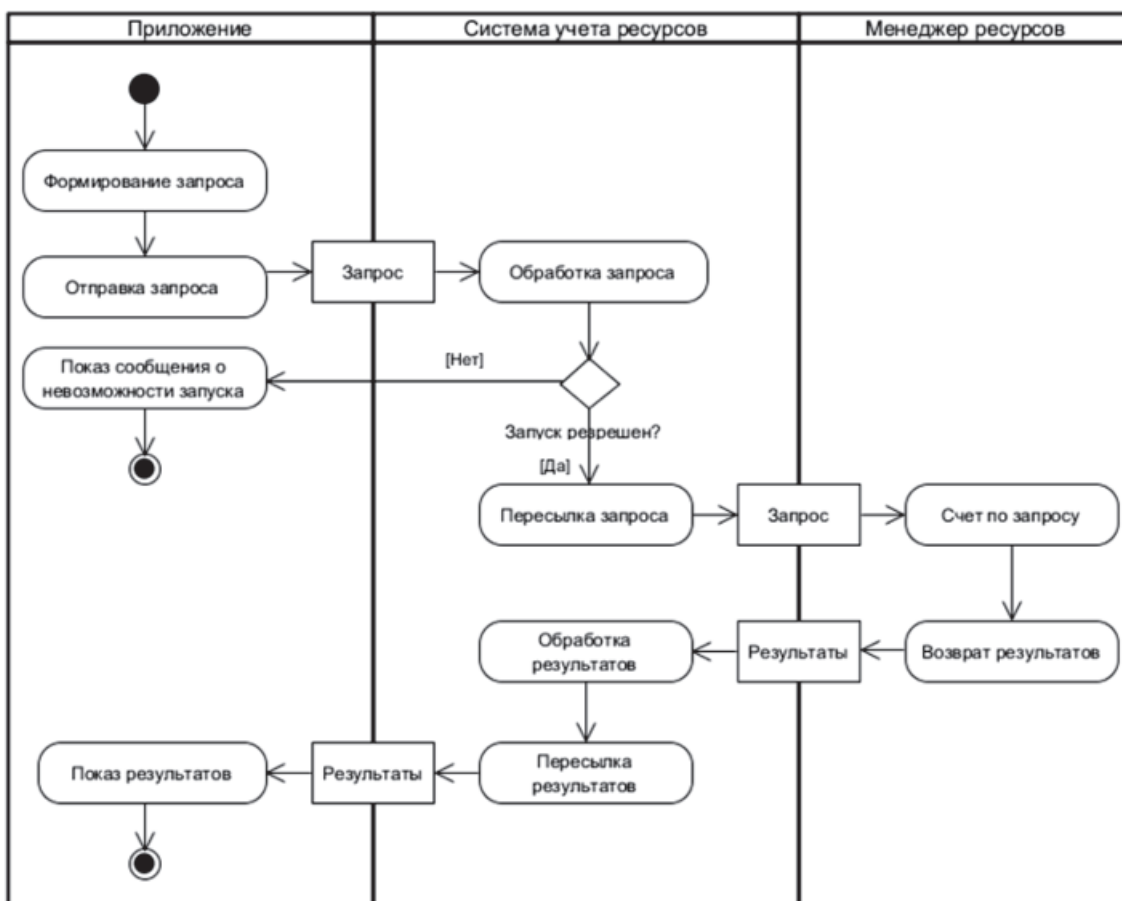


Рис. 4. Запуск приложений в модифицированной концепции

Средства поддержки исследовательских проектов

Вместе с появлением системы учета ресурсов в концепции лабораторий вводится понятие *Исследовательского проекта*.

Исследовательский проект можно представить как совокупность работ, проводимых пользователями лаборатории в рамках одного научного исследования с помощью конкретного набора прикладных

приложений. При этом один пользователь может состоять в нескольких проектах.

Исследовательский проект в рамках лаборатории предназначен для координации его участников, предоставления им средств для эффективной работы над исследованием (форум, wiki-раздел, блог проекта) и контроля за программными и аппаратными ресурсами, израсходованными за время выполнения проекта.

При заведении нового проекта в лаборатории пользователь, создавший его, становится Руководителем проекта и ему доступны все средства по управлению проектом, в том числе участниками и приложениями.

При создании проекта руководителю необходимо указать:

1. Название и его краткое описание, где приводятся основные цели и задачи проекта.
2. Уровень доступа к проекту.
3. Принцип добавления новых участников.
4. Уровень видимости проекта.

После заведения проекта в лаборатории руководителю предоставляется доступ к панели администрирования, где определяются дальнейшие характеристики проекта, необходимые для работы системы учета ресурсов.

Изначально руководителю необходимо определить состав участников проекта и их статуса в зависимости от выбранных настроек при создании. Участниками обязаны быть только зарегистрированные пользователи лаборатории.

Следующим этапом становится определение приложений лаборатории, которые необходимы для проведения исследования, и количество процессорного времени¹ для каждого приложения.

¹ Суммарное время работы всех вычислительных ресурсов (рассчитывается как количество часов счета, умноженное на количество вычислительных ядер, на которых этот счет производился).

