

GPFS-WAN 與 Gfarm 檔案系統之建置與分析

鄭國陽

國家高速網路與計算中心

Email: kuoyang@nchc.org.tw

摘要

在網路與儲存設備迅速發展，許多網路檔案系統隨之因應而生，文中探討多個檔案系統，如 NFS、SRB、iRODS、Lustre、Gfarm 及 GPFS-WAN，在評估資料格網環境應具備的條件，如擴充性、安全性及穩定性，此外，對外組織的認同度及實際使用的狀況皆列入考量重要因素。最後，選擇 Gfarm 與 GPFS-WAN 作為重要指標，進行實際建置、效能測試及整體評估以提供使用者在建置資料格網環境作為參考。實驗結果所示，在實際複製檔案的傳輸速度 GPFS-WAN 的表現優於 Gfarm。在整體評估上，Gfarm 具有檔案複本功能在未來大規模的資料格網環境應取得相對的優勢。在建置過程發現，GPFS-WAN 需預留磁區來建立檔案系統，因此可能需要重新配置硬碟或新增硬碟及 GPFS 目前為商業軟體，使用者需要作成本及未來開發的考量。

關鍵詞: Gfarm、PRAGMA、GPFS-WAN、TeraGrid

Abstract

The network technologies and storage devices are developing quickly and the technical devices are quite cheap. Therefore, users have a large amount of storage space, Data Grid can collect distributed storage device to share to other users. In this paper, we discuss several file systems that consist of NFS, SRB, iRODS, Lustre, Gfarm and GPFS-WAN. In Data Grid, we select the specific file system to administrate the storage devices. We must consider the file system that obtains scalability, security and stability. Finally, we choose two file systems that consist of Gfarm and GPFS-WAN to be our evaluative target. Gfarm and GPFS-WAN are used to build up large-scale dataset storage system that consists of PRAGMA and TeraGrid, respectively. In this paper, we use Gfarm and GPFS-WAN to create a Data Grid environment and evaluate the performance of data transmission, respectively.

As results, the GPFS-WAN's performance of data transmission is better than Gfarm. The data transmission is one of the factors in Data Grid environment. For instance, the system obtains replication, open source software, easy to install and maintain that is important to be a great Data Grid environment, these important information will be discuss in this paper.

Keywords Gfarm、PRAGMA、GPFS-WAN、TeraGrid

1. 緒論

在科技發展的年代中，資訊相關技術發展迅速，不論在硬體設備及軟體開發皆有大幅的成長。在資訊技術中，網路及硬碟設備其傳輸速度與儲存空間成長迅速，眾多網路檔案系統隨之因應而生，本文探討數個檔案系統，如 NFS、SRB、iRODS、Lustre、Gfarm 及 GPFS-WAN。

把網路檔案系統推廣至大規模的資料儲存空間，並且透過網路集結分散於各地的檔案於單一平台以供使用者存取。因此 Data Grid 的技術即日趨受到重視。Data Grid 不僅作單純的檔案共享與管理，在學術界與業界進行大量運算的過程，需要使用大量的資料搬移，因此 Data Grid 與 Computing Grid 的結合是提昇整體執行效能重要的議題之一。綜合上述得知網路檔案系統之重要性，乃成為本篇文章探討主要動機。

在本文中，第 1 章介紹格網技術重要性，第 2 章內容主要是探討網路檔案系統相關技術，第 3 章主要是介紹 Gfarm 與 GPFS-WAN 技術與建置，第 4 章進行效能評估與整體評量，結論與未來展望為第 5 章。

2. 相關議題研究與探討

在網路迅速發展下，許多網路檔案系統隨之因應而生，在叢集電腦的架構下，存在著許多不同的檔案系統，一般使用者最常見的應是 NFS (Network File System)，NFS 不論在架設和管理方面皆兼具便利及迅速的特點。NFS 主要功能是提供使用者透過網路傳輸的方式來完成檔案共享的目標。首先，NFS Server 必須設定共享的主機、目錄及存取權限。在使用者方面，需透過 NFS 特定的 mount 指令，把遠端的 NFS Server 上的共享資料目錄掛載在本地端的指定目錄上。隨之，使用者在掛載的目錄上即可共享遠端 NFS Server 上的資料以達到共享的目的。

如果採用 NFS 作為計算格網環境下的網路檔案系統，可能會面臨到檔案傳輸的負載量皆集中於 NFS Server 作處理，因此 NFS Server 將受限於網路傳輸及集中處理的問題。當使用者數量上升其效能則隨之下降，且缺乏檔案容錯及檔案修復的功能，由此可知 NFS 適用於叢集電腦規模的環境，在格網環境下仍存在著許多問題需克服。

