

Лекция
04.04.2019

DBA2 Theory. p5.

Асинхронная репликация MariaDB (GTID)

Ильшат Каразбаев
руководитель группы DBA
АО ТК Центр

Немного обо мне

Вместе со своей командой администрирую:

СУБД MySQL, Mariadb, galeracluster, Postgres

Главный по базам в ТК Центр

Повестка дня:

1. Вводная
2. Как работает репликация
3. Преимущества репликации
4. Структура GTID
5. Настройки репликации
6. GTID + Galera
7. Несколько примеров
8. Механизм работы
9. Литература

Вводная

На прошлой лекции мы ознакомились с классической репликацией.

Сейчас мы поговорим про GTID репликацию

Как работает репликация

На мастере все изменения БД записываются в события бинарного лога

Реплика читает события бинарного лога

Реплика может быть мастером одновременно для следующих реплик в каскадной репликации

GTID это новое событие в бинарном логе в группе событий.

GTID уникален в группе серверов и не меняется при записи в группу событий на реплике, это позволяет легко сменить мастера на другой сервер из группы серверов.

Преимущества репликации

Легко сменить мастера, так как GTID внутри одного `gtid_domain_id` не меняется на серверах

Состояние репликации записывается надежно. В старом стиле состояние записывалось в `relay-log.info` независимо от изменения данных и при авариях на реплике эта запись могла сломаться. При использовании GTID изменения пишутся в таблицу. Надежная запись касается только транзакционных DML, но не для DDL и нетранзакционных DML.

Структура GTID

GTID выглядит следующим образом:

0-1-10

0 - domain id, 32 bit unsigned integer

1 - server id, 32 bit unsigned integer

10 - sequence number, 64 bit unsigned integer, монотонно возрастает

GTID domain id

Внутри одного gtid domain id события бинарного лога пишутся последовательно и это позволяет избежать конфликтов при применении изменений на реплике

Если используется multi source репликация порядок событий бинарного лога отличается на реплике от порядка на мастерах. Они различны в различных gtid domain id, но события одного gtid domain id становятся последовательными

Реплика запоминает GTID позицию последнего изменения и отправляет значение на мастер, мастер начинает отправлять события бинарного лога на реплику начиная со следующей позиции

Настройка репликации по GTID

Настраиваем репликацию по позиции бинарного лога в старом стиле

Делаем

```
CHANGE MASTER TO MASTER_USE_GTID = slave_pos;
```

или

```
CHANGE MASTER TO MASTER_USE_GTID = current_pos;
```

Также возможно просто выбрать позицию на мастере явно и назначить на реплике

На мастере

```
SELECT BINLOG_GTID_POS("master-bin.000001", 600);
```

На реплике

```
SET GLOBAL gtid_slave_pos = "0-1-2";
```

```
CHANGE MASTER TO master_host="127.0.0.1", master_port=3310, master_user="root", master_use_gtid=slave_pos;  
START SLAVE;
```

Current_pos vs. slave_pos

current_pos - последнее изменение в GTID domain

slave_pos - последнее среплицированное изменение

с current_pos нельзя изменять данные, так как репликация остановится и сломается

с gtid_slave_pos можно остановить репликацию, изменить данные и переназначить gtid_slave_pos, потом стартовать репликацию

MariaDB GTID + Galera

Значения GTID не реплицируются writeset-ами галера кластера

Разные значения GTID могут генерироваться на нодах галера кластера, либо могут рассогласоваться в процессе записи на разные ноды.

Для того, чтобы избежать рассогласования рекомендовано включить wsrep
GTID mode on

Также важно писать в один момент времени только на одну ноду.

MariaDB GTID + Galera

Включить wsrep gtid mode:

