

資訊整合軟體解決方案
2003 年 3 月

DB2 Information Management Software



資訊整合

分散式存取與資料合併

作者 *Dr. Barry Devlin*
IBM Software

目錄

2	簡介
3	IBM 對整合資訊的願景
3	其他整合資訊的方法
5	分散式存取與資料合併
7	比較資料聯合與資料合併
10	成功地結合資料聯合與資料合併的方法
11	資料聯合與資料合併的工具
14	結合工具集以解決商業問題
14	改善客服中心的績效
15	擴展資料倉儲來源
15	支援分散式作業系統網路
16	結論

簡介

今天的企業對統一掌控資訊的需求日增，整個企業中已經儲存非常大量的資料，但是卻極為零碎和分散。為了有效地快速回應，企業使用者都必須能透通地使用資料，而無論資料儲存的位置及方式，且毋需擔心資料的時效性、一致性和安全性。他們需要實際有用的資訊。他們要求 IT 基礎結構能周延地處理如何將資料整合成有用的資訊。

今天，這種對資訊整合的需求並不容易滿足。一些必要功能並不存在。其他可以使用的技術仍在萌發當中。可用的功能以不同的技術類型分散於許多產品中。

企業通常使用擷取、轉換、載入 (ETL) 以及複製的傳統整合方法。這些資訊整合方法在本文中稱為資料放置 (placement) 或合併 (consolidation)，即在使用資訊前先於自己的電腦上予以合併來整合這些資訊。另外有一種新興的方法，稱為分散式存取或企業資訊整合 (EII)，讓使用者能夠在資料的原始位置直接進行存取。

本白皮書的目的就是提供一種高階指南，以涵蓋這些領域內存在的功能、提供的產品、使用的方式以及如何協同工作。IBM 針對資訊整合的整體願景中已清楚描述這些資訊，於本白皮書的第一節予以概述。

接著，本白皮書描述並對照兩種主要的資訊整合方法，即分散式存取以及資料放置或合併，詳述何時較適用何種方法，以及它們如何互相搭配。

最後，本白皮書說明目前的 EII、ETL 和複製等產品如何適用於資訊整合的架構內，介紹 IBM 及其合作夥伴在這些領域裏的產品，並對這些產品之間的關聯予以定位。一些範例說明何時需要混合使用 EII、ETL 以及複製工具來解決商業需求。

IBM 對資訊整合的願景

今天，除了那些最簡單的商務工作以外，其他的商業工作都需要使用企業多年來從各種資料來源所建立的資訊。這些資料來源或許是本機的、遠端的，或存在於公司內部網路、連外網路或網際網路上。資料可以採用任何格式加以儲存，關聯式或非關聯式資料庫、文字檔、非結構化內容儲存等。資料也可以是目前的或是某個時間點的副本。通常，使用者不僅需要對這些資料來源進行讀取，而且也需要寫入資料。

這個複雜，事實上也在不斷改變的資訊環境，對企業使用者、應用程式，以及必須維護和管理企業的 IT 人員來說，成為重大的挑戰。簡單來說，IBM 對於資訊整合的願景，就是要大幅減少甚至完全消除這些問題。

基本的資訊整合原則是：使用者應當能夠看到他們使用的所有資料，就如同這些資料是位於一個單一資料來源一樣。資訊整合技術掩蓋了從各種不同的位置、語法、格式及存取方法來存取資料時所衍生的複雜性。例如，使用一種標準的語言，如結構化查詢語言 (SQL) 或是透過 XQuery 的 XML，或是標準的 Web 服務或內容 API，資訊整合中介軟體即能讓使用者或應用程式依照自己所需透過地存取資訊，而不用在意實際的運作。

其他資訊整合的方法

提供一個整合式資訊的目標可以透過兩種方式實現，這兩種方式可分開或合併運用。資訊整合的主要方法是：

1. 透過資料存取或資料聯合 (federation) 提供分散式資料存取。
2. 將資料移動到一個更有效或更易於存取的位置 — 資料合併 (consolidation) 或放置 (placement)。

