

УДК 004.7

## Применение систем управления базами данных как сервиса в сложных информационных системах

Д.С. Синюков, А.Д. Данилов\*

Воронежский государственный технический университет, ул. 20 лет Октября,  
84, Воронеж, 394006, Россия

\*E-mail: csit@bk.ru

**Аннотация.** Для снижения эксплуатационных расходов во многих случаях базы данных мигрируют с выделенных частных машин в центры обработки данных, расположенные у провайдеров облачных услуг. Подобное использование облачного сервиса, когда пользователи арендуют аппаратную инфраструктуру и могут запускать на ней любые приложения и службы, известно, как «инфраструктура как услуга».

**Ключевые слова:** базы данных, миграция, облачный сервис, провайдер

## Application of database management systems as a service in complex information systems

D.S. Sinyukov, A.D. Danilov\*

Voronezh State Technical University, 84, 20 let Oktyabrya str., Voronezh, 394006,  
Russia

\*E-mail: csit@bk.ru

**Abstract.** To reduce operating costs, in many cases databases migrate from dedicated private machines to data centers located with cloud service providers. Such use of a cloud service, when users rent a hardware infrastructure and can run any applications and services on it, is known as "infrastructure as a service".

**Keywords:** databases, migration, cloud service, provider

## 1. Введение

Облачные службы, использующие телеметрические данные о пользовании базами данных, сегодня стали повсеместны и широко доступны [1], [2]. Следуя этому современному подходу, целью настоящей работы является изучение целесообразности использования моделей и экспертных систем, базирующихся на полученных данных, для решения вышеуказанных проблем в облачных средах. Основной целью является разработка системы, способной практически в реальном времени автоматически решать проблемы, возникающие на уровне детализации запросов.

Для снижения эксплуатационных расходов во многих случаях базы данных мигрируют с выделенных частных машин в центры обработки данных, расположенные у провайдеров облачных услуг. Подобное использование облачного сервиса, когда пользователи арендуют аппаратную инфраструктуру и могут запускать на ней любые приложения и службы, известно, как «инфраструктура как услуга» (IaaS).

Наибольших успехов в развитии данного направления достигли такие провайдеры, как Microsoft [3] и Amazon [4]. Они ввели такое понятие как «база данных как услуга» (DBaaS), которая представляет собой службу, размещающую серверы баз данных в облачной инфраструктуре. При данном подходе пользователю не нужно беспокоиться об обслуживании аппаратной части серверов, а также об установке и обновлении программного обеспечения.

## 2. Цель исследования

Целью исследования является обоснование применения систем управления базами данных как сервиса в сложных информационных системах.

### 2.1. Azure SQL

Облачная база данных Microsoft Azure SQL является одним из первых решений DBaaS [3]. Она спроектирована как высокодоступная и надежная система. По сравнению с готовым аппаратным решением, сервер и база данных, размещенные в среде Azure SQL позволяют клиентам использовать самую последнюю версию программного обеспечения и все новые возможности сервера SQL. Помимо этого, в Azure SQL реализованы дополнительные службы, обеспечивающие автоматическое резервное копирование, возможность восстановления резервных баз данных за последние 30 дней, географическую репликацию для сценариев аварийного восстановления, а также

богатую телеметрию, которая отслеживает все аспекты, относящиеся к базе данных и серверу.

За счет огромных масштабов Azure SQL, собранная телеметрия носит более объективный характер в отличие от локального сервера. В зависимости от важности некоторых телеметрических данных записи делаются с разной периодичностью, так наиболее важные данные собираются с точностью до секунды, в то время как редко меняющаяся информация, как и метаданные базы данных, собираются не чаще 2-3 раз в сутки. Все собранные данные хранятся в соответствии со стандартами общего регламента по защите данных и другими государственными регулятивными документами в структурированном виде, как в "холодном", так и в "горячем" хранилище в центрах обработки данных. Горячее хранение используется для сценариев оповещения, когда крайне важно получить доступ к данным с минимальной задержкой. Запрос на горячее хранение выполняется в течение ограниченного периода времени, что позволяет получать результат практически в реальном времени.

## **2.2. Наука о данных и дата-майнинг**

В настоящее время для изучения огромных объемов данных широко используются методы добычи данных (Data-mining) и науки о данных (Data science). С появлением интернета, облачных сервисов и телеметрии, наука о данных и интеллектуальный анализ данных приобретают все большее значение для понимания поведения пользователей. Основная идея этих дисциплин заключается в том, чтобы использовать наблюдаемые модели поведения и соотносить их с определенными пред- или пост-мероприятиями, которые проводит пользователь конкретного сервиса. При использовании масштабируемых решений возрастает потребность в поиске и устранении неисправностей, а также сортировке сложных систем с минимальными задержками. В таких системах соответствующий анализ данных и эффективная реализация научных моделей играют очень важную роль [1].

