

應用格網技術於數位典藏系統

張瑞雄

王智敏

林春福

黃甯園

林皓嶽

國立東華大學資訊工程系

國立東華大學資訊工程系

國立東華大學資訊工程系

rschang@mail.ndhu.edu.tw

d9221007@em92.ndhu.edu.tw

{d9321010, m9321049, m9321004}@em93.ndhu.edu.tw

摘要

數位典藏技術用以保存歷史文物，資料檔案文件或影音記錄等資料。對於龐大的影音資料，傳統的方法是以單一儲存硬體設備建置網站與資料庫，當儲存空間無法負荷時則必需擴充硬體設備。此外對於各種不同的典藏應用均採各別獨立的網站建置，沒有一個整合性的系統存在，一個典藏檔案可以分散儲存、有效管理與搜尋檢索的系統。格網計算(Grid Computing)可被視為下一代的分散式網路的應用，於格網系統運作環境下，我們將串連許多儲存設備形成一個大型的儲存系統以作為數位典藏技術所需的資料儲存空間。基於上述的理念，本論文擬結合數位典藏技術、全球資訊網服務與格網技術，建構完成一數位典藏格網的應用架構，提供文物與大型檔案數位典藏的環境。

關鍵字

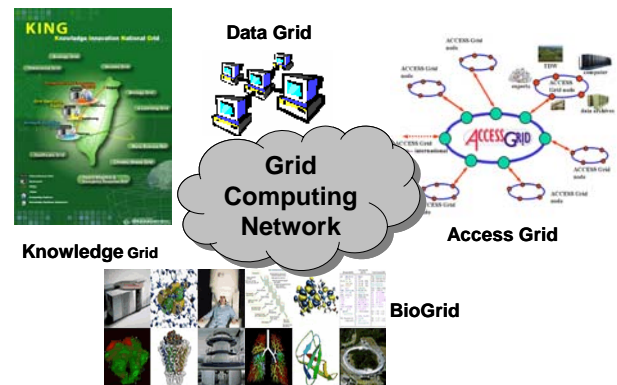
數位典藏、格網、全球資訊網服務、資料備份(replica)

1. 前言

隨著文明的進步，人類運用了許多方法紀錄歷史資料，最早期的石刻、羊皮、竹簡、布帛，直到紙張發明後，才配合書寫工具大量記載流傳。此後，科技進步，影音設備、無線設備、網路與電腦更便利了人類的生活。值得深思的是，舊資訊技術記錄的內容諸如歷史文物，資料檔案文件或影片記錄等，這類文物往往易隨著時間而損壞，長久保存不易，該如何利用近代的科技技術將其重新呈現再運用，於是掀起了一股數位典藏技術的研究。近年來，數位典藏已成為國家大力推行的研究計劃。自民國八十五年起，中央研究院開始推動數位典藏計畫，進行檔案數位化典藏，至目前已經完成戰後台灣經濟發展相關資料庫、實業部及經濟部商業司公司登記資料庫(1926-1949)等[1][2]。國科會亦自民國91年1月1日起開始推展「數位典藏國家型科技計畫」[3][4]，將國家珍貴文物典藏數位化，藉以保存文化資產，融合人文與資訊科技，此外亦配合政府所推動之

數位台灣計畫，提升典藏技術研究的發展，並透過各種加值的應用，創造豐富的產業商機。

網路是先進國家日常生活必備的工具，網路的普及率和使用率，其帶來的便利性和安全性，亦已成為衡量國家發展潛力的重要指標之一。政府利用網路施政、以及網路教學等已是現代化的常態。而學術研究網路，具有教育性、實驗性、前瞻性與未來性之特質，為將來國內網路商業化及制度化之先鋒部隊，扮演著國家推動知識經濟之基礎命脈。目前，「格網」(Grid)技術已是國際間網路應用於創新科技之主流，所謂「格網」即是透過寬頻網路連結分散之相關資源：如儲存設施與知識庫、計算設施、儀器設備、軟體、以及專業人力等整合為一共享之應用平台。此即符合資源有效整合與共享、互利之合作式趨勢。



