

企業 ODS 資料運用

陳靖國*
朝陽科技大學資訊管理系
jkchen@cyut.edu.tw

李維哲
朝陽科技大學資訊管理系
s10154601@cyut.edu.tw

摘要

為了解營運狀況與決策分析，企業主利用 ERP 的資料來產生各式各樣的報表。若干資訊系統例如資料倉儲系統、線上分析處理系統等，都能夠滿足管理者對資訊取得的需求。但是建置這些系統無論成本或是時間都相當多，對中小企業而言是緩不濟急又負擔不起。本研究利用資料複寫(Data replica)和觸發程序(Trigger)，建置一個可營運的資料儲藏所(Operational Data Store, ODS)，很快就可以讓企業管理者快速取得所需要的資訊。

關鍵詞：ODS、資料複寫、觸發程序

Abstract

In order to understand the operating conditions and decision analysis, business owners take advantage of ERP data to generate a variety of reports. Some information systems such as data storage systems, online analytical processing systems, are able to meet the demand for information made by the manager. But build these systems, whether or time costs are considerable, for SMEs is being slow and unaffordable. In this study, data replication and trigger procedures, data storage by (Operational Data Store, ODS) build an operations quickly so that business managers can quickly obtain the required information.

Keywords: ODS, Data replica, Trigger

1. 前言

企業資訊的取得有三個方式。第一是從 ERP 系統上取得，大都是屬於 OLTP 的內容，包括進出貨統計、出貨單據、傳票、生產資料或是分析性的報表。但這類型報表的生成是利用最底層的交易資料表產生，並不適合執行長時間的交易分析。第二是從 OLTP 的交易性資料中取得資訊，經過彙總整理之後產生的彙總資料表，可以再進一步做分析而得到有用的資訊。大都經過一段期間的整理，才能產生各類型的彙總資料表。這取決於企業對這些資料表的需求迫切性，彙總資料表產生後，企業的資訊人員會依據管理者的需求撰寫不同的報表，使用這些資料表進行分析。但是這些報表仍屬於傳統類別的分析報表，以銷售分析為主，對於高階管理者在決策分析上雖有幫助，但終究屬於傳統條列式分析報表。

第三是藉由資料倉儲裝置來產生資訊。先將各種不同型態的來源資料經過提取、轉換、和匯入步驟，存放至資料倉儲伺服器，再用商業智慧軟體產生報表或查詢結果。此類型資訊的查詢與報表的功能可以同時存在，是一個相當方便的解決方案，但是需要相當多的時間及高昂成本。如果建置一個可營運的資料儲藏所(Operational Data Store, ODS)，將資料彙總，一樣可滿足管理者需求，但時間少成本低。所以本研究提出以資料複寫和觸發程序這兩種技術[7][8]，建立一個提供即時資訊的 ODS，以提供管理者做到報表控管的需求。

2. 相關文獻

2.1 資料複寫

資料複寫有兩個優點。

- (a) 改善並提高應用系統的延展性和可用性，一方面把複寫資料當備份，另一方面將複寫資料利用轉換程序來產生其他的資料表加以利用。
- (b) 將不同類型資料，透過資料複寫，把資料存放在同一個資料庫中。

以下說明資料複寫概念，資料複寫的架構中有三種角色，即發行者(Publisher)、散發者(Distributor)、和訂閱者(Subscriber)，如圖 1 所示。這三種角色的功能說明如下[1]。

- (1) **Publisher**：資料複本的發行者，是資料來源者，簡單說就是出版者的角色。其功能是維護來源資料以及確保發行集的資料可以使用。
- (2) **Distributor**：負責資料發佈、傳送到各個訂閱者的伺服器上。其功能為散佈系統資料、儲存散發歷程紀錄的資料。
- (3) **Subscriber**：從發行者那裡取得所需資料。訂閱方式有發行訂閱及提取訂閱兩種。前者就是由發行者主動傳送發行資料給訂閱者。後者則是依提取者的時程要求，來提取發行資料。

