

Tarantool Clusters Federation

Описание функциональных
характеристик

Оглавление

Термины и сокращения	3
1. Общие положения	5
1.1. Наименование системы	5
1.2. Разработчик системы	5
1.3. Назначение документа	5
2. Функциональные характеристики	6
2.1. Цели и назначение	6
2.2. Ключевые функции	6
2.3. Сценарии использования	6
2.4. Проблематика построения отказоустойчивых геораспределенных систем	6
2.5. Описание архитектуры и концепции	7
2.5.1 Растянутые кластеры	8
2.5.2 Изолированные кластеры	9
2.6. Функции, выполняемые системой	10

● Термины и сокращения

В настоящем документе используются сокращения, перечисленные в таблице ниже.

Сокращение	Расшифровка
БД	База данных
ВМ	Виртуальная машина
ОС	Операционная система
ПИКД	Подсистема идентификации и контроля доступа
ПО	Программное обеспечение
СУБД	Система управления базами данных
ACL	Access Control List, список управления доступом, который определяет, кто или что может получать доступ к объекту, и какие именно операции разрешено или запрещено выполнять субъекту
API	Application Programming Interface - Аппаратно-программный интерфейс. Описание способов (набор классов, процедур, функций, структур или констант), которыми одна компьютерная программа может взаимодействовать с другой программой
JavaScript	Мультипарадигменный язык программирования. Поддерживает объектно-ориентированный, императивный и функциональный стили
JSON	JavaScript Object Notation – текстовый формат обмена данными, основанный на JavaScript
TCF	Tarantool Clusters Federation – система построения катастрофоустойчивых конфигурации кластеров Tarantool Enterprise Edition.
URL	Uniform Resource Locator, унифицированный указатель ресурса – система унифицированных адресов электронных ресурсов, или единообразный определитель местонахождения ресурса. Используется как стандарт записи ссылок на объекты в Интернете

Сокращение	Расшифровка
Lua	Процедурный динамически типизированный модульный язык с автоматическим управлением памятью

● 1. Общие положения

1.1. Наименование системы

Полное наименование системы: Программа для ЭВМ «Tarantool Clusters Federation».

Условное наименование системы: «Tarantool Clusters Federation», Система, TCF.

1.2. Разработчик системы

Полное наименование: Общество с ограниченной ответственностью «ВК Цифровые Технологии».

Сокращенное наименование: ООО «ВК Цифровые Технологии»

1.3. Назначение документа

Настоящий документ входит в комплект эксплуатационной документации по «Tarantool Clusters Federation» и описывает функциональные характеристики Системы.

● 2. Функциональные характеристики

2.1. Цели и назначение

Tarantool Clusters Federation — продукт, который позволяет выстроить катастрофоустойчивую конфигурацию из двух или более кластеров Tarantool Enterprise Edition. TCF обеспечивает скорость работы, приближенную к реальному времени, и отказоустойчивость до 99,99%. Кластеры Tarantool Enterprise Edition Master-Replica находятся в нескольких (двух и более) разных дата-центрах.

2.2. Ключевые функции

Ключевые функциональные возможности Системы:

- Передача данных из одного кластера в другой, с использованием механизма Change Data Capture
- Ручное переключение активного кластера
- Автоматическое переключение активного кластера

Более детальное описание представлено в следующих разделах документа

2.3. Сценарии использования

Сценарии использования – последовательность действий, которые может осуществлять Система при взаимодействии с конечным пользователем или другой системой. Система поддерживает следующие сценарии:

- Ручное переключение активного кластера
- Автоматическое переключение активного кластера

2.4. Проблематика построения отказоустойчивых геораспределенных систем

TCF является механизмом переключения кластеров и синхронизации данных и используется для построения Mission Critical систем хранения данных, для которых необходимо выполнение следующих требований:

- гарантия постоянной доступности данных
- обеспечение отсутствия потерь данных
- минимизация времени отклика для Клиента

При построении геораспределенных систем хранения данных могут возникать нештатные ситуации, которые негативным образом сказываются на соблюдении обозначенных выше требований. К таким нештатным ситуациям могут быть отнесены:

- ошибки при проведении регламентных работ на узлах распределенной системы (изменение схемы хранения данных)
- чрезвычайные ситуации, приводящие к полной недоступности одного из узлов распределенной системы (выход из строя ЦОД или нарушение сетевой связанности)
- ошибки при разработке прикладного ПО / SDK, приводящие к полной недоступности кластера во всех ЦОДах

