

Tarantool Queue Enterprise

Описание функциональных
характеристик

Версия 1.0.0 от 12.07.23

Оглавление

Термины и сокращения	3
1. Общие положения	4
1.1. Наименование системы	4
1.1.1. Полное наименование и условное наименование системы	4
1.1.2. Полное наименование разработчика	4
1.1.3. Сокращенное наименование разработчика	4
1.2. Цели, назначение и области использования	4
1.2.1. Назначение системы	4
1.2.2. Цели создания системы	5
2. Основные технические решения	6
2.1. Концептуальная архитектура решения	6
2.2. Требования к режимам функционирования	7
2.3. Функциональные требования	8
2.4. Нефункциональные требования	9
2.4.1. Синхронный API	9
2.4.2. Интерфейс для Подписок	9
2.4.3. Tarantool Data Queue Master	10
2.4.4. Tarantool Data Queue Replica	10
2.4.5. Требования к режиму работы	11
2.5. Требования к диагностированию работы	11
2.5.1. Метрики	11
2.5.2. Логирование	12
2.5.3. Алертинг	13
2.6. Требования к инфраструктуре	15

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
TQE	Tarantool Queue Enterprise - система распределенных очередей на базе in-memory платформы Tarantool Enterprise.
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
Lua	Процедурный динамически типизированный модульный язык с автоматическим управлением памятью

1. Общие положения

1.1. Наименование системы

1.1.1. Полное наименование и условное наименование системы

Полное наименование системы: «Tarantool Queue Enterprise».

Условное наименование системы: «Tarantool Queue Enterprise», или Система, или TQE.

1.1.2. Полное наименование разработчика

Общество с ограниченной ответственностью «ВК Цифровые Технологии».

1.1.3. Сокращенное наименование разработчика

ООО «ВК Цифровые Технологии».

1.2. Цели, назначение и области использования

1.2.1. Назначение системы

Tarantool Queue Enterprise (далее по тексту TQE) представляет собой распределенную систему очередей событий/сообщений для решений, подразумевающих высокую нагрузку, в сфере финтеха, real-time маркетинга, логистики, телекома и др., реализованную на базе in-memory платформы Tarantool Enterprise.

Система позволяет создавать очереди сообщений, подписываться на них и получать сообщения в рамках своих бизнес-процессов в условиях высокой нагрузки.