wsrep_gtid_mode=ON

wsrep_gtid_domain_id должен быть идентичен на всех нодах

log_slave_updates - включить

log_bin должен быть идентичен на всех нодах

gtid_domain_id не равен **wsrep_gtid_domain_id**

gtid_domain_id должен быть различен на всех нодах при

wsrep_sst_method=RSU и **wsrep_on=OFF**

Смена мастера

Очень просто

```
STOP SLAVE;
```

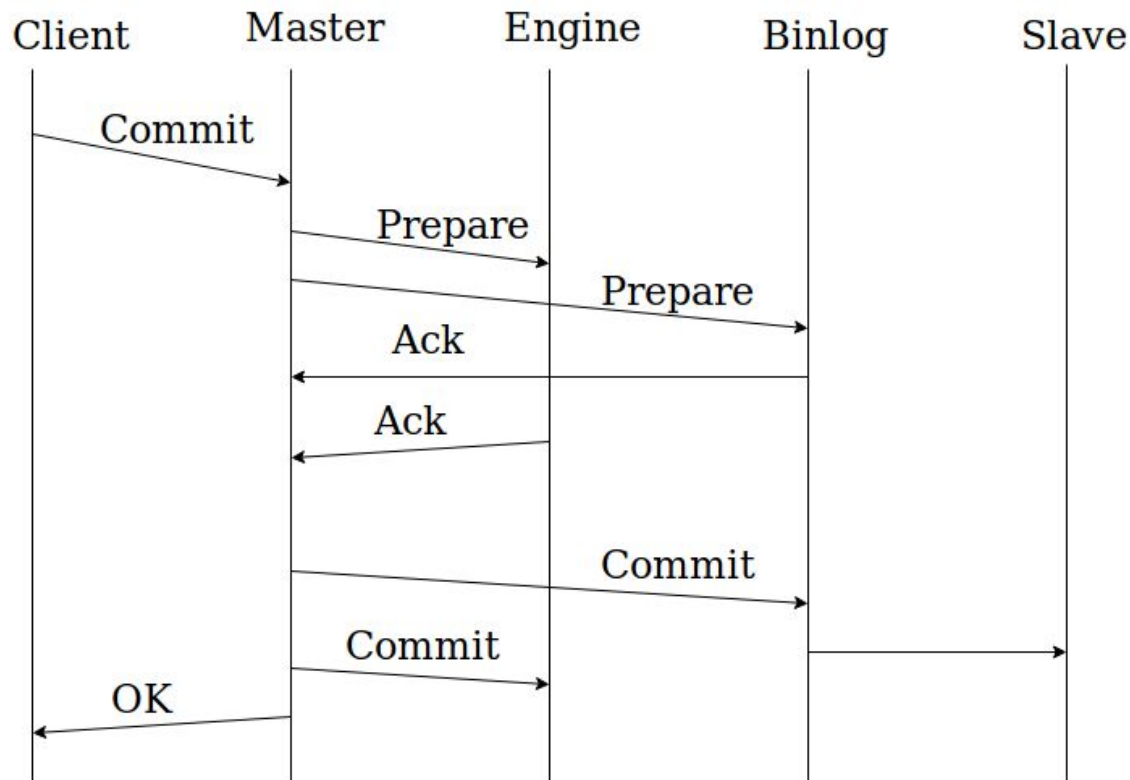
```
CHANGE MASTER TO master_host='192.168.0.121';
```

```
START SLAVE;
```

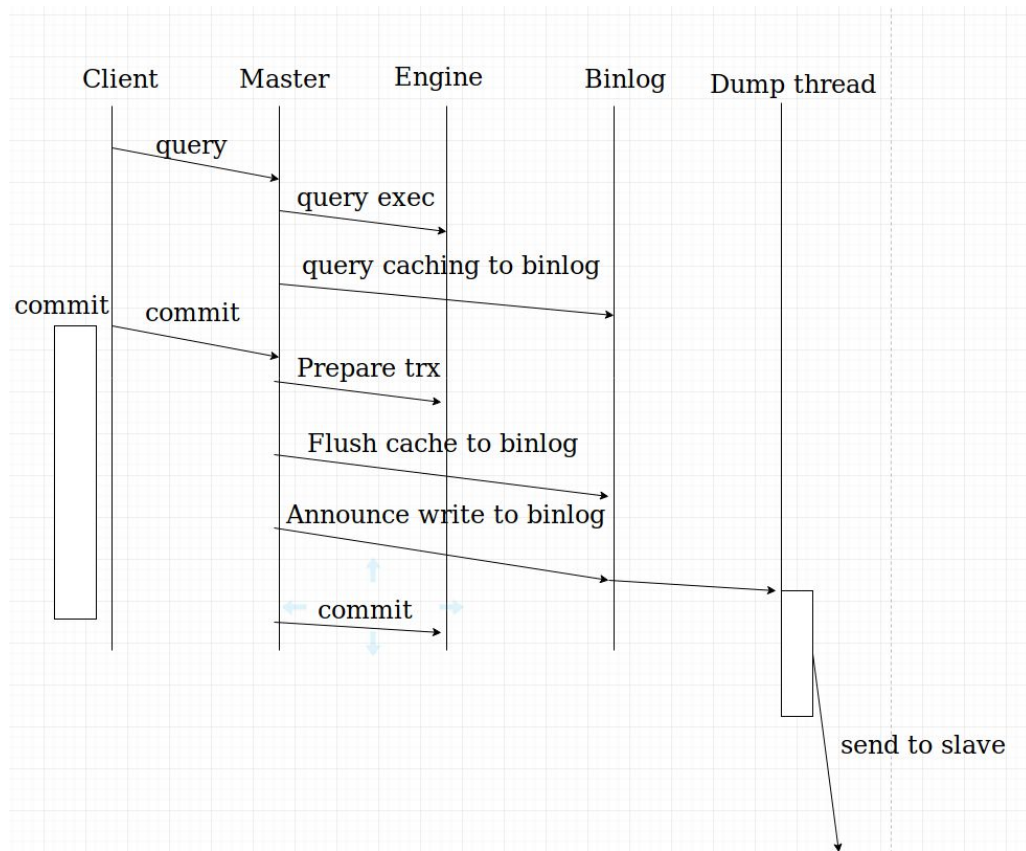
Старт репликации до определенной позиции

```
START SLAVE UNTIL master_gtid_pos = <GTID position>
```