這些功能集結起來便構成了資訊整合所需要的核心部份。最簡單的說法是，資料聯合是在一個地點執行查詢，然後將其適當的部份發送出去並按照資料的所在位置及其格式來處理資料。分散式存取是屬於企業資訊整合 (EII) 的技術類別。而另一方面，資料放置則預先將位於不同位置的資料放置於同一個地方，如此一來，使用者查詢就不需要分散。該方法相當於擷取、轉換和載入 (ETL) 以及複製功能。

兩種方法都需要廣泛且大量的共同支援的功能。如果沒有可以確保資料完整性的對應及轉換功能，就不可能使用分散式存取或資料放置兩種方法。此外，根據企業的需求，相同的資料在某些情形下需要合併，而在某些情形下卻需要聯合。因此，為了維持企業資料的一致性，兩種情況都需要一套共同的轉換和對應功能。

對應功能讓使用者瞭解資料各個部份之間的關係，並能夠判定如虛擬表格中的欄位實際對應到不同資料來源的屬性。在一個更複雜的例子中，對應功能讓一個查詢或應用程式中的「客戶」與一個資料組合產生關聯，這些資料的一部份位於本機的 IBM DB2® Universal Database，另一些則位於遠端的 Oracle 資料庫，還有一些則位於 IBM Lotus® Notes® 資料庫中。

而轉換這種功能則是透過不同表示法之間的對應，實際轉換及結合相關的資料。例如，一個簡單的轉換能將數字轉換為字串，而更複雜的轉換則包含將階層式資料轉換為關聯式資料，甚至根據更為精密的商業規則來進行轉換。

快取功能提供一個暫時的資料儲存，利用透過地儲存結果集的本機副本來改善資料聯合的效能。從資料放置的觀點來看，它只是一個可能包含一些遠端資料來源的轉換副本，而且可能需要加以管理。因此，快取是另外一個介於兩種整合模式之間的關鍵鏈結，使得一種模式的優點可以支援另一種模式的缺點。

對應、轉換和快取功能有賴於它們操作環境的詳細描述。這類描述包括商業含義、關係、位置、技術格式等。簡言之，就是中間資料 (metadata)。這種中間資料必須是廣泛且一致的，並且從一個整合專案的探索和定義到聯合式查詢的運作，都必須是有用的。一組廣泛且邏輯一致的中間資料，無論是位於單一的實體儲存區或是分散於多個儲存區，都是任何資訊整合願景的根基。

分散式存取與資料合併

分散式存取和資料合併可以單獨使用，而且從這種觀點來看是最容易理解的。然而，這兩種方法都彼此相互提供關鍵的支援功能，因而能夠發揮完整的資訊整合功能。

*分散式存取或 EII 方法。*分散式存取的功能基礎是聯合查詢。在最簡單的情況下，想想一個需要從兩種不同的資料庫（例如 DB2 和 Oracle）聯結資料的應用程式。從應用程式的觀點來看，這是一個包含兩個資料表聯結的簡單查詢。然而在背景處理時，聯合元件瞭解到這些資料表中的其中一個位於本機的 DB2，而另一個則指向遠端的 Oracle 資料庫。顯然，系統必須建立兩個查詢，一個適用 DB2 資料表，而另一個則適用 Oracle 資料表。接著，必須決定最有效的方法來執行實際聯結，取決的因素包括個別查詢結果集的相對大小、兩套系統的處理能力、兩者之間的連接速度等等。

當然，完整的聯合查詢所需的功能遠比上面所述的複雜。絕大部份的資料並不是存放在關聯式資料庫內。這類資料可透過各種語言或應用程式加以存取，而它們（語言或應用程式）極少提供 SQL 中所提供的豐富功能。因此，資料聯合必須模擬所需的功能，參與查詢最佳化，將子查詢推至資料來源，協調不同的存取機制，並且在需要時管理安全性及交易完整性。資料聯合並不限於唯讀存取，許多業務工作也要求分散式寫入存取。以其最完整的形式而言，聯合查詢必須處理所有這些需求。

