

Миграция баз данных MySQL в облачную среду Google Cloud SQL и возможности её применения в сфере образования

Зигмундс Аустрович Булиньш
аспирант, кафедра программного обеспечения прикладных компьютерных систем
Рижский технический университет,
ул. Межа, 1/3, г. Рига, LV-1048, Латвия,
Zigmund.Bulinsh@gmail.com

Аннотация

Облачные вычисления – это хороший способ повысить производительность предлагаемого сервиса без инвестиций в новую инфраструктуру, обучение персонала или приобретения программного обеспечения. Эта технология расширяет потенциальные возможности существующих информационных систем. В последние годы облачные вычисления выросли из хорошего бизнес концепта до одной из самых запрашиваемых индустрий информационных технологий. В статье присутствует краткий обзор функциональности новой платформы облачной базы данных Google Cloud SQL на базе СУБД MySQL версии 5.5. Рассматриваются технические нюансы, потенциальные проблемы и риски связанные с миграцией существующих баз данных MySQL на новую платформу, а так же отмечается возможность применения данного сервиса в образовательных целях в вузах и других схожих учебных заведениях.

Cloud computing is a good way to raise productivity of offered service without investments into new infrastructure, training of the personnel or software acquisition. This technology expands potential possibilities of existing information systems. In recent years cloud computing grew from good business concept to one of the most demanded industry in information technologies. At article there is a review of functionality of a new platform of the cloudy Google Cloud SQL database on the basis of version 5.5 MySQL DBMS. Technical nuances, potential problems and risks connected with migration of the existing MySQL databases on a new platform are reviewed. Possibility of application of this service in the educational purposes in higher education institutions and other similar educational institutions is noted.

Ключевые слова

MySQL, облачная база данных Google Cloud SQL, миграция баз данных, обучение СУБД;
MySQL, Google Cloud SQL, database migration, DBMS education.

Введение

Облачные вычисления – это динамический метод увеличения производительности сервиса или возможностей, без вложений в новую инфраструктуру, обучение персонала или лицензирование программного обеспечения [1]. Они расширяют возможности существующих информационных систем. В последние годы облачные вычисления выросли из хорошего бизнес концепта до одной из самых быстро развивающихся индустрий информационных технологий.

В конце 2011 года компания Google анонсировала новый сервис – облако Google Cloud SQL. Веб-служба позволяет создавать, формировать и использовать

реляционные базы данных, которые физически размещаются в облачной среде Google. Сервис обеспечивает полное обслуживание баз данных пользователей сервиса, освобождая их от дополнительных затрат времени и ресурсов. На сегодняшний день (апрель 2012) данный сервис находится в ограниченном (необходимо специально запрашивать доступ к данному сервису, заполняя соответствующую электронную форму), предварительном испытательном режиме.

Новый сервис вызывает интерес по двум направлениям, описываемым в данной статье – вопрос миграции существующих баз данных на облачную, и использование возможностей Google Cloud SQL для создания виртуальных баз данных, используемых для обучения студентов.

Облако Google Cloud SQL

Облако Google Cloud SQL – веб-служба, которая позволяет создавать, формировать и использовать реляционные базы данных со своими приложениями на основе App Engine приложений[2]. Это полностью управляемое обслуживание, которое поддерживает и управляет базами данных, позволяя разработчикам сосредоточиться на реализации приложений и необходимых услуг.

Предлагая функциональность базы данных MySQL, сервис позволяет легко переместить свои данные, приложения и услуги в облако и из него. Это позволяет повысить мобильность данных и обеспечивает более быстрый выход на рынок предоставления услуг, потому что есть возможность быстро изменить масштабируемость существующей базы данных.

Чтобы гарантировать исправную работу критических приложений и сервисов, облако Google Cloud SQL реплицирует данные в разные географические области для обеспечения высокой доступности данных [3].

Основные особенности данного сервиса [3]:

1. Легкость в использовании – богатый графический интерфейс пользователя учитывает создание, формирование, управление и контроль баз данных;
2. Полностью управляемый сервис – никаких забот о репликации, настройке и других похожих задачах базы данных. Все эти задачи выполняются сервисом;
3. Высоко доступный – чтобы соответствовать критическим потребностям доступности современных приложений и сервисов, такие функции как репликация в многократные географические области уже встроены, таким образом, сервис остаётся доступен, даже если центр данных становится недоступным.