圖(一)、格網應用的環境。

數位典藏通常牽涉到大量資料的存取，同樣的一份典藏檔案也希望能夠儲存到不同的地方以增加其安全性和使用性，因此一個能夠儲存、取用和管理分散式數位典藏系統是非常重要的。另外目前由於電腦的普遍，將分散在各地的電腦系統資源整合起來將可發揮巨大的力量，格網系統(Grid Computing System)的目的就是希望能有一套系統來好好利用分散在各地的各種

資源，所以本論文的目的就是希望將格網系統的技術應用在數位典藏上，讓分散式的數位典藏變的更可靠，更方便，也更有效率。格網目前可提供下列已趨成熟之專業應用：如圖(一)所示，諸如生物資訊、遠距教學、醫療資訊、生態觀測、防救災資訊等，國內所發展的格網系統，如國家高速網路與計算中心推廣的 Taiwan Unigrid 環境，其也可以和先進國家的格網系統接軌，推動國際合作與科技外交，使台灣發展成為亞太創新、知識流之重鎮。

2. 數位典藏與格網系統

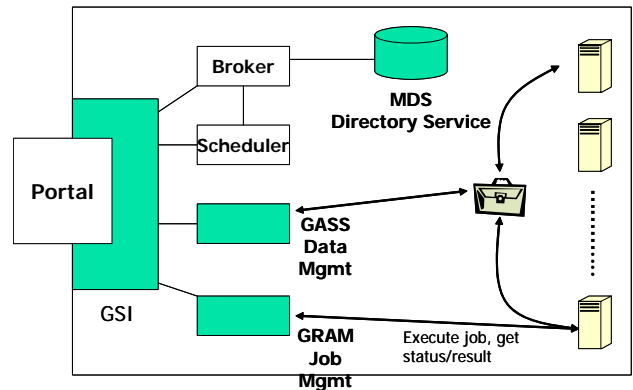
為了永續保存國家各種文物豐富之收藏，提供學術界研究，以及提供給更多人可以欣賞到珍貴的歷史文物，數位典藏的技術發展是有其必要性。數位典藏的研究隨著網際網路的發展逐漸受到大家的重視，各先進國家無不熱烈支持有關的各項計畫。我國有許多機構在從事珍貴藏品數位化的計畫，例如數位博物館專案計畫、國家數位典藏機制專案。將資源數位化並呈現於全球資訊網上是首要的工作，但是更重要的是：必須針對這些資源的特性加以組織與整理，才能有效地檢索、使用。

在數位典藏的數位資料中，不乏有大型檔案容量的一些數位資料[11][12]，要如何去有效管理這些巨型資料，以及如何去分散處理、儲存、複製這樣龐大的資料量，是需要一個完整且具安全性的系統平台來做整合，目前在分散式架構下的格網系統，已經逐漸發展成一個成熟的分散式運算的系統平台，可以提供在數位典藏下如此大型且繁雜的資料的儲存與檢索和存取等工作。將數位典藏的資料複製與管理，轉至格網系統上是很值得探討的一個架構與研究方向。

在目前格網系統的研究中，Globus[5]組織提供 Globus Toolkit 這個工具讓建置高運算格網與格網相關程式變的容易許多。如圖(二)所示，Globus Toolkit 提供了三個基本的服務，分別是 GRAM、MDS、GridFTP。其中，GRAM 為 Globus Resource Allocation Manager，是一資源管理協定，會自行轉換使用者需求及所需資源，成為區域電腦中可以了解的指令。MDS 為 Globus Metacomputing Directory Service 以 LDAP 的訊息主從架構為基礎，提供了建置高運算格網所需的工具。MDS 使用 LDAP 協定作為統一的橋樑，使大量的不同的系統成員可以查詢系統資料，且可以建造有關在許多組織不同系統間資源資訊的資料庫。GridFTP 為資料傳輸協定，提供高效能、安全及健全的資料傳輸機制。此外，在安全方面，這些協定均使用 GSI 安全協定(X.509 Proxy Certificates)，以確保安全認證。

