

# 第一章

## 導論

### 1.1. 研究背景

善用運算資源一直是電腦發展演化過程中一項重要的命題。從早期的用戶端/伺服器端架構模式（Client/Server）到最近的分散式運算架構（Distributed Computing），目的都是希望讓可供利用的公用運算能力（Utility Computing）可以整合在一起，形成一個虛擬組織（Virtual Organization, VO），而虛擬組織裡面的電腦節點可以運用這些可資利用的公用運算能力來完成計算任務。

所謂的公用運算能力包含各項電腦運算元件，其中最主要的便是 CPU 運算能力，主記憶體以及儲存設備的儲存能力。透過分享虛擬組織中閒置的這些運算能力，許多需要大量電腦資源的複雜計算問題可以獲得解決。每一台個人電腦因為虛擬組織的分享機制而取得遠大於本身固有的運算資源，所能完成的計算工作也隨之增加。

早期個人電腦在主從式架構下透過網路共同分享伺服器上的資源，並有條件的存取其他用戶端的儲存空間。主從式架構雖然立意良好，在早期提供了資料分享的管道，但卻有著以下的缺點：1) 隨著用戶端數量的增加，伺服器端的負荷不斷加重，因而我們不得不隨之增加伺服器的數量，以免在伺服器端形成所謂的熱點，也就是服務的瓶頸（Hot-spots）[4]。瓶頸的出現將會使伺服器端的服務品

質 ( Quality of Service, QoS ) 下降，喪失原先希望透過分享公用運算能力的好意。長期下來，更有可能因為負擔過重的伺服器發生所謂的單點崩潰 ( Single-point crash, 或稱為單點故障, Single-point failure )，而導致虛擬組織的癱瘓。2) 根據以往的觀察，我們曉得電腦節點常常因為某些因素的影響，例如突然的離線或是已經無法再負擔更多的計算，而導致被交付的任務無法被完成。過去的主從式架構中，這個負責計算任務的角色往往就是伺服器。一旦伺服器上的計算任務無法完成，所有需要這個計算結果的用戶端將全部進入等待狀態，如此一來就很容易形成所謂的計算瓶頸，輕則拖慢整體運算速度，重則很有可能造成伺服器的當機。

為了解決如上面所描述的計算瓶頸問題，分散式運算架構開始興起並廣為應用。在分散式運算的架構下，我們可以進一步的將分享公用運算資源的概念擴大，透過網路的连接，把運算任務切割，分配到多台電腦上運算。參與運算任務的電腦節點可以是大型的主機或是伺服器，也可以是一般的個人電腦；可以都是同一區域內的電腦，也可以是分散在世界各個角落。擔任計算的電腦節點間由非單一網路路徑相連，相互傳遞、整合、分享計算結果，並因為計算任務的切割，每一個電腦節點都可以只負責獨立的運算任務。這樣的作法使得負責某一項運算任務的電腦即使發生故障，無法繼續運算，也不會影響整體的運算結果以及服務品質。如此便可以大幅降低單點崩潰所帶來的影響。圖 1.1 到圖 1.3，分別呈現目前兩個較成熟且被廣泛應用的分散式運算技術分別為「網格運算技術 ( Grid Computing )」及「對等網路運算技術 ( Peer-to-Peer Computing, P2P )」，我們將兩種不同的技術詳述如下。

