

以 ETL 方式建構半導體業 B2B 系統之方法

張永昌 范金鳳

元智大學資訊工程系

chang.pepper@gmail.com

csfanc@saturn.yzu.edu.tw

摘要

半導體業上下游廠商的企業流程(Business processes)資訊包括訂單、工單、生產、入庫、出貨、請款對帳單等。目前多數此類企業流程使用 B2B 平台、點對點開發方式。但此作法面臨兩大挑戰：(1)對外導入時，各客戶有各自不同的 B2B 系統檔案交換規格與需求；(2)對內，因應客戶規格所需要的資料將橫跨各異質性環境系統。所以，供應鏈 B2B 垂直整合與廠內多元系統的水平整合愈顯重要、挑戰亦愈高。本研究針對此議題，建構一個全方位企業流程資訊整合系統的架構，設計了企業流程資訊通用之 E-R Model Data Mart 結構(schema)，運用 ETL 實作整合異質性系統，能有效地處理半導體企業供應鏈資訊交換與企業內系統整合的問題。初步實作結果顯示此方法可成功地增加資訊分享效益、並提升服務之深度及廣度。

關鍵詞：半導體產業，企業間電子商務，企業應用系統整合。

Abstract

The business processes among the supply chain in semiconductor companies include product order, work order, production, inventory, delivery, bill, etc. Currently most companies use the Peer to Peer B2B platform to perform information transfer and integration. However, this approach faces two major challenges: (1) Externally, different customers may have different B2B system format for exchange. (2) Internally, the data required by the customer may spread in different heterogeneous systems. Therefore, it is essential to integrate B2B supply chain vertically for external partners, and horizontally across internal systems.

This research develops an integrated framework for business process in semiconductor industry. It first designs generic E-R model Data Mart schema for business processes, and then uses an ETL tool to implement the integration of heterogeneous systems so as to effectively handle the supply-chain information exchange. The preliminary case study results show that this approach is successful in enhancing information sharing and IT (Information Technology) service.

Keywords: Semiconductor industry, B2B, enterprise application integration

1. 前言

半導體產業變化急遽、分工專業、競爭激烈，半導體工廠必須達成快、準、好、省的經營方針，因此企業需強化和客戶與供應商之間緊密的關係，以提升競爭力。企業利用網際網路進行與合作夥伴之間企業流程(Business processes)的資訊，如訂單、工單、生產、入庫、出貨、請款、對帳單等狀態資訊溝通，使參與的夥伴間皆能即時獲得相關訊息，以充分協同作業。

過去尚未有眾所接受的 B2B 整合標準，客戶往往自定標準，造成各廠商引用標準之紛雜。導入 B2B 時面臨內外的兩大挑戰：(1) 對外導入時因各家客戶有各自不同的 B2B 檔案規格與需求，將面臨一大挑戰；(2) 對內因應客戶規格所需要的資料橫跨各異質性環境系統，又是另一大挑戰。本研究針對此議題，發展半導體封裝廠企業流程之 B2B 系統導入的資料模組與架構技術。

過去半導體導入 B2B 系統是透過點對點方法，大部份只是單純的將客戶的資料匯進來與匯整資料給客戶，如此方式會造成 B2B 系統變成複雜的網狀架構，往往導致成本過高與整體效率過低。所以本研究發展了整合異質系統的共用資料模組與架構的技術，規劃設計整合資料模組的 Data Mart 概念，並運用開放原始碼的 ETL (Extract-Transform-Load) 工具來實作完成整合異質性系統的建構方法，可以彈性且較低成本，來達到系統的一致性、擴充性、及執行效率等好處。

2. 相關背景及研究

2.1 資料市集(Data Mart)

資料市集(Data Mart，以下縮寫為 DM)可視為資料倉儲 (Data Warehouse，縮寫為 DW)縮影或元件。資料市集的重點是只支援某些特定的部門需求或單一目的要求；面向某個特定的主題。若將企業組織內或多個部門的 DM 整合，可發展成完整企業體的資料倉儲 DW，以提供企業重要決策者完整且廣泛的訊息。