如圖(二)所示，網格平台入口介面(Portal)提供一個整合性的網站入口，以作為資訊整合，資訊交換以及提供使用者存取虛擬機構所共享資源的介面。利用網

路無遠弗屆的連通特性，建立虛擬機構的入口網站將可以達到此一目的。使用者可透過入口介面監看格網的狀態，也可以透過入口介面提送工作至 Broker，經由 Broker 的運作，查詢 MDS 資料庫，選擇適合的機器執行(Scheduler)，並將結果回傳給使用者。



圖(二)、開發格網應用服務的元件圖[6]。

數位典藏與格網雖為兩種不同的技術，但卻息息相關，可以相互配合應用。就其整合的優點我們歸納以下幾點：

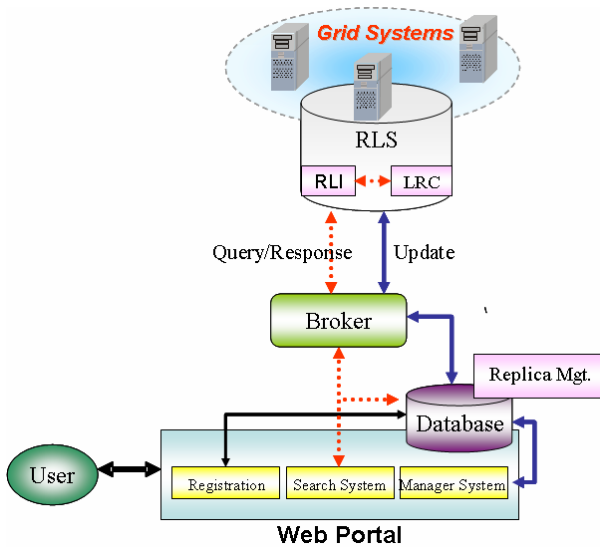
- (1) 超大型的儲存系統：格網技術將異質性平台的儲存設備聚集形成一個超大型的儲存系統，可提供數位典藏資料所需的儲存空間，建置所需的資料庫。
- (2) 安全性：Globus Toolkits 為現今格網環境建置的標準工具，其內部元件 GSI (Grid Security Infrastructure) 提供了良好的認證、授權及資料加密的安全機制。GSI 透過使用公鑰密碼系統，特別是公鑰/私鑰和 X.509 憑證，建立了格網安全的基礎。
- (3) 可靠性：透過複製(Replication)管理的機制，資料可被複製儲存於不同的站台，以確保資料不會因站台硬體損壞而導致資料損毀。同時，Globus Toolkits 亦提供 Reliable Transfer Service (RFT)，一個可靠的資料傳輸方式。
- (4) 檔案取得快速：GridFTP 為 Globus Toolkits 提供的服務之一，其支援檔案平行傳送，加快檔案傳送的時間，亦是格網環境的優異之處。

3. 系統架構

在數位典藏與格網系統主要可分為二階段進行研發：

第一階段：格網系統的建置、資料庫的建置以及 Web-based 服務的建置。使用者在通過註冊後可登入檢索系統瀏覽資訊，經由 Web-based 服務連結格

網系統的資料庫以取得檢索的資料。在格網系統的建置中採用 Globus Toolkits 建置，並發展 Broker 的功能以分散資料儲存、檢索與資料收尋的工作。格網環境中資料庫的建置則採用 Globus Toolkits 中 Replica Location Service (RLS)的建置方式，建立一個階層式的資料庫。Web-based 服務則提供使用者一方便的操作界面，包括使用者註冊，檢索系統以及管理者系統，如圖(三)所示。



