

GoldenGate – 一個即時性的異質資料庫複製工具

隨著科技日新月異的進步，RDBMS 引擎的功能及效能是愈來愈強悍了。但即使如此，無論是 Oracle、DB2、Sybase 還是 MSSQL，若要它們在處理日常 OLTP 交易之外，還要額外進行其他諸如報表查詢、統計分析等工作的話，相信再好的系統也負擔不了。因此今天在講到資料庫管理的時候，便會經常談到一個概念：建立一個分離的報表查詢系統(Report System)。

以目前主流的技术來說，最常見的就是以資料複製(Data Replication)的方法將 Production DB 複製一份到另一個資料庫，使之成為企業內專供報表等其他系統查詢使用。而用來進行複製的方式則是廠商們百家爭鳴、各出奇招，各有各的特色。用戶們要使用哪一種工具進行來 Replication，並沒有一定的答案，全看該產品的特色是否能真正契合用戶的需求。本文之所以會針對 Oracle 公司的 GoldenGate 軟體做介紹，主要也就是認為此產品在建立所謂 Report System 上，比其他工具具有更多優勢與特色。茲說明並比較如下：

Storage Level Replication

Replication 直接透過硬體(不論是 Physical Storage 或 Logical Volume)來進行，其實是個相當好的主意。雖然成本較為昂貴，但由於他是直接進行硬體層的複製，完全不必去分析資料庫邏輯架構、或執行了哪些 DML 與 DDL 指令，省下大量的 CPU 運算，效能甚佳。這個優勢在面對數十、數百 TB 以上的資料量時尤為顯著。但比較大的缺點是，絕大多數這一類利用儲存設備來進行的 Replication，複製出來後的資料庫都是封閉的；也就是說，當他在進行複製時，其他 AP 也好、報表程式也罷，都無法連結到這個複製出來的資料庫，導致使用者最大的需求：『建立報表系統』成為不可能。因此儘管硬體層級的 Replication 有這麼多優勢，我們還是不得不採取其他的措施。

Database Level Replication

另一種主要的 Replication 方式是從資料庫層級來進行的。這種複製工具會嘗試著去捕捉、分析所有發生在資料來源端(多半就是 Production DB)的 DML 或 DDL，並將他們在資料目的端(也就是報表系統的 DB)重新執行一次，間接的使得來源與目的資料庫達成一致，以提供使用者正確的資料來產生報表。常見的此類工具包括有 Oracle Stream、Oracle Data Guard、DSG Real Sync、Quest SharePlex、以及 GoldenGate 等等...

但即使是此類工具，有的還是會發生上述『複製端無法開啓應用』這樣的狀況。例如大家所熟悉的 Data Guard 就是如此，當他正在複製資料時，是無法開放給其他系統連結或查詢的。

而除了複製端資料庫的開放性這個問題外，通常使用者們在建立報表查詢專用資料庫時還會注意到的是資料篩選的問題。在 Production DB 裡面，很可能並不是每一個資料表都會被查

詢系統使用到，如果完全未經篩選的就把整個 DB 複製過來，往往造成很多軟硬體、甚至網路資源的浪費。

另外一個 DBA 們經常會列入建置 Replication 考量重點的是資源競爭的問題。通常在建置資料庫系統的時候，該配給多少記憶體資源、會耗用多少 CPU 時間、以及將來可能的頻寬消耗等，一定都經過相當嚴謹的規劃與計算。但忽然多了一個 Replication 系統進來，他會不會因此跟資料庫競爭那些 CPU、RAM、Bandwidth 之類的 Resource 呢？我們希望最好不會，但實務上你會很失望的發現，相當多的 Replication 工具是會去占用資料庫本身資源的。這將來或多或少會對資料庫的運作與效能造成衝擊。

再來要考慮的則是資料的即時性。某些報表有強烈的即時資料需求，那使用者所選用的複製工具就必須盡量做到 Real Time 或 Near Real Time。可惜不少複製工具卻因為某些技術上的原因無法提供這樣的即時性。例如，某些工具的運作方式是要等到資料來源端做出 Commit 或 Rollback 等動作時才會將整個 Transaction 的交易 Apply 到目的端資料庫，因此造成不少延遲；有些還甚至只能等到線上的 Redo Log 成爲 Archived Log 後才能複製到目的端，即時性更加的不足。報表系統的查詢如果是建立在這樣的複製工具所產生的資料庫上，相信會嚴重的影響到報表的完整性、以及決策的正確性。

最後還有一個很重要的問題，那就是是否支援異質資料庫間彼此複製的問題。截至目前為止，市面上常用的 Replication 工具其實絕大部分都無法支援異質資料庫之間彼此的複製，當然，ETL 工具不在此限，但企業內的資料庫卻很可能包含多種異質資料庫，例如 Oracle、Sybase 之類，所以 Replication 工具如果能夠提供這樣的異質整合功能的話，相信對企業絕對是非常有價值的。

GoldenGate 的優勢

從上面兩段對目前主要複製工具的介紹來總結的話，可以看出通常會影響報表查詢系統的重點因素大概有：資料庫的開放性、複製資料的篩選、對原資料庫的資源競爭、資料即時性、以及異質資料庫的整合性等等。有些工具可以解決前幾個問題，有些工具可以排除後幾個問題，但都不像 GoldenGate，根本就沒有這些惱人的 Issue。這其中當然牽涉到很多 GoldenGate 運作的架構與原理，本文不擬在此深入討論。讀者若有興趣，可自行透過 Oracle 網站對該產品的介紹或相關技術研討會了解其細節。

前些日子眾所矚目的 iPhone 4 終於上市了，台灣聽說已經有人花了四萬多元買了水貨。相信購買者本人絕不認爲四萬多元很貴，因爲該款手機所提供的像是照相、行動上網、視訊等各項功能都能符合他的需求；但對某些不用多媒體功能、不需要行動上網、或不需要照相的人來說，花四萬多塊錢實在是沒有必要。所以到底那隻手機好、還是不好，並沒有絕對的答案，全看它能不能符合使用者的需要。買錯手機最多浪費幾萬塊，若因爲事前的評估不當而錯買了

July 10 M-Power eNew

本篇文章版權爲倍力資訊股份有限公司所有，未經書面同意，嚴禁複製、轉載

一個無法符合企業需求的 Replication 工具，所浪費的資源可不是一隻小小的手機可比的了。能不慎重嗎？