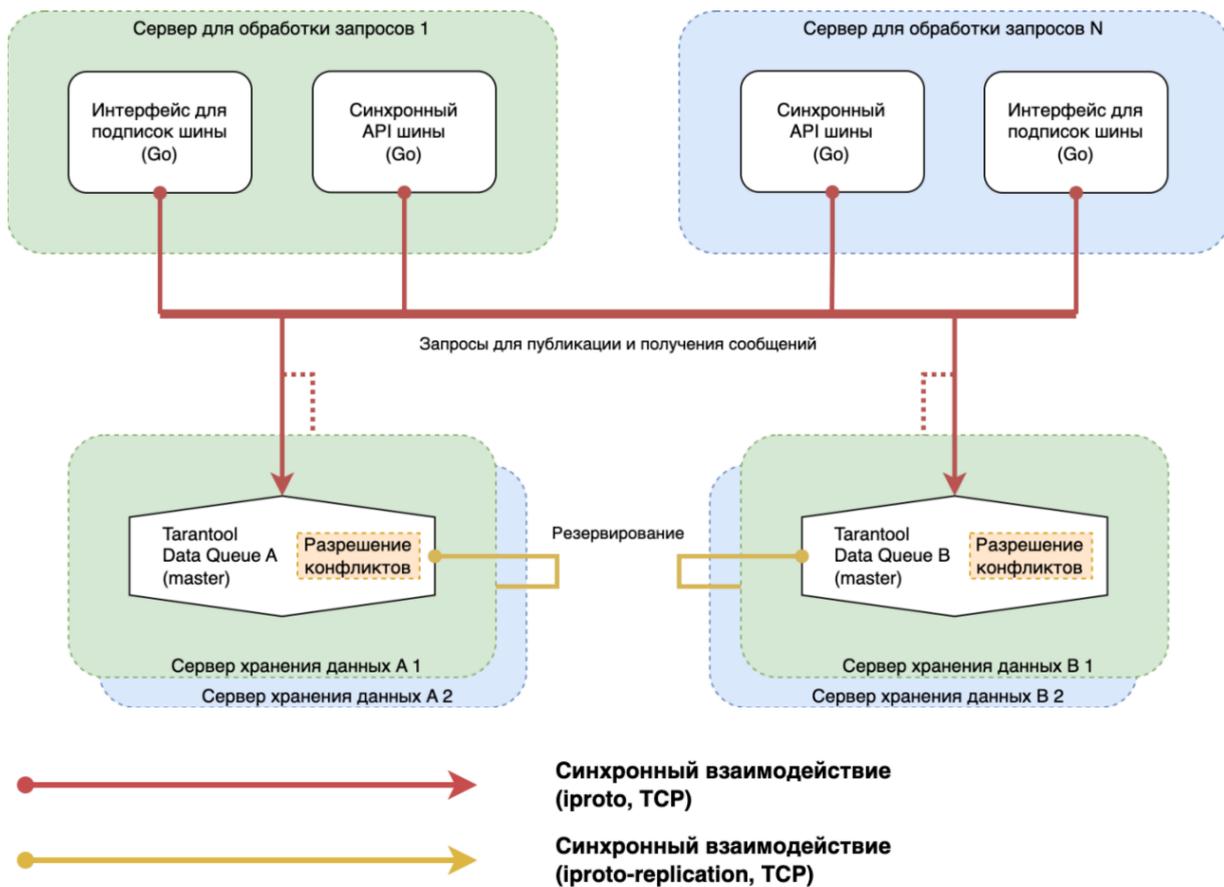
Система позволяет обрабатывать большие объемы данных с высокой производительностью и надежностью.

1.2.2. Цели создания системы

Предоставление возможности создавать и использовать очереди событий/сообщений для распределения высокой нагрузки.

2. Основные технические решения

2.1. Концептуальная архитектура решения



Архитектура TQE состоит из следующих основных компонентов:

- Tarantool Data Queue - хранит данные очередей:
 - Master - основной инстанс Tarantool (для записи);
 - Replica - резервный(-е) инстанс(-ы) Tarantool (для чтения).
- Синхронный API – обеспечивает доступ к публикации сообщений в очередь;

- Интерфейс для подписки – обеспечивает доступ к чтению сообщений из очереди.

2.2. Требования к режимам функционирования

TQE функционирует круглосуточно: 7 дней в неделю, 24 часа в сутки.

В зависимости от состояния компонентов TQE выделяются следующие основные режимы функционирования:

- Штатный режим;
- Аварийный режим функционирования;
- Режим обслуживания.

В штатном режиме функционирования TQE обеспечивает непрерывное функционирование сервисов и выполнение всех функций с заданными параметрами производительности.

Аварийным считается режим, при котором TQE или одна из его подсистем является полностью недоступной.

Для перевода TQE из аварийного в штатный режим функционирования необходимо выполнить комплекс мероприятий по восстановлению работоспособности, приведенный в документе «Эксплуатация экземпляра программного обеспечения».

В режиме обслуживания TQE обеспечивает:

- Проведение технического обслуживания, требующего полной или частичной остановки компонентов TQE, но не оказывающего существенного негативного влияния на основную работу пользователей.
- Модернизацию аппаратно-программного комплекса.
- Установка обновлений системного и прикладного ПО.

Общее время режима обслуживания TQE не должно превышать двух часов в сутки и должно приходиться на нерабочие часы основного числа пользователей.

2.3. Функциональные требования

- Создание и управление очередями данных;
- Поддержка различных типов данных;
- Гибкие настройки для управления очередями (выбор полей для дедубликации сообщений, управление шардированием);
- Публикация сообщений в очередь Tarantool Queue Enterprise;
- Получение сообщений из очереди Tarantool Queue Enterprise.