本研究利用到 DM 常用的星狀綱要(Star Schema)與雪花綱要(Snowflake Schema) [1]。星狀綱要(Star schema)，是一種簡單的資料庫設計，有一個主要物件在中央(Fact table)，包含著大量的資料但不重複，其輻射狀地連接其他物件，看似星狀，

故稱之為星狀綱要。

雪花綱要(Snowflake Schema)與星狀綱要類似，差別只在於雪花綱要將各維度資料表正規化(Normalized)，將重複的地方分別以獨立的表格儲存，因此形成多階的資料結構狀態，資料表增加；在查詢時，會因為需連結多個表而影響效率(如圖 1 所示)。

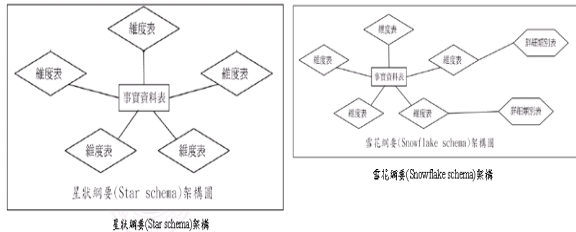


圖 1 DM 資料的結構

2.2 ETL(Extract-Transform-Load)

ETL 即資料數據抽取 (Extract)、轉換 (Transform)、裝載 (Load) 的程序。在整合系統架構建置過程中，資料整合轉換(ETL)是最花費時間與人力，也是導入中關鍵要素之一。本研究在 B2B 系統導入中，使用 ETL 工具使企業應用易於更改資料，ETL 程序方式可將不同的系統與來源資料檔案或資料庫所擷取的資料，依據企業之需求及 ER-Model 的設計加以轉換、整理、清除不需要的資料，將洗淨後的資料轉至目的地 DM 資料庫中，以保留在企業內以利後續使用。

IT(Information Technology)系統資料與 DM 資料除了在記錄不同產品成長階段所留下來歷史資料，還包括使用者所產生的大量資料，及對外部所取得的資料，這些資訊可能來自不同的資料平臺，或一些特定的檔案格式。隨著時間與系統增加，而累計資料越來越多，尤其是在 DM 的資料中心，如此大量資料將造成系統變慢。本研究提出的策略也可針對此部份也一併處理，透過整合的模組與 DM 架構，使用 ETL 的方式並依據訂定的規則來規劃資料保留與封存策略。

2.3 資料交換現況

早期電子資料交換 EDI(Electronic Data Interchange)[2]，透過雙方定義好的規定和格式交換，其後演變成透過網際網路利用 XML[3]格式來傳送資料，但面臨到各企業間各自有不同的商業字彙，有不同的交易規格和不同的 XML。軟體間轉換格式無統一標準，以致於有後續相關的國際電子商務標準，如 RosettaNet[4]與 ebXML[5]等，這些都是以 XML 為基礎來制定電子商務標準。下面簡述 RosettaNet 標準化 B2B 標準[6] 和 ebXML 國際標準[5]。

RosettaNet 協會針對全球資訊、電子元件、及半導體業者發展了產業供應鏈商務標準，旨在形成

共同的電子商務語言。RosettaNet B2B 標準具有上層商業處理與下層技術服務的整體架構。依其結構分為三塊：(1)詞典(Dictionary)：定義交易夥伴與產品描述字典的標準語言，將供應鏈中組織運作的術語標準化。(2)執行架構標準 (RosettaNet Implementation Framework, RNIF)：規範了傳輸、通信協定、安全、路由及訊息包裝的方式。(3)作業程序標準 (Partner Interface Processes, PIP)：將商品資訊之收集、維護，產品資訊加以分類、標準化、及電子化。