資料合併或 ETL/ 複製方法。資料合併或放置是整合資訊的傳統方法；不同於資料聯合，它是將資料移動到查詢。人們總認為它比資料聯合簡單，因為資料合併會建立資料的第二份本機副本，並視需要預先處理，因此減少大量資料操作以及在使用者查詢內進行遠端存取的需要。由於資料合併在使用者查詢的關鍵時間以外操作，所以也允許大量及複雜的資料轉換，進一步解決了資料的清除、語法及時間一致性等問題。因此也展現了不同層次的複雜性。在最簡單的情形下，它是以手動方式啟動資料庫卸載以及目標系統的載入作業。在最複雜的情形下，它可能牽涉到自動、即時、多向地同步化數個遠端系統上的資料庫。目前在大部分的情況下，它是介於兩者之間。

資料合併的一個關鍵考量是，從來源向目標傳送資料時所能容忍的最長延遲時間。企業通常需要指定資料副本的更新程度。比方說，對於資料倉儲，其更新頻率可以是每天或每週，而資料合併的延遲很容易延長為數個小時。而另一種極端情況則需要近乎即時的資料，例如股票市場交易系統，在資料合併上則要求最短的延遲時間。

在資料合併中，決定最小延遲時間的兩個最重要因素是：需要轉換的複雜度以及傳送的資料量。這些因素導致兩種合併資料的互補方法。ETL針對較大的資料量作了最佳化處理，且通常與更複雜的轉換有關；而資料複製則是強調個別資料記錄的傳送，且通常限制為較簡單的轉換。

比較資料聯合與資料合併

資料聯合及資料合併實際上是相當類似的概念。兩者都要請求和接收與使用者互動的資料庫實體範圍以外的資料。主要的不同在於從資料庫請求資料與將資料傳送到資料庫的時間。藉由資料聯合，兩者（請求和傳送資料）發生於使用者提出請求之後。對於資料合併，兩者（請求和傳送資料）都發生於使用者提出請求之前，事實上，定義傳送就會發生一次請求，而接著可能發生許多次傳送。

從使用者的觀點，或是就其使用的應用程式觀點來看，資料聯合與資料合併的作用是相反的。資料聯合直接從原始來源隨機整合必要的資訊，並且只有在使用者決定所需的資訊後才會運作。因此，它必須在使用者或請求的應用程式可以接受的時間範圍內傳回結果。資料合併在使用者查詢前先行操作，因而有更多的時間執行必要的處理。然而，它顯然必須事先瞭解需要哪些資料，而且，由於必須建立資料的第二份副本，因此比資料聯合需要更多的儲存空間。

有時候會建議採用資料聯合以提供使用者對任何資料進行直接的、非計劃性的存取，因為聯合式查詢允許資料隨時予以組合。這是很危險的迷思，因為這種不受限制的存取會為使用者和 IT 系統帶來重大的問題。對於使用者來說，如果他們試圖組合那些在意義和結構上不一致的資料，則資料聯合為了避免顯著的效能及語法問題，事實上比資料合併需要更為嚴謹的分析、建模、控制及規劃。像資料倉儲以及透過資料放置所構成的作業型資料儲存 (operational data store) 這類的合併資料儲存體便能夠避免此類問題，因此將繼續扮演重要的角色。

只要將這些納入考量，就很容易知道資料聯合或資料合併何時才是整合資訊的適當方法。

聯合或分散式查詢在以下的情況下為適當的方法：

- 必須即時或近乎即時存取快速變動的資料。
為快速變動的資料建立副本¹是相當昂貴的，而且在此過程中總會有些延遲。而透過資料聯合，可在查詢中直接存取原始資料和進行聯結。但是，必須考慮存取原始資料時所涉及的效能、安全性、可用性及隱私等方面的問題。
- 需要立即、直接寫入原始資料。
當需要插入或更新資料時通常不建議使用資料副本，因為在原始資料及其副本之間可能會產生資料完整性的問題。即使能夠使用雙向資料合併工具，也需要複雜的兩段式鎖定方案。不過，通常會禁止直接寫入另一個應用程式的資料庫。因此，一般會建議透過聯合式查詢經由 API 呼叫該應用程式來提出任何的更新請求。
- 在技術上很難使用來源資料的副本。
當使用者需要存取多種異質資料和內容時，也許很難將所有結構化和非結構化的資料放在單一的本機副本中。或者，當來源資料具有非常特殊的結構，或是與其他資料來源相依時，可能就無法明確地查詢資料的本機副本。在這種情況下，建議存取原始資料。