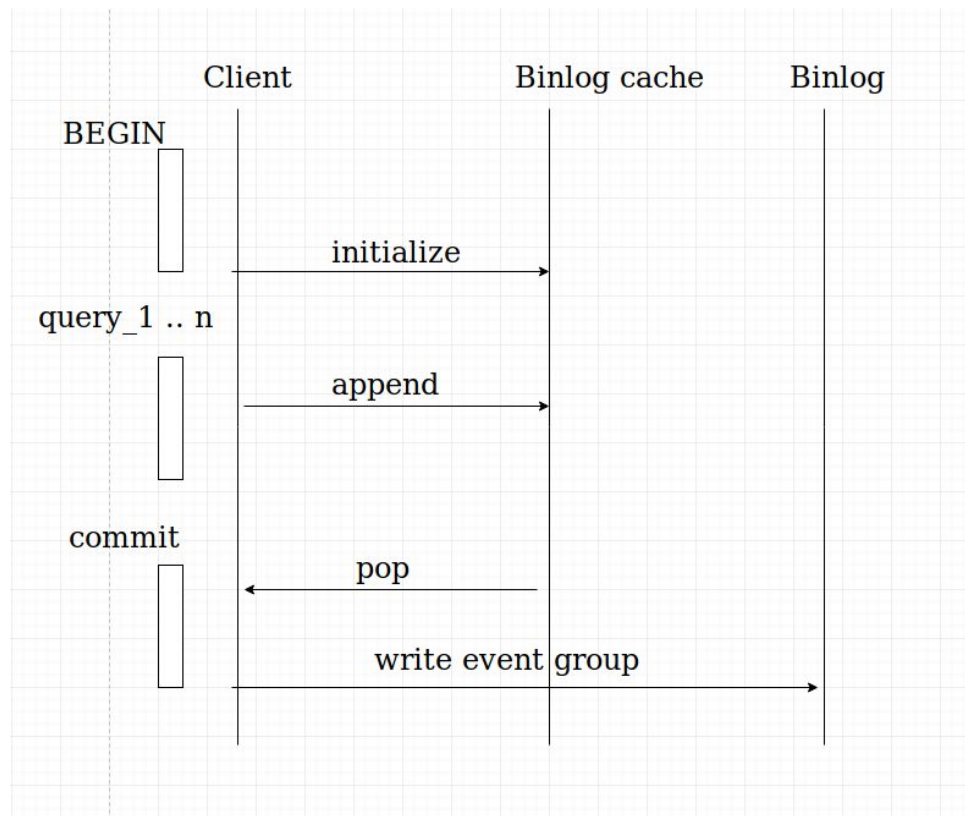
Репликация как двухфазный коммит



Механизм репликации



Механизм репликации



Виды event group

Transactional DML

DDL

Non-transactional DML

Запись event group

Запись event group GTID включает в себя

domain_id

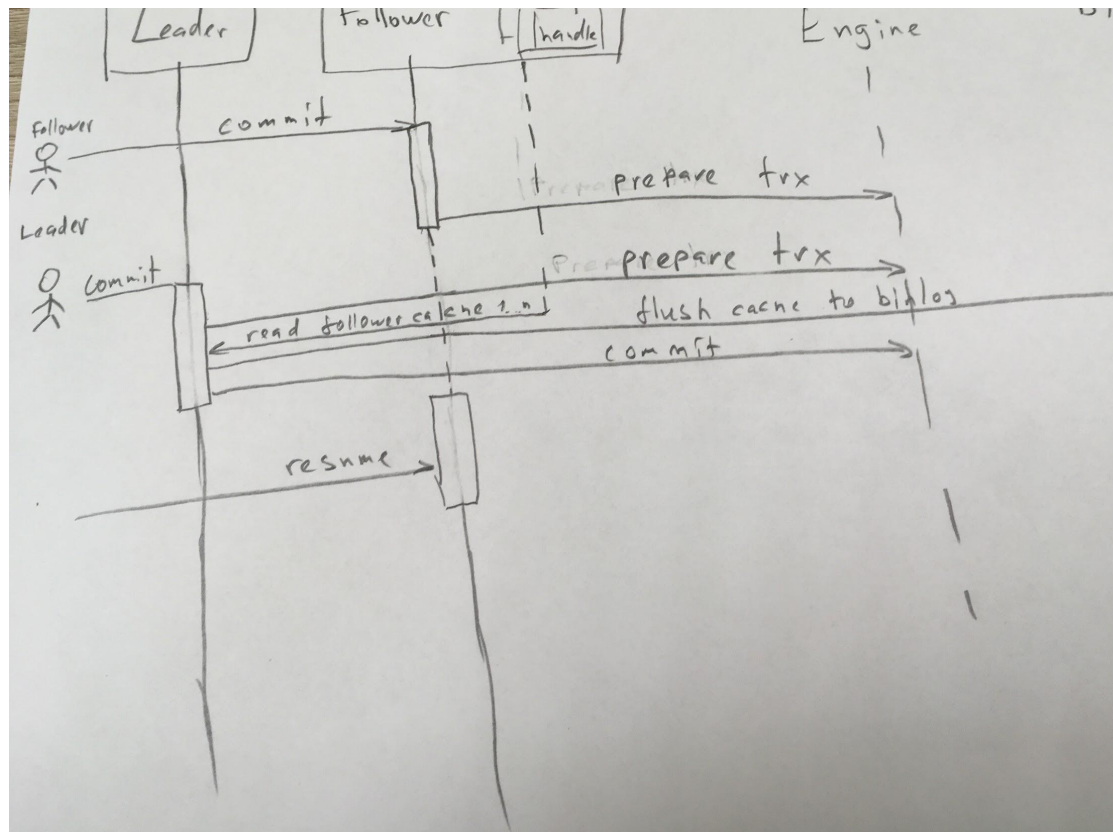
server_id

seq_no

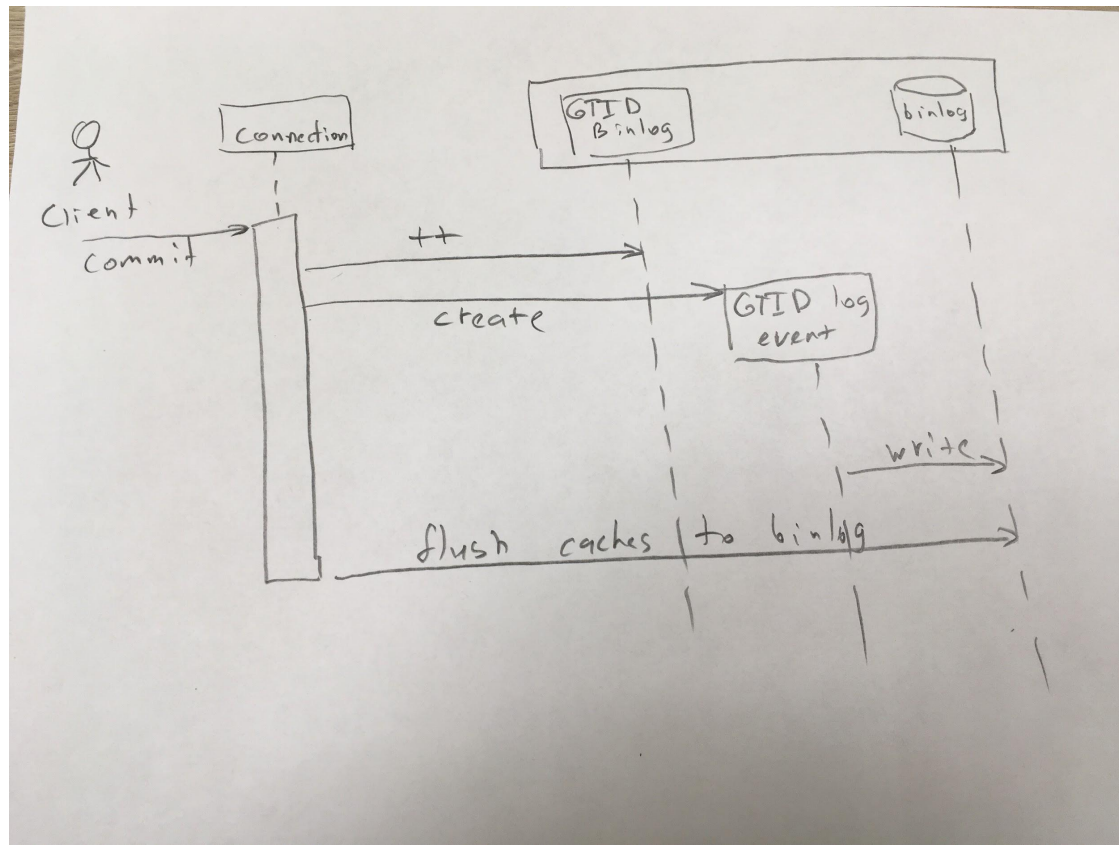