Ограничения Google Cloud SQL и вопросы миграции

Облако Google Cloud SQL представляет собой СУБД MySQL, размещённую в облаке. Сервис обеспечивает все такие же функции, которые предлагает данная СУБД, но с несколькими ограничениями. Главные ограничения облачного сервиса следующие [3]:

- Размер отдельного экземпляра базы данных ограничен 10 гигабайтами;
- Пользовательские функции не поддерживаются (UDF);
- Не поддерживается репликация;
- Не поддерживаются вставки (плагины);
- Не поддерживаются функции, работающие с локальными файлами (LOAD DATA INFILE, LOAD_FILE() и другие).

Так как размер экземпляра ограничен, значит, данный сервис применим в основном только для приложений малого и среднего бизнеса (как было заявлено в описании сервиса Google Cloud SQL). Так же это ограничение может коснуться мультимедийных и других приложений, которые хранят в базе данных бинарные данные и, как правило, это использует много дискового пространства.

Если существующее приложение использовало пользовательские функции СУБД (UDF, не путать с хранимыми процедурами), то возможно миграция будет усложнена проблемой нахождения замены используемым библиотекам пользовательских функций. К примеру, в данный момент СУБД MySQL 5.5 не поддерживает расчёт медианы вектора, а также другие статистические функции [4], поэтому часто данные функции реализуются именно как пользовательские (так как они многократно производительней, чем хранимые процедуры – исполняемый код выполняется в разы быстрее, чем интерпретируемый).

Отсутствие поддержки репликации можно считать второстепенным недостатком, так как Google Cloud SQL сам реализует этот механизм и следит за корректной его работой.

Больше проблем при миграции вызовет, конечно, отсутствие поддержки вставок (плагинов), если функциональность существующей системы основывалась на этой возможности СУБД (к примеру, могла быть реализована аутентификация пользователей через домен или какие-либо другие более сложные плагины).

Ещё одно ограничение, которое не было указано – это отсутствие поддержки удалённых таблиц (federated table storage engine). Если приложение использовало данный механизм, то придётся видоизменять архитектуру приложения при переносе его в облако Google Cloud SQL. Вероятнее всего, конечно, что новая платформа как раз решает проблему, из-за которой были использованы удалённые таблицы, если нет, то будет необходимо найти другое решение проблемы в новой среде.

Ещё одно техническое ограничение облачной СУБД заключается в том, что платформа предназначена для приложений с низким и средним уровнем интенсивности записи данных, то есть приложения с интенсивной записью данных не будут работать эффективно в среде Google Cloud SQL, так как используется механизм репликации [3].

В контроле доступа к экземплярам СУБД появляется новый слой – слой проекта Google API. То есть для получения доступа к базе данных недостаточно знания кодов доступа к СУБД, необходимо так же иметь доступ к проекту, в пределах которого создана СУБД. Проект и доступ к нему контролирует получатель услуг посредством Google API Console в учётной записи Google.

Технологические риски миграции в облако

С ростом количества пользователей и компаний, которые хранят свои данные в «облаках», всё чаще и чаще возникают вопросы о безопасности [5]. Несмотря на всю активность вокруг облачных вычислений, клиенты бизнес сектора всё ещё не желают размещать свои системы в облачных средах. Безопасность – главная причина, которая задерживает стремительное развитие рынка облачных вычислений [6]. Так же вопросы приватности данных и проблемы защиты продолжают влиять на рынок облачных вычислений [5, 7]. По результатам последнего опроса IDC, 74% [8] технических директоров и менеджеров информационных технологий назвали безопасность как главный вызов, который останавливает их от адаптации существующих систем в модель облачного сервиса [9].