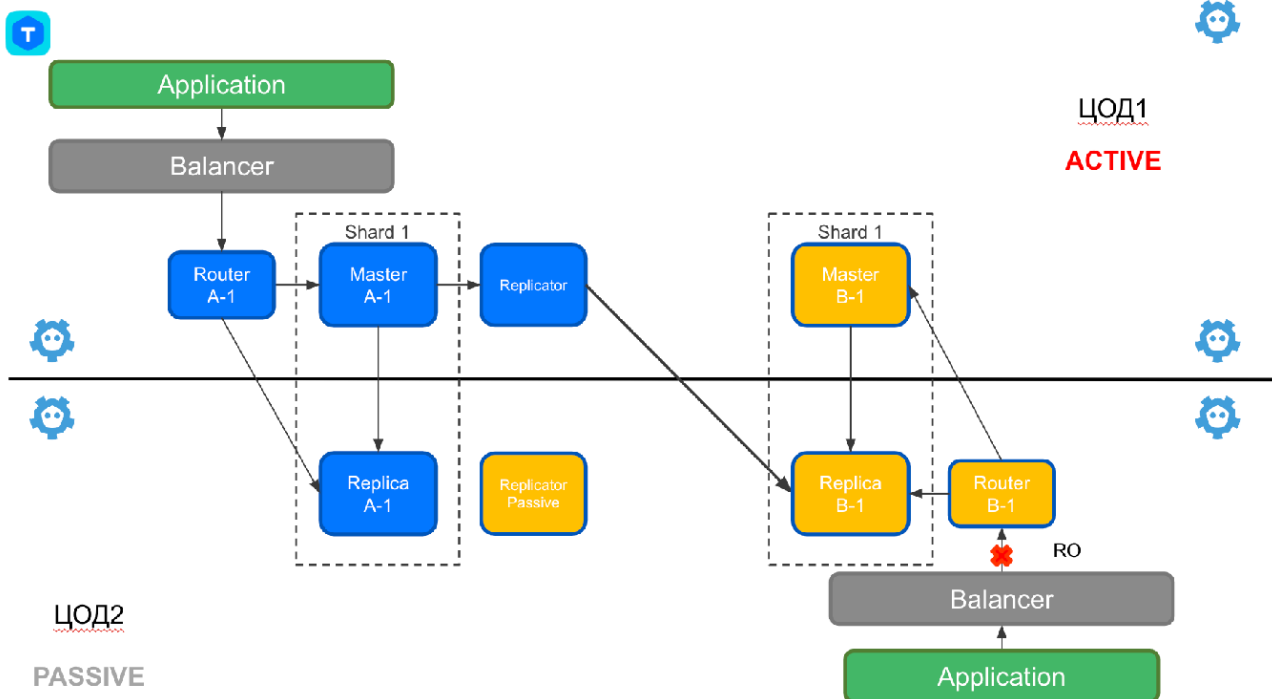
Для минимизации негативных последствий от возникновения нештатных ситуаций, обозначенных выше, требуется механизм синхронизации нескольких независимых кластеров и переключения между ними для обеспечения бесперебойного доступа к данным.

2.5. Описание архитектуры и концепции

Типовая архитектура TCF включает в себя минимум 2 кластера Tarantool Enterprise Edition, работающих под управлением Tarantool Cartridge.

Рассмотрим два типовых варианта – 2 растянутых кластера и 2 изолированных кластера.