RosettaNet 標準因使用方式複雜、牽涉技術繁複，給企業『高單價、難導入、維護與學習成本高』的印象，以致許多企業仍然裹足不前。

ebXML (Electronic Business using eXtensible Markup Language) 是由聯合國 (UN/CEFACT) 與 OASIS (The Organization for the Advancement of Structured Information Standards) 兩大國際組織於 1999 年 11 月共同發起了一項名為 ebXML 的計畫，企圖建構一個全球互通性之電子交易基礎架構和電子商務標準。其使用流程如下：參與 ebXML 式企業間電子商務的交易夥伴將其商業處理細節及能力記載在一份稱為「合作協定細節」(Collaboration Protocol Profile, CPP) 的 XML 文件中，並公佈在註冊服務處，以供其他有興趣的交易夥伴查詢。透過 CPP 的查詢，ebXML 交易夥伴透過協商達到共識，協商後的結果紀錄在另一份稱為「合作協定契約」(Collaboration Protocol Agreement, CPA) 的標準 XML 文件中，以做為雙方商業交談的依據。

ebXML 最大困難在於企業需與每一客戶取得協定契約 CPA，來實現交易夥伴間的 B2B 導入；每個不同的客戶，CPA 可能都不一樣。客戶也需要導入 ebXML，而且對於企業內整合效益不大。

3. 研究設計與方法

本研究針對半導體後段封裝廠 B2B 系統企業流程，提出了一個模組化的架構，建置共用 E-R Model 結構來達到 B2B 系統企業流程之整合，並透過 ETL 工具實作。研究的模組架構圖，如圖 2 所示，在半導體上游各家晶圓廠與下游封裝廠資訊，皆透過 ETL 匯整到下游封裝廠企業 DM 內，並共用 E-R Model 結構表。本研究各步驟在下面子節中敘述。

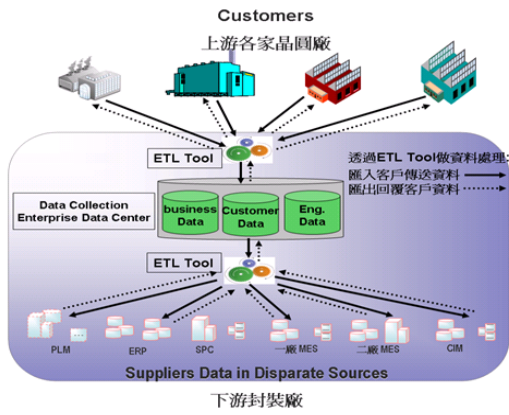


圖 2 上游與下游廠 B2B 系統與對內系統整合架構

3.1 步驟一：需求規格分析

本研究首先分析 B2B 系統在半導體上游晶圓廠與下游封裝廠的企業間通用性的 B2B 企業流程。此企業流程 B2B 系統需求(Business needs)一般包括了訂單資訊、產品資訊、工單資料、下線生產資訊、WIP (Work In Process) 在製品流程資訊、成品入庫、出貨、工程資訊、與帳單資訊等傳遞資訊(如圖 3 所示)。若以 UML 使用者案例說明，主動者(actor)為上游晶圓廠 B2B 系統，相關的使用者案例為下訂單、訂單異動、工單通知，在下游封裝廠 B2B 系統角色，相關使用者案例為訂單回覆、在製品 WIP、相關生產 Event 通知、及月初的帳款訊息等。

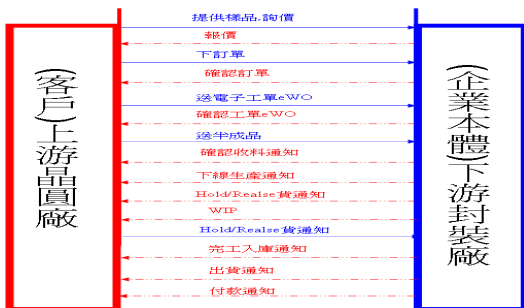


圖 3 半導體 B2B 系統企業流程資訊傳遞需求分析

至於交易流程細節方面，上游晶圓廠與下游封裝廠間 B2B 系統及企業本體內部流程皆以產品批號 (LOT_NO) 為控管單位，所以系統相關資料傳送及企業流程皆以產品批號 (LOT_NO) 為出發控管。交易流程如圖 4 所示。從上游客戶開立工單給下游封裝廠通知開使生產，庫存收到上游廠晶圓半成品，在 ERP 系統上做收料確認，記錄客戶批號與數量等資訊、生管依據主生產排程 MPS (Master Production Schedule) 開始用 MES (Manufacture Execution System) 系統投入下線生產製造，此時 MES 即開始產生企業本體的生產批號，在生產製造中有各種的生產製造狀況，如生產中 Run、等待 Wait、報廢 Scrap、客户要求終止生產 Terminate、暫停與釋放