SRB (Storage Resource Broker) [5] 整合不同儲存設備並透過網路傳輸的方式把不同的檔案儲存設備集結成一個大規模的檔案共享系統，實現所謂的資料格網(Data Grid)。SRB 主要架構是透過一個 Metadata Catalog (MCAT)及數個 SRB Server 組合而成。每個 SRB Server 代表著分散於各地的檔案儲存設備，透過 MCAT Server 把所有分散於各地的 SRB Server 集結成一個大規模的檔案共享系統，使用者可利用指令 (Shell Command)、網頁 (MySRB) 或針對 Windows 開發的應用程式 (inQ) 來進行共享檔案的存取。由於 SRB 並非 Open Source (開放原始碼) 的程式，在許多使用及管理方面嚴重缺乏彈性而不利於整合。SRB 本身為非開放原始碼軟體，在軟體取得及開發上將面臨許多問題，譬

如在未來功能最佳化部份無法自行開發及修正及面臨與計算格網整合的問題。

近年 SRB 團隊重新開發一個 Open Source 專案軟體稱之為 iRODS (i Rule Oriented Data System)[3]，iRODS 利用模組化的方式改善原有 SRB 在開發上的限制。iRODS 使用 Rule Engine 在集結許多不同的組織時，讓每個不同組織可針對組織內的需求來制定其管理方式及策略以解決整合及管理上的問題。iRODS 克服許多 SRB 的問題，目前存在有檔案權限管理、檔案傳輸安全性的相關問題需追蹤評估及測試。

檔案系統為提昇傳輸速度及速率，隨之發展出平行檔案系統的概念，在 Cluster File System Inc. 公司開發的 Lustre 平行檔案系統 [4]，由於 Lustre 是 Open Source 開放原始碼專案的檔案系統，因此 Lustre 是廣泛被使用者採用的平行檔案系統。在 Lustre 檔案系統中，提出一個 Object 概念，Object 是一個邏輯集合，包含有屬性、存取方式及位元組。

Lustre 平行檔案系統主要組成的元件有 Metadata Server、Object Storage Server 及 Client。Metadata Server 的功能主要是管控目錄、metadata 資料一致性、資料掛載點、Object 與檔案相對應關係等。Object Storage Server 主要是配置 Object data 與 blocks 的資訊。Client 則負責把 Lustre 檔案系統掛載到本地端來進行存取。

在 Lustre 平行檔案系統中，每個 Lustre 的節點都必須使用 Lustre 專屬的核心來開機。在 Lustre 平台下，使用者在掛載檔案系統的方式與 NFS 相同採用一對一的掛載方式來完成。此外 Lustre 支援多種隊同的網路設備，譬如 Myrinet、Infiniband、Quadrics 及常用的 Ethernet。Lustre 具備開放原始碼、平行檔案處理及適用於多種網路設備，不論是在新增功能、修改錯誤，使用者皆可自行取得原始碼進行開發及修補的動作。

3. Gfarm and GPFS-WAN File System

在資料格網的環境下，存在著許多的檔案系統及架構，在第 2 章相關研究與探討的部份有提及四種不同的網路檔案系統分別為 NFS、SRB、iRODS 及 Lustre，這四種檔案系統皆作為建置資料格網平台的選擇之一，NFS 及 Lustre 主要是適用於叢集電腦環境下，且分別存在著潛在問題，在 NFS 方面，有擴充性及網路集中在伺服器端而造成傳輸速率的問題。在 Lustre 使用專屬的核心，在既有的作業系統建置 Lustre 可能在變動核心的過程會造成原有系統的運作。在 SRB 方面，本身為非開放原始碼軟體，導致在使用及未來開發無法自行控制的問題。在 iRODS 方面，克服許多 SRB 的問題，目前存在有檔案權限管理、檔案傳輸安全性的相關問題需追蹤評估及測試。

縱觀上述的四種檔案系統，各自都具備其優缺點，但以現實考量，一個檔案系統是否可實際運作於資料格網環境，必須有通盤性的考量如擴充性、安全性及穩定性等。因此，我們選出 2 種檔案系統作為實驗的主軸，分別是 Gfarm File System[1, 7, 8, 9]及 GPFS-WAN[2]。選擇這 2 套檔案系統主要的原因有擴充性、安全性及穩定性的考量，在擴充性方面，Gfarm 應用在 PRAGMA(The Pacific Rim Application and Grid Middleware Assembly)、Particle Physics 的實驗上及 GPFS-WAN 在 TeraGrid[6]的大規模測試，表示此 2 套檔案系統具備其實用性，因此，透過本次的實驗及分析來評估其優缺點及實用性以供使用者在選擇資料格網檔案系統作為建置評估之參考。

3.1 Gfarm File System

Gfarm 檔案系統是一個開放原始碼軟體(Open Source Software)的網路分享檔案系統(Network Shard File System)，適用於一般的叢集(Cluster)及格網(Grid)的大規模運算及儲存空間環境。

Gfarm 是參考 Grid Datafarm 架構且針對大量資料(Petascale)運算的服務進行開發與研究。Gfarm 是屬於分散式檔案系統，集結數個不同檔案空間來達到檔案分享的目的，並且採用檔案複本的方式來提昇檔案傳輸速度及容錯機制。在 Gfarm 檔案系統的架構如圖 1 所示，架構中有三個主要因素，分別為 Metadata Server、File System Node 及 Gfarm Client。