¹「副本」這個術語之用法在此處及以下的討論中並不嚴謹。它也許不是暗指實際的副本，為了易用性、效能或其他因素的考量，在資料合併的過程中，資料可能已經轉換。

- 複製資料的成本超過遠端存取資料的成本。
查詢遠端資料集的相關效能影響和網路成本必須與儲存資料的多份副本所需的網路、儲存及維護成本相比較。在某些情況下會明顯傾向選用資料聯合的方法，例如：
 - 當原始資料的資料量過大而不應該進行複製時
 - 當資料不常使用而不應該進行複製時
 - 使用的資料量太少或其比例無法預期時
 - 從許多遠端或分散的地點存取資料（意味具有多份副本）時。
- 製作來源資料的副本是非法或禁止的。
基於安全性、隱私及授權的限制，建立由另一個組織所控制的來源資料的本機副本，或是建立位於網際網路上來源資料的本機副本也許是不切實際的，但是允許以個別查詢進行聯合式存取。
- 事先不知道使用者的需要。
允許使用者直接立即地存取所需的資料顯然是有利於資料聯合的論點。不過，這可能是一個危險的策略，因為使用者可能會建立一個「莫明其妙的查詢」，而對來源系統及網路效能造成負面的影響，並降低回應的速度。除此之外，由於組織內部目前關於資料儲存的語法不一致，因此這類的查詢可能傳回完全沒有意義的結果！

在許多方面，資料合併的論點正好與資料聯合的論點相反，但也有一些其他特定的涵義：

- 必須唯讀存取相當穩定的資料時。
如果使用者不想看到發生於來源資料上的每一次變更，則定期建立資料來源的副本可以隱藏資訊不斷進行的變化。
- 使用者需要歷史資料或趨勢資料時。
歷史資料和趨勢資料極少從作業型資料來源取得，但是可以透過資料合併的過程經過一段時間的累積來建立。這是一個非常普遍的資料倉儲需求。

- **資料存取效能或可用性凌駕所有需求。**
使用者總是希望快速地存取資料，因而有必要製作一個本機的、預先處理的資料副本。正如在資料倉儲環境中所見，這些查詢可能很複雜，或者需要一個對歷史性或趨勢資料的多維檢視。因此，資料合併是資料倉儲的基本技術。不過，正如 IBM 白皮書《Information integration-Extending the data warehouse (資訊整合－資料倉儲的延伸)》中所述，對於諸如存取即時資料等特定需求而言，可以有效地利用聯合式查詢。
- **使用者的需求是重複並且可以事先預測的。**
當使用者的查詢是明確定義、重複且僅存取已知的資料來源子集時，建立本機副本以利本地的存取和使用便是合理的做法。特別是當一些特定使用者需要檢視的資料與來源中的儲存資料明顯不同時尤然。在這種情況下，資料合併能夠建立一份特別的資料副本，該副本能夠妥善執行，並讓使用者更容易瞭解。
- **轉換或所需的聯結很複雜或耗時。**
如果聯結或轉換作業很複雜或很耗時，那麼基於可能造成效能低落與高昂成本的情況下，並不建議將聯結或轉換當作使用者查詢的一部份來執行。在這種情況下，透過資料合併來建立資料副本便較為合理。

成功地結合資料聯合與資料合併的方法

有些時候，結合資料聯合與資料合併的優點是合理的作法。

第一種情形是：聯合式查詢能夠潛在地利用資料合併功能。即便如前所述的最簡單的情況下，有時候聯合查詢就是無法運作。網路效能或可用性等問題可能會阻止查詢的執行。就策略性和技術的角度來看，允許應用程式直接存取遠端資料是不切實際的作法。或者說，那些基礎資料的結構可能會將這類查詢的效能減緩到無法接受的程度。

在這種情況下，為了支援這種需求，資料聯合可以使用資料合併來建立或管理快取資料。今天，系統設計師能事先明確地指定或定義這些資料快取。在將來，資訊整合技術能夠在必要時自動設定它們。顯然，這種方法在某種程度上限制了聯合查詢，因為快取的資料不完全是最新的，而且，如果快取不完全支援雙向同步的話，或許寫入存取也會受到限制。不過，從其他方面來看，在資料聯合底層使用資料合併確實擴展了資料聯合所能夠提供的解決方案。