Hold/Release 等，後續經產品檢驗，入成品庫、出貨、以及帳款等都是以產品批號去做系統資料交易，呈現產品以批號的交易流狀態的生命週期(如圖 4 所示)，後續整合 B2B 系統的 Data Model 設計也將用產品批號 (LOT_NO) 為控管單位。

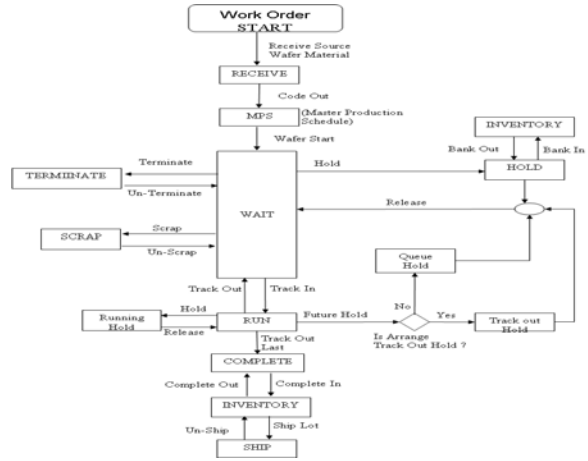


圖 4 生產批號狀況交易圖

3.2 步驟二：資料模型設計開發

此 B2B 系統在整合異質資料庫上有著不同於以往的設計與規劃，不同於傳統直接使用各系統的 OLTP (On-Line Transaction Processing) 資料，而是將從不同來源的資料，以半導體產業(產品生產批號)的角度將資料反正規化，匯整成星狀網要架構 (Star schema)。

此研究將以半導體產業(產品批號)的需求為出發，將所有資料的資訊分為以下兩大類，即不會變動的事件檔(Event Data)，與一直變動現況檔(WIP Data)。其中訂單資訊、訂單回覆、收料通知、工單、下線生產通知、成品完工入庫通知、送貨通知、出貨通知、及最後月初帳單通知這一些都歸納於不會變動事件檔(Event Data)。另一類如在生產製造中客戶會想知道目前當下在製品(WIP)資訊的相關資訊，因為此資料內容會一直變動。以上這資訊都須依據客戶的需求於第一時間傳送資訊給客戶。另有一種報表資訊是客戶生產完所要的即產品生產的歷史資料(Lot history Report)從收料，到下線生產，生產中包含各製造站點的進站(Track In)與出站(Track Out)時間與相關資訊如良品數量、報廢數量、異常時的 Hold/Release 等，最後成品入庫與出貨資訊。當然最重要需要生產批的基本資料檔供參考。故在設計 Data Model 時可以，以反正規方式分成三類的 DM 星狀網要架構：(1)事件檔(Receipt Note 收料資訊、Wafer Start 開始下線生產資訊等)。(2)現況檔(WIP 在製品資訊)。(3)基本資料檔(PO (Purchase Order) 訂單資訊、WO(Work Order)工單資訊)(如表 1 所示)。

表 1 產品生產批號分類設計表

	事件檔	現況檔	基本資料檔
(PO)訂單資訊			V
(WO)工單資訊			V
(Receipt Note)收料資訊	V		
(Wafer Start)開始下線生產資訊	V		
(WIP)在製品資訊		V	
(SCRAP)報廢資訊	V		
(Inventory)成品入庫資訊	V		
(Delivery note)送貨單資訊	V		
(Shipping)出貨資訊	V		
(Bill)帳單	V		