При миграции базы данных в среду облака Google Cloud SQL появляются новые вызовы безопасности и доверия к поставщику услуг. Вызовы безопасности в основном связаны с новыми элементами инфраструктуры и платформы [9]:

1. Функциональные атаки на элементы облака – этот тип атаки связан с многослойностью «облаков», главный принцип безопасности – общий уровень безопасности системы равен безопасности самого слабого звена;
2. Атаки на клиента – так как доступ и контроль сервиса Google Cloud SQL осуществляется посредством web-браузера, то этот вид атаки становится особенно актуальным. Вид атаки включает в себя такие методы, как Cross Site Scripting (XSS), DoS атаки, перехват веб-сессий, кража паролей, использование уязвимостей переполнения буфера, «человек по середине» (man in the middle) и другие [10];
3. Угрозы виртуализации – так как платформой для облачных элементов, как правило, служит виртуальная среда, то атака на систему виртуализации угрожает всему облаку в целом;
4. Комплексная угроза «облака» – этот тип угрозы связан с обработкой облака как единого целого и поиском мошенничества и других нарушений в «облачной» структуре [11, 12];
5. Миграция виртуальных машин – уязвимости, связанные с переносом виртуальных машин из одного места в другое [13].
6. Атаки на гипервизорⁱ – вмешательство в работу гипервизора или её нарушение может позволить одной виртуальной машине получить доступ к ресурсам другой – сетевому трафику, хранимым данным. Это может привести даже к выгрузке виртуальной машины из сервера [13,14];
7. Атаки на системы управления – вмешательство в систему управления может привести к невидимым виртуальным машинам, блокировке некоторых машин и подмене слоёв или элементов в облаке на преступные [15].

В данном контексте данные угрозы распространяются в большей степени на поставщика сервиса, нежели на пользователя Google Cloud SQL, но это не освобождает пользователя от потенциальных рисков.

Вопрос доверия состоит в том, что данные размещаются в среде, принадлежащей и контролируемой третьими лицами, за действиями которых нет возможности следить, что не применимо для государственных учреждений или систем с конфиденциальными данными (для использования в среде Google Cloud SQL необходимы изменения вида хранения данных в СУБД).

Доступные средства для работы с сервисом

В отличие от стационарной СУБД MySQL, в Google Cloud SQL нет возможности подключиться к базе данных на прямую из любого компьютера. Подключение осуществляется либо посредством web-браузера (см. **Рис. 1**), либо используя специальную программу-инструмент Command Line Tool (см. **Рис. 2**).

ⁱ В вычислительной технике, гипервизор, по-другому называется системой виртуальных машин (VMM, virtual machine manager).

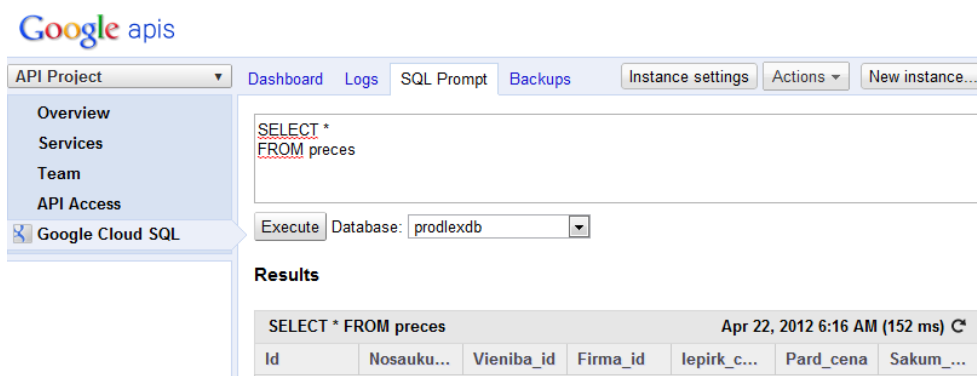


Рис. 1. Доступ через веб-браузер.