(1)Metadata Server 負責儲存 Gfarm 內部檔案資訊、使用者資訊、建立使用者、建立檔案系統及管理 Gfarm 使用者及檔案系統。Gfarm 提供 2 個不同使用者認證方式，認證方式分別為 Sharedsecret Authentication 與 GSI Authentication (Grid Security Infrastructure)，讓管理者自行選擇合適的方式來使用。其中 GSI Authentication 適用於 Grid 環境，使用者透過憑證即可通行於各個不同的主機及檔案系統。

(2)File System Node 當 Metadata Server 採用 GSI Authentication 認證時，每個 File System Node 必須向公認組織許可的合法機構申請主機憑證，合法機構如國家高速網路與計算中心之國家高速網路憑證維運中心 <http://ca.goc.nchc.org.tw/nchcca/index.php>。Gfarm 管理者把主機憑證的資訊匯入 grid-mapping 完成後，連結 File System Node 與 Metadata Server，讓原先本地端的儲存空間得以分享至 Gfarm 檔案系統上，讓每個 Gfarm 使用者共同分享使用。

(3)Gfarm Client 當 Metadata Server 採用 GSI Authentication 認證時，Client 必須向公認組織許可的合法機構申請個人憑證，合法機構如國家高速網路憑證維運中心 <http://ca.goc.nchc.org.tw/nchcca/index.php>。當使用者取得個人憑證後，向 Gfarm 管理者申請新增使用者，Gfarm 管理者把使用者憑證的資訊匯入 grid-mapping 完成後，使用者即可透過個人憑證來使用 Gfarm 的檔案系統。

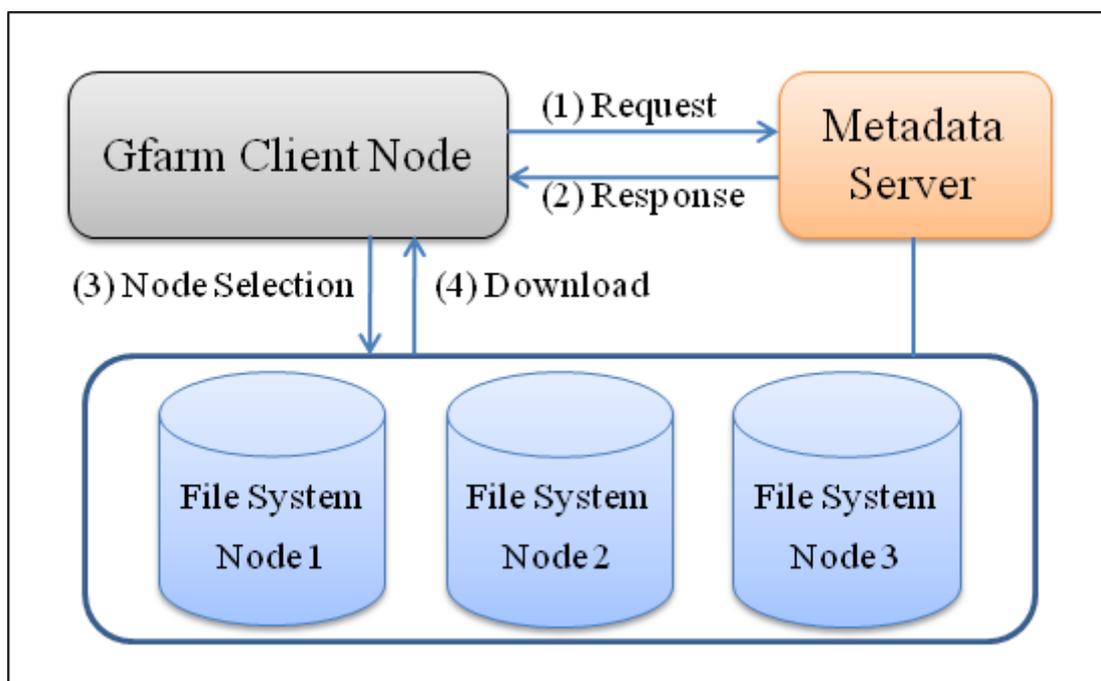


圖 1. Gfarm 流程圖

Gfarm 的運作流程如圖 1 所示，在確認使用者憑證已置入指定位置後，並且確認 grid-proxy-init 已經成功，即可登入 Gfarm 執行檔案系統相關指令的測試及使用。

首先，步驟(1)使用者向 Metadata 提出請求，例如 gfls 指令，步驟(2)Metadata Server 對使用者的請求作出適當的回應，當使用者送出 gfls 指令，Metadata Server 立即列出目前 Gfarm 檔案系統內的檔案資訊。

