

進擊的網管Jay

不負責DevOps

首頁 關於我

Galera Cluster不存在同步延遲？不，Galera Cluster到處都是同步延遲

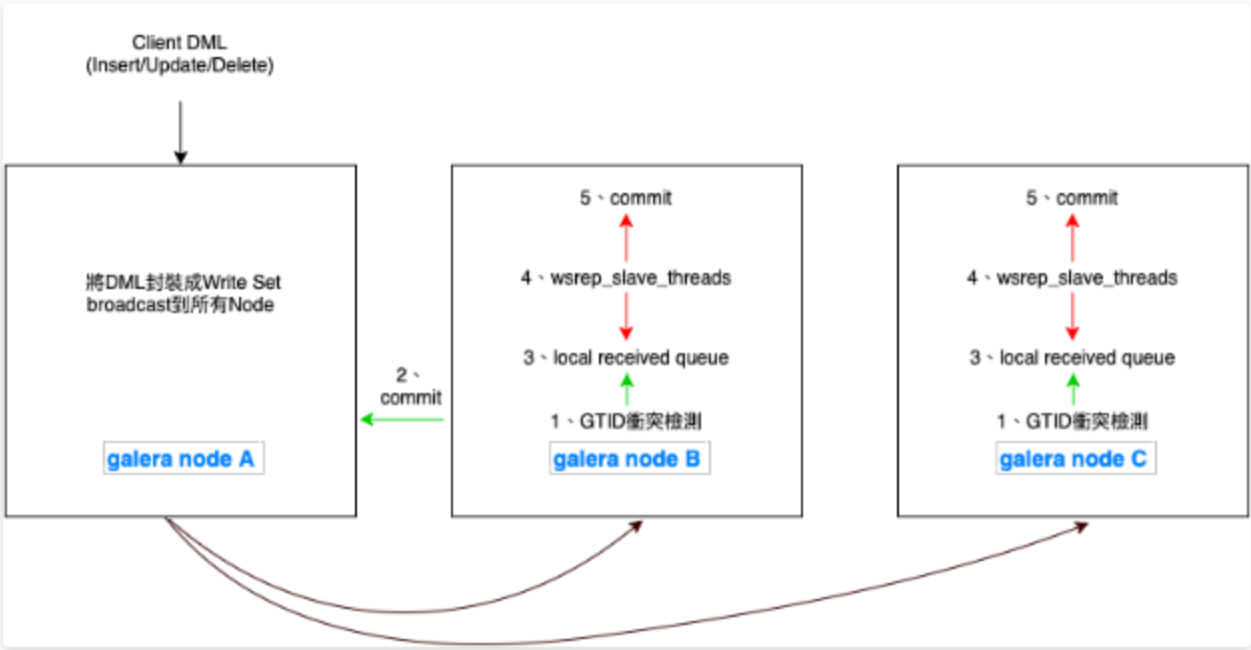
發佈留言

Views: 3207

這陣子因為資料量過大的問題，導致我們MariaDB的Replication架構一直掉資料，所以我們決定開始評估Galera Cluster，查了一堆Blog，都說到Galera Cluster是 Multi-master，每一個Node都可以讀寫，且沒有資料同步延遲問題.....自己試用下來，才發現誤會大了，原來Galera Cluster到處都是同步延遲.....

Galera Cluster的同步流程

Galera Cluster採用的是虛擬同步，又叫邏輯同步，指的是一筆交易在一個Node上執行成功後，保證它在其它Node上也一定會被成功執行，但不保證能即時同步。我們由下圖簡略說明Galera Cluster是如何實現同步機制的，當Client端有DML請求(Insert/Update/Delete)落在node A時，node A會將DML封裝成Write Set，並廣播到node B及node C，node B及node C收到Write Set後，進行衝突檢測，如有資料衝突，則所有Node都rollback，如果沒有資料衝突，先通知node A進行commit，node B及node C則將資料寫入queue，在由wsrep_slave_threads讀取queue的資料後進行commit，由此同步流程看得出來，資料同步時勢必會延遲。



為什麼Galera Cluster不採用全同步而是虛擬同步?

Galera Cluster其實可以採用全同步模式，只要將wsrep_sync_wait參數設為1即可，但要考量到一個情況，全同步代表的是每一個Node都要確實寫入資料後，才可以接著下一筆資料寫入，但當cluster裡的某一台Node負載較重時，全同步會導致已完成寫入作業的Node要等待尚未寫入作業的Node，此時會導致整體寫入作業變慢，這樣就等於喪失Multi-master可以同時多個Node寫入資料的優點，所以Galera Cluster預設不採用全同步模式

如何掌控Galera Cluster的同步延遲，使其同步延遲減至最小

- 1、調整wsrep_slave_threads參數：有多少支執行緒可以同時由queue讀取資料及commit，建議為cpu的1~1.5倍
- 2、調整flow control參數：避免任一Node交易時落後其它Node太多，用於協調每個Node，保證commit的速度可優於queue的增長速度

分類: [MariaDB](#)，發佈日期: 2020 年 5 月 2 日，作者:

[← 使用MariaDB Thread Pool實現DB端的連接池](#) [使用Suricata搭配ELK實現入侵偵測系統IDS →](#)

發佈留言

發佈留言必須填寫的電子郵件地址不會公開。必填欄位標示為 *

留言 *

顯示名稱 *

電子郵件地址 *

個人網站網址

☐ 在瀏覽器中儲存顯示名稱、電子郵件地址及個人網站網址，以供下次發佈留言時使用。

發佈留言

本站採用 [WordPress](#) 建置