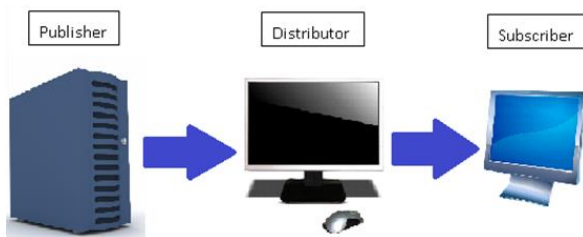


圖 1、複寫服務的架構圖

複寫資料內容有兩種：

- (a) 發行項：發行者所發行的資料內容，包括特定資料表的資料，檢視表、觸發程序。
- (b) 發行集：一個以上的發行項所構成的集合。發行集是複寫的基本單位，也是訂閱的基本單位。

圖 2 描述複寫拓樸架構，說明資料伺服器與資料副本之間的關聯性，伺服器之間資料流動方式的邏輯。還有數項複寫處理(代理程式)負責發行者與訂閱者之間的資料移動與複製。

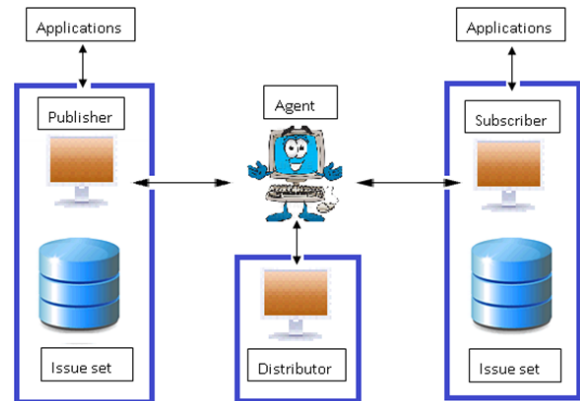


圖 2、複寫拓樸

再來說明如何在伺服器上設定複寫功能及發行資料。然後設定散發者，並建立、啟用訂閱者。最後將資料同步，即可依照所設定的方式進行資料複寫的工作。在建立複寫之前必須在資料庫資料夾底下，建立名為複寫資料的資料夾並設定使用者權限及安全性設定。然後按照下列步驟設定複寫相關資訊。

- (1) 設定散發者：主要是設定散發者的機器、並設定快照集資料夾，並設定散發資料庫的名稱、檔案的資料夾。
- (2) 建立發行集：可以透過新增發行集精靈，選擇欲發行之資料庫並選擇發行集的模式(例如交易式發行集)，同時設定快照集代理程式排程，然後給予發行集名稱即可。
- (3) 設定資料集的訂閱工作：透過新增訂閱精靈，選擇欲訂閱之發行集名稱，並選擇複寫傳送方式的基礎，例如發送訂閱或提取訂閱。在此要設定訂閱者及訂閱的資料庫，並設定合併式代理程式安全性及將訂閱的排程同步即可完成。
- (4) 驗證：資料複寫經過上述步驟後，最重要的是驗證上述步驟是否正確。透過資料庫複寫下的選項「啟動複寫監視器」

驗證步驟是否正確。

在[6]提到複寫可分為兩種類型，一是從一個資料伺服器複寫資料到另一個資料伺服器，二是在資料伺服器和用戶端之間做資料的複寫。前者是用來支援下列五點的需求：

- (1) 改善資料延展性或可用性
- (2) 建立資料倉儲和產生報表
- (3) 整合多個站台資料
- (4) 整合異質性資料庫的資料
- (5) 卸載批次處理等。

後者則是應用在下列三點的需求：

- (1) 企業與行動使用者交換資料
- (2) 消費者零售店面 POS 應用程式的資料複寫模式
- (3) 整合多個站台資料。

同時建議使用者在選定複寫模型時，必須要考慮資料發行者、散發者、及訂閱者資料庫的規劃。複寫模型不外乎是點對點、中央發行者、中央發行者與遠端散發者、中央訂閱者及發行訂閱者等幾種關係。

在[3]中探討如何將量販店分支機構銷售資料同步到總公司。即使網路中斷，各分店亦可正常作業，並於網路恢復時，再將資料傳回總公司。執行複寫之前應仔細規劃以提高資料的一致性，降低網路頻寬需求以避免日後處理異常狀況的工作。設計複寫機制時要注意下列各點。

- (1) 複寫資料是否需要更新或由誰來更新。
- (2) 關於一致性、自主權、與延遲時間的資料散發需求。
- (3) 複寫環境包含商務使用者、技術基礎建設、網路與安全性的特性。
- (4) 複寫方式與複寫選項。
- (5) 複寫拓撲及其如何搭配何種複寫型別的方式。