反過來說，資料聯合又能對資料合併提供何種好處呢？任何資料合併工具面臨的挑戰之一是其必須交換資料的來源及目標之數量和種類。因此，一般來說，資料合併工具通常會針對可用來源和目標的子集加以最佳化。將聯合功能用來支援資料合併可擴展這些來源和目標集，同時能夠預先聯結多個來源的資料。然而必須犧牲效能，才能在交換過程中增加可存取之來源或目標的數量，而此方法僅適用於有限的資料數量。

資料聯合與資料合併的工具

IBM 及其事業夥伴提供了包含資料聯合及合併功能的多項工具。然而，您應該瞭解沒有任何單一產品能夠提供上述的所有功能。一些產品是針對某種類型的功能，而其他產品則專精於某些領域，而且有些功能類型仍在開發當中。

EII

EII 工具提供對遠端資料的分散式存取，並對應到先前所述的分散式存取或聯合查詢功能。這類工具已變得日益重要，因為分散式電子商務由於唯讀及讀寫的目的，需要廣泛且立即存取分散的資料。

IBM DB2 Information Integrator 及其前身產品 IBM DB2 DataJoiner 和 IBM DB2® Relational Connect 都提供 EII 功能。DB2 Information Integrator 允許定義及查詢多種不同且分散的資料庫及檔案格式的整合性檢視，它們包括：

- DB2、Informix、Oracle、Sybase、Microsoft SQL Server 和 Teradata 資料庫以及 ODBC 資料源。
- XML、Microsoft Excel、OLE DB 和文件檔案。
- Web 服務、訊息佇列和應用程式資料來源。
- IBM Lotus® Extended Search 來源（內容儲存庫、LDAP 目錄、Web、電子郵件資料庫、企業內容等）。

在 EII 中，快取功能是一項非常重要的先決條件。快取通常表示延遲時間很小或完全沒有延遲的自動化資料傳送，在來源與目標快取之間僅有最小的轉換。然而，從長遠的觀點來看，這個概念要求雙向複製資料的能力，也需要解決介於來源資料庫和目標快取之間的衝突。Materialized Query Table (MQT) 便是目前已經可用的一種快取基本形式，它儲存查詢的結果以備日後重複使用。DB2 Information Integrator 可以使用 MQT，但是這些快取的設定和管理仍然需要手動處理。

DB2 Information Integrator for Content 及其先前的版本 IBM Enterprise Information Portal 也提供 EII 功能，但是更傾向內容整合，並採納 IBM Content Manager 的程式設計模式。它可以搜尋及存取的資料來源包括 Content Manager 產品及其它內容儲存體、Lotus 資料庫、關聯式資料庫，以及可以用 Lotus Extended Search 搜尋到的、範圍廣泛的內容。它還包括一個完備的資訊採礦功能，該功能使用網絡搜尋 (web crawling) 和文字採礦演算法將非結構化內容予以組織結構。而且，DB2 Information Integrator for Content 還提供進階工作流程應用程式，幫助企業提高生產力、縮短生產時間以及改進內部的溝通與協同作業。

擷取、轉換和載入 (ETL)

ETL 工具的特色在於有效地轉換大量資料的能力，通常在來源和目標之間進行精密的轉換。這些工具一般用來載入資料倉儲和資料超市 (data mart)，這種過程通常對時間延遲沒有很高的要求，然而，為了讓資料適用於一般使用者，因而需要複雜的資料轉換。

DB2 Warehouse Manager 主要是在 DB2 系列產品內提供 ETL 功能。它使用所有的 SQL 功能、Web 服務，還加上一組豐富、預先內建的轉換功能，提供在來源與目標之間進行轉換，並採用一個代理程式方法，以便在各個節點上分派處理程序。藉由鏈結到 DB2 Information Integrator 並作為其中一個資料來源，IBM DB2 Warehouse Manager 亦提供一個能夠潛在地使用資料聯合功能的 ETL 工具的例子。