步驟(3)倘若使用者已選擇特定檔案，可先透過 gfwhere 來查詢檔案所儲存的實際位置。步驟(4)選定檔案所屬的主機後，即可進行檔案下載的動作。由於 Gfarm 是採用檔案複本的機制來提昇檔案傳輸速度與維護資料的安全性，在產生檔案複本的方式分為 2 種，分別為把檔案複本隨機產生於 Gfarm 內的檔案系統，另一種則是可指定把檔案複本產生於特定的主機上，透過產生複本的方式可平衡每個檔案系統的負載並且可提昇整體的檔案傳輸速度。

Gfarm 開發者為提供讓使用者便利的方式，在執行檔案上傳與下載的動作時，利用 Gfarm-FUSE 來取代繁複的指令動作，把上傳與下載檔案的動作簡化為一般使用者透過 Gfarm-FUSE 把 Gfarm 檔案系統直接掛載至使用者的本地端來執行，解決使用者對於 Gfarm 指令使用上的問題，並且經由設定後於每次開機後自動把 Gfarm 檔案系統掛載到本地端，節省使用指令上繁複的步驟。

在一般的 Network Shared File System(網路分享檔案系統)存在著檔案權限及檔案擁有者的潛在問題。在 Gfarm 檔案系統可依不同使用者來給予適當的使用權限以達到保護每個擁有者所屬的檔案安全性。使用上如同在主機上，每個使用者在檔案的存取上有著不同的權限，在檔案分享的過程可保持適當的隱私及安全性，讓使用者在存取檔案的過程中，避免機密檔案與資料外流之虞。

3.2 GPFS-WAN File System

IBM 於 1993 年開發的 GPFS (General Parallel File System) 在諸多叢集檔案系統中，傳輸效能及速度一直都是檔案系統的指標性代表，近年來 IBM 公司仍不遺餘力的持續發展，並且和 San Diego Supercomputer Center (SDSC) 進行技術的交流及合作，目標是為提供一個透明的高性能介面，讓使用者在進行存取遠端資料時，無須考慮資料所存在位置以及是否在同一個檔案系統。雙方共同合作為 Data Grid(資料格網)創造出一個嶄新的平台，稱為 TeraGrid，使用的技術仍是以 IBM 開發出來的 GPFS 作為主軸，進而利用網際網路進行各個檔案系統間的資料傳遞媒介，最後共同命名為 GPFS-WAN。

GPFS-WAN 在 TeraGrid 上實現是一個 700 TB 的檔案系統，從分散於全部 TeraGrid 的多種計算和視覺化資源以供所有節點進行存取。TeraGrid 透過 GPFS-WAN，讓使用者在分散於不同平台的資源上執行準備工作、計算任務及進行視覺化的處理。

GPFS 檔案系統在 LAN 的環境下的架構如圖 2 所示，每台主機分別含有四個元件，分別為 NSD、GPFS、Linux 與 Application。在架構中包含三個 GPFS 節點，其中二個節點為 NSD (Network Shared Disk，網路共享磁碟)，其功能為貢獻 Local Disk 儲存空間，讓 NSD Server 集中管理及共享。另外，NSD Server 其主要功能是透過 Private IP 及 GPFS 指令 `mmcrcluster` 來建立主機間連結及設定，把數個 NSD 獨立的 Local Disk 整合為一個檔案共享系統。

叢集電腦環境利用區域網路及 GPFS 來整合，讓同一個叢集電腦環境使用者可利用 GPFS 檔案系統來存取其他 GPFS 主機的共享資料。其中 NSD 在分享空間是以一個硬碟磁區為單位。因此，在建立 GPFS 檔案系統時，

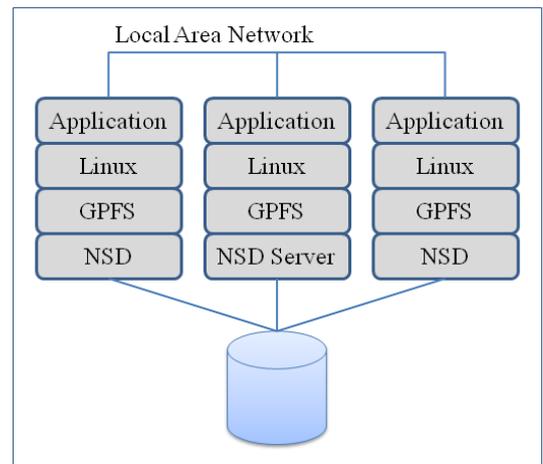


圖 2: GPFS 架構圖

至少需預先保留一個硬碟磁區讓 GPFS 來建立網路共享磁碟或者新增一個新硬碟以供使用。

GPFS 檔案系統有高效能的檔案存取速度、安全性及資料一致性，因此 IBM 開發團隊讓原本在 LAN 下執行的 GPFS 採用 TCP/IP 網路傳輸來連結叢集電腦環境的網路傳輸檔案系統，GPFS 把原本數座叢集電腦藉由 WAN (Wide Area Network) 的連結來達成 GPFS-WAN 的檔案系統架構。