Демонстарция

binlog.sql

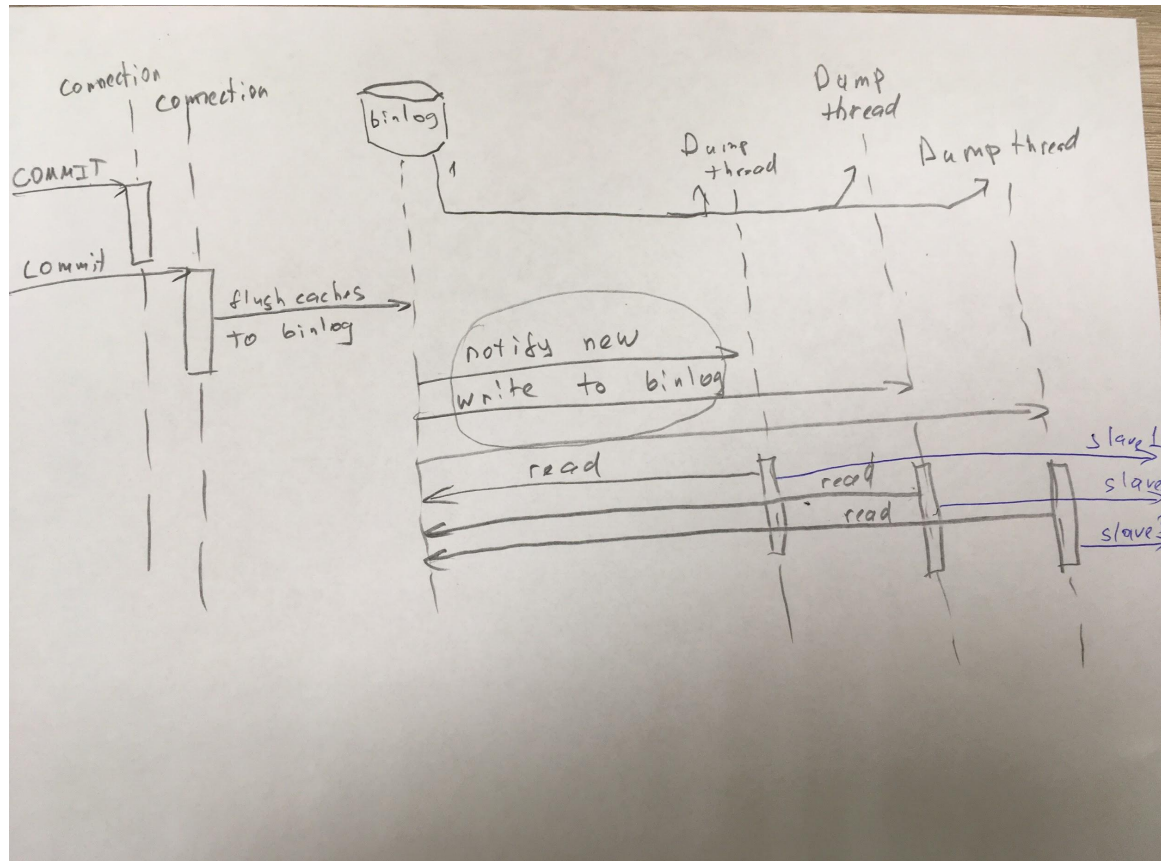
Binlog group commit



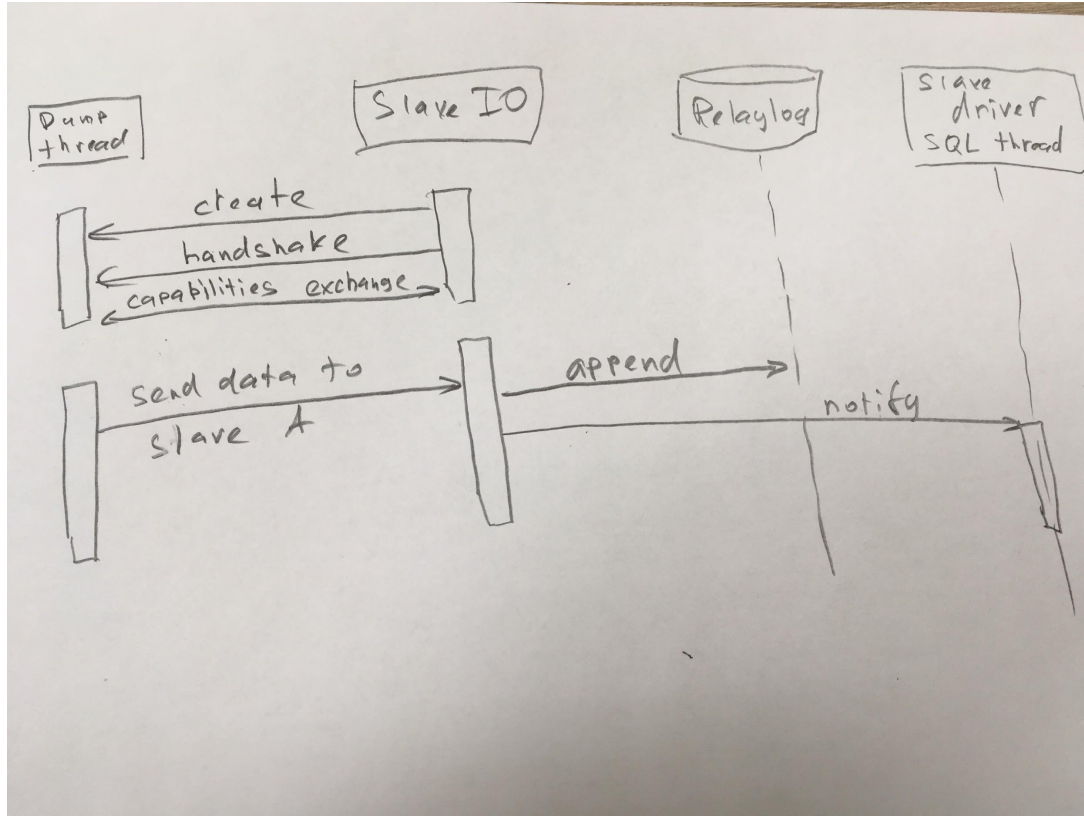
Запись GTID



Доставка событий бинлога. Dump thread



Прием событий бинлога. IO thread



Темы следующих занятий

На следующих занятиях рассмотрим

multi source replication

parallel replication

Вопросы?

Все анонсы здесь:

- telegram чат: t.me/mariadb_course

Материалы курса:

- видео: https://www.youtube.com/channel/UCGsmu6YDpcR_kWcXzeQkWrA
- слайды лекций и примеры: [git@github.com:barazbay/mariadb_course.git](https://github.com:barazbay/mariadb_course.git)

Меня можно найти:

- vk, instagram: barazbay
- twitter: karazbay

Литература

1. <https://mariadb.com/kb/en/library/gtid/>
2. <https://mariadb.com/resources/blog/enabling-gtids-for-server-replication-in-mariadb-server-10-2/>
3. <https://mariadb.com/kb/en/library/parallel-replication/>
4. <https://vimeo.com/258533271>
5. <https://mariadb.com/kb/en/library/using-mariadb-gtids-with-mariadb-galera-cluster/>