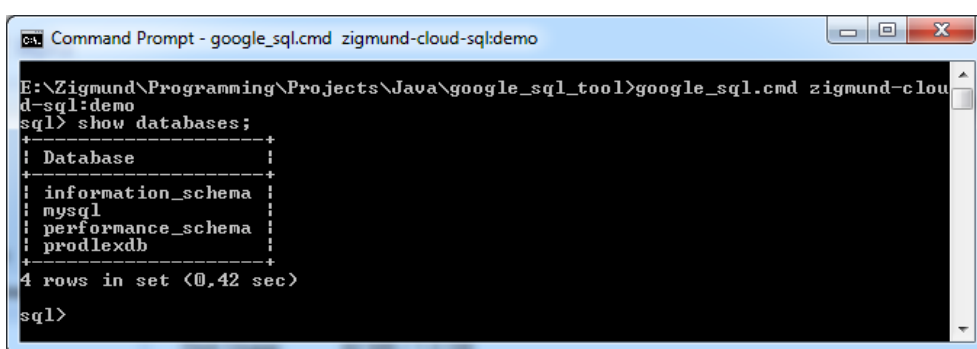


Рис. 2. Доступ через программу-инструмент.

Дополнительно ещё возможно использовать приложение Squirrel SQL [16], которое фактически использует вышеупомянутую программу-инструмент для запуска запросов на Google Cloud SQL.

Для того чтобы подключиться к базе данных Google Cloud SQL с определённого компьютера, необходимо генерировать ключ доступа в настройках учётной записи сервиса и позже подать его программе-инструменту при первом использовании – этот механизм позволяет получателю сервиса контролировать доступ к экземплярам баз данных.

Применения в сфере образования

Учитывая ранее описанные возможности и функции новой платформы Google Cloud SQL, появляется отличная возможность применения данной среды для обучения предмету баз данных. Легкость настройки, создания и управления облачными экземплярами СУБД облегчают работу преподавателя и сводят время, необходимое для подготовки урока, к минимуму. Создание нового экземпляра виртуальной СУБД занимает менее 1 минуты, а так же резервных копий и их последующее восстановление не требует долгого времени. Возможности привязки учётных записей к проекту Google Cloud SQL так же могут способствовать более удобному и быстрому удалённому обучению технологий СУБД – обучаемые могут выполнять задания на созданном преподавателем экземпляре, а преподаватель, в свою очередь, незамедлительно видит действия обучаемых.

Для проведения курса обучения предмету СУБД существует возможность создать изолированные схемы (БД) для каждого студента, где тот может выполнять практические задания преподавателя. При этом студентам нет необходимости иметь доступ к компьютеру, на котором установлена СУБД MySQL – нужно только интернет соединение. Так же местонахождение студента может быть удалённым, и задания могут быть выполнены в удобных для обучаемого времени и месте. В свою очередь, преподаватель имеет информацию о состоянии работ всех обучаемых и может в любой момент осуществить проверку, и при необходимости помочь обучаемым, изучив информацию в конкретной схеме студента. Создавая материалы для учебного курса, нужно придерживаться жизненного цикла и этапов разработки электронных учебных курсов, которые описаны детальней в статье Л.В. Зайцевой [17].

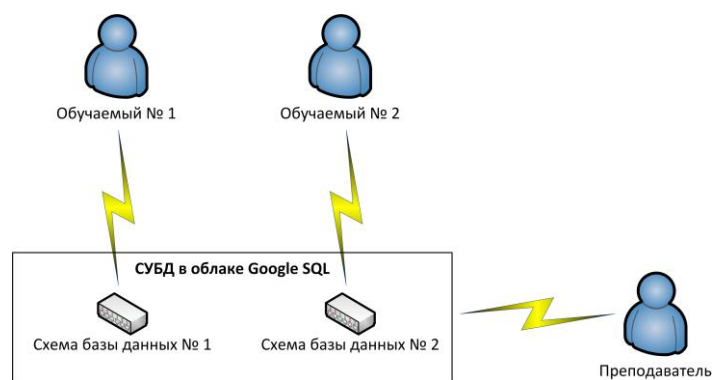


Рис. 3. Схема обучения

Заключение и рекомендации

С выходом нового сервиса Google Cloud SQL, предоставляющего облачную версию СУБД MySQL 5.5, появляется много возможностей для миграции существующих систем на новую архитектуру со сравнительно небольшими потерями при переходе. Проводя анализ возможности перехода на облачную модель базы данных, нужно учесть некоторые факторы риска, описанные в этой статье, а так же убедиться, что существующая система соответствует ограничениям сервиса и не превысит их в ближайшее время.