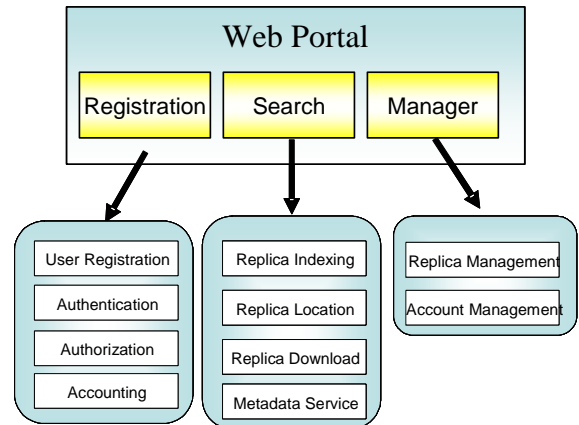
圖(三)、數位典藏與格網技術應用之架構圖。

第二年階段：透過第一階段所建置的環境，在第二階段重點將針對大型檔案的下載及複製等傳輸機制。為了提高系統檔案的可信度(Reliability)與一致性(Coherence)，我們研發檔案複製的機制與更新的方法。當檔案產生儲存入資料庫後，依照使用者存取檔案的頻率與格網環境各 Sites 頻寬的資訊，我們研發檔案複製的機制。而更新的方法則依據檔案的大小以及更新的頻率而有不同的選擇。當檔案在進行下載時，為了加快檔案的傳輸，我們可以透過 Replica Selection 排程選擇的方法，採用平行下載的方式，從多個傳輸率較快的站台分區塊取得所要的檔案。此階段為格網方面的研究，我們也發表了相關的論文成果[13][14][15]。

結合此二階段計劃之成果，將可研製一套大型儲存系統，高可信度、高傳輸率的數位典藏格網應用環境。在本論文中，我們致力於第一階段系統的建置。此外於研究過程中，我們將發掘未來於數位典藏與格網技術所可能面臨的問題及其解決的技術與方法，以作為未來數位典藏與格網實用系統架設與運作之重要參考。

以下我們就數位典藏格網系統第一階段，描述其系統環境與子系統(圖(四))：

- (1) Replica Location Service (RLS)
- (2) 會員註冊系統 (Registration)
- (3) 檢索系統 (Search)
- (4) 管理系統 (Manager)



圖(四)、Detail Architecture of Web Service System。

以下我們就上述子系統做詳細的說明：

3.1 Replica Location Service

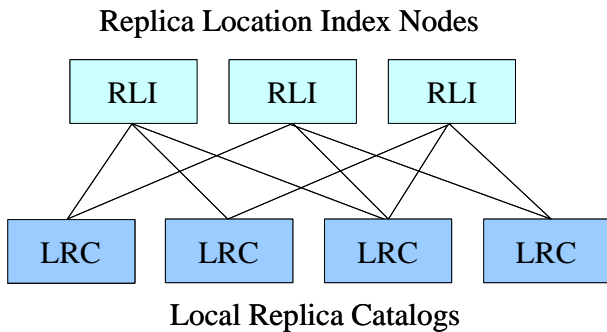
數位典藏資料的儲存、複製和檢索系統的製作中，主要是以 Globus Toolkits 來建構格網系統環境，其中 Globus 亦提供了一階層式 Replica Location Service (RLS)的資料庫系統[5]，做為分散的檔案查詢與儲存的管理系統。RLS 提供一種對映的方法，使得每筆資料的資訊能夠從邏輯名稱(Logical File Name, LFN)對映到實體目標名稱(Physical File Name, PFN)。目標名稱可能代表著實體資料的真正位置 (Physical Locations)，或者只是另一個 RLS 的入口 (Entry)，對映到另一個層次的資料邏輯名稱。