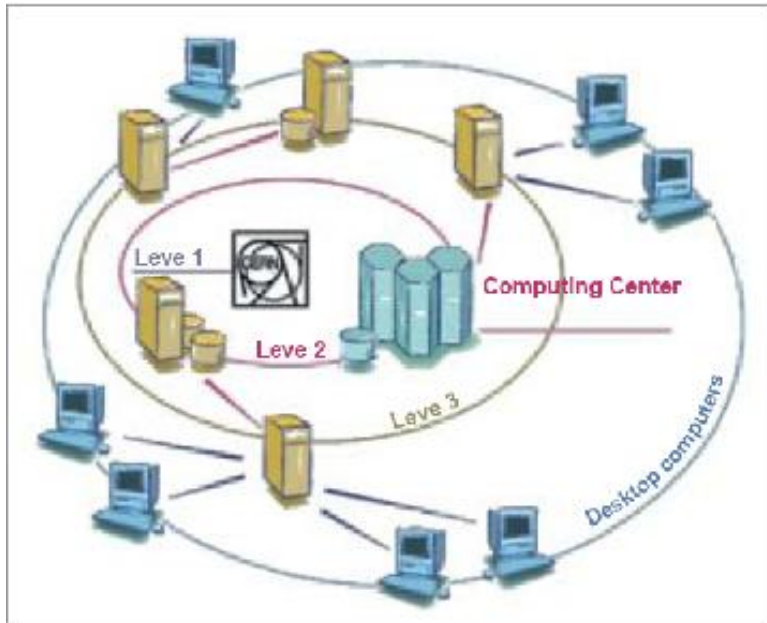


圖 1.1，資料網格中的層級式資料存取架構 [3]

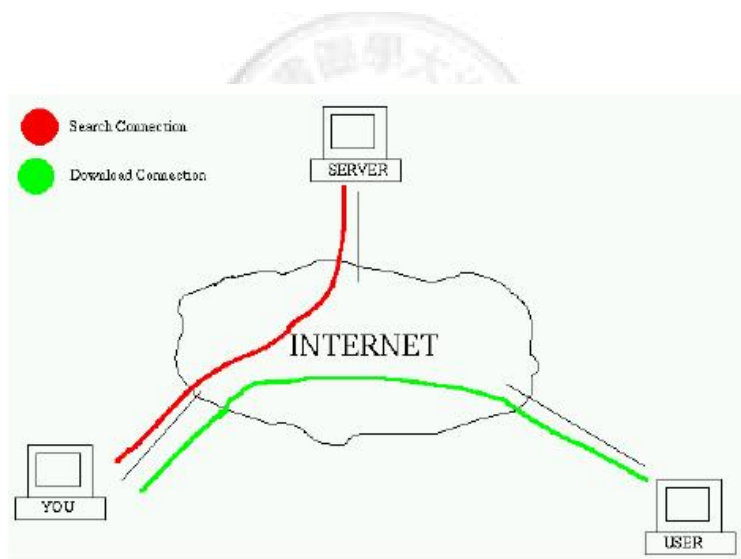


圖 1.2，集中式 P2P 架構，Napster Model P2P 示意圖 [5]

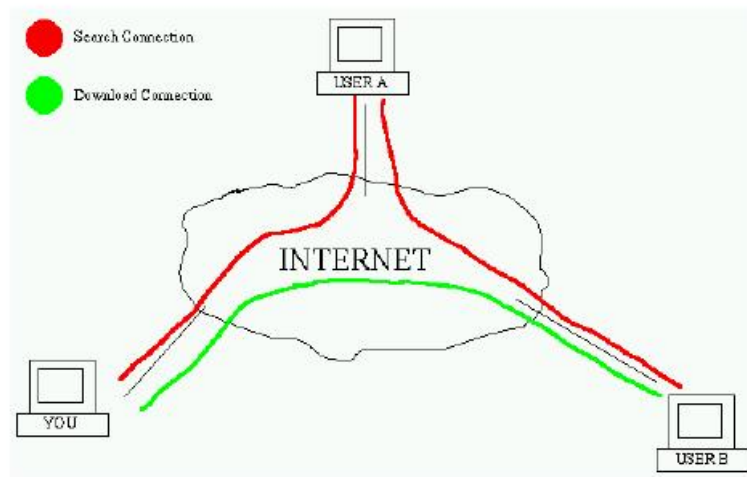


圖 1.3，分散式 P2P 架構， Guntella Model P2P 示意圖 [5]