在建置 GPFS-WAN 的架構前，必須將獨立的 GPFS Cluster 建置完成後，隨之把數個 GPFS Cluster 串接成一個 GPFS-WAN 的環境。數個 Cluster 進行串接時，Cluster 的 GPFS Server 必須具備 Public IP，Cluster 之間透過 TCP/IP 來進行連結與互動，GPFS Server 使用 Public IP 與遠端 Cluster 進行連結後，並且使用 Private IP 與 Cluster 下的 GPFS Node 進行資料共享以達到 GPFS-WAN 網路資料共享的目標。

在每個 Cluster 各自建立 GPFS 檔案系統完成後，隨之如何使用 TCP/IP 網路傳輸通訊協定來連結數個不同 Cluster 的 GPFS 檔案系統成為極為關鍵的因素。在連結數個不同 Cluster 的 GPFS 檔案系統需循下列步驟來完成設定。GPFS-WAN 之架構圖如圖 3 所示。

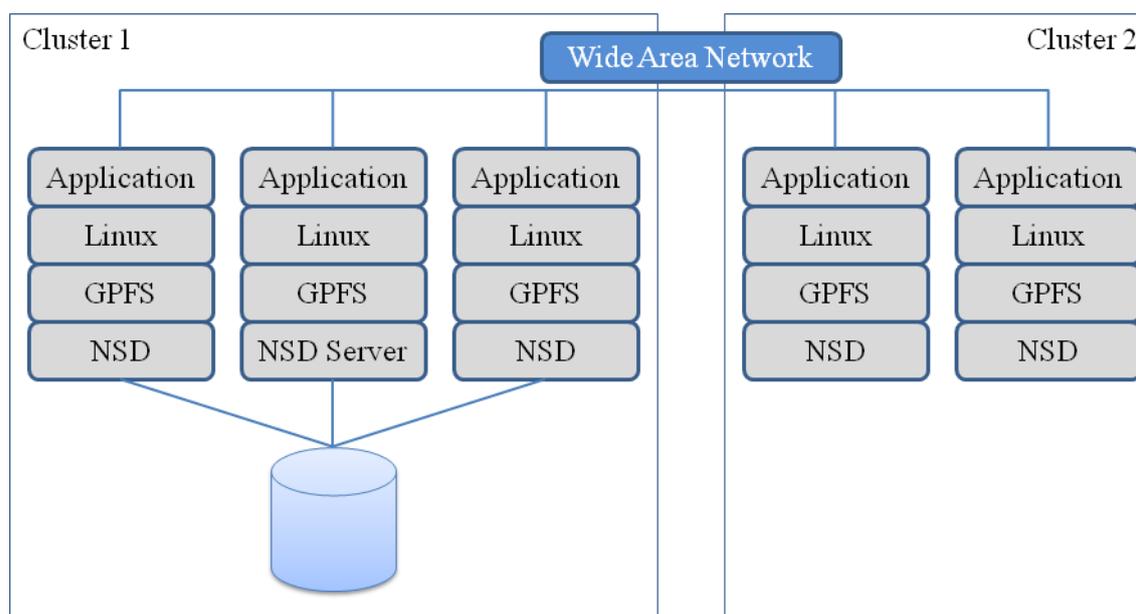


圖 3. GPFS-WAN 架構圖

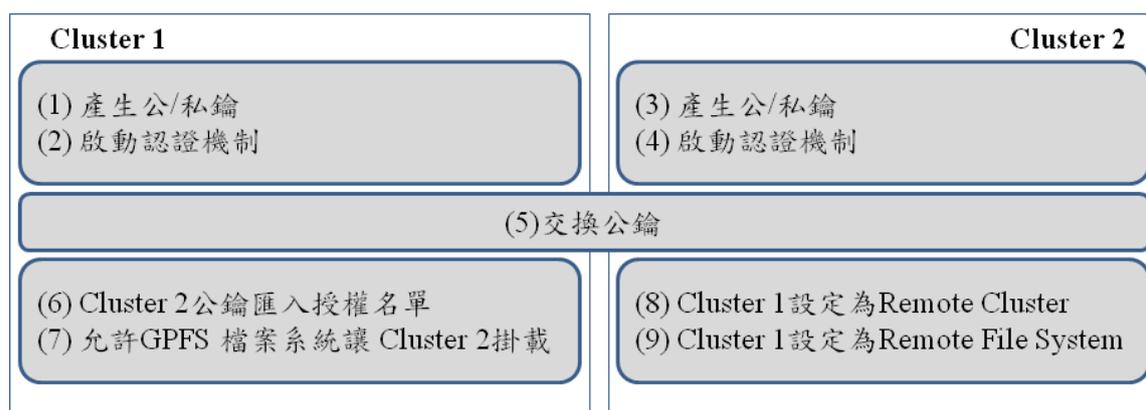


圖 4. GPFS-WAN 流程圖

2 個建置完成的 GPFS 檔案系統分別為 Cluster 1 和 Cluster 2，把 2 個不同 Cluster 的 GPFS 檔案系進行連結與溝通的步驟如圖 4 所示。建立一個 GPFS-WAN 總共需要 9 個步驟，(1)至(4)的步驟是 Cluster 1 和 Cluster 2 分別需要在各自的主機上產生用來溝通認證使用的公鑰和私鑰，隨之主機啟動認證機制，讓 GPFS 檔案系統採用公鑰和私鑰的認證機制來驗證 GPFS 檔案系統的存取權限。第(5)個步驟，2 座 Cluster 必須交換公鑰且把對方的公鑰儲存於/var/mmfs/ssl 資料夾下。