IBM 的事業夥伴 Ascential Software 提供的 Ascential DataStage 產品，能夠支援來自各種異質資料來源的大量 ETL。其中有些產品版本能夠在 IBM 390®、Microsoft® Windows® NT 和 Unix® 平台上執行。DataStage 對資料轉換提供廣泛的支援，同時，其最新的增益功能也提供資料品質和資料清除方面的支援。

兩種產品傳統上都專注於資料倉儲市場，但是，資料倉儲正逐漸朝向更即時且分散式的資料需求。因此，這些工具可以視為資訊整合更大範圍的一部份，提供了廣泛的資料合併功能。尤其是在任何大規模的資訊整合解決方案中，這些廣泛的轉換和大量資料處理能力是極為關鍵的。

複製

透過捕捉來源資料庫的異動，並將這些異動傳送及套用到目標系統，複製工具的目的在於提供從來源到目標的低延遲傳送。由於專注於減低時間延遲，轉換功能常常在複雜性方面有所限制。複製常用於填入作業型資料儲存 (ODS)，或定期在一個中央作業型資料庫及其多個遠端子集之間散佈資料，該資料庫能夠獨立地在資料更新之間進行操作。

IBM DB2 DataPropagator™ 在 DB2 系列產品中提供複製功能，而 DB2 Information Integrator 則在 DB2、Informix、Oracle、Sybase、Microsoft SQL Server 等異質的關聯式資料庫之間提供複製功能。利用整套的 SQL 功能、Web 服務和預儲程序，DB2 DataPropagator 和 DB2 Information Integrator 兩者即可支援資料轉換。

目前越來越多的 ETL 工具廠商開始提供複製功能，通常是在「異動的資料捕捉」(changed data capture) 條件下提供。以 ETL 產品為例，複製工具亦能提供資訊整合的部分資料合併功能。在這裏所要強調的是解決立即或者近乎即時的資料傳送需求的能力，而非複雜的轉換或大量資料處理的能力。

結合工具集以解決商業問題

探索資訊整合的概念性功能並介紹現今的資料聯合及資料合併的工具後，可以開始看一些例子，這些例子說明這些工具是如何搭配使用來解決商業難題。

改善客服中心的績效

在 IBM 白皮書《Information integration-Extending the data warehouse (資訊整合－資料倉儲的延伸)》中，討論了在客服中心環境中可以取得的優勢，就是允許查詢跨越提供歷史資訊的資料倉儲，以及紀錄目前狀態的作業系統，來進行資料聯合。在本例中，當需要時，資料聯合允許從作業環境中存取最新的帳戶資訊。然而，在這種環境中，便難以保證查詢的效能。

經過一段時間後，便可明顯看出從作業系統請求某些資訊項目時具有一定的規律性。例如，一些大客戶總會定期查詢付款狀態。而這些資料並不值得以近乎即時的方式儲存於資料倉儲中，建立一個每隔幾分鐘便要更新的 ODS 便能快速地存取這些資料。DB2 DataPropagator 可以用基於 DB2 的作業系統的資料來定期更新 ODS。同時，DB2 Information Integrator 的複製功能能夠從其他的關聯式作業系統中執行同樣的工作，這些作業系統也就成為聯合查詢的另一種資料來源。

擴展資料倉儲來源

DB2 Warehouse Manager 通常在完備的 DB2 資料倉儲環境中提供 ETL 功能。DB2 Information Integrator 則擴展其規模，以處理從非 IBM 關聯式資料庫的原生資訊擷取。

支援作業系統的分散式網路

一個常見於零售和其他分散式環境中的系統模式，包含了一個中央主作業系統和一些於遠端位置執行的間歇性連接的衛星系統。在這種作業模式中，中央系統在某些時間向衛星系統傳送更新資料，隨後，衛星系統將更新的資料傳回到中央資料庫。例如，在零售環境中，中央系統在清晨所有商店開門之前，便將價格及促銷更新資料送至所有的分店，而在當天結束營業以後，各分店將銷售資訊送回中央系統。這種模式的特性是，雖然更新是雙向的，但是由於資料流是在不同的時間朝不同的方向傳送，而且更新作業所影響的是資料庫的不同部份，因此，不可能發生衝突。

