

Облачная платформа для создания вычислительных веб-сервисов на базе инструментария MathCloud

Сухорослов О.В. (к.т.н., с.н.с. ИППИ РАН)

oleg.sukhoroslov@gmail.com

Аннотация

В докладе рассматривается облачная платформа mc2, предназначенная для развертывания и интеграции вычислительных веб-сервисов. Предлагаемый подход, основанный на модели PaaS, позволяет радикально упростить процедуру создания вычислительных сервисов и их совместное использование в рамках сервис-ориентированных научных сред. Создаваемая облачная платформа основана на разработанном ранее инструментарии MathCloud. Подробно рассматриваются архитектура и принципы реализации отдельных компонентов платформы.

Расширенные тезисы доклада

Современные научные исследования неразрывно связаны с проведением сложных расчетов и использованием высокопроизводительных вычислительных ресурсов. При этом эффективность исследований напрямую зависит от наличия и доступности вычислительных приложений для решения той или иной задачи. В настоящее время накоплен большой багаж подобных приложений, в том числе библиотек численных методов, прикладных вычислительных пакетов, вычислительных моделей и т. д. Наличие готового ПО позволяет исследователям в большинстве случаев избежать трудоемкой реализации программного кода и сконцентрироваться на решаемой задаче.

Вместе с тем, остается ряд проблем, затрудняющих повторное использование имеющихся приложений. Для того, чтобы воспользоваться приложением, исследователю необходимо самостоятельно произвести установку и настройку приложения на своем компьютере или вычислительном ресурсе. В случае появления новой версии приложения, данные действия необходимо повторить снова. Кроме того, в случае параллельных приложений исследователю также необходимо освоить механизм запуска приложения на доступном ему вычислительном ресурсе. Зачастую возникает ситуация, когда исследователь может грамотно сформулировать задачу, но не обладает необходимой квалификацией для выполнения описанных действий.

Для решения указанных проблем применяются высокоуровневые среды [1-3], предоставляющие исследователям удаленный доступ к вычислительным приложениям через проблемно-ориентированные интерфейсы. Как правило, данные среды реализуются в виде веб-порталов, а работа с приложениями осуществляется через веб-браузер. Недостатком подобных решений является невозможность или сложность подключения новых приложений пользователями. Подобная возможность была бы очень востребована среди квалифицированных исследователей, желающих предоставить коллегам доступ к своим разработкам. Кроме того, существующие решения не предоставляют удобных средств интеграции приложений друг с другом и внешними приложениями. Это затрудняет решение часто встречающихся на практике задач, требующих использования сразу нескольких приложений.

В докладе рассматривается облачная платформа mc2, ориентированная на решение описанных проблем. В основе предлагаемого подхода лежит сервис-ориентированная архитектура, в рамках которой вычислительное приложение представляется в виде доступного по сети сервиса. Платформа mc2 реализует модель облачных вычислений «платформа как сервис» (Platform as a Service), заключающуюся в предоставлении интегрированной платформы для разработки, тестирования, развертывания и поддержки

некоторого класса приложений как услуги на основе концепции облачных вычислений. В данном случае речь идет о предоставлении платформы для создания, публикации и интеграции вычислительных веб-сервисов.

В отличие от существующих решений на базе веб-порталов, предлагаемая платформа позволяет обеспечить возможность интеграции приложений друг с другом и внешними приложениями на уровне программных интерфейсов соответствующих сервисов. При этом также предусматривается возможность создания и настройки пользовательского веб-интерфейса сервиса. Также в отличие от существующих решений, пользователи платформы могут самостоятельно создавать на основе имеющихся приложений новые вычислительные сервисы и предоставлять доступ к ним другим пользователям. Это позволяет раскрыть имеющийся потенциал сервис-ориентированного подхода для научной коллаборации и публикации результатов исследований [4]. Наконец, в отличие от существующих средств создания веб-сервисов [5-7], платформа mc2 доступна через веб-браузер, без необходимости установки дополнительного ПО на компьютер пользователя.

Важной особенностью платформы mc2 является то, что, вместо предоставления собственной вычислительной инфраструктуры для выполнения приложений, платформа предлагает разработчикам сервисов готовые средства для интеграции с внешними вычислительными ресурсами. Это позволяет избежать затрат, связанных с созданием соответствующей инфраструктуры, и максимально задействовать потенциал имеющихся в распоряжении пользователей вычислительных ресурсов.