Переход на облачную базу данных даёт высокую масштабируемость системы в контексте доступности, так как поставщик услуги обеспечивает распределение данных между многими географическими регионами, что ускоряет доступ к данным конечным пользователям.

Другое полезное применение данного сервиса может быть развито в области обучения студентов курсу СУБД. Такой подход ликвидирует необходимость в конфигурации многих компьютеров для работы в классах, а так же даёт возможность удалённого обучения, как это рассказывалось ранее.

В дальнейшем автор планирует развить тему облачных баз данных на сравнение различных поставщиков услуг облачных баз данных и функций, которые они предоставляют.

Литература

1. Identification of a company's suitability for the adoption of cloud computing and modeling its corresponding Return on Investment. Subhas Chandra Misra, Arka Mondal. *Mathematical and Computer Modeling*, ScienceDirect.
2. Google Inc., Google App Engine — Google Developers. <https://developers.google.com/appengine/> (дата обращения: 22.04.2012).
3. Google Inc., Google Cloud SQL — Google Developers. <https://developers.google.com/cloud-sql/> (дата обращения: 22.04.2012).
4. Oracle Corporation, MySQL 5.5 Reference Manual 12.16.1 GROUP BY (Aggregate) Functions. <http://dev.mysql.com/doc/refman/5.5/en/group-by-functions.html> (дата обращения: 22.04.2012).
5. *Journal of Network and Computer Applications*. A survey on security issues in service delivery models of cloud computing. S. Subashini; V. Kavitha. December 2010.
6. Cloud computing — The business perspective. Sean Marston, Zhi Li, Subhajyoti Bandyopadhyay, Juheng Zhang, Anand Ghalsasi. *Decision Support Systems*, ScienceDirect.
7. Danger in the clouds. Steve Mansfield-Devine. URL: <http://www.webvivant.com/dangers-in-the-cloud.html> (дата обращения: 22.04.2012).
8. Clavister. Security in the cloud, Clavister White Paper. URL: http://www.it-wire.nu/members/cla69/attachments/CLA_WP_SECURITY_IN_THE_CLOUD.pdf (дата обращения: 22.04.2012).
9. З.А.Булиньш, Ю.О.Лавенделс, В.С.Шитиков, «Безопасность в облачных вычислениях», В: Труды Международной научно-методической конференции «Информатизация Инженерного Образования», Москва, Россия, 10-11 апреля 2012, Издательский дом МЭИ, 2012, ISBN 978-5-383-00747-1, с. 149-152.
10. Cloud security defense to protect cloud computing against HTTP-DoS and XML-DoS attacks. Ashley Chonka, Yang Xiang, Wanlei Zhou, Alessio Bonti.
11. Cloud Security – безопасность облачных вычислений, угрозы и технологии защиты - Аналитика – Статьи - Anti-Malware.ru. URL: <http://www.anti-malware.ru/node/2333> (дата обращения: 22.04.2012).
12. Security Risks in the Cloud – Reality, or A Broken Record? Ted Kritsonis. *Infosecurity*, 8 (2011) 20-23. doi:10.1016/S1754-4548(11)70005-5.
13. Secure virtualization for cloud computing. Flavio Lombardi, Roberto Di Pietro. *Journal of Network and Computer Applications*, ScienceDirect.
14. Cloud computing – A new business paradigm for biomedical information sharing. Arnon Rosenthal, Peter Mork, Maya Hao Li, Jean Stanford, David Koester, Patti Reynolds. *Journal of Biomedical Informatics*.
15. Identifying the security risks associated with governmental use of cloud computing. Scott Paquette, Paul T. Jaeger, Susan C. Wilson. *Government Information Quarterly*, ScienceDirect.
16. Universal SQL Client. URL: <http://squirrel-sql.sourceforge.net/> (дата обращения: 22.04.2012).
17. Зайцева Л.В. Технология разработки адаптивных электронных учебных курсов для компьютерных систем обучения // Международный электронный журнал "Образовательные технологии и общество (Educational Technology & Society)" – 2008. - V.11. - №1. - С.400-413. - ISSN 1436-4522. URL: <http://ifets.ieee.org/russian/periodical/journal.html>