Функции модуля Синхронный API:

- Публикация сообщений:
 - Обработка запросов на публикацию сообщений.

Функции модуля Tarantool Data Queue:

- Публикация сообщений;
- Выбор нового Tarantool Data Queue мастера в случае отказа действующего;
- Получение сообщений:
 - Отдача сохраненной информации в Интерфейс для подписок.

Функции модуля Интерфейс для подписок:

- Получение сообщений:
 - Отдача данных из Tarantool Data Queue потребителям.

2.4. Нефункциональные требования

Режим функционирования (штатный):

- 24/7, предусматривается технологическое окно;
- SLA = 99.7%;
- RPO = до 24 часов при наличии ежедневного бэкапирования;
- RTO = 2 часа.

Режим обслуживания: в технологическое окно.

2.4.1. Синхронный API

Механизмы отказоустойчивости:

- избыточность по вычислительной мощности

Объем потери данных: нет потерь.

Время восстановления: 2 часа.

Время простоя: таймаут соединения.

2.4.2. Интерфейс для Подписок

Механизмы отказоустойчивости:

- избыточность по вычислительной мощности

Объем потери данных: нет потерь.

Время восстановления: 2 часа.

Время простоя: таймаут соединения.

2.4.3. Tarantool Data Queue Master

Механизмы отказоустойчивости:

- избыточность по данным (репликация);
- механизм автоматического переключения трафика внутри и между компонентами Tarantool.

Объем потери данных: нет потерь.

Время восстановления: 2 часа.

Время простоя: недоступность на запись новых событий в компоненты TQE, находящихся в отказавшем ЦОДе на конфигурируемый таймаут (на практике, обычно устанавливается 10с для избежания преждевременных переключений).

2.4.4. Tarantool Data Queue Replica

Механизмы отказоустойчивости:

- избыточность по вычислительной мощности;
- избыточность по данным (репликация);
- механизм автоматического переключения трафика внутри и между компонентами Tarantool.

Объем потери данных: нет потерь.

Время восстановления: 2 часа.

Время простоя: простой отсутствует.

2.4.5. Требования к режиму работы

TQE должен поддерживать масштабирование до конфигурации, обеспечивающей соответствие следующим характеристикам:

Характеристика	Значение
Количество сообщений в сутки	60 млн
Непрерывная работа	Не менее 24 часов
Пиковая интенсивность потока входящих сообщений	100 000
Среднее время отклика запроса, не более	100 мс
Максимальная длительность пиковой нагрузки	600 секунд

2.5 Требования к диагностированию работы

2.5.1. Метрики

Метрики собираются и отдаются с помощью библиотеки metrics. Перечень и описание метрик, которые предоставляют компоненты ПО на базе Tarantool приведен в документации Tarantool, размещенной по адресу https://www.tarantool.io/ru/doc/latest/book/monitoring/metrics_reference/.

Реализуется сбор метрик по следующим параметрам:

- **Rate** - количество запросов или задач, которые сервис обработал или принятых/отправленных сообщений. Пример: количество запросов к методу API. Метрика типа Counter. В момент запуска сервиса значение метрики равно 0. Значение метрики увеличивается на 1 после обработки каждого запроса к API или обработки каждой задачи.
- **Errors** - количество запросов или задач, обработка которых привела к ошибке. Пример: количество вызовов API подачи заявки, завершившихся с ошибкой авторизации. Метрика типа Counter. В момент запуска сервиса значение метрики равно 0. Значение метрики увеличивается на 1 после обработки каждого запроса, сообщения или задачи, который завершился ошибкой.