同時作者也認為複寫機制可以支援日益增

加分散性、移動性的可攜式營運方式，且可共享跨異質性平台及資料庫的資訊。

在[4]中提到資料複寫技術是一種高效率、安全又簡單的資料共享方法。對於分散式的異質資料庫系統，提供資訊整合的方法。原始資料修改後，透過及時複寫手法可以使副本資料和正本一致。該研究強調有效率的資訊同步分享平台，是全球運籌管理重要的成功因素。本文認為這也是企業整合資源必需考量的重要因素，而且本研究將朝向多個站台資料整合的實作。

許、雷兩人在論文[2][3][4][5]中提到資料複寫機制的應用，值得企業在資料管理時參考，其重點描述如下。以商品的配銷管理為案例，提出運籌管理的各項服務。以資料庫複寫技術為核心達成資料共享，並連接企業營運總部與經銷商之間的資料。資料複寫可以協助分散式交易得到以下幾點好處。

- (1) 允許多個站台保持相同資料的副本。
- (2) 使用者可以在網路離線時處理資料副本，然後在網路連線後，將異動資料傳送出去。
- (3) 使用複寫資料作為備份，在伺服器故障時提供複本資料。
- (4) 各銷售據點可選擇不同的複寫方式、不同的同步頻率、不同的網路架構，使連線成本降到最低。

2.2 觸發程序

觸發程序[15]是當包含有觸發程序的資料表中發生異動的指令時，就啟動在這個資料表中的預存程序，亦即針對特定的資料表在發生異動的指令動作時，指定固定執行一連串的程序。在[16]一書中提到，觸發程序有下列幾項的優點：

- (1) 利用 Trigger 做資料表的異動紀錄。
- (2) 可以維持各種整合限制條件。
- (3) 一般使用者感覺不到指令程序被執行，透過此可以簡化使用者的某部分的工作。

(4)在分散式資料庫上，由於資料可能被重複儲存在不同的資料庫站台上，可以透過觸發程序以維持個站台的資料一致性。

(5)觸發程序可以用來加總欄位值得總數，或是可以透過其計算出某一個欄位之值，使之可以即時且正確。

(6)可以簡化管理的程序，例如增加一筆出貨單內容時，此時透過觸發程序可以對該出貨單產品的存量，執行扣除的動作指令。

2.3 商業智慧

1989 年首次出現商業智慧(Business Intelligence)這個名詞[9]，Moss[17]認為商業智慧不是一個產品，也不是一個系統，是個蒐集並整合交易性資料，然後透過決策支援應用工具，讓企業容易存取商業資訊的架構。根據 1996 年 Greene 為商業智慧所下的定義：「商業智慧是將與企業運作有關的重要性資訊蒐集整理，用來管理目前和將來的商業運作環境的一個過程」[14]。在[11]中提到知識管理必須藉由商業智慧及協同技術才可以獲得效益，並提出知識管理的組成架構為營運系統，資料倉儲，商業智慧應用部分如圖 3 所示，並指出商業智慧是一個複雜的階段而且不是一蹴可成的。

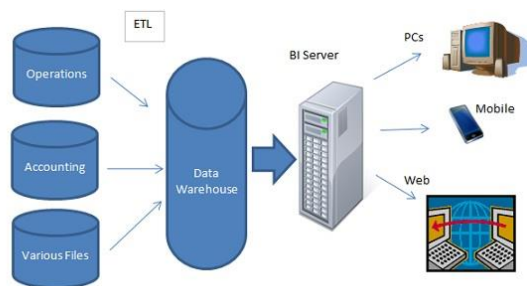


圖 3. 商業智慧框架圖。

針對商業智慧的實做，本研究是採用類似於圖 4 [18]的方式。資料來源端可能有 Excel、ORACLE、或是 ERP 資料，經過 ETL (SSIS) 建立資料倉儲，再進一步透過 SSAS 的功能，建立商業智慧的 OLAP，最後使用不同的前端工具軟體來展現資料，提供管理經營者需求的分析資料。