В рамках платформы mc2 предусмотрен следующий набор компонентов:

- Средства описания и развертывания вычислительных веб-сервисов, поддерживающие преобразование существующих приложений в сервисы с унифицированным интерфейсом;
- Среда выполнения вычислительных веб-сервисов, обеспечивающая трансляцию запросов к сервисам в запуск вычислительных заданий на внешних ресурсах, контроль выполнения заданий и обработку полученных результатов;
- Набор адаптеров для различных типов вычислительных ресурсов и инфраструктур, обеспечивающих интеграцию платформы с внешними вычислительными ресурсами и запуск заданий на данных ресурсах от имени пользователей;
- Механизм безопасности, поддерживающий аутентификацию пользователей, контроль прав доступа к сервисам, ресурсам и данным пользователей, а также защиту передаваемых по сети данных;
- Средства интеграции вычислительных веб-сервисов, поддерживающие наглядное описание интеграционных сценариев на основе workflow и других подходов с последующим преобразованием созданных сценариев в новые сервисы;
- Интерфейсы прикладного программирования (API) и командной строки, предназначенные для интеграции платформы и сервисов с внешними приложениями.

В настоящий момент ведется поэтапная разработка данных компонентов в рамках рабочего прототипа платформы, доступного в режиме онлайн для разработчиков и первых пользователей платформы. Подобный подход позволяет ускорить тестирование и внедрение новой функциональности за счет обратной связи от пользователей.

В качестве технологической основы платформы mc2 используется разработанный ранее инструментарий MathCloud [8]. Данный инструментарий предназначен для построения сервис-ориентированных научных сред и включает средства развёртывания, публикации и композиции вычислительных сервисов. Инструментарий MathCloud базируется на общепризнанных подходах, стандартах и технологиях, таких как архитектурный стиль REST, протокол HTTP, формат JSON и платформа Java. В рамках инструментария разработан унифицированный интерфейс доступа к проблемно-ориентированным вычислительным

сервисам. Создана среда выполнения сервисов, реализующая описанный унифицированный интерфейс и поддерживающая быстрое преобразование существующих приложений в сервисы. Разработана система управления сценариями на основе workflow-подхода для поддержки композиции сервисов при решении прикладных задач.

В рамках платформы mc2 предполагается сделать наработки инструментария MathCloud доступными широкому кругу пользователей, не требуя при этом развертывания и сопровождения компонентов инструментария на машинах пользователей. Для этого проводится модификация данных компонентов для функционирования в режиме облачных вычислений, например, реализация многопользовательского режима, разграничения прав доступа, динамического развертывания сервисов и т. д.

Кроме переноса инструментария MathCloud в облачную среду, планируется дальнейшее развитие имеющихся средств, в первую очередь в плане гибкого и эффективного использования имеющихся в распоряжении пользователей вычислительных ресурсов. Если в рамках инструментария MathCloud сервис статически связан с одним вычислительным ресурсом, то в рамках платформы mc2 планируется обеспечить поддержку динамического выбора ресурса для выполнения запроса к сервису, а также использования набора ресурсов (комбинированной вычислительной инфраструктуры). Другим важным направлением развития является расширение спектра поддерживаемых вычислительных ресурсов, в том числе за счет ресурсов облачных сервисов и систем добровольных вычислений.

Список литературы

1. Kacsuk, P., Sipos, G.: Multi-Grid, Multi-User Workflows in the P-GRADE Portal. *Journal of Grid Computing*, Vol. 3, No. 3-4, Springer, pp. 221-238 (2005)
2. McLennan, M., Kennell, R.: HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering. *Computing in Science and Engineering*, 12(2), pp. 48-52, March/April (2010)
3. Afgan, E., Goecks, J., Baker, D., Coraor, N., Nekrutenko, A., Taylor, J.: Galaxy - a Gateway to Tools in e-Science. In: K. Yang, Ed. (ed) *Guide to e-Science: Next Generation Scientific Research and Discovery*, pp. 145-177. Springer (2011)
4. Foster, I.: Service-Oriented Science. *Science*, vol. 308, no. 5723, pp. 814–817 (2005)
5. Delaitre, T., Kiss, T., Goyeneche, A., Terstyanszky, G., Winter, S., Kacsuk, P.: GEMLCA: Running Legacy Code Applications as Grid Services. *Journal of Grid Computing* Vol. 3. No. 1-2, pp. 75-90 (2005)
6. Krishnan, S., Stearn, B., Bhatia, K., Baldridge, K. K., Li, W., Arzberger, P.: Opal: Simple Web Services Wrappers for Scientific Applications. In: *IEEE Intl. Conf. on Web Services (ICWS)* (2006)
7. Krishnan, S., Clementi, L., Ren, J., Papadopoulos, P., Li, W.: Design and Evaluation of Opal2: A Toolkit for Scientific Software as a Service. In: *2009 IEEE Congress on Services (SERVICES-1 2009)*, pp.709-716 (2009)
8. О.В. Сухорослов. Реализация проблемно-ориентированных вычислительных сервисов в среде MathCloud // Параллельные вычислительные технологии (ПаВТ'2011): труды международной научной конференции (Москва, 28 марта – 1 апреля 2011 г.) [Электронный ресурс] – Челябинск: Издательский центр ЮУрГУ, 2011. – 730 с. (с. 288–299) – ISBN 978-5-696-04090-5