採用 RLS 的主要原因如下：

- (1) 對於分散式系統而言，RLS 能避免單點失敗的情況發生，且增加較大的可靠度，也提供較佳的負載平衡 (Load Balancing)、效能 (Performance)以及可塑性 (Scalability)。
- (2) RLS 的實作技術是使用開放原始碼 (Open Source)的關聯式資料庫。
- (3) RLS 能夠嚴謹的區分複製的資訊以及其他型態的 Metadata，與原本的複製品目錄 (Replica Catalog)不同，RLS 並不包含關於邏輯收集 (Logical Collection)的資訊，而是假設這類的資訊是由其他 Metadata 服務所儲存。

如圖(五)所示，格網環境中 RLS 架構主要是由一些 LRC 與 RLI 所組成，我們可以挑選部分站台扮演 RLI

的角色或是 LRC 的角色，也可以同時具有兩種身份，用來記錄系統中檔案與複製品的實體位置資訊。以下就 RLS 架構以及重要的元件與機制做詳細的介紹：



圖(五)、階層式 Replica Location Service

(1) Local Replica Catalogs (LRCs)

本地端目錄紀錄儲存系統內任意命名的邏輯檔案名稱 (Logical File Name, LFN) 與實體檔案名稱 (Physical File Name, PFN) 之間的對映。

(2) Replica Location Indices (RLIs)

每一個 RLI 都會利用索引 (Index) 的方式紀錄如何從 LFN 對映到 LRC 的方法，RLI 的索引數量大小、重複性資料的多寡以及 LRC 更新時間的切割長短，都將會影響建立出來的索引結構，而不同的索引結構會有不同的執行效能。

(3) Soft State Maintenance of RLI State

LRC 透過「Soft State Protocols」將目前的狀態資訊傳送到 RLI 作紀錄，RLI 會週期性的更新狀態的資訊，以記錄最新的狀態資訊，由圖(五)可看出 RLI 與 LRC 之間的關係。

(4) Compression of State Updates (Optional)

LRC 在傳送「Soft State Updates」給 RLI 之前可以決定是否使用「Bloom Filters」的壓縮方法，利用傳送概括的資訊取代傳送完整的資訊，以達到資訊壓縮的目的，RLI 可以透過壓縮的技術降低系統儲存資訊所需的成本。

(5) Membership and Partitioning Information 的維護

藉由發展 Membership 的機制來紀錄與維護分散式系統中有關於 LRC 和 RLI 的靜態資訊 (static information)，並且利用 OGSA 的機制來實作對於服務的註冊以及服務時間的管理。

3.2 註冊系統

用於控制使用者存取數位典藏格網系統的管理程式。在數位典藏系統或是格網系統中，身分認證與存取控制是系統中不可缺少的一部分。當使用者要求註冊時，系統會將使用者的註冊資訊送 Portal 上的 MySQL 中紀錄。目前註冊系統帳號管理與格網 CA 使用者帳號管理是分開獨立的。原由是，在格網系統下我們採用統一帳號密碼登入，若非格網系統成員之使用者欲藉由 Portal 使用本系統之服務，其也必需取得 CA 認可之帳號才可使用格網系統提供之服務，但若其取得 CA 認可帳號密碼後便可藉由遠端操控程式 (如 putty) 任意的登入使用所有格網系統上的機器，此一行為在安全性考量上稍有不足，因此我們採用 web 介面製作一使用者註冊系統，使用者必須通過 Portal 上的註冊程序取得使用者帳號密碼，並用此帳號密碼登入本系統。使用者登入後，在 Portal 後端我們再以 CA 認可的帳號密碼跟 CA 服務端進行 Proxy-init 取得系統的使用權，如此可避免系統上真正帳號密碼洩漏之問題。

以下就註冊系統子項做細部的說明：

- (1) User Registration：使用者註冊資訊，取得登入的帳號密碼憑證。
- (2) Authentication：驗證使用者的登入資訊，以防止不當的使用存取。
- (3) Authorization：對於使用者與管理者給予不同的授權認證，分別對於數位典藏格網系統也有不同的存取權限。
- (4) Accounting：當考慮到數位典藏格網系統是收費性質的商業系統，Accounting 可以作為計費的一個機制。