其中現況檔，例如生產批現況資料，主要解決上游晶圓廠要求下游封裝廠提供即時的生產現況資料。本研究設計以半導體產業(產品生產批號)的角度將資料反正規化，故在 E-R Model 架構的中間建立一事實資料表 (fact table) 來呈現封裝廠生產批即時資料現況，故建立生產批即時在製品現況檔 (B2B_LOT_WIP_C)，其欄位資訊包括在製品時間的生產批號(LOT_NO)，在那一廠(SITE_ID)與站點(STAGE)，目前良品(WAFER_QTY)、不良品(LOSSQTY)等，都可以詳細記錄匯整後提供給上游晶圓廠。為了更加快速所以建立相關索引參考資料，包含批號報廢記錄(B2B_LOT_SCRAP_H)、異常原因明細記錄(B2B_LOT_ERR_H)、晶圓明細檔記錄(B2B_Lot_WaferContent_H)，這些資訊都與生產現況有關(如圖 5 所示)。

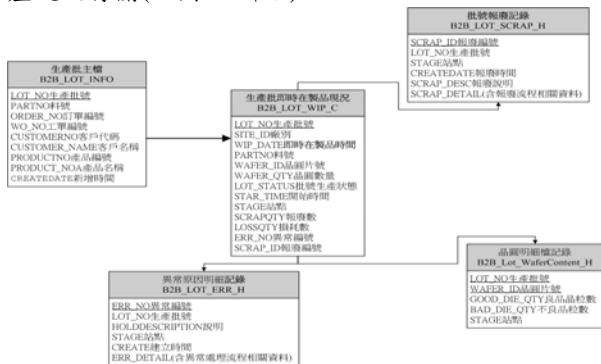


圖 5 WIP 生產批現況資料之 E-R Model

至於事件檔，例如生產批事件，主要記錄上游晶圓廠要求下游封裝廠提供產品的生產事件的資料通知。如圖 6 所示。本研究設計在生產批事件 E-R Model 架構中間建立生產批過帳交易記錄表 (B2B_LOT_STEP_H) 來呈現封裝廠生產批歷史資料與相關事件，其欄位資訊包括產品生產批號(LOT_NO)在那一廠(SITE_ID)與站點(STAGE)，開始進站時間(STAR_TIME)與有多少數量進站點(INPUTQTY)，完成後出站時間(END_TIME)與有多少數量出站點(GOODQTY)，生產中有無異常或報廢及其數量(SCRAPQTY)與說明，都詳盡記錄下游封裝廠的產品生產歷程，最後則建立維度參考資料表，包含批號報廢記錄(B2B_LOT_SCRAP_H)、異

常原因明細記錄(B2B_LOT_ERR_H)、晶圓明細檔記錄(B2B_Lot_WaferContent_H) 等生產歷史資料有關(如圖 6 所示)。

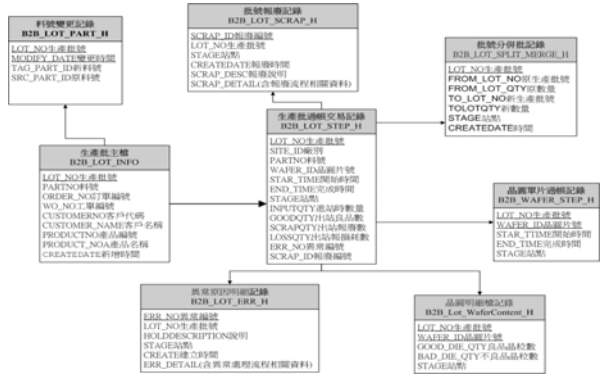


圖 6 EVENT 生產批事件資料 E-R Model

至於基本資料，此部份的來源依據上游晶圓廠提供的主資料表與明細資料表其間的主鍵 PK(Primary Key) 與外來鍵 FK(Foreign Key) 為工單編號 WO_NO，其主檔欄位資訊包含上游供應商相關資訊與此工單總數量(B2B_EWO_M.WAFER_QTY)與工單時間(B2B_EWO_M.WO_DATE)，工單明細欄位資訊包含工單型態(B2B_EWO_D.WO_TYPE)，(如量產、實驗或工程)等 (如圖 7 所示)。