網格（Grid）是一種利用散佈在網際網路（Internet）上的電腦資源來完成計算任務的分散式運算技術。用戶端透過向網格網路登記使用的方式，分享已加入網格網路的公用運算資源；而登記之後的用戶端也將會提供出本身的閒置時間，成為可供其他已登記的用戶端利用的公用運算資源。伺服器端扮演著工作分派與協調的角色。伺服器分割運算任務並將被切割的運算任務指派給可提供運算時間的用戶端協助完成。網格發展至今致力於幾項重要的目標 [3]：

- 1) 更進一步地分享公用運算能力：使用者可利用的資源將遠遠超越任一個超級電腦中心容量，而且公用運算資源的分享不再侷限於單純的儲存空間，而是更進一步地可以存取任何地點的應用程式與資料。
- 2) 安全存取：運算資源共享是網格的一大優勢，所以網格運作環境必須確保資源共享的安全性。藉由授權（Authorization）、認證（Authentication）與稽核（Accounting）三『A』整合機制決定誰有哪些權限可以使用何種網格資源，以及那些資源被授與哪些權限給誰使用等不同資源間的使用政策。
- 3) 節省成本：個人電腦並非長時間處於忙碌狀態，而是大部分的時間都處在「閒置」（Idle）或是「處理低優先權」任務狀態，而不會使用全部的計算資源。如果我們能將網路上每一個電腦節點的計算資源加以精算並且整合利用，將運算任務做正確的切割與分派，便可以用少數的電腦節點完成更多的運算任務，節省成本。

對等網路運算技術 ( Peer-to-Peer Computing, P2P ) 所倚賴的是對等網路中參與者 ( Peer ) 自願提供的計算能力以及頻寬來完成運算任務。所有參與者登入及登出 P2P 系統時，均會自動地將分享資訊傳送到管理 P2P 網路的中央伺服器，並由中央伺服器集中管理維護，以保持分享資訊的最新狀態 [3] 。因為智慧財產權的爭議， P2P 系統除了維持「分享」的核心價值外，並建立「非集中化」特色，所謂的非集中化其實便是將在原本 P2P 網路架構裡中央伺服器中的工作轉移到集線器 ( Hub ) 上，而一般的使用者扮演葉節點 ( Leaf Node ) 的角色。這裡所謂的「集線器」是指在 P2P 分享網路中屬於較高分享等級的成員，持有較多的分享主機與檔案資訊；而葉節點是指 P2P 網路中一般的分享成員，透過不同的方式來得到「鄰近集線器」的分享清單。不同的 P2P 分享網路系統的非集中化程度有不同的集線器數量，而不同的集線器數量則影響 P2P 網路上具有檔案分享連線能力的主機清單資訊是否為最新或最正確，意即影響所謂的內容可獲得性 ( Content Availability ) 。 P2P 將所有用戶視為平等的同級節點，每一個參與的節點將自身的運算任務從鄰近的集線器開始依次詢問可協助完成運算任務的可能，並且在本身處於閒置狀態時提供自己的運算能力給其他需要的節點。這樣的作法讓 P2P 擁有了運算資源接近無限的可能，也就是說當有新節點加入時，不論所能提供的運算能力是大是小，整個系統的公用運算能力都會往上提升，這是傳統通過單一伺服器來達到分享運算能力的主從式架構所不能實現；再者 P2P 網路可以通過在多節點上複製數據增加排除單點故障的耐用性，使系統不易出現單點崩潰的現象。

Grid 網路與 P2P 分享網路的異同點在於下列幾項：

- 1) Grid 網路下的運算資源與服務需透過註冊與認證才可以使用，而 P2P 網路則是透過參與者在登入與登出時主動更新集線器上的分享資訊。相較之下，比起 P2P 分享網路， Grid 更能有效掌握可利用資源的數量與耐用性