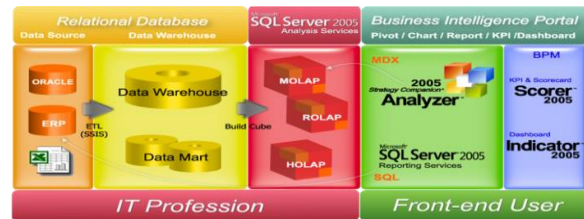


圖 4. MS SQL 商業智慧架構圖。

在[10]中提到經營管理者的決策一般可分為長期決策與短期決策兩種。資料倉儲因為有儲存大量歷史資料與彙總資料，適合提供長期決策用的分析資料。相反地短期決策的需求則需要近期的、最新的、詳細的交易資料，所以 ODS 的概念與架構，比較適合支援短期決策的需求。ODS 的基本概念與資料倉儲相似，以主題導向、整合的、揮發性、現階段、和細部資料為主要儲存對象，這些資料經過進一步的彙整即可載入資料倉儲。ODS 與 DW 有三點不同。

1. 作業資料儲存庫中的資料具有揮發性，亦即資料是可以被異動的，可以執行新增、修改、刪除等操作。
2. 作業資料儲存庫所儲存的為最新的資料，任何應用程式觸發的資料異動都會更新至作業資料儲存庫，其更新的頻率甚至可以是每次交易完成即更新一次。
3. 作業資料儲存庫中的資料是詳細的交易資料，而不是彙總的資料，假如作業資料儲存庫中含有彙總資料，則必須是動態彙總資料，亦即每完成一筆交易，必須重新計算彙總資料。

由上述可以知道 ODS 是儲存到目前為止最新的資料，以支援管理者進行不同層次的決策分析，ODS 相較於 DW 是較小型的架構，但其基本導入概念是相同的，仍需整合不同的資料來源，執行 ETL 的程序將資料導入 ODS。

在[11]中提到 ODS 具有部分 DW 的主題導向及整合特性，但是 ODS 仍舊可以在作業環境中維持它的即時性，也就是資料可以隨時被異動。對企業而言，應該只有一個集中的 ODS 來達成該企業所需要的資料整合。該文也指出 ODS 是屬於第三正規化的型態資料庫，並帶出 Inmon 對 ODS 所

下的定義：「具有主題導向的、整合性的、時間變動性的、即時性的資料儲存庫」[19]。ODS 資料變動的頻率可分為三級。第一級是交易檔產生交易時，及時的異動資料。第二級是資料每隔 2~4 小時會被匯入 ODS 資料庫。第三級是以日(24 小時)為期間將資料彙總轉入 ODS 資料庫。ODS 資料也可將已經整理好的資料，轉入企業的資料倉儲資料庫中做更進一步的應用。

3. 研究方法

處理方法分為三個階段描述如下。第一階段：從 ERP 資料庫將資料複製至 BI 資料庫中。亦即當銷售資料發生異動後，立刻透過觸發程序產生 ERP 銷售分析表。同時針對銷售客戶產品等資料做 ETL 程序，將銷售資料表變成單一筆紀錄的資料表。以 ERP 資料庫主機為發行者及散發者，對 BI 資料庫進行散發資料的動作，而 BI 資料庫也將被設定為訂閱者的角色，並以發行訂閱作為此次的訂閱模式，資料複製的模式設定為交易式複製。同理也將客戶及產品資料發生異動時，將資料複製至 BI 資料庫主機之客戶及產品基本資料。

第二階段：規劃彈性報表的設定功能。報表的設定及管理最主要重點是在設定報表代號種類、傳入的參數條件、報表列印時的模板、報表所需要的另類參數、報表產生模板時需要傳入的標題欄位說明設定、設定該報表下拉時的報表代號及所需要傳遞的參數、和最後該報表可使用的人員代號。

第三階段：設計報表查詢資料模組功能。資料查詢及設計功能說明步驟如下：

1. 選擇報表大類。
2. 選擇該報表大類之可用之報表代號。
3. 依據設定報表查詢參數，輸入查詢範圍。
4. 依據報表代號之 SQL 語法，配合顯示第一層報表，並在此功能上設計往下展開及往上合併，若不選擇展開角度時，以往下展開之第一筆資料代替。
5. 最後設計列印按鈕選項，將此資料依設定之可選擇模板，彈出選擇可以列印之