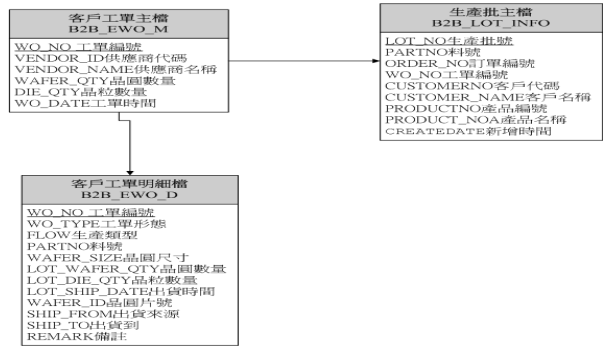


圖 7 生產批基本資料之 E-R Model

3.3 步驟三：ETL 工具並評估

B2B 於整合需求與所需技術可分為三類：應用系統的整合(企業內部異質系統之間的整合)、組織內流程的整合(企業內部異質且分散的商業流程整合)、商業社群的整合(整體供應鏈上下游整合)，其所需要的整合技術各有不同，在對內與外部資料傳輸與轉換技術，介面整合與附加服務技術都不同。本研究研議透過開放原始碼 ETL 工具[7,8,9]支來援系統建構 Data Mart。理想的 ETL 工具需要有下列特性：在資料傳輸技術上 ETL 工具可以支援系統(FTP 或 ODBC 等)、在資料轉換技術上可支援不同資料格式並可處理轉換成所需要的檔案格式(如

XML 或 Excel 格式)、附加服務可外掛自行開發或外購模組(如 Java 或已編譯執行檔)、在商業流程支援技術上可透過 Data Mart 當中介溝通橋樑。根據這些原則,本研究最後選定了 Pentaho kettle ETL[9] 工具來實作。

3.4 步驟四：資料庫與系統架構

半導體產業 B2B 系統整合的資料流程圖如圖 8 所示。如企業將依據規範透過 FTP Server 存放客戶檔案,並透過 ETL System 取得檔案將之處理清理轉換匯入 Data Mart,各系統依據需要透過 ETL System 取得所需要的資料,如 ERP 與 MES 皆需要客戶資料,Work Order 工單,透過 ETL System 被 MES 取得所有生產的資料提供給 MPS 主生產計畫排定生產的下線時間,當中有相關的生產事件也會透過 ETL System 寫入到 Data Mart 給相關系統運用,則 MES 系統產品報廢與原因資料皆需給廠內 ERP 系統與客戶 B2B 系統知悉,產品經檢驗符合品質要求後做成品入庫流程,此舉將成品入庫資訊透過 ETL 寫入到 Data Mart,以便提供給 ERP 取用做後續出貨動作,最後依據客戶需求將 Data Mart 的資訊透過 ETL System 處理後傳送到客戶(如圖 8 所示)。

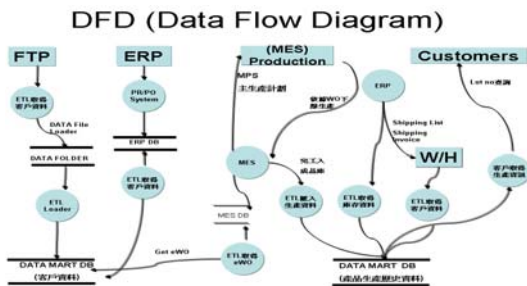


圖 8 設計資料流程圖

本研究建立兩組 DM 匯總資料庫,對外各上游晶圓廠提供下游封裝廠不同來源客戶資料表的匯總資料庫與企業本體內不同系統來源的產品生產相關資料表匯整資料庫,再透過四部份不同目的 ETL System 來處理匯整與傳送資料(如圖 9 所示)。

步驟1. 依據 DFD 及建立 Data Mart。

- 不同來源上游晶圓廠客戶資料表的匯總資料庫,客戶相關資料 DM(如工單基本資料)。
- 封裝廠企業本體內不同系統來源產品生產資料匯總資料庫,產品生產相關資料 DM(如事件、在製品資料)。

