

# 行政院國家科學委員會專題研究計畫 期中進度報告

## 協同計算之網路資源管理(2/3)

計畫類別：個別型計畫

計畫編號：NSC91-2213-E-002-041-

執行期間：91年08月01日至92年07月31日

執行單位：國立臺灣大學電機工程學系暨研究所

計畫主持人：顏嗣鈞

計畫參與人員：陳建良、簡演龍、許華祐、康師誠、柯信豪

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 92 年 5 月 29 日

# 行政院國家科學委員會專題研究計畫執行進度報告

## 協同計算之網路資源管理(2/3)

計畫編號：NSC 91-2213-E-002-041

執行期限：90年8月1日至93年7月31日(全程)

(第二年: 91年8月1日至92年7月31日)

主持人：顏嗣鈞 台大電機系

計畫參與人員：陳建良、簡演龍、許華祐、康師誠、柯信豪 台大電機系

### 一、中文摘要

協同計算 (Collaborative Computing) 是一個術語，我們將會用它來取代一般傳統群體軟體 (groupware) 的觀念，而群體軟體就我們所知，是指在非同步方面合作，以電子郵件、討論區及一些相似性的軟體幫助我們以電子手段交換資訊。協同運算的觀念可以將一些網路的資源分享出來給所需要的人使用，而這些資源是可以即時方式所取得。

本研究的主要目的在發展一套跨資源的分散式資源搜尋與管理系統，並結合智慧型軟體代理人進行實作。在第一年，本團隊完成一些理論上的分析，將網路上資源量化並以 ER 派區網路 (ER Petri net) 語言來描述網路上各種不同的資源以及資源上的限制...等，加上引進了配對演算法 (matchmaking) 解決一些變動環境極大的網路資源分配問題，網路資源的配置可以藉由配對演算法所造成的配對樹中就可以反映出一些特性。第二年，本團隊規劃了一個以 Linux 為主軸的系統來模擬協同運算環境，利用修改一些公開共享軟體並將理論部份實做化，完成我們所提出的資源管理服務，這些我們會在第二部份中說明，並進一步在第三部份中簡述系統分析結果與討論。

本團隊未來將針對所設計的系統，做更進一步的測試與評估，並期以此系統，驗證更深一層的學理研究，提供更多樣的資源管理服務。

**關鍵詞：**協同計算、智慧型代理人、派區網路、資源分配、配對演算法。

### Abstract

Collaborative computing, nowadays, is gradually used to replace the traditional concept of "groupware." Generally speaking, groupware focuses on the asynchronous aspect of human collaboration such as email, discussion lists, etc. which helps people to electronically exchange information. Collaborative computing is a term describing a variety of resources where people need it; more importantly, people are able to use

these resources in the real time aspect.

The main purpose of our project is to develop a distributed resource-discovery and resource-manage system. In order to managing heterogeneous resources more intelligently, our implementation will combine software agents as a whole. Beginning in the first year, we have surveyed and analyzed currently worldwide activities about resource management. By introducing the ER Petri net, we established our system model to further describe the resource language and its limitation. When a job tells its requirement, matchmaking algorithm provides us the way to allocate resources with respect to their different availability. During second year, we implemented our system architecture based on PCs with RedHat Linux 8.0. With some open-source softwares, we designed three kinds of services: monitor, scheduling, and matchmaking. We would show some details and also analyze them in the second and third part of the report.

In the future, our team will estimate and evaluate our system. Further more, we will be able to verify the system and supply more versatile resource management services.

**Key words:** Collaborative computing, software agent, ER Petri net, matchmaking algorithm, monitor, scheduling.

## 二、計畫目標與規劃

第二年開始了系統架構的程式設計，目前已規劃並積極實作了三項協同系統上的資源管理服務～

### 1、監控的服務 (Monitor)

- 監控主要是對在同一個區域網路底下的分散式系統，進行各項資源的資訊掌握。
- 監控重點～
  - 顯示目前手邊可用的資源情況 (資源狀態、網路狀態、中央處理器、記憶體、...等)。
  - 顯示目前手邊已被使用的資源負載情況 (工作分配、佔具資源、...等)。
  - 顯示目前網路頻寬的使用分析。
- 監控實作規劃～
  - 設計 Linux 上的工作環境。

- 建構一個資料庫供作資訊存取。
- 設計一個智慧型代理人去動態的更新分散式環境底下的相關資訊。
- 動態探測區域網路內的使用狀態。
- 將開發一視窗環境，顯示上述實作。
- 硬體測試環境～
  - 目前以三台電腦搭配 RedHat Linux 8.0 架構一個小型區域網路，進行監控服務的測試環境。
- 軟體支援～
  - 利用 QT desinger 設計 Linux 上的工作環境。
  - 在測試環境下，目前規劃在各用戶端上設置 spy 程式，藉由各 spy 程式向主系統更新資料庫的方式，以達到監控的目的。
  - 目前資料庫系統使用 MySQL，未來規劃能夠移植到功能較為強大的 Oracle 上。

