

Galera IST: что показывает PMM и что существует на самом деле

Aurélien LEQUOY · March 21, 2026

GALERA MARIADB IST PMM MONITORING



Миф о дашборде PMM для Galera IST

Если вы используете Persona Monitoring and Management (PMM) для мониторинга кластера Galera на MariaDB, вы, вероятно, заметили панель под названием «IST Progress» или «IST Receive». Она показывает... ничего. Пустые строки, значения N/A, плоские графики.

Это не баг отображения. **Переменные, которые запрашивает PMM, просто не существуют на MariaDB 10.6.**

Напоминание: IST vs SST

Когда узел Galera присоединяется к кластеру после отключения, возможны два механизма синхронизации:

- **SST (State Snapshot Transfer):** узел-донор отправляет полную копию набора данных. Медленно, ресурсоёмко, потенциально блокирует донора. Речь идёт о минутах или часах в зависимости от объёма.

- **IST (Incremental State Transfer)**: донор отправляет только недостающие writeset-ы из GCache. Быстро, легко, от нескольких секунд до нескольких минут.

Разница критична в production. IST за 20 секунд незаметен для пользователей. SST за 45 минут может спровоцировать инцидент.

Тест: MariaDB 10.6.23 + sysbench

Для документирования реального поведения мы подняли кластер Galera из 3 узлов на MariaDB 10.6.23 и запустили непрерывную нагрузку sysbench:

```
sysbench oltp_read_write --tables=10 --table-size=100000 \  
--threads=16 --time=600 --db-driver=mysql run
```

Во время нагрузки мы остановили узел 3 на 30 секунд, затем перезапустили. Результат:

- **188 516 writeset-ов** накоплено в GCache донора
- Узел 3 присоединился к кластеру через **IST за 20-25 секунд**
- Никакого прерывания сервиса на узлах 1 и 2

Лог MariaDB на узле 3 подтверждает:

```
[Note] WSREP: Receiving IST: 188516 writesets, seqnos 1045632-1234148  
[Note] WSREP: IST received: 85a4c3e2-xxxx  
[Note] WSREP: 3.0 (node3): State transfer from 0.0 (node1) complete.
```

IST сработал идеально. Теперь посмотрим, что об этом говорит PMM.

Что пытается прочитать PMM

Дашборд Galera в PMM v2 запрашивает следующие переменные для отслеживания прогресса IST:

```
SHOW GLOBAL STATUS LIKE 'wsrep_ist_receive_seqno_start';  
SHOW GLOBAL STATUS LIKE 'wsrep_ist_receive_seqno_current';  
SHOW GLOBAL STATUS LIKE 'wsrep_ist_receive_seqno_end';
```

На MariaDB 10.6.23 все три запроса возвращают **пустой результат**. Переменные не существуют.

```
MariaDB [(none)]> SHOW GLOBAL STATUS LIKE 'wsrep_ist%';  
Empty set (0.001 sec)
```

Это не ошибка конфигурации. Эти переменные специфичны для Percona XtraDB Cluster (PXC) и никогда не были реализованы в провайдере Galera от MariaDB.

Ловушка SST-обёртки

Дополнительная деталь вносит путаницу: даже во время IST MariaDB вызывает скрипт SST-обёртки (`wsrep_sst_mariabackup` или `wsrep_sst_rsync`). Лог, таким образом, содержит строки вроде:

```
WSREP: Running: 'wsrep_sst_mariabackup --role donor ...'
```

Оператор, читающий эти логи, мог бы заключить, что идёт полный SST. В действительности **обёртка инициализируется, затем обнаруживает, что догоняние будет выполнено через IST**. Реальный трансфер инкрементальный.

Как обнаружить IST на MariaDB

Поскольку переменные PMM не существуют, нужен альтернативный подход. Три метода:

1. Парсинг лога MariaDB

Сигнатура IST в логе однозначна:

```
[Note] WSREP: Receiving IST: <N> writesets, seqnos <start>-<end>
```

Простой grep по error log даёт ответ мгновенно. Это самый надёжный метод.

2. Наблюдение за `wsrep_local_state_comment`

Во время IST переменная `wsrep_local_state_comment` присоединяющегося узла проходит через:

```
Joining → Joined → Synced
```

Если этот переход занимает менее 30 секунд для активного кластера, это с высокой вероятностью IST. SST на наборе данных в десятки гигабайт занял бы значительно больше времени.

3. Проверка GCache

Переменная `wsrep_local_cached_downto` на доноре показывает старейший `seqno`, ещё доступный в GCache. Если `seqno` отключённого узла больше этого значения, IST возможен:

```
-- На доноре
SHOW GLOBAL STATUS LIKE 'wsrep_local_cached_downto';
-- Результат: 1045000

-- Если отключённый узел был на seqno 1045632 → IST возможен
-- Если отключённый узел был на seqno 8000000 → SST обязателен
```

Что делает PmaControl

PmaControl комбинирует все три метода для автоматического обнаружения и классификации трансферов Galera:

1. **Непрерывный мониторинг** `wsrep_local_state_comment` — обнаружение перехода в состояние `Joining`
2. **Парсинг лога MariaDB** — извлечение строки `Receiving IST` с количеством `writeset-ov`
3. **Временная корреляция** — измерение времени между `Joining` и `Synced`

Результат отображается на дашборде PmaControl с чётким различием: IST (зелёный бейдж, длительность в секундах) vs SST (оранжевый бейдж, расчётная длительность в минутах).

В отличие от PMM, PmaControl не зависит от переменных, существующих только в PXC. Подход через парсинг логов работает на **всех версиях MariaDB Galera** начиная с 10.1.

Ключевые цифры

Метрика	Наблюдаемое значение
Тестируемая версия	MariaDB 10.6.23 Galera

Метрика	Наблюдаемое значение
Нагрузка sysbench	16 потоков, oltp_read_write
Накопленных writeset-ов	188 516
Длительность IST	20-25 секунд
Переменные PMM <code>wsrepist*</code>	Не существуют
Обнаружение PmaControl	Автоматическое через парсинг логов

Рекомендации

1. **Не полагайтесь на дашборд Galera в PMM** для отслеживания IST, если вы на MariaDB — панели останутся пустыми
2. **Щедро размерьте GCache** (`gcache.size=2G` минимум) для максимизации шансов IST после краткого отключения
3. **Централизуйте логи MariaDB** — это источник истины для трансферов Galera
4. **Используйте PmaControl** для мониторинга Galera, который реально работает на MariaDB, а не только на PXC

Заключение

Дашборд PMM для Galera IST создан для Percona XtraDB Cluster. На MariaDB Galera он показывает пустоту — не потому что IST не работает, а потому что запрашиваемые переменные не существуют.

IST прекрасно работает на MariaDB 10.6. Нужно просто знать, куда смотреть: в логи, а не в фантомные переменные.