步驟2. 建構 ETL System 主機來處理資料。

- 建立 ETL System 主機匯整上游來源晶圓廠客戶資料到 Data Mart,並轉提供相關資訊給工廠生產使用。
- 建立 ETL System 主機匯整封裝廠企業本體內不同系統來源的產品生產資訊,並依據客戶需

求轉拋給客戶。

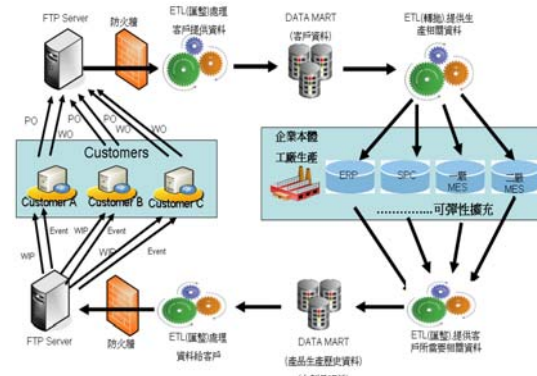


圖 9 上游晶圓廠與下游封裝廠整體系統架構

3.5 資料保存與歸檔

透過 ETL 處理資料建構 Data Mart (DM) 會隨著時間與客戶數量增加會導致效能低落且成本負擔增加,故將透過 ETL 處理,並依據符合政府電子法規明確規範資訊保留原則,企業會清楚知道哪些資訊應該保留多久,哪些資訊應該要刪除,以及應該要如何保留資訊。流程說明:

- 以 IT 基礎設施角度,將企業儲存底層分級,制定保留原則,目前規劃 Disk to Disk to Tape 備份架構。
- 以客戶批號(LOT_NO)來查詢稽核生產歷程資料的角度重新檢視各類型資料的保存期限。

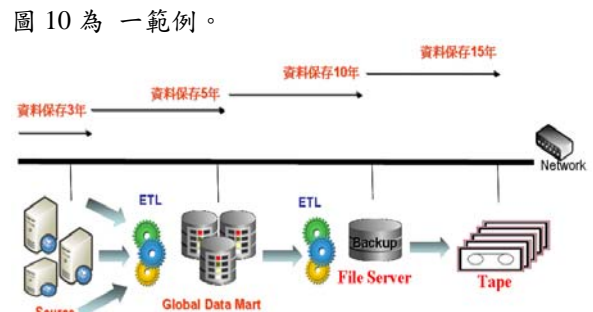


圖 10 資料保留與封存架構

4. 案例實作與效率評估

為應用上章之方法,我們以一實際個案評估所提方法之有效性。個案為半導體供應鍊的上游晶圓廠與下游晶圓封裝廠的 B2B 資料交換為例,提供兩個面向實作:(1)從上游晶圓廠提供 WO(Work Order)電子給下游晶圓封裝廠。(2)下游晶圓封裝廠匯整企業內各系統生產製造中資訊,提供在給上游晶圓廠在製品 WIP(Work In Process)報告[10]。圖 11 顯示部份實作畫面,初步結果令人滿意。

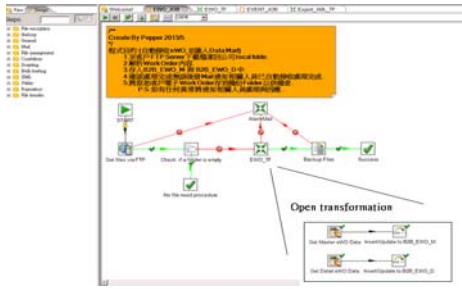


圖 11 ETL 實作取得客戶 FTP 工單資料流程

最後，我們評估傳統點對點的 B2B 方式與本研究整合方式，本研究方式好處如下：

1. 點對點的資料交換往返數減少(如圖 12 所示)。
2. 從開發週期來看，傳統點對點的開發約 4 週，使用整合的架構約 1~2 週，故可縮短開發週期 50%~75%。
3. 從系統維護來看，每年維護傳統點對點分散式 B2B 系統的人力，需 3 個人力，使用標準整合的架構每年用於維護人力約 2 位人力。
4. 從開發品質來看，自行開發點對點 B2B 系統資料分散在各個異質系統，通常是根據系統的當前客戶與開發需求與模組人員各自規劃，應用系統每增加一個新的模組，這一類需求相關系統常常需要隨之進行相應改動。而使用資料整合的架構及清楚 E-R Model 策略，可以有效保證系統品質，並減少資料不一致與錯誤。