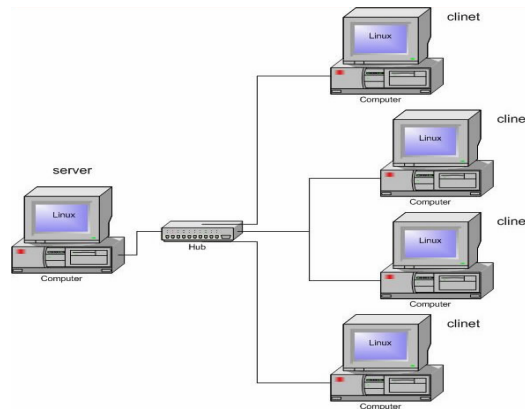
## 2、排程的服務 (Scheduling)

- 排程針對我們所建構的分散式系統環境，進行 job 的排程，並利用自我學習的訓練，提供更具效率的排程機制。
- 排程重點～
  - 進行各項可能的負載模擬。
  - 利用目前較為成熟的排程方式，進行負載分配的環境設計。
  - 分析並評估各項排程的優勢。
  - 利用 AI 的概念，使得本系統會學習最佳的排程機制。
- 排程實作規劃～
  - 設計一個工作模擬器，模擬各項可能的工作。
  - 搭配使用上面提到的資料庫系統，進行模擬負載的存取。
  - 程式實做出幾種常見的排程演算法機制。
  - 設計+整合+評估，利用圖形或統計的方式，展現效能評估。
  - 設計人工智慧機能，學習並自動應變可能的即時變化。
- 硬體測試環境：同服務 1。
- 軟體支援～
  - 沿用服務 1 的架構。
  - 關於排程的工作分派，目前將利用 PVM (Parallel Virtual Machine) 作為其工具。

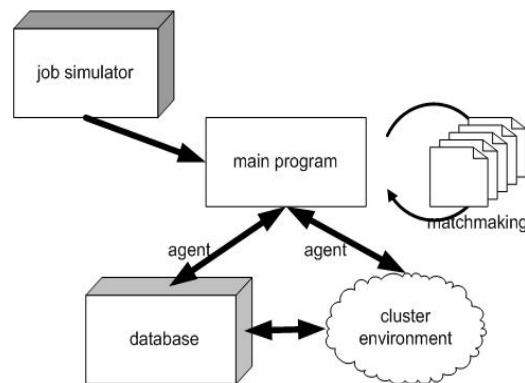
## 3、配對的服務 (Matchmaking)

- 配對是資源管理中常運用的一種資源分配方法，利用廣告配對的想法，將{供}與{需}作最佳的組合。
- 配對重點～

- 可使用的資源，藉由公告周知為系統提供選擇性。
- 為排程進來的工作做一個合適的資源配對組合。
- 一對一、一對多、多對多的不同配對組合。
- 配對組合最佳化。
- 配對實作規劃～
  - 利用資料庫結構，提供可用資源資訊。
  - 針對所需要的工作，進行適當切割。
  - 設計演算法尋找合適資源。
  - 修改參數，變化環境。
  - 最佳化的學習。
- 硬體測試環境：同服務 1。
- 軟體支援：同服務 1 與 2。



圖一、硬體結構。



圖二、系統運作示意圖。

### 三、計畫分析與討論

根據本團隊的研究，資源管理中首重各種資源之性質定義，針對其性質將本系統設計成可以妥善管理各種資源，我們可以對所有資源進行『保留』、『配置』

以及『使用』等動作，利用資料庫的架構妥善規劃，使個別資源都可以在不影響整個系統的前提下，獨立地被加入或移出本系統。藉由工作模擬器的設計，產生多樣式的工作群，此時，工作排程與資源分配便扮演著舉足輕重的關鍵。在兼顧效率與公平的考量下，以智慧型代理人為工具，動態更新各項資源即時資訊。隨著各種不同需求的提出或改變，資源必須能夠被隨時動態重分配並立即使用，本系統主要利用配對演算法，這是資源管理中常運用的一種資源分配方法，利用廣告配對的想法，將{供}與{需}作最佳的組合。分散式的環境下，本系統必須具備足夠的容錯能力，可以適時地偵測並且更正錯誤。