第(6)個步驟，Cluster 1 把 Cluster 2 的公鑰匯入授權名單作為主機認證之檢驗依據。第(7)個步驟，Cluster 1 開啟本地端的 GPFS 檔案系統讓遠端經過主機認證的 Cluster 2 來掛載使用。第(8)個步驟，Cluster 2 把 Cluster 1 設定為遠端 Cluster 且把 Cluster 1 的公鑰匯入授權名單。第(9)個步驟，Cluster 2 把 Cluster 1 允許遠端掛載的 GPFS 檔案系統設定為遠端檔案系統。最後，Cluster 2 透過 mmmount 的指令即可把遠端 Cluster 1 的 GPFS 檔案系統掛載至 Cluster 2 的本地端，即完成 GPFS-WAN 建置。

4. 效能評估與測試

每個 File System Node 必須使用 Public IP 與 Server 進行連結。本實驗環境共有 7 個 File System Node，在配置主機方面，新竹設有 Node 1 至 Node 5 共 5 個 File System Node，此外 Node 6 與 Node 7 則置放於台中。把檔案系統分別置於異地的主要原因是為達到資料格網目標，把分散於各地的資料透過網際網路的傳輸媒介來達到資料匯集及共享的目的。

4.1 主機系統配置

在主機系統配置，如表 1 所示，在作業系統上使用 2 種不同的核心版本進行測試，在 CPU 位元數分別採用 32 位元及 64 位元來進行相容性的測試來符合實際資料格網需整合異質性系統的目標。

4.2 Gfarm 建置與設定

在 Gfarm File System 實驗環境如圖 5 所示。為達到資料格網分散共享資源的目標，在實驗中配置 7 個 File System Node，其中 Node 7 作為 Metadata Server 來管理 Gfarm File System 的使用者及 File System Node。在 Metadata Server 須具備 OpenSSL library 與 OpenLDAP library，此外在 OpenLDAP library 與 PostgreSQL library 擇一使用。在認證方面，Gfarm 是採用 GSI Authentication 需要參照 Globus library。

在 Gfarm File System Node 必須申請合法的主機憑證及指定特定目錄為 Gfarm 共享目錄，指定的共享目錄不得為共用目錄。Gfarm 檔案系統的優點為使用已存在的磁碟空間，而不須額外預留磁區。

在使用者設定方面，使用者必須向 Gfarm 管理者申請開啟使用者帳號及必須向合法認證組織申請使用者憑證才得以使用 Gfarm 檔案系統。

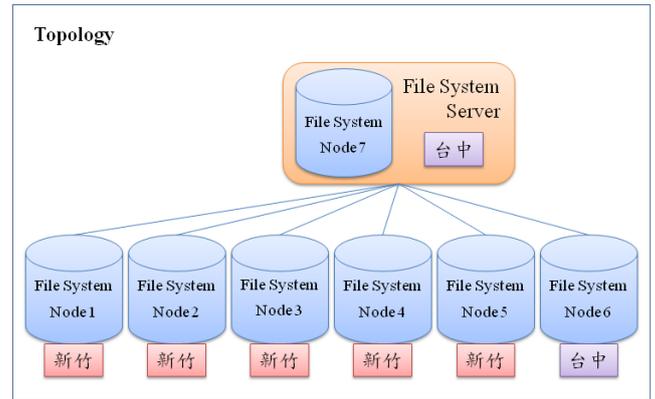


圖 5. 資料格網實作環境

表 1. 主機系統配置

主機	Node 1 ~ Node 5	Node 6、Node 7
作業系統	CentOS 5.3	CentOS 5.2
CPU 位元數(Bits)	32	64
核心版本	2.6.18-128.1.10.el5	2.6.18-92.el5
Gfarm 版本	gfarm-2.2.0	gfarm-2.2.0
FUSE 版本	fuse-2.7.4	fuse-2.7.4
Gfarm-FUSE 版本	gfarm2fs-1.1.0	gfarm2fs-1.1.0
GPFS 版本	gpfs-3.2.1-7.i386	gpfs-3.2.1-3.x86_64