假設現在中央資料庫執行 DB2 而衛星系統執行的是 Oracle。DB2 Information Integrator 捕捉在中央 DB2 上的價格和促銷資料的異動，並將這些異動套用到 Oracle 資料庫的價格和促銷表格上。POS系統使用這些基本表格開始一天的營運。當營業結束時，所有的銷售資料透過各種方式匯總，而所有這些異動會由 DB2 Information Integrator 的複製功能加以捕捉，並套用到中央系統。

在這種情況下，複製與聯合功能共同運作以使這個分散式環境同步化。除此之外，也能利用 DB2 Information Integrator來幫助業務分析人員追蹤產品銷售。分析人員僅能存取中央系統上的彙總資訊，但是如果他們需要深入研究細節資訊時，DB2 Information Integrator 能夠將他們的查詢順暢地傳送到放置詳細資訊的 Oracle 資料庫。

結論

資訊整合的目標：在明日的世界裏，企業使用者將不需要知道、也不必在乎資料的格式或其位置。相反地，他們所需要的資訊將在他們的掌握之中，是一種能夠立即理解和使用形式。系統和儲存體、內容和格式、一致性與效能等的複雜問題將全部交由資訊整合中介軟體來管理。這不但使企業使用者與基層的複雜系統完全隔離開來，它還會簡化甚至消除一部份的複雜性。它將會自動管理這種環境，讓 IT 脫離過度的繁瑣工作。

這種突破不可能一夜之間實現。不過，目前的資訊整合工具確實提供了一個良好的開端。DB2 Information Integrator 的聯合功能為企業使用者或請求的應用程式建立了一種單一、整合的資訊儲存外觀。簡言之，它將原始的位置和格式隱藏在標準查詢介面之後。

但是我們都知道，我們不可能也不希望一直將資料保持原來的格式或保存在原來的位置。可用性、簡潔性及效能，所有這些都要求在某些時候，資料必須要從其原始形態加以移動或複製、進行鏡射或轉換，無論是臨時的還是永久的。今天的 ETL 和複製工具都提供了這種功能的基礎。

DB2 Warehouse Manager 和 Ascential DataStage 之類的 ETL 產品強調在複雜的轉換中進行單向的大量資料傳送。如 DB2 DataPropagator 和 DB2 Information Integrator 這類複製產品支援資料庫異動或是具有些微轉換之交易的有效傳送。資料庫快取技術，例如在 DB2 Universal Database 和 DB2 Information Integrator 中的 Materialized Query Table (MQT)，便從另一方面展現了需要建立遠端資料的本機副本之需求。

正如我們先前所說的，無論單獨使用或者共同使用，現在的 EII、ETL 以及複製工具都可以有效地運用於各種資訊整合方案。每種工具都提供了有用且必要的功能，該功能是多數資訊整合實作的一部份。

單憑任何一種工具是無法完全實現資訊整合的願景，但是，每一種工具都是一個重要的環節。當我們擴展並組合這些工具時，我們終究能落實 IBM 所謂的真正整合式資訊環境的願景。

詳細資訊

請聯絡 IBM 業務代表或「IBM 事業夥伴」；或請造訪我們的網站，網址為：
ibm.com/software/data/integration



台灣國際商業機器股份有限公司

台北市敦化南路一段 2 號 12 樓
市場行銷處：0800-016-888 按 1
技術諮詢熱線：0800-000-700

© Copyright IBM Corporation 2003

2003 年 3 月，保留所有權利

DB2、DB2 Universal Database、DataPropagator、DataJoiner、IBM、IBM 標誌、Lotus、Lotus Notes 和 OS/390 是 International Business Machines Corporation 在美國及其他國家或兩者兼有的的商標或註冊商標。

Microsoft 和 Windows 是 Microsoft Corporation 在美國或其他國家（或兩者）的註冊商標。

UNIX 是 Open Group 在美國以及其他國家的註冊商標。

其他公司產品或服務名稱可能是其他所有者的商標或服務標誌。

本出版品中所提及的 IBM 產品或服務，並不代表 IBM 會在設立分公司的國家提供其中所有引述的產品或服務。