使用者可以透過搜尋介面將自己可用資源和想用資源公告在伺服器上，此搜尋介面是由 QT 所建構而成並可達到跨平台性。之後藉由伺服器中管理資源排程演算法來決定資源該如何分配使用，接下來如何讓各個電腦之間進行溝通就變成一個重要的因素，還有將工作傳送給不同的電腦來進行計算的任務是本團隊所需要考慮的，故我們以 PVM (Parallel Virtual Machine) 這個軟體來進行測試，PVM 軟體可以使各個電腦具有互相溝通計算之能力。因此我們利用 PVM 來達成平行運算的功能。當系統發生錯誤時，從最接近現在的檢查點重新開始工作，可是協同式計算環境中因為資源使用者並不擁有該項資源，所以資源很可能會被收回，或是工作中的某個伙伴發生當機，使得工作必須中斷，這樣子的中斷通常很突然且難以處理，故這也是我們嘗試要去解決的問題。協同計算環境是一個極易發生錯誤的環境，這樣的環境對安全性更是一大考驗，加強容錯能力與防止駭客入侵實在是很重要的課題。最後可以透過系統監測，指的是資源與使用者的使用率，使用方式，計算行為等等，這些資料經過整理統計之後可以供作我們調整系統設計及管理策略之用。

本團隊已經將系統架構建構出，並開始著手修改一些相關軟體，嘗試利用分配演算法和分類廣告 (classified advertisement) 的方法，使用由上至下回溯演算法 (top-down backtracking algorithm) 來安排資源分配的問題，如此可以解決一些變動環境極大的網路資源分配問題，使其達到我們預期的要求。利用資源監督程式知道資源如何的被使用及分配，但是安全性和系統發生錯誤的考量是我們所缺乏的，期待在第三年中能夠試著提出有效率的解決之道。

#### 四、計畫成果自評

根據本計畫的進度，本團隊對第二年研究的統整如下：

- 研究資訊統整。
- 制訂完整可行之系統架構。
- 研究設計三項資源管理服務。

- 監控。
- 排程。
- 配對。
- 制訂富有調整彈性之協定。
- 演算法的設計與運用。
- 整合發展必要的所需軟體。
- 檢查點與安全性的探討。

在第三年的初步規劃中，本團隊將針對所設計的系統，做更進一步的測試與評估，使本系統於效能上，能有一個量化的資源管理機制。並期以此系統，驗證更深一層的學理研究，使本系統不僅僅是具備高效能，更同時於跨平台性和穩定度上，亦有相對等的表現。最後，試著規劃與提供更多樣的資源管理服務。

## 五、參考文獻

- [1] D. Angulo, I. Foster, C. Liu, and L. Yang. *Design and Evaluation of a Resource Selection Framework for Grid Applications*. Proceedings of IEEE International Symposium on High Performance Distributed Computing (HPDC-11), Edinburgh, Scotland, July 2002.
- [2] G. A. Geist, J. A. Kohl and P. M. Papadopoulos, *PVM and MPI: a Comparison of Features*, *Calculateurs Paralleles* Vol. 8 No. 2 (1996).
- [3] I. Foster and C. Kesselman. *"The Grid: Blueprint for a New Computing Infrastructure"*, Morgan Kaufmann, San Fransisco, 1999.
- [4] Jim Basney and Miron Livny, *"Deploying a High Throughput Computing Cluster"*, in *High Performance Cluster Computing*, Rajkumar Buyya, Editor, Vol. 1, Chapter 5, Prentice Hall PTR, May 1999.
- [5] Jim Basney and Miron Livny, *"Managing Network Resources in Condor"*, Proceedings of the Ninth IEEE Symposium on High Performance Distributed Computing (HPDC9), Pittsburgh, Pennsylvania, August 2000, pp 298-299.
- [6] J. L. Peterson. *Petri Net Theory and the Modeling of Systems*. Prentice -Hall International, 1981.
- [7] MPI Forum. *MPI: A message- passing interface standard*. International Journal of Supercomputer Application, 8 (3/4) : 165-416, 1994
- [8] Rajesh Raman, Miron Livny and Marvin Solomon, *"Resource Management through Multilateral Matchmaking"*, Proceedings of the Ninth IEEE Symposium on High Performance Distributed Computing (HPDC9),

Pittsburgh, Pennsylvania, August 2000, pp 290-291.

- [9] Rajesh Raman, Miron Livny and Marvin Solomon, *"Matchmaking: Distributed Resource Management for High Throughput Computing"*, Proceedings of the Seventh IEEE International Symposium on High Performance Distributed Computing, July 28-31, 1998, Chicago, IL.
- [10] V. S. Sunderam. *PVM: A framework for parallel distributed computing*. Concurrency: Practice & Experience, 2 (4), 1990.