(Robust) 。

- 2) Grid 網路下的運算資源與服務需透過註冊與認證才可以使用，代表 Grid 伺服器必須維護已註冊運算資源的正確性與可用性；P2P 網路則沒有所謂的認證機制，單純倚賴提供分享的節點的「道德心」。相較之下，架設 Grid 網路的維護成本較高，且由於認證的難易度較高，所以整個網路架設難易度也較高。
- 3) Grid 網路下參與資源分享的電腦節點透過三 A 機制（授權 Authorization、認證 Authentication 與稽核 Accounting）來掌管分享的運算資源，分享資源的過程較為嚴謹；P2P 網路的葉節點在登出登入系統的時候，向鄰近的「集線器」提供分享的資訊。這種非集中化的分享方式，雖然沒有通過認證稽核等等安全機制步驟，但是卻較富有彈性變化。
- 4) 最後，Grid 與 P2P 網路架構都缺少信賴機制。當葉節點提供分享資訊後，Grid 與 P2P 網路會在認證伺服器或是集線器上建立一份可利用的資源索引，當葉節點或是其他伺服器需要服務的時候，依照這份資源索引去找提供相對服務資源的電腦節點。索引本身並不會提供服務的評價，只會列出任何有認證、註冊、分享的計算資源，而任何提供相同資源的節點之間並無好壞的差別，也就是索引本身無法分辨服務的好壞。信賴模型（Trust model）的建立有助於將可服務的品質做出區分，利用「口耳相傳」的評價機制，讓建立起的索引更有參考價值，加速得到良好服務的速度，增加完成任務的數量。

Grid 和 P2P 都致力於提供良好的公用運算服務來協助完成計算任務，各有各的優缺點。由於可以截長補短，從事分散式運算技術的人開始研究 Grid&P2P 混合式計算模式的可能性 [1]。但是截至目前，Grid 以及 P2P 分享網路均不具備「信賴模型」機制，這一點讓「取用散佈在分散式架構下的服務」這樣的動作，有了很高比例的風險，容易遭受劣質服務的欺騙。本論文嘗試將基礎信賴機制 [15] 融入分散式分享網路中，並在同時具有 Grid 工作分派以及 P2P 自主

提供公用運算資源的模型上，以社會科學的研究為基礎定義一個新的信賴模型，讓在 Grid&P2P 的架構下，具有弱耦合（Loose coupling）特性的電腦節點可以降低發生節點詐欺的機率及詢問任務所花費的成本，並有效的讓整體參與計算任務的節點發揮趨良逐劣的功效，提升整體的服務品質。

## 1.2. 研究目標

本論文的研究目標在提出一套新的信賴模型，並以此為基礎模擬一個整體服務品質或是整體公共生產力可經由評價機制自我提升，且沒有詐欺行為的小世界 [33]。「小世界」一詞，最早出現在心理學家米爾格倫（Stanley Milgram）所提出的六度分離理論（Six degrees，1971）之中，本論文引用鄧肯·華茲（D. J. Watts）的「Six Degrees: The Science of a Connected Age」（六個人的小世界，2004）的解釋 [33]。我們不難從中發現，如果允許模型中的個體採用多重社交向度，他們相對就很容易在廣大的網路中尋得選定的目標，也就是該網路的「類聚性」 [33] 便會很高。也就是說，可搜尋網路的依存條件主要並不在類聚參數或是社交向度的數目，而是只要個體比較容易認識與自己相仿的人，便可以更容易的認識另一個與自己相仿的人。而且很重要的是，因為模型中的個體含有多重社會面向，那麼不限任何地點，不限原本存在任何連結路徑，個體只需擁有網路的「區域」（Local）資訊，便可以很快地找到通往另一個相仿的人的路徑，這樣的架構便是所謂的小世界模型。

我們將「相仿」定義為「信賴度一樣良好」或是「口碑一樣良好」的服務運算資源。假設個體本身在小世界中信譽良好或是口碑良好，當此個體需要另一個信譽良好的人幫忙時，根據「類聚性」，此一個體會較容易從它可信賴的朋友口中詢問到擁有優良服務的人，自然而然便會愈來愈容易尋得幫忙，整個

