

論文名稱：以網格技術擷取即時股市資料

總頁數：99

以支援資料倉儲之建置

校(院)所組別：中國文化大學商學院資訊管理研究所

畢業時間及提要別：96 學年度第 2 學期碩士學位論文提要

研究生：曾建銘

指導教授：李中彥

論文提要內容：

近年來資訊科技的進步產出了數量龐大的資料，並散佈於網路上，而這些資料經過分析整理後，有些相當適合作為決策支援之用，例如：股市交易資訊、期貨資訊、原物料價格資訊、匯率資訊...等等，但這些資訊通常變動快速，雖然可以容易的從網路上取得，但往往因為擷取以及處理資料的速度不夠快速，而造成決策失去其時效性因而不夠精確。

本研究以現有設備為基礎，設計出一個網格技術的系統架構來解決上述問題，集結多台個人電腦的資料處理能力成為一台虛擬的超級電腦，大幅提升整體資料處理的能力，減少擷取與處理大量網路即時資料所需要花費的時間；此外，隨機式工作排程、備份、資料完整性檢查等機制，更使本系統的可靠度與安全性大為提昇。

本研究建置了資料倉儲，在設計上力求簡潔與效能最佳化，讓處理完成的資料可以快速的存入資料倉儲中作決策支援分析。

最後，本研究所實作出的實驗數據中，證明以此系統架構進行資料擷取與處理的工作，與單台個人電腦相比，可大幅減少約 70% 左右的時間，搭配設計良好的資料倉儲，相當適合作為決策支援之用。

關鍵字：網格運算，資料擷取，動態資料

Realtime Stock Data Retrieving Based on Grid Technology for Constructing Data Warehouse

Student: Chien-Ming Tzeng

Advisor: Prof. Chong-Yen Lee

Chinese Culture University

ABSTRACT

In recent years, huge amounts of data appeared on the internet. Some of these are frequently changed and necessary for decision making such as stock information, raw material information, etc. A decision making (or support) system uses these kinds of data needs to retrieve these data from internet in rational time interval.

This research presents a system model based on grid technology to solve this problem using existing equipments. The model can not only retrieve huge amounts of data by using many PCs in a satisfactory time period, but can aggregates processing powers of many computers into a virtual super computer to finish job easily which is impossible to be done for a single PC.

The implementation in this research proved that about 70% of time saved between the model used and a single-computer mode. Beside, this model combined several different mechanisms like dynamic job scheduling, data backup, data integrity check, ..., etc. To ensure the system model more robust.

Key Words: grid computing, data retrieving, dynamic data

誌謝辭

首先誠摯的感謝指導教授李中彥老師，老師悉心的教導使我得以一窺網格運算領域的深奧，並透過不時的討論以指點我正確的研究方向，使我在這些年中獲益匪淺；此外，也要感謝黃謙順與王美慈兩位老師，在論文初審時給我許多寶貴的意見，讓我的研究可以更加的完善。

另外要感謝聖文學長不厭其煩的回應我關於程式開發的相關問題，且總能在我迷惘時為我解惑，也感謝儒霆同學的幫忙，有他的幫忙，我的論文才能更趨完整，也恭喜我們順利走過這幾年。當然林杰學弟也不能忘記，在系統開發的過程中，提點了我相當多財經相關知識，有他的協助才能將系統順利完成。

另外也感謝在資訊中心的各位教職員，在這幾年中提供我許多研究所需要的資源，並持續不斷的給我支持與鼓勵；也感謝工讀的學弟妹，在我的研究生涯中不時帶給我歡笑，給我持續不斷的動力作研究。

最後，謹以此文獻給我摯愛的雙親。

內 容 目 錄

中文摘要	iii
英文摘要	iv
誌謝辭	v
內容目錄	vi
表目錄	viii
圖目錄	ix
第一章 緒論	1
第一節 研究背景與動機	2
第二節 研究目的	5
第三節 研究範圍與限制	8
第四節 研究流程	9
第五節 章節架構	11
第二章 網格運算相關技術之探討	14
第一節 網格的介紹	14
第二節 網格的運作模式	17
第三節 網格與其他平行計算技術之差異	22
第四節 網格技術應用之相關研究	23
第三章 資訊安全相關技術之探討	25
第一節 資訊安全簡介	25
第二節 密碼學簡介	27
第三節 對稱式與非對稱式加解密	31
第四節 雜湊函數	39
第四章 資料倉儲相關文獻之探討	42
第一節 資料倉儲的介紹	42
第二節 ETL 之探討	47

第五章	系統架構	50
第一節	架構說明	50
第二節	資料特性	56
第三節	使用網格的考量	58
第四節	運作流程與理論	60
第五節	架構特色	62
第六章	研究方法與系統實作	71
第一節	研究方法	71
第二節	系統建置	72
第三節	系統實作	79
第四節	與不同架構的比較	85
第七章	結論與建議	87
參考文獻		91



表 目 錄

表 6-1	最好結果(網格環境)	83
表 6-2	最差結果(網格環境)	83
表 6-3	平均結果(網格環境)	83
表 6-4	最好結果(一般架構)	88
表 6-5	最差結果(一般架構)	88
表 6-6	平均結果(一般架構)	88



圖 目 錄

圖 2-1	網格架構	19
圖 3-1	資訊安全原則	25
圖 3-2	ECB的運作方式	32
圖 3-3	CBC的運作方式	33
圖 3-4	雜湊函數	39
圖 3-5	加密雜湊函數	40
圖 3-6	雜湊訊息確認碼的產生方式	41
圖 4-1	Informaster系統	43
圖 4-2	ETL的三階段過程	48
圖 5-1	系統架構圖	50
圖 5-2	資訊套包內容	53
圖 5-3	模式化後的Star Schema	54
圖 5-4	金鑰分配示意圖	68
圖 5-5	檢查資料完整性的運作機制	69
圖 6-1	一般系統架構圖	85