2.5.1 Растянутые кластеры



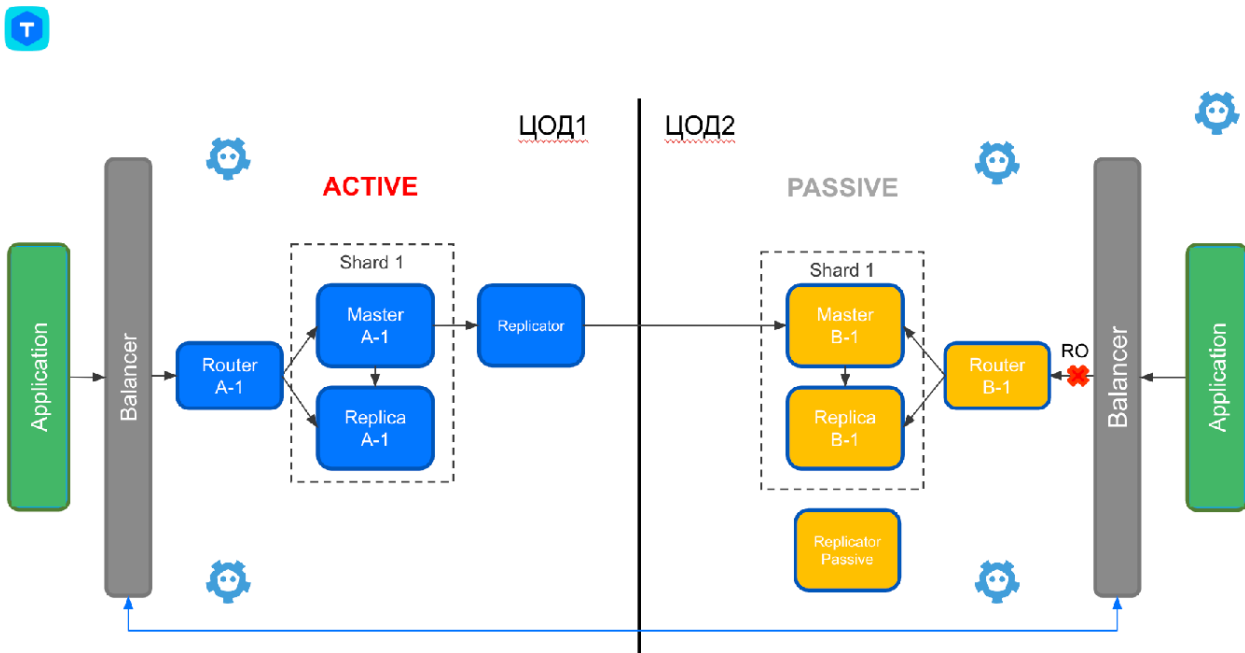
На схеме представлен пример размещения 2 кластеров Tarantool Enterprise Edition вместе с TCF в 2-х ЦОД. Растянутый кластер подразумевает размещение Master и Replica узлов в различных ЦОД.

Решение состоит из:

- Application – приложение клиента, которое работает с данными, хранящимися в кластере
- Balancer – программное или аппаратное решение, используемое для распределения (балансировки) всех запросов приложения по кластеру с целью оптимизации использования вычислительных ресурсов
- Router – узел Tarantool Enterprise Edition с установленной ролью Cartridge-router
- Master/Replica - узел Tarantool Enterprise Edition с установленной ролью Cartridge-storage используется для непосредственного хранения данных.
 - Master – обеспечивает операции чтения и записи данных.
 - Replica – обеспечивает только операции чтения данных
- Etcd (представлен на схеме логотипом) – узел etcd используемый для консистентного хранения состояния кластеров (Active/Passive)

- Replicator – часть Tarantool Clusters Federation обеспечивающая отправку данных в соседний кластер по протоколу gRPC

2.5.2 Изолированные кластеры



На схеме представлен пример размещения 2 кластеров Tarantool Enterprise Edition вместе с TCF в двух геоураспределенных ЦОД. Изолированный кластер подразумевает размещение Master и Replica узлов в одном ЦОД.

Решение состоит из:

- Application – приложение клиента, которое работает с данными, хранящимися в кластере
- Balancer – программное или аппаратное решение, используемое для распределения (балансировки) всех запросов приложения по кластеру с целью оптимизации использования вычислительных ресурсов
- Router – узел Tarantool Enterprise Edition с установленной ролью Cartridge-router
- Master/Replica - узел Tarantool Enterprise Edition с установленной ролью Cartridge-storage используется для непосредственного хранения данных.
 - Master – обеспечивает операции чтения и записи данных.
 - Replica – обеспечивает только операции чтения данных

- Etcd (представлен на схеме логотипом) – узел etcd используемый для консистентного хранения состояния кластеров (Active/Passive)
- Replicator – часть Tarantool Clusters Federation обеспечивающая отправку данных в соседний кластер по протоколу gRPC

2.6. Функции, выполняемые системой

Основная функция системы – возможность переключить активный кластер, сообщив при этом балансировщику информацию о том, в какой кластер необходимо отправлять запросы на чтение и запись из приложения клиента. Эта функция может использоваться для реализации следующих сценариев:

- Безопасное обновление кластеров
- Безопасное обновление схемы данных
- Переключение в момент возникновения аварийной ситуации