4.3 GPFS-WAN 建置與設定

GPFS-WAN 實驗環境如圖 5 所示。在建置 GPFS-WAN 檔案系統時，必須把各自獨立的 GPFS 建置完成，當 GPFS 共享檔案系統時，需預留磁區以供 NSD(Network Shared Disk)來使用。倘若系統未預留空白磁區，則有 2 種方式可解決，首先是把硬碟重新切割後，預留一個空白磁區供 GPFS 使用，或另外新增一個硬碟以供 GPFS 來使用。因此建置 GPFS 時，應預先保留空白磁區以避免這類問題產生。此外，在建置過程 GPFS 使用的傳輸埠為 1191，必須開通此傳輸埠才能得以順利達成 GPFS 檔案系統的設定。獨立的 GPFS 檔案系統測試完成後，隨之在設定 GPFS 檔案系統間，每個獨立的 GPFS 檔案系統的 Server 必須使用 Public IP 來進行傳輸與溝通。同一個 GPFS 檔案系統內的節點可使用 Private IP 來溝通與連結。

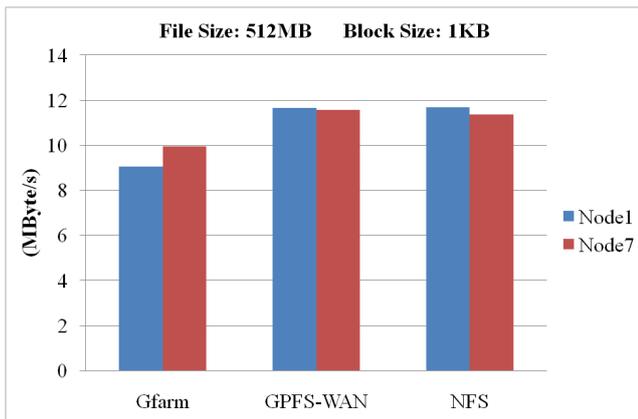


圖 6(a)

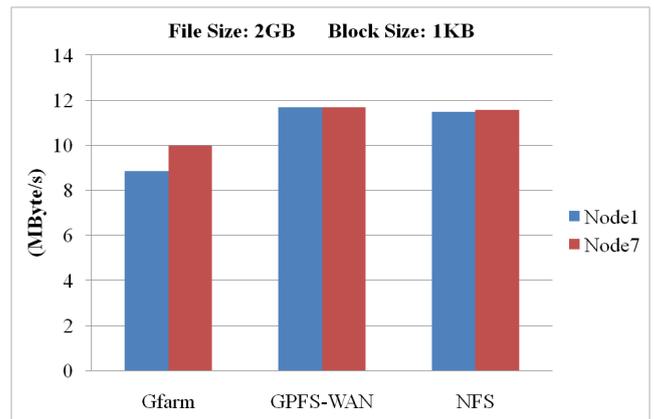


圖 6(b)

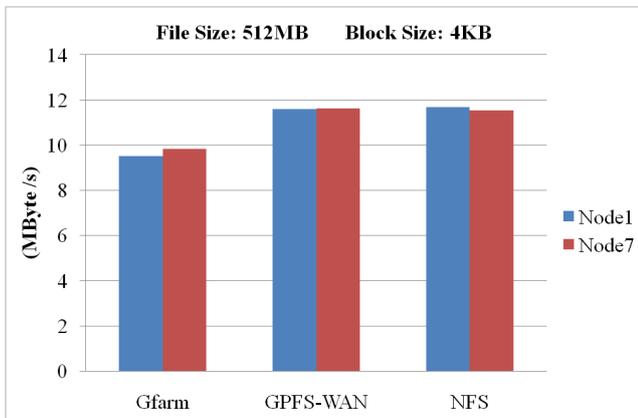


圖 6(c)

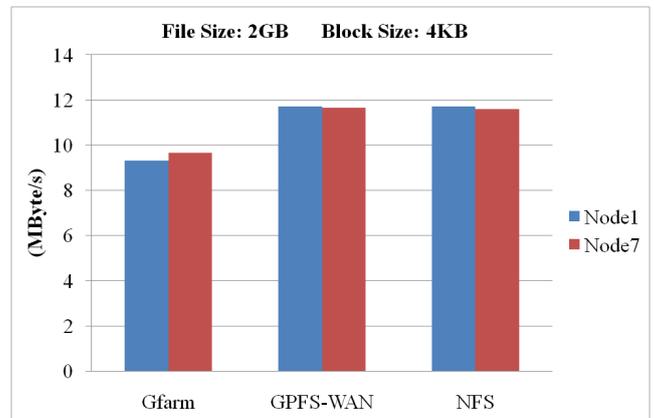


圖 6(d)

4.4 實驗數據與分析

在實驗內容分別在 Gfarm 及 GPFS-WAN 環境建置完成後，本實驗把特定的測試檔案分散於實驗環境中，隨之，讓使用者登入系統進行檔案複製的動作，分別在複製檔案的過程量測其實際檔案傳輸速度。本實驗的測試檔案分別為 512MB 及 2GB 的檔案，Block Size 分別為 1KB 及 4KB。實驗結果如圖 6(a)、(b)、(c)及(d)所示。在實驗過程中，量測三種不同的檔案系統分別為 Gfarm、GPFS-WAN 及 NFS。