На основе введенных руководителем данных система учета ресурсов сохраняет профиль проекта в собственной базе данных. Затем система в режиме реального времени начинает процесс обработки запросов на получение доступа к прикладному программному обеспечению.

Запрос на счет содержит информацию, которую участник проекта отправляет системе учета ресурсов, чтобы осуществить запуск выбранного приложения. В запросе указывается:

- Имя (идентификатор) пользователя.
- Проект, в котором состоит пользователь.
- Название и версия приложения, из которого осуществляется запуск.
- Количество вычислительных узлов и время, на которое они требуются.
- Входные данные и прочая метаинформация, зависящая от приложения.

По информации, полученной из запроса на счет, система учета ресурсов последовательно осуществляет проверку:

- Является ли пользователь участником проекта,
- Есть ли у проекта доступ к запрашиваемому приложению,
- Есть ли у проекта необходимое количество доступного процессорного времени для осуществления счета для данного приложения.

В случае если на каком-то из этапов проверка не пройдена, система учета ресурсов ограничивает доступ приложению и извещает пользователя о невозможности запуска. В противном же случае перенаправляет задание на счет менеджеру ресурсов.

Всю информацию об осуществленных запусках приложений на вычислительном оборудовании система учета ресурсов хранит в собственной базе данных в Таблице запусков (таблица). В таблицу заносится информация о каждом осуществленном запуске после его окончания.

Формат таблицы запусков

Идентификатор запуска	Название проекта	Имя пользователя	Название приложения	Количество потраченного процессорного времени
-----------------------	------------------	------------------	---------------------	-----------------------------------------------

Из данных, хранящихся в таблице запусков, формируется финальный отчет по проекту. В отчет входит информация обо всех проведенных запусках в рамках проекта каждым из его участников. Таким образом, администрация лаборатории имеет возможность оценить количество израсходованных ре-

сурсов по проекту и выставить счет за его проведение.

Заключение

Подводя итог вышесказанному, следует отметить, что с введением системы учета ресурсов в концепцию виртуальных инфор-

мационно-вычислительных лабораторий существенно расширяется класс устанавливаемых программных продуктов, а администрации лаборатории предоставляются средства контроля над ресурсами.

Помимо этого у пользователей появляется возможность проведения сложных исследовательских проектов, состоящих из большого количества людей. Это осуществляется за счет предоставления им средств контроля над расходами ресурсов внутри проекта и механизма формирования финальных отчетов.

Список литературы

1. McLennan M., Kennell R., HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering // *Computing in Science and Engineering*. – 2010. – № 12(2). – P. 48–52.
2. Klimeck G., McLennan M., Brophy S.P., Adams G.B. III, Lundstrom M.S. NanoHUB.org: Advancing Education and Research in Nanotechnology // *Computing in Science and Engineering*. – №10(5), 2008. – P. 17–23.
3. Гайсарян С.С. и др. Университетский кластер: интеграция образования, науки и индустрии // *Открытые системы*. 2010. – № 5. – С. 46–49.
4. Федосин М.Е. Создание виртуальных информационно-вычислительных лабораторий на основе технологической платформы UniHUB // *Системы управления и информационные технологии*. – 2012. – № 3.1(49). – С. 175–178.
5. Федосин М.Е. Введение системы учета ресурсов в концепцию виртуальных информационно-вычислительных лабораторий // *Перспективы развития информационных*

технологий: IX ежегодная международная научно-практическая конференция. –Новосибирск, 2012. – С. 64–66.

References

1. McLennan M., Kennell R. HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering, *Computing in Science and Engineering*, 12(2), 2010, pp. 48–52.
2. Klimeck G., McLennan M., Brophy S.P., Adams G.B. III, M.S. Lundstrom Nanohub.org: Advancing Education and Research in Nanotechnology, *Computing in Science and Engineering*, no. 10(5). 2008, pp. 17–23.
3. Gaisaryan S.S., etc. Universitetskij klaster: integracija obrazovanija, nauki i industrii, *Otkrytye sistemy*, no. 5, 2010, pp. 46–49.
4. Fedosin M.E. Sozdanie virtual'nyh informacionno-vychislitel'nyh laboratorij na osnove tehnologicheskoy platformy UniHUB, *Sistemy upravlenija i informacionnye tehnologii*, no. 3.1(49), 2012, pp. 175–178.
5. Fedosin M.E. Vvedenie sistemy ucheta resursov v koncepciju virtual'nyh informacionno-vychislitel'nyh laboratorij, IX ezhegodnaja mezhdunarodnaja nauchno-prakticheskaja konferencija «Perspektivy razvitija informacionnyh tehnologij» (IX Annual International Scientific Conference Perspective so development of information technologies), Novosibirsk, 2012, pp. 64–66.

Рецензенты:

Гейфман Е.М., д.т.н., профессор, генеральный директор, ЗАО НПК «Электровыпрямитель», г. Саранск;

Гуляев И.В., д.т.н., профессор, декан факультета электронной техники, НИ МГУ им. Огарева, г. Саранск.

Работа поступила в редакцию 29.12.2012.