小世界的總體服務品質也會愈來愈好，而這也正是本論文的最終研究目標。所以在這樣的前提下，我們的研究觀察將會著重在以下幾個重點上：

- 1) 信賴路徑：我們定義「從發出服務需求的個體，到尋得相對應服務的提供者之間所連結的路徑」為「信賴路徑」。信賴路徑的建立不僅僅要快，而且最終尋求得到的服務必須要在小世界中有優良的評價或口碑，無任何詐欺或惡意破壞的行為。我們了解分散式分享網路具有可搜尋性，本論文將漫無目的的廣布搜尋加以進化，透過親朋好友的推薦並排列信賴度評比，從中選擇品質較好且可以立即服務的節點。如此一來便可降低搜尋服務的成本與風險，提升整體小世界的生產力。
- 2) 轉手路徑：被發出服務需求的個體，請求計算資源服務的提供者若無法完成計算任務，在計算任務的成本以及自身的信譽度考量下，應將任計算任務讓渡給與自己擁有提供相同服務能力且評價優良的友好節點。此時從讓渡開始到尋得協助完成計算任務的節點為止，其間所連成的路徑我們定義為「轉手路徑」。與信賴路徑不同的是，當需要服務的個體挑選可服務的節點時，此時所有可提供相同服務的節點處於「競爭」( Competition ) 的狀態；當服務指定給其中之一，倘若計算任務的成本已無法再忍受增加時( 很有可能之前在建立信賴路徑就已花費過多的時間導致成本累加過多 )，此時接受指定任務的節點為了維持自身在小世界中的聲望，勢必尋求其他友好且提供相同服務的節點的幫忙，這時候便稱為「合作」( Cooperation ) 狀態。本論文所研究的新信賴模型仰賴節點之間的合作與競爭，以促進整體小世界公共生產能力 [2][25] 。
- 3) 公共生產力：根據上述所定義的「信賴路徑」以及「轉手路徑」，我們將藉以建立起分散網路上的新信賴模型。我們預期新信賴模型將提升整體小世界的公共生產力，也就是所有的計算資源均被善加利用，且所有小世界中的節點基於善意，共同完成所有小世界中的運算任務。我們若將小世界視為一個黑箱 ( Black Box )，則公共生產力的提升便可以理解為「任務完成率或是滿意度的提升」。本論文將會關注在這個議題上，並且藉此判斷本文所提出之新信賴模型比其他信賴模型有著更多效能改進。



4) 排擠效應：我們關心基於善意的優良節點所擁有的計算資源是否都可以被善加利用，促進整體小世界的公共生產力；同樣的我們也關心懷有惡意的詐欺節點出現在小世界中時，是否可以因為我們提出的新信賴模型，很快的被「指認」出來，並將其孤立在小世界之外。我們將一個懷有惡意的節點，從出現在小世界之中，到被孤立於任何被詢問服務的可能之外的現象稱為「排擠效應」(Supplanting effect)。如同「公共生產力」一般，「排擠效應」亦為本論文賴以評判本文所提出的新信賴模型比其他信賴模型有著更多效能改進的依據。

本論文在稍後的討論中將會關注在這四個議題上，我們會先從建立「信賴路徑」以及「轉手路徑」逐步探討，最後我們以分析「公共生產力」、「排擠效應」以及「社會學研究上分析人際關係指標」[29][30]三個評量值來區分本文所提出的新信賴模型與一般信賴模型的異同。

### 1.3. 本論文之章節架構

本論文的以下討論概略分成幾個章節。第二章的部份，我們介紹本文所提出之新信賴模型跟目前相關研究的不同處。在第三章中，我們將探討本論文所引用的社會學的觀點。在第四章中，會詳細描述以社會學觀點，配合 Grid&P2P 混合模型所設計的演算法與公式。而在第五章中，我們會概略的介紹以這些設計的演算法與公式所實作出來的信賴模型模擬器。最後在第六及第七章中嘗試著以設計出來的模擬器做相關信賴度實驗以及未來研究的目標。