3.3 檢索系統

以目前的檢索技術而言，提供查詢的方法不外為：全文檢索或進階檢索。全文檢索可以不必對資料做描述，但檢索效能較差，對於無文字的影像、聲音或視訊資料而言，全文檢索技術也無法派上用場，所以，進階檢索以人工描述 Metadata，建立欄位化的書目資料，對於數位圖書館而言，是絕對必要的工作。因此，在一數位圖書館或博物館等系統中，Metadata 扮演重要的角色。Metadata 的訂定需由瞭解使用者需求及資訊屬性入手，同時也需考慮到資訊系統之間的互通性，所以相關標準的採用非常重要，如國際間已發展出來的格式：Dublin Core (Dublin Metadata Core Element Set) [7]、EAD (Encoding Archival Description) [8]、FGDC (Federal Geographic Data committee) [9]、GILS (Government Information Locator Service) [10]、以及 metadata 的交換語言，如 SGML/XML/HTML 等。

Metadata 在檢索的時候扮演很重要的角色，所以我們藉由設計 Metadata 的欄位，來定義對檢索時所需的檢索條件。所謂的檢索條件，也就是對資料庫中的資料內容，依照資料的特性做分類的依據。

設計的檢索條件如下：

- (1) 檔案類型：使用者可以依照所需要的檔案副檔名類型做為檢索條件，檢索使用者所需要的檔案類型，如.txt與.mpg等。
- (2) 檔案建立日期：使用者可以藉由檔案建立日期欄位作為檢索條件，檢索最新更新更新的檔案。
- (3) 檔案大小：使用者可針對檔案大小搜尋符合條件的檔案。
- (4) 檔案關鍵字：使用者可鍵入關鍵字搜尋比對資料庫中 Metadata 內容符合的檔案。
- (5) 檔案類別：對典藏的檔案歸類索引，例如：文學類、歷史類、電影類等，讓使用者可以依類別做為檢索條件，加快檢索的速度，以及檢索的準確性。

使用者可以使用全文檢索，針對單一檢索條件，查詢出所有符合該檢索條件的所有檔案。使用者也可以利用進階檢索條件，也就是多個檢索條件，更準確且更快速的找出所需要的檔案。

3.4 管理系統

- (1) Replica Management：提供一個易於管理的介面，讓管理人員可以透過此介面，對格網各站台中的典藏資料作管理。新增、刪除、修改查詢的指令，都可以在 Replica Management 中完成。
- (2) Account Management：此一管理介面提供管理者管理目前典藏系統中，所有使用者與管理者的帳號設定。

4. 系統實作

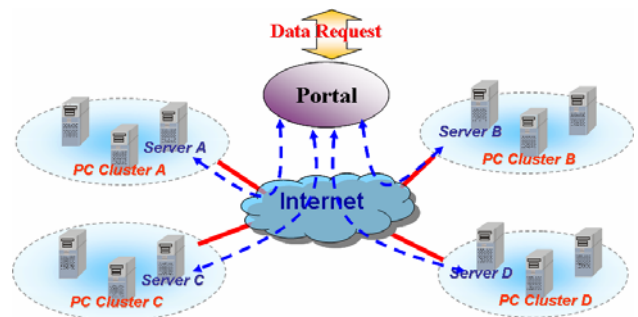
4.1 格網系統

格網環境的建置目的在提供一個穩定及高效能的格網環境，我們建置安裝的軟體套件如下：

- (1) 作業系統: Linux
- (2) 資料庫: MySQL
- (3) 格網中介模組: Globus Toolkit 3.0.2
- (4) 監督程式: SCMS 2.3
- (5) 程式開發程式: MPICH-G2 1.2.5

本系統環境建置在 Taiwan Unigrid 系統下，採用 Linux 作業系統，連結 MySQL 資料庫，安裝 Globus 提供的工具，將參與 Taiwan Unigrid 計劃成員所提供之電

