

Tarantool Column Store

Инструкция по установке экземпляра
продукта

Версия 1.0.1 от 23.10.23

1. Введение

1.1. Область применения

Tarantool Column Store (далее по тексту TCS) представляет собой распределенное блочно-колоночное хранилище данных для решений, подразумевающих комплексные аналитические запросы с группировками и расчетом итоговых значений (агрегатов), в сфере финтеха, real-time маркетинга, логистики, телекома и др., реализованное на базе in-memory платформы Tarantool Enterprise.

Настоящее руководство предназначено для администраторов TCS в частной установке.

Область применения настоящего документа распространяется на все подсистемы, модули и разделы TCS в частной установке.

1.2. Краткое описание возможностей

TCS представляет собой распределенное блочно-колоночное хранилище данных осуществляющее хранение данных не в виде строк, а в виде колонок. Такое хранение позволяет оптимизировать хранение данных для выполнения комплексных аналитических запросов с группировками и расчетами итоговых значений. Оптимизация осуществляется за счет снижения нагрузки на устройства хранения и обработки данных. Снижение же в свою очередь достигается за счет того, что при выполнении запроса на чтение фактически с диска происходит чтение только тех полей, которые фигурируют в запросе, в то время как в реляционных базах данных (записи хранятся построчно) запись считывается с диска и обрабатывается целиком.

Дополнительно можно отметить, что поколоночное хранение данных существенным образом влияет на возможность компрессии данных ввиду того, что данные в одной колонке как правило являются однотипными, чего нельзя сказать о данных, хранящихся в строке. Кроме того, компрессия может достигаться за счет применения ограничений, накладываемых предметной областью на хранимые данные.

Запись данных в TCS может осуществляться как построчно, так и блоком данных, состоящим из заданного количества строк (по умолчанию - 1024 записи). При этом происходит

автоматическое разбиение записываемой строки на колонки согласно определенной схеме данных.

Получение данных из TCS может осуществляться при помощи SQL запросов, отправляемых в виде POST запросов в storage инстанс Tarantool, на базе которого построен TCS.

Часто используемые SQL запросы могут быть сохранены в виде PREPARE STATEMENT с использованием параметров, указываемых при последующем выполнении данного SQL запроса. Сохранение и последующий вызов сохраненного SQL запроса осуществляется при помощи отправки соответствующего POST запроса в storage инстанс Tarantool, который осуществляет непосредственное хранение данных.

1.3. Требования к квалификации

Системные администраторы TCS отвечают за комплексную настройку инфраструктуры TCS, предоставляемой в качестве услуги конечным пользователям, устранение неисправностей, сбор диагностической информации и эскалацию неисправностей производителю аппаратной или программной составляющей TCS.

Системные администраторы выполняют в рамках работы с TCS следующие функциональные обязанности:

- Управление хранилищем данных (создание, удаление, настройка конфигурации, настройка доступа к данным).
- Обслуживание системного и прикладного программного обеспечения системы.
- Управление пользователями (создание, удаление, выдача прав).
- Резервное копирование и восстановление данных.
- Производство регламентных работ и анализ их результатов.

К системным администраторам предъявляются следующие требования:

- Навыки системного администрирования Linux.
- Прохождение обучения от компании ООО «ВК ЦИФРОВЫЕ ТЕХНОЛОГИИ» по использованию компонентов, разработанных вендором.
- Навыки работы со службами TCS, в том числе способность самостоятельно осуществлять.

- Мониторинг программными средствами.
- Изменение настроек TCS.
- Сбор диагностических данных и эскалацию сложных проблем по гарантийным обязательствам производителя.

Для обеспечения функционирования системы в круглосуточном режиме (24/7) необходимо наличие как минимум двух системных администраторов.

2. Установка и настройка Tarantool Column Store

2.1. Описание дистрибутива

TCS передается для проведения экспертной проверки в виде комплекта, содержащего следующие файлы:

- `tarantool_column_store-0.9.4.x86_64.tar.gz` – архив с TCS
- `config.json` – конфигурационный файл (содержит колоночную схему данных)

2.2. Требования к окружению

Минимальные системные требования, предъявляемые к рабочей станции, на которой осуществляется экспертная проверка:

- ОС - Linux (Ubuntu, Debian, CentOS, ALTLinux, AstraLinux, RedOS)
- 4 Core 2.6ГГц +
- 4Гб RAM

2.3. Подготовка к проведению экспертной оценки

Подготовка TCS к проведению экспертной оценки состоит из следующих шагов:

- Развернуть архив.

```
tar -zxvf tarantool_column_store-0.9.4.x86_64.tar.gz
```

- Перейти в директорию с разархивированным приложением.

```
cd tarantool_column_store
```

- Запустить приложение.

```
./tt start  
./tt cartridge replicaset setup --bootstrap-vshard
```

- Проверить запущенные инстансы. Все инстансы должны иметь статус RUNNING.

```
./tt status
```

- Загрузить конфигурацию приложения.

```
curl -X 'PUT' 'http://localhost:8081/admin/config' --data-binary @../config.json
```

- Перезапустить все работающие инстансы.

```
./tt restart -y
```

- Проверить успешное применение конфигурации:

- Подключиться к storage инстансу Tarantool.

```
./tt connect tarantool_column_store:s1-1
```

- Проверить наличие спейса `cols` и записей в нем.

```
box.space.cols:len()
```

- Выполнить команду.

```
Ctrl + D
```

2.4. Проверка работоспособности функций TCS

Проверка работоспособности функций TCS в рамках проведения экспертной оценки описывается в документе:

- Tarantool Column Store_v.1.0.1_Эксплуатация экземпляра программного обеспечения, предоставленного для проведения экспертной проверки

2.5. Поддержка

В случае возникновения проблем при установке TCS для проведения экспертной проверки вопросы могут быть направлены техническим специалистам:

- Алексей Андреев (al.andreev@corp.mail.ru)