

Galera IST: co PMM pokazuje vs co naprawdę istnieje

Aurélien LEQUOY · March 21, 2026

GALERA MARIADB IST PMM MONITORING



Mit dashboardu PMM dla Galera IST

Jeśli używasz Percona Monitoring and Management (PMM) do nadzoru klastra Galera na MariaDB, prawdopodobnie zauważyłeś panel zatytułowany "IST Progress" lub "IST Receive". Wyświetla... nic. Puste linie, wartości N/A, płaskie wykresy.

To nie jest błąd wyświetlania. **Zmienne, które PMM odpytuje, po prostu nie istnieją na MariaDB 10.6.**

Przypomnienie: IST vs SST

Gdy węzeł Galera dołącza do klastra po rozłączeniu, możliwe są dwa mechanizmy synchronizacji:

- **SST (State Snapshot Transfer):** węzeł donor wysyła pełną kopię zestawu danych. Wolny, kosztowny, potencjalnie blokujący donora. Mowa o minutach do godzin w zależności od wolumenu.

- **IST (Incremental State Transfer)**: donor wysyła tylko brakujące writesety z GCache. Szybki, lekki, od kilku sekund do kilku minut.

Różnica jest krytyczna na produkcji. IST trwający 20 sekund jest niewidoczny dla użytkowników. SST trwający 45 minut może spowodować incydent.

Test: MariaDB 10.6.23 + sysbench

Aby udokumentować rzeczywiste zachowanie, postawiliśmy klaster Galera z 3 węzłami na MariaDB 10.6.23 i uruchomiliśmy ciągłe obciążenie sysbench:

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

Podczas obciążenia zatrzymaliśmy węzeł 3 na 30 sekund, a następnie go zrestartowaliśmy. Wynik:

- **188 516 writesetów** nagromadzonych w GCache donora
- Węzeł 3 dołączył do klastra przez **IST w 20-25 sekund**
- Brak przerwania usługi na węzłach 1 i 2

Log MariaDB węzła 3 potwierdza:

```
[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 zadziałał doskonale. Zobaczmy teraz, co PMM na ten temat mówi.

Co PMM próbuje odczytać

Dashboard Galera PMM v2 odpytuje następujące zmienne, aby śledzić postęp 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';
```

Na MariaDB 10.6.23 te trzy zapytania zwracają **pusty wynik**. Zmienne nie istnieją.

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

To nie jest brak konfiguracji. Te zmienne są specyficzne dla Percona XtraDB Cluster (PXC) i nigdy nie zostały zaimplementowane w providerze Galera MariaDB.

Pułapka SST wrappera

Dodatkowy szczegół sieje zamęt: nawet podczas IST, MariaDB wywołuje skrypt SST wrapper (`wsrep_sst_mariabackup` lub `wsrep_sst_rsyc`). Log zawiera więc linie takie jak:

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

Operator czytający te logi mógłby stwierdzić, że trwa pełny SST. W rzeczywistości **wrapper inicjalizuje się, a następnie wykrywa, że nadrabianie nastąpi przez IST**. Faktyczny transfer jest inkrementalny.

Jak wykryć IST na MariaDB

Ponieważ zmienne PMM nie istnieją, potrzebne jest alternatywne podejście. Trzy metody:

1. Parsowanie logu MariaDB

Sygnatura IST w logu jest jednoznaczna:

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

Prosty grep na error logu daje natychmiastową odpowiedź. To najbardziej niezawodna metoda.

2. Obserwacja `wsrep_local_state_comment`

Podczas IST zmienna `wsrep_local_state_comment` dołączającego węzła przechodzi przez:

```
Joining → Joined → Synced
```

Jeśli ta zmiana trwa mniej niż 30 sekund dla aktywnego klastra, jest to najprawdopodobniej IST. SST na zbiorze danych o rozmiarze kilkudziesięciu gigabajtów trwałby znacznie dłużej.

3. Weryfikacja GCACHE

Zmienna `wsrep_local_cached_downto` na donorze wskazuje najstarszy seqno wciąż dostępny w GCaCache. Jeśli seqno odłączonego węzła jest wyższe od tej wartości, IST jest możliwy:

```
-- Na donorze
SHOW GLOBAL STATUS LIKE 'wsrep_local_cached_downto';
-- Wynik: 1045000

-- Jeśli odłączony węzeł był na seqno 1045632 → IST możliwy
-- Jeśli odłączony węzeł był na seqno 800000 → SST obowiązkowy
```

Co robi PmaControl

PmaControl łączy te trzy metody, aby automatycznie wykrywać i klasyfikować transfery Galera:

1. **Ciągłe monitorowanie** `wsrep_local_state_comment` — wykrywanie przejścia w stan `Joining`
2. **Parsowanie logu MariaDB** — ekstrakcja linii `Receiving IST` z liczbą writesetów
3. **Korelacja czasowa** — pomiar czasu między `Joining` a `Synced`

Wynik wyświetla się w dashboardzie PmaControl z wyraźnym rozróżnieniem: IST (zielona plakietka, czas w sekundach) vs SST (pomarańczowa plakietka, szacowany czas w minutach).

W przeciwieństwie do PMM, PmaControl nie zależy od zmiennych istniejących tylko na PXC.

Podejście przez parsowanie logów działa na **wszystkich wersjach MariaDB Galera** od 10.1.

Liczby do zapamiętania

Metryka	Zaobserwowana wartość
Testowana wersja	MariaDB 10.6.23 Galera
Obciążenie sysbench	16 wątków, oltp_read_write
Nagromadzone writesety	188 516
Czas IST	20-25 sekund
Zmienne PMM <code>wsrepist*</code>	Nie istnieją
Wykrywanie PmaControl	Automatyczne przez parsowanie logów

Rekomendacje

1. **Nie polegaj na dashboardzie Galera PMM** do śledzenia IST, jeśli korzystasz z MariaDB — panele pozostaną puste
2. **Wymiaruj GCache** hojnie (`gcache.size=2G` minimum), aby zmaksymalizować szanse na IST po krótkim rozłączeniu
3. **Centralizuj logi MariaDB** — to źródło prawdy dla transferów Galera
4. **Używaj PmaControl** do nadzoru Galera, który naprawdę działa na MariaDB, nie tylko na PXC

Podsumowanie

Dashboard PMM dla Galera IST jest zaprojektowany dla Percona XtraDB Cluster. Na MariaDB Galera wyświetla pustkę — nie dlatego, że IST nie działa, ale dlatego, że zmienne, które odpytuje, nie istnieją.

IST działa doskonale na MariaDB 10.6. Trzeba po prostu wiedzieć, gdzie szukać: w logach, nie w zmiennych-widmach.