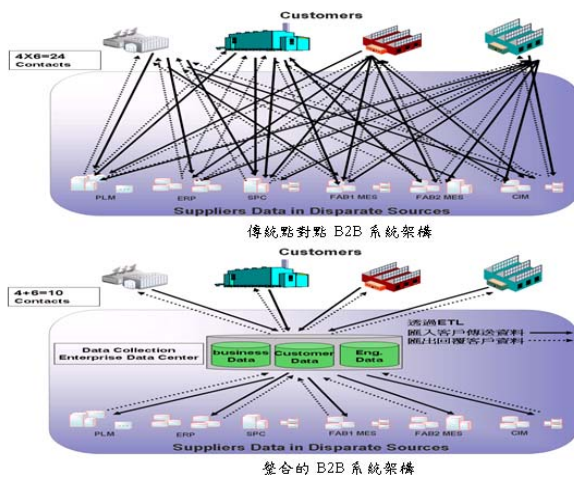


圖 12 上游晶圓廠與下游封裝廠整體系統架構

5. 結論與未來展望

半導體業的企業流程 B2B 系統傳統使用點對點的方法，此方法面臨到上下游垂直整合與企業本體內多元系統的水平整合挑戰；此傳統方法往往無法滿足企業的需求並導致開發成本過的現象，需要撰寫大量的程式碼並進行長時期的系統維護。本研

究針對半導體封裝廠的企業流程 B2B 系統設計通用星狀網要架構的 E-R Model (生產批現況資料 E-R Model 等)，整合相關企業流程資料並建立 Data Mart，使半導體上游晶圓廠與半導體下游封裝廠的企業流程資料都可以含蓋，且可以減少 B2B 系統需要不斷與各個廠內系統連線取得資料，降低企業內系統負載。本研究透過開放原始碼 ETL 工具來實作整合異質系統與資料，可快速導入 B2B 系統，且使封裝廠只需短時間，即可建構完整與上游廠企業流程與資訊整合。企業內部流程發生變化後，也不會影響到上游晶圓廠，對於封裝廠內部系統而言也可解決各異質應用系統的難題，亦容許內部應用系統延伸。當然 DM 的資料中心隨著時間與系統增加，而累計資料越來越多，將造成系統變慢。本研究提出的策略也可針對此部份也一併處理，透過整合的模組與 DM 架構，使用 ETL 的程序方式並依據訂定的原則來實作資料保留與封存。總結本方法效率有：節省系統維護成本、整合系統導入完成後，不易受企業流程變動影響。未來將在資安的管理、認證、授權上進一步探討。

參考文獻

- [1] 許智誠，資料倉儲的資料建制，中央大學資訊管理學系，2012。
- [2] EDI 資訊服務網，<http://edi.csd.org.tw/index.htm>。
- [3] XML 簡介，<http://www.xml.org.tw/Plan/XML>。
- [4] 林玉凡，EB 前瞻產品-企業應用程式整合(EAI)產品之探討，資策會電子商務應用推廣中心-FIND 研究群，2001。
- [5] ebXML.org, "ebXML Technical Architecture Specification," <http://www.ebxml.org/specs/>。
- [6] 資策會電子商務應用推廣中心，"RosettaNet Overview," <http://www.rosettanettaiwan.org.tw>, 2002。
- [7] 資策會 EB 研究小組，製商整合資訊網企業應用整合，<http://ciis.cdri.org.tw/>。
- [8] T. Jun, C. Kai, F. Yu, T. Gang, "The Research & Application of ETL tool in Business Intelligence Project," May, 2009。
- [9] The Best Open-Source ETL Tool, <http://www.ask.com/explore/opensource-etl-tools-3962>。
- [10] Yen-Yun Feng, "Development of WIP Monitoring Process Applications in Semiconductor Engineering Chain," May 2008。