資料格網主要目的是把分散於各地的檔案匯集在同一個平台上，若非檔案系統有產生檔案複本的功能，實際上檔案仍然分佈於各自的本地端的檔案系統。因此，為達到實驗的公平性，在三種檔案系統中，惟有 Gfarm 檔案

具備檔案複本的功能，然而在本實驗過程中，Gfarm 即不啟動其檔案複本的功能。

本篇論文主要探討內容以 WAN 的網路架構之檔案系統量測與分析。因此，分別把測試檔案存放在 Node1 及 Node7 的檔案系統，其主要目的是讓檔案分散於台中與新竹兩地，以利於使用者透過網際網路來進行異地檔案傳輸。

在實驗過程為避免使用者登入的主機即是複製測試檔案所存在的主機，避免使用者在複製檔案過程，未經由網際網路傳輸而直接透過本地端網路傳送進行複製的動作。因此，限定使用者從其它節點(Node 2、3、4、5、6)登入進行測試以避免產生不公平的情況。

依實驗結果所示，在不同檔案大小(512MB、2GB)與 Block Size(1KB、4KB)，其

測試結果三種檔案系統以 GPFS-WAN 檔案系統整體效能最佳。

實驗結果 6(a)、(b)、(c)及(d)所示，NFS 與 GPFS-WAN 結果相近，皆領先 Gfarm 檔案系統。Gfarm 雖然在單一檔案傳輸速度不敵 GPFS-WAN 及 NFS。在大規模的資料量下，倘若納入檔案複本功能，Gfarm 檔案傳輸效能極可能優於其它的檔案系統。

本實驗測試為單一檔案透過不同檔案系統進行複製檔案的動作來量測檔案系統實際檔案傳送的速度。因此，實驗結果只能呈現出單純的檔案傳輸的效能，但在評估資料格網環境尚有許多因素必須考量以作為整體規劃重要的參考，其它重要評量因素如表 2 所示。

5. 結論與未來展望

在本文中，探討數個檔案系統，如 NFS、SRB、iRODS、Lustre、Gfarm 及 GPFS-WAN，在經過慎審的評估資料格網所需的特點，如擴充性、安全性及穩定性的考量，此外，外界組織接受度及實際使用的狀況皆列入考量因素。

最後選擇 Gfarm 與 GPFS-WAN 作為重要指標，進行實際建置、效能測試及整體評估，在效能部份，GPFS-WAN 優於 Gfarm，在整體評估中，發現 GPFS-WAN 為商業軟體、需預留磁區以及不提供檔案複本機制的潛在問題，使用者若有預算或未來功能開發的考量，則可考慮使用 Gfarm 檔案系統來達到資料格網集結分散資料的效果。

Gfarm 檔案系統的開發團隊仍持續開發新功能來達成不同研究領域及對象的需求，Gfarm 也是一個開放原始碼專案，使用者可自由進行修正問題及開發新功能的動作。

GPFS 目前仍為商業軟體，倘若 IBM 把 GPFS 開放原始碼讓廣大的民眾使用，相信對資料傳遞的技術發展具有極為正面的影響。

表 2. Gfarm 與 GPFS 整體評量

	Gfarm	GPFS-WAN
Open Source	Yes	No
主機問驗證機制	GSI Authentication	Public/Private Key
使用者驗證機制	GSI Authentication	GPFS Server
檔案系統特定Port	600	1191
Public IP	Yes	Yes
預留磁區	No	Yes
檔案複本功能	Yes	No
查詢檔案儲存位置	Yes	No
動態新增檔案系統節點	Yes	Yes

6. 參考文獻

- [1] Gfarm <http://datafarm.apgrid.org/>
- [2] GPFS <http://www.ibm.com/us/en/>
- [3] iRODS <https://www.irods.org/index.php>
- [4] <http://www.sun.com/software/products/lustre/>
- [5] http://www.sdsc.edu/srb/index.php/Main_Page
- [6] <http://www.teragrid.org/userinfo/data/gpfswan.php>
- [7] Shohei Nishida, Nobuhiko Katayama, Ichiro Adachi, Osamu Tatebe, Mitsuhsa Sato, Taisuke Boku, Akira Ukawa, "High Performance Data Analysis for Particle Physics using the Gfarm file system", Journal of Physics: Conference Series, 119, 062039, 2008.
- [8] Atsuko Takefusa, Osamu Tatebe, Satoshi Matsuoka, Youhei Morita, "Performance Analysis of Scheduling and Replication Algorithms on Grid Datafarm Architecture for High Energy Physics Applications", Proceedings of the 12th IEEE International Symposium on High Performance Distributed Computing (HPDC-12), pp.34-43, 2003.
- [9] Yusuke Tanimura, Yoshio Tanaka, Satoshi Sekiguchi, Osamu Tatebe, "Performance Evaluation of Gfarm Version 1.4 as a Cluster Filesystem", Proceedings of the 3rd International Workshop on Grid Computing and Applications (GCA 2007), World Scientific Publishing Co. Pte. Ltd., pp.38-52, 2007.