## **3. Методы и материалы исследования**

Развитие технологий и применение отраслевых инноваций на практике изменили рабочий процесс, связанный с устранением неисправностей в системах баз данных. Использование удаленного аппаратного оборудования существенно облегчает создание и развертывание нового программного обеспечения, использующего базу данных. Традиционно независимые поставщики программного обеспечения развертывают

отдельные копии приложений в оперативной среде. В настоящее время большинство таких вендоров переходят на другую модель. Они размещают комбинированный программный сервис, который работает для их клиентов коллективно, в публичной облачной среде, используя вместо этого тип "платформа как услуга". Таким же образом, ответственность за обслуживание теперь передается от конечного клиента поставщику услуг компьютерного облака. Однако, в то же время, вводится новый уровень абстракции, что серверы баз данных больше не доступны для разработчиков и администраторов баз данных. Это существенно уменьшает некоторые возможности поддержки производительности на должном уровне.

Другим очень важным аспектом является масштаб «платформы/базы данных как службы». Традиционно поиск и устранение неисправностей в базе данных выполнялся для каждого клиента и для каждого конкретного случая. В настоящее время такая высокотребовательная задача в облачном масштабе требует, как значительных вычислительных, так и инженерных мощностей.

#### 4. Полученные результаты

Системы управления реляционными базами данных превратились в крупные, часто очень сложные программные продукты. Эти системы позволяют пользователям проектировать произвольно сложную логику, что иногда приводит к низкой производительности приложений. Ранее такие проблемы решались путем сочетания хорошего выбора дизайна реляционных баз данных во время разработки, а также путем устранения неполадок с производительностью администраторами баз данных в процессе разработки. Однако, более сложные системы часто требуют более глубоких человеческих знаний для их отладки и настройки. Имея в виду, что каждое приложение отличается, поэтому для того, чтобы эффективно управлять такими системами, чтобы они могли работать адекватно, как правило, требуется ручная настройка. Эксплуатационные расходы на крупномасштабную СУБД составляют почти 50% от общих инвестиций, в то время как администраторы СУБД тратят около 25% своего времени на ручную настройку.

Существует несколько коммерческих решений, которые обеспечивают расширенную поддержку мониторинга через набор пользовательских интерфейсов, позволяющих администратору базы данных отслеживать поведение базы данных. Эти решения основаны на расширении возможностей администраторов или консультантов

баз данных для эффективного управления производительностью систем. Тем не менее, многие проблемы с производительностью выходят за рамки навыков среднего администратора баз данных. Даже для высококвалифицированных администраторов может потребоваться несколько часов, чтобы определить проблему, найти правильное решение и применить соответствующие исправления. При применении к системам хостинга общедоступных облачных баз данных с миллионами баз данных и приложений, размещенных в одной инфраструктуре, существующие методы просто не могут справиться с этой задачей.

## 5. Выводы

Для того, чтобы отличить реальные проблемы от проблем, вызванных со стороны пользователя, требуется исследование со стороны инженеров поддержки, которое занимает значительное время и ресурсы в облачной среде. Было бы полезно иметь такую систему, которая могла бы различать эти ситуации и давать соответствующие советы инженерам поддержки, чтобы сократить время и стоимость эксплуатации. Следовательно, раннее обнаружение проблем и надлежащее оповещение позволяют пользователям сосредоточиться на исправлении фактической проблемы.

В области поиска и устранения неисправностей с реляционными базами данных существует несколько коммерческих решений, предоставляемых Idera, Sentry, Dell, Amazon и Oracle.

## Список литературы

1. Cao, L. Data Science: Challenges and Directions / L. Cao // Communications of the ACM. – 2017. – 60(8). – Pp. 59-68.
2. Wang, C. Performance troubleshooting in data centers: an annotated bibliography / C. Wang, K. Soila, J. Tan, L. Hu, M. Kutare, M. Kasick, K. Schwan, P. Narasimhan, R. Gandhi // ACM SIGOPS Operating Systems Review. – 2013. – 47(3). –Pp. 50-62.
3. Bernstein, P. Adapting Microsoft SQL server for cloud computing / P. Bernstein, I. Cseri, N. Dani, N. Ellis, A. Kalhan, G. Kakivaya, D. Lomet, R. Manne, L. Novik, T. Talus // 2011 IEEE 27-th Int. Conf. on Data, Hannover, Germany. – 2011. – Pp. 1255-1263.
4. Sivasubramanian, S. Amazon dynamoDB: a seamlessly scalable non- relational database service / S. Sivasubramanian // 2012 ACM SIGMOD Int. Conf. on Management of Data, Scottsdale, Arizona, USA. – 2012. – Pp. 729-730.