報表格式。

4. 實做與測試

為驗證所提方法可行，選擇一家中等規模的食品企業合作，導入方法實做與測試。這個企業總部位於台北，在台中、高雄、南投設有分支機構，所有的分支機構以網路連接，構成一個 VPN 網路。實驗結果顯示，以觸發程序做為資料 ETL 的轉換程式至第三正規化資料表，確定可行且結果正確無誤。搭配資料複製方法將資料複製至另一個資料庫主機，是一種簡單快速又安全的做法，且不需要特別撰寫程式，可以達到即時的轉換資料。不必透過以預存程序搭配資料庫排程功能，定時將資料移轉至另一個資料庫主機及時又方便。

5. 結論

本研究以觸發程序搭配資料複製機制，將資料移轉至另一個主機。利用資料庫本身的安全機制，不僅免去撰寫程式所帶來的不便，同時透過以預存程序搭配資料庫排程功能，定時將資料移轉至另一個主機所帶來的資料不即時性的困擾。需求的設備成本低廉，增設資料庫主機成本也不高，但卻可以建立廉價的 ODS，及時提供企業管理者所需要的資料數據。

致謝

這篇論文是科技部計畫(MOST103-2632-E-324-001-MY3)研究成果的一部份，在此感謝科技部經費支持這個計畫的研究。

參考文獻

- [1]王鴻儒，SQL Server 2005 資料庫設計建置管理實務，金禾出版社，2005。
- [2]許湧銘、雷欽龍、林鼎國，”資料庫複製支援協同商務與協同作業效能評估”，第一屆創新與管理學術研討會，實踐大學企業管理研究所，民國 93 年 12 月 17

- 日，頁 1-14，2004。
- [3] 許湧銘、雷欽隆、王文楷、黃蕙樺、林淑美，”利用資料探勘與複寫為量販店提昇銷售之績效”，第六屆產業資訊管理學術暨新興科技實務研討會，輔仁大學資訊管理學系，民國 94 年 11 月，2005。
- [4] 雷欽隆、許湧銘，”全球運籌管理資料分享使用資料庫複寫之研究”，2007 全球化暨國際企業研討會，靜宜大學國際企業學系，民國九十六年四月二十六日，2007。
- [5] 許湧銘、雷欽隆、謝曉雯、張智誠、吳水淇，”以商業智慧提昇協同商務績效”，2006 電子商務與數位生活研討會，國立台北大學主辦，民國 95 年 2 月 12-14 日。
- [6] 陳信固、王威然譯，Microsoft SQL SERVER 2005 最佳化和維護資料庫系統管理員解決方案，精誠資訊股份有限公司，2007。
- [7] 陳靖國、盧福雄，”利用資料複寫機制所建立的 CTI 與 ERP 之間整合的模式”，2011 年 4 月 AIT2011 研討會。
- [8] 陳靖國、盧福雄，”利用資料複寫機制及觸發程序整合企業外部系統”，2011 年 5 月 ICIM 研討會。
- [9] 王茁，「商業智慧」，初版，博碩文化股份有限公司，台北，2005。
- [10] 夏則智，”線上分析處理之彈性報表設計方法論”，國立中山大學資管研究所博士論文 93 年元月。
- [11] 林子平，”商業智慧元件化建置法之研究”育達商業技術學院資訊管理研究所碩士論文 94 年六月。
- [12] 曾勝麟，”商業智慧應用於客戶信用風險控管之研究-以 N 公司為例”大同大學資訊經營研究所碩士論文 98 年七月。
- [13] 黃達業、林容竹編譯，「金融機構管理精要」，Saunders, A. and Cornett, M.M. 著，美商麥格羅希爾國際股份有限公司台灣分公司，台北，民國九十三年七月。
- [14] 梁定澎，「決策資源系統與企業智慧」，初版，智勝文化事業有限公司，台北，民國九十五年一月。
- [15] 謝邦昌、鄭宇庭、蘇志雄，”SQL Server 2008R2 資料採礦與商業智慧”，基峯資訊，2011.01 初版。
- [16] 曾守正、周韻寰，”資料庫系統之理論與實務”，華泰文化事業股份有限公司，2007 02 第二版。
- [17] Moss, L.T., and Atre, S., Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications, Addison Wesley, Boston, MA, 2003
- [18] 美商睿智公司
http://web.analyzer.com.tw/product_overview.html。
- [19] Inmon, W.H. Building the Operational Data Source, 2nd ED., John Wiley, New York, 1999.
- [20] 維基百科全書
<https://zh.wikipedia.org/wiki/ETL>。