Автор:
12.09.14 20:02 -

После года разработки доступен релиз распределённой БД Apache Cassandra 2.1 , относящейся к классу noSQL-систем и рассчитанной на создание высокомасштабируемых и надёжных хранилищ огромных массивов данных, хранимых в форме ассоциативного массива (хэша). Код проекта написан на языке Java и распространяется в рамках лицензии Apache 2.0.

Особенности новой версии:

- Проведена значительная работа по увеличению производительности. По сравнению с прошлым выпуском выигрыш может достигать 50%;
- В язык CQL3 (Cassandra Query Language) добавлены средства для определения пользовательских типов (UDT , User Defined Types), реализованы типы кортежей и добавлена возможность использования вторичных индексов для коллекций;
- Реализована концепция инкрементального восстановления, при которой сохраняется состояние ранее восстанавливаемых sstable, а новые операции восстановления проводятся только для новых областей. Подобный подход позволяет существенно ускорить процесс и снизить нагрузку на систему в условиях часто выполняемых операций восстановления. Запустить инкрементальное восстановление можно командой "nodetool repair -par -inc", для ручного управления статусом восстановления подготовлена утилита tools/bin/sstablerepairedset;
- Применения метода инкрементального замещения в упакованных SSTables;
- Увеличена эффективность работы кэша строк;
- Поддержка размещения хранимых в памяти таблиц (memtables) вне области кучи JVM;
- Расширены возможности инструментов для стресс-тестирования;
- Возможность использования экранированных кавычками идентификаторов в именах триггеров;
- Предлагаемое по умолчанию местоположение файлов с данными и логами изменено на $CASSANDRA_HOME/data/data, $CASSANDRA_HOME/data/commitlog, $CASSANDRA_HOME/data/saved_caches и $CASSANDRA_HOME/logs. Изменено наименование директорий с данными SSTable (например, ks/cf-5be396077b811e3a3ab9dc4b9ac088d/);
- Поддержка платформы Windows переведена в разряд готовых к промышленному использованию.

БД Cassandra объединяет в себе полностью распределённую hash-систему Dynamo, обеспечивающую практически линейную масштабируемость при увеличении объема

данных. Cassandra использует модель хранения данных на базе семейства столбцов (ColumnFamily), отличающуюся от систем подобных memcachedb, которые хранят данные только в связке ключ/значение, возможностью организовать хранение хэшей с несколькими уровнями вложенности.

Для упрощения взаимодействия с БД поддерживается язык формирования структурированных запросов CQL (Cassandra Query Language), на первый взгляд напоминающий SQL, но существенно урезанный по функциональности. Например, можно выполнять только простейшие запросы SELECT с выборкой по определённому условию, но без поддержки сортировки и группировки. Добавление и обновление данных производится через единое выражение UPDATE, операция INSERT отсутствует (если записи нет, при выполнении UPDATE она создаётся). Из возможностей можно отметить поддержку пространств имён и семейств столбцов, создание индексов через выражение "CREATE INDEX". Драйверы с поддержкой CQL подготовлены для языков Python, Java (JDBC/DBAPI2) и JavaScript (Node.js).

Cassandra относится к категории хранилищ повышенно устойчивых к сбоям: помещаемые в БД данные автоматически реплицируются на несколько узлов распределённой сети или даже равномерно распределяются по нескольким дата-центрам. При сбое узла, его функции на лету подхватываются другими узлами. Добавление новых узлов в кластер и обновление версии Cassandra производится на лету, без дополнительного ручного вмешательства и переконфигурирования других узлов.

Изначально проект был разработан в недрах компании Facebook и в 2009 году передан под покровительство фонда Apache. Промышленные решения на базе Cassandra, способные обрабатывать тысячи запросов в секунду, развернуты для обеспечения сервисов таких компаний и организаций, как Adobe, CERN, Cisco, IBM, HP, Comcast, Disney, eBay, Netflix, Sony, Rackspace, Reddit и Twitter. Наиболее крупный кластер серверов, обслуживающих единую БД Cassandra насчитывает более 400 машин и используется для хранения более 300 Тб данных.

Read more http://www.opennet.ru/opennews/art.shtml?num=40572