- **Duration** - время, которое сервис затратил на обработку запроса или задачи. Метрика типа Histogram. В момент старта сервиса значение метрики по всем заданным интервалам гистограммы равно 0. Значение метрики, в соответствующем диапазоне гистограммы, увеличивается после обработки каждого запроса на 1. Нижняя граница нижнего интервала гистограммы - 0, нижняя граница верхнего интервала гистограммы - SLA, установленное на время обработки запроса или задачи. Используется не более 9 интервалов в гистограмме. Пример задания интервалов для мониторинга времени обработки сообщения: (0, 1], (1, 2], (2, 3], (3, +∞). Для расчета Duration не используется тип метрик Summary.
- Разделение расчета по методам и типам задач реализуется с использованием дополнительных меток (label).
- Рекомендуемые к использованию метки:
 - method - для методов API;
 - task_type или job_type для обработчиков задач.

2.5.2. Логирование

Логирование в TQE выполняет следующие функции:

- Запись логов в файл;
- Переконфигурирование во время исполнения: изменение уровня логирования, изменение пути к файлу с логом. Применение конфигурации логирования не требует перезагрузки компонентов ПО;
- Поддержка ротации логов. Для этого все компоненты ПО должны реализовывать обработку сигнала операционной системы HUP. По получении сигнала HUP компонент прекращает запись в файл лога, закрывает указатель на этот файл и открывает запись в файл лога по пути, указанному в конфигурации.
- Все компоненты записывают лог в формате JSON, пример:

```
{“timestamp”: “timestamp_value 1”, “message”: “log entry 1”}
{“timestamp”: “timestamp_value 2”, “message”: “log entry 2”}
```

Для всех операций должны регистрироваться следующие сведения:

- название операции;
- время операции в формате UTC с точностью до миллисекунд.

Что логируется:

- События жизненного цикла: старт, остановка, изменение конфигурации.
- Ошибки, которые возникают во время работы.
- Взаимодействия между клиентом и TQE:
 - входящие/исходящие запросы;
 - результат выполнения запросов.

Уровень логирования VERBOSE обеспечивает расширенный уровень логирования, на котором в лог записывается:

- тело и заголовки запросов и ответов API;
- сообщения из очередей;
- сообщения, получаемые из потоков (stream) данных.

2.5.3. Алертинг

Для реализации оповещений о возникновении нештатных ситуаций в функционировании TQE предусматривается алертинг. Алертинг реализуется средствами Zabbix с помощью механизмов Triggers и Notifications.

Для настройки правил оповещения для каждого правила описываются:

- метрика, на основе значения которой, будет рассчитываться алерт;
- пороги срабатывания алерта: значение метрики, при достижении/превышении которого отправляется оповещение;

- краткое описание рекомендуемых к выполнению действий для администратора ПО при срабатывании оповещения.

Перечень оповещений, значения порогов срабатывания оповещений формируются и дополняются в процессе и по результатам проведения тестирования, а также дополняется в процессе эксплуатации.

Оповещения настраиваются для:

- пороговых значений утилизации ресурсов сервера:
 - Например, занятое место на диске;
- пороговых значений утилизации ресурсов компонентами:
 - Например: объем данных в Tarantool достиг 80% выделенной процессу Tarantool памяти;
- пороговых значений следующих параметров работы TQE:
 - Количество запросов;
 - Количество ошибок в единицу времени;
 - Пороговых значений SLA на установленное время обработки запросов компонентами.
- пороговых значений времени прохождения заявок через компоненты ПО.

Общие характеристики оповещений:

- оповещения срабатывают при наступлении заданного условия:
 - Достижения порогового значения метрик;
 - Аномальные отклонения от ожидаемых значений метрик;
 - Полное отсутствие метрик.
- при нормализации значения метрики, на которую сработало оповещения, срабатывают оповещения о нормализации метрики.

2.6. Требования к инфраструктуре

При планировании комплекса технических средств необходимо учитывать следующие требуемые характеристики TQE:

- Количество входящих запросов в секунду (RPS) на публикацию сообщений в очередь;
- Количество входящих запросов в секунду (RPS) на чтение объекта из очереди;
- Средний размер объекта – определяет требования к сетевым картам и количеству ядер CPU на серверах клиента;
- Совокупный объем хранилища в байтах – определяет требования к объемам физического хранилища серверов.