

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ НЕРЕЛЯЦИОННЫХ И РЕЛЯЦИОННЫХ СИСТЕМ УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ

Новогрудская Р.Л., Добров Ю.А.

Институт телекоммуникационных систем КПИ им. Игоря Сикорского, Украина

E-mail: rinan@ukr.net, yuridobrov@gmail.com

A benchmark comparison of NoSQL and SQL DBMS

The paper is devoted to comparing the performance of the most popular database management systems (DBMS) of nonrelational and relational data models. The big trend of Big Data has had a great impact on the development and spread of NoSQL databases. It is assumed that they are significantly ahead of traditional relational databases for performing data access operations. However, is it so? The purpose of this work is to perform benchmarks for measuring the time for create, read, update and delete (CRUD) operations for different volumes of data. As the results presents, just a few instances of NoSQL: MongoDB, Redis and Couchbase, really work faster than the traditional classical RDBMS MSSQL Server. CouchDB and RavenDB lost to the MSSQL Server in all tests. Cassandra and Hypertable were controversial in this comparison - they turned out to be slower than RDBMS with the exception of operations for writing large portions of data.

Потребность в хранении и обработке больших объёмов неструктурированных данных привела к созданию новых методик и технологий в области баз данных, такими являются NoSQL решения – системы управления базами данных (СУБД), которые отошли от традиционной реляционной модели данных, предложенной Эдгаром Коддом в 1969 году.

СУБД – это набор прикладных программных и лингвистических средств общего или специального назначения, обеспечивающих управление созданием и использованием баз данных.

На сегодняшний день можно выделить четыре основные категории NoSQL-хранилищ:

1. Key-value (ключ — значение). Является большой хеш-таблицей, в которой допустимы только операции записи и чтения данных по ключу.

2. Document-oriented (документо-ориентированные). Коллекции структурированных документов. Позволяют осуществлять выборку по различным полям документа, а также модификацию частей документа.

3. Column (колоночные). Таблицы с колонками и строками. Отличаются от SQL баз данных переменным количеством колонок для разных строк на уровне одного документа, каждая строка имеет уникальный ключ. Можно рассматривать данное решение как хеш-таблицу хеш-таблицы, первым ключом является ключ строки, вторым — имя колонки. При задании вторичных индексов возможны выборки по значению в колонке, а не только по ключу строки.

4. Graph (графовые). Данный тип БД специально предназначен для хранения математических графов: узлов и связей между ними.

Сравнительный анализ производительности нереляционных и реляционных СУБД проводился с целью оценки времени, необходимого для операций записи, чтения, изменения и удаления для разных объёмов данных. Для такой оценки было выполнено ряд тестов. Для замеров были взяты представители трёх вышеперечисленных категорий: MongoDB, RavenDb, CouchDB, Couchbase (документо-ориентированная), Cassandra, Hypertable (колоночные), Redis (ключ-значение). В качестве РСУБД был взят Microsoft SQL Server.

Реализация сравнения производительности была выполнена с использованием .NET, языка программирования С# и официальных драйверов для подключения к серверам тестируемых баз данных, взятые из NuGet Gallery. Для снятия показателей затраченного времени был использован .NET класс Stopwatch из пространства имён System.Diagnostics.

Тестовыми данными во всех базах являются автоматически сгенерированные строковые пары типа ключ-значение. Стоит отметить, что все результаты являются средним значением времени для трёх операций для каждой базы данных.

Результаты эксперимента показаны в сводных таблицах.

Табл. 1. Время на запись (мс)

| База данных | Количество операций | | | |
|--------------|---------------------|------|-------|--------|
| | 10 | 100 | 10000 | 100000 |
| MongoDB | 58 | 76 | 2543 | 19876 |
| RavenDb | 566 | 1904 | 81290 | 738650 |
| CouchDB | 67 | 379 | 66478 | 912010 |
| Couchbase | 69 | 71 | 959 | 8382 |
| Cassandra | 120 | 238 | 9805 | 89341 |
| Hypertable | 56 | 198 | 10945 | 115321 |
| Redis | 34 | 65 | 980 | 13015 |
| MSSQL Server | 27 | 111 | 16732 | 219567 |

Табл. 2. Время на чтение (мс)

| База данных | Количество операций | | | |
|--------------|---------------------|-----|-------|--------|
| | 10 | 100 | 10000 | 100000 |
| MongoDB | 7 | 19 | 1020 | 9871 |
| RavenDb | 122 | 511 | 47200 | 423300 |
| CouchDB | 21 | 186 | 18670 | 17500 |
| Couchbase | 13 | 21 | 780 | 7000 |
| Cassandra | 111 | 318 | 19200 | 220190 |
| Hypertable | 56 | 102 | 3103 | 78450 |
| Redis | 5 | 46 | 439 | 4899 |
| MSSQL Server | 13 | 44 | 312 | 16749 |

Табл. 3. Время на обновление (мс)

| База данных | Количество операций | | | |
|--------------|---------------------|------|-------|--------|
| | 10 | 100 | 10000 | 100000 |
| MongoDB | 76 | 83 | 3186 | 25963 |
| RavenDb | 614 | 2104 | 96790 | 836363 |
| CouchDB | 98 | 459 | 69783 | 110569 |
| Couchbase | 78 | 97 | 1460 | 9754 |
| Cassandra | 159 | 258 | 9976 | 94378 |
| Hypertable | 78 | 259 | 13875 | 167456 |
| Redis | 45 | 98 | 3950 | 50478 |
| MSSQL Server | 31 | 118 | 19678 | 317856 |

Табл. 4. Время на удаление (мс)

| База данных | Количество операций | | | |
|--------------|---------------------|-----|-------|--------|
| | 10 | 100 | 10000 | 100000 |
| MongoDB | 2 | 25 | 19655 | 17656 |
| RavenDb | 76 | 766 | 84904 | 776390 |
| CouchDB | 78 | 556 | 64789 | 704345 |
| Couchbase | 5 | 16 | 989 | 7493 |
| Cassandra | 33 | 97 | 13807 | 94778 |
| Hypertable | 17 | 106 | 10478 | 124569 |
| Redis | 6 | 9 | 9 | 56 |
| MSSQL Server | 9 | 45 | 2654 | 27690 |

Как показывают результаты экспериментов, реляционная база данных MSSQL Server по-прежнему является высокопроизводительным решением, которая уступает лишь немногим NoSQL СУБД в плане временных затрат на операции доступа к данным.

Литература

1. Yishan Li, Sathiamoorthy Manoharan, "Communications, Computers and Signal Processing", August 2013.
2. Haleemunnisa Fatima, Kumud Wasnik, "International Journal of Advances in Electronics and Computer Science", March 2016.
3. Abdullah Talha Kabakus, "Journal of King Saud University – Computer and Information Sciences", 2 July 2016.