

Galera IST : Ce que PMM Affiche vs Ce qui Existe Réellement

Aurélien LEQUOY · 21 mars 2026

GALERA MARIADB IST PMM MONITORING



Le mythe du dashboard PMM pour Galera IST

Si vous utilisez Percona Monitoring and Management (PMM) pour superviser un cluster Galera sur MariaDB, vous avez probablement remarqué un panneau intitulé "IST Progress" ou "IST Receive". Il affiche... rien. Des lignes vides, des valeurs N/A, des graphiques plats.

Ce n'est pas un bug d'affichage. **Les variables que PMM interroge n'existent tout simplement pas sur MariaDB 10.6.**

Rappel : IST vs SST

Quand un noeud Galera rejoint le cluster après une déconnexion, deux mécanismes de synchronisation sont possibles :

- **SST (State Snapshot Transfer)** : le noeud donateur envoie une copie complète du dataset. Lent, coûteux, bloque potentiellement le donateur. On parle de minutes à heures selon le volume.

- **IST (Incremental State Transfer)** : le donateur envoie uniquement les writesets manquants depuis le GCache. Rapide, léger, quelques secondes à quelques minutes.

La différence est critique en production. Un IST de 20 secondes est invisible pour les utilisateurs. Un SST de 45 minutes peut provoquer un incident.

Le test : MariaDB 10.6.23 + sysbench

Pour documenter le comportement réel, nous avons monté un cluster Galera 3 noeuds sur MariaDB 10.6.23 et lancé une charge sysbench continue :

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

Pendant la charge, nous avons arrêté le noeud 3 pendant 30 secondes, puis redémarré. Résultat :

- **188 516 writesets** accumulés dans le GCache du donateur
- Le noeud 3 a rejoint le cluster via **IST en 20-25 secondes**
- Aucune interruption de service sur les noeuds 1 et 2

Le log MariaDB du noeud 3 confirme :

```
[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.
```

L'IST a fonctionné parfaitement. Maintenant, voyons ce que PMM en dit.

Ce que PMM essaie de lire

Le dashboard Galera de PMM v2 interroge les variables suivantes pour tracer la progression 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';
```

Sur MariaDB 10.6.23, ces trois requêtes retournent **un résultat vide**. Les variables n'existent pas.

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

Ce n'est pas une omission de configuration. Ces variables sont spécifiques à Percona XtraDB Cluster (PXC) et n'ont jamais été implémentées dans le provider Galera de MariaDB.

Le piège du SST wrapper

Un détail supplémentaire sème la confusion : même lors d'un IST, MariaDB appelle le script SST wrapper (`wsrep_sst_mariabackup` ou `wsrep_sst_rsync`). Le log contient donc des lignes comme :

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

Un opérateur qui lit ces logs pourrait conclure qu'un SST complet est en cours. En réalité, **le wrapper s'initialise puis détecte que le catch-up se fera par IST**. Le transfert réel est incrémental.

Comment détecter un IST sur MariaDB

Puisque les variables PMM n'existent pas, il faut une approche alternative. Trois méthodes :

1. Parser le log MariaDB

La signature IST dans le log est sans ambiguïté :

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

Un simple grep sur le error log donne la réponse instantanément. C'est la méthode la plus fiable.

2. Observer `wsrep_local_state_comment`

Pendant un IST, la variable `wsrep_local_state_comment` du noeud qui rejoint passe par :

```
Joining → Joined → Synced
```

Si cette transition prend moins de 30 secondes pour un cluster actif, c'est très probablement un IST. Un SST sur un dataset de plusieurs dizaines de gigaoctets prendrait beaucoup plus longtemps.

3. Vérifier le GCACHE

La variable `wsrep_local_cached_downto` sur le donateur indique le plus ancien seqno encore disponible dans le GCACHE. Si le seqno du noeud déconnecté est supérieur à cette valeur, l'IST est possible :

```
-- Sur le donateur
SHOW GLOBAL STATUS LIKE 'wsrep_local_cached_downto';
-- Résultat : 1045000

-- Si le noeud déconnecté était au seqno 1045632 → IST possible
-- Si le noeud déconnecté était au seqno 800000 → SST obligatoire
```

Ce que fait PmaControl

PmaControl combine ces trois méthodes pour détecter et classier automatiquement les transferts Galera :

1. **Surveillance continue de `wsrep_local_state_comment`** — détection du passage en état `Joining`
2. **Parse du log MariaDB** — extraction de la ligne `Receiving IST` avec le nombre de writesets
3. **Corrélation temporelle** — mesure du temps entre `Joining` et `Synced`

Le résultat s'affiche dans le dashboard PmaControl avec une distinction claire : IST (badge vert, durée en secondes) vs SST (badge orange, durée estimée en minutes).

Contrairement à PMM, PmaControl ne dépend pas de variables qui n'existent que sur PXC.

L'approche par log parsing fonctionne sur **toutes les versions de MariaDB Galera** depuis la 10.1.

Les chiffres à retenir

Métrique	Valeur observée
Version testée	MariaDB 10.6.23 Galera
Charge sysbench	16 threads, oltp_read_write
Writesets accumulés	188 516

Métrique	Valeur observée
Durée IST	20-25 secondes
Variables PMM <code>wsrep/ist*</code>	N'existent pas
Détection PmaControl	Automatique via log parsing

Recommandations

1. **Ne vous fiez pas au dashboard Galera de PMM** pour le suivi IST si vous êtes sur MariaDB — les panneaux resteront vides
2. **Dimensionnez le GCache** généreusement (`gcache.size=2G` minimum) pour maximiser les chances d'IST après une déconnexion brève
3. **Centralisez les logs MariaDB** — c'est la source de vérité pour les transferts Galera
4. **Utilisez PmaControl** pour une supervision Galera qui fonctionne réellement sur MariaDB, pas seulement sur PXC

Conclusion

Le dashboard PMM pour Galera IST est conçu pour Percona XtraDB Cluster. Sur MariaDB Galera, il affiche du vide — non pas parce que l'IST ne fonctionne pas, mais parce que les variables qu'il interroge n'existent pas.

L'IST fonctionne parfaitement sur MariaDB 10.6. Il faut simplement savoir où regarder : dans les logs, pas dans des variables fantômes.