腦串連成一共通的平台，如圖(六)所示。在架設的過程中，此系統必需統一 CA 的認證，各 Grid 機器的使用者，其 CA 認證必須經由一服務端(NCHC)統一認證，服務端提供客服端認可之 CA 憑證。客服端取得服務端之 CA 簽章後，透過此簽章產生三個服務申請的檔案(usercert_request.pem、hostcert_request.pem 及 ldapcert_request.pem)回送至服務端，等候服務端簽發回應後產生的檔案(usercert.pem、hostcert.pem 及 ldapcert.pem)。在收到簽發回應後，將此檔案存放到指定的位置即通過認證，而客服端方可透過 Globus 的程序使用系統上的資源。在系統中，我們採用 SCMSWeb 監督格網系統的運作及其效能。並使用 MPICH-G2 程式開發工具，此程式開發工具是一套 C 語言為基礎且可以使用在格網環境中的程式開發工具，藉由 MPICH-G2 我們可以使用 Globus Toolkits 的服務，並將工作分配到多台機器上運作。



圖(六)、Taiwan Unigrid 環境。

當格網環境建構完成時，若要使用 Globus 所提供的服務時，首先必需使用 CA 認可之使用者帳號密碼登，取得 Proxy 的許可方可使用格網系統中的資源。使用者透過 Globus 提供的 globus-job-run 或 globus-url-copy 等指令送發工作至系統上的機器執行。

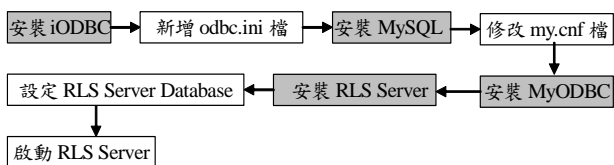
4.2 RLS 系統

RLS 安裝限制必需使用 MySQL 版本 4.0.1 之後的版本，圖(七)為選擇的軟體及安裝的步驟。當 RLS 服務啟動後，透過 RLS 的分散式註冊 (Distributed Registry) 機制，能夠紀錄資料複製品的所在位置，進而從邏輯的檔案名稱找出一個較合適的實體檔案位置，以提供使用者滿意的服務與適合的資源。但是 RLS 無法保證複製的資料是否一致，或者檔案的名稱是否唯一，因此，使用 RLS 的前提是要有較高一層的服務，必須能夠支援 RLS 所無法保證的這些功能。

當使用者登入數位典藏格網系統執行欲檢索資料時，會經由 Portal 資料庫比對檢索的條件，取出 Metadata 的資訊得知檔案的 LFN，而後再透過 RLS 的服務取得檔案的實體位置。一筆檔案會因為複製分散

在格網系統中而查詢到多筆資料結果，經由 Broker 的挑選可以選出較佳的檔案連結。此外管理者在進行檔案分享新增的過程中，必需先對 Metadata 的描述新增一筆資料，而後再經由 Broker 挑選將檔案儲存在某一站台，並根據此站台對映的 LRC Server 新增一筆 LFN 與 PFN 的記錄。在新增資料對應之後 LRC Server 會定期將對應表的更新傳到指定的 RLI Server 上，以便於達到階層式架構的優點：容錯與負載平衡。而 Broker 的運作本身是使用 globus_rls_client command line 來對 LRC 作操作，rls_client 的指令，常用的有 Create, Add, Query, Delete 這些命令可以對 LRC Server 做一些基本的操作。

iODBC	libiodbc-3.52.2
MySQL	mysql-4.0.24
MyODBC	MyODBC-3.51.11
RLS Server	globus-rls-server-2.1.5



圖(七)、RLS 安裝流程

在格網環境中，系統可以透過 Globus Toolkit 的功能，對遠端機器下達資料庫查詢指令。而在 RLS 的架構中，所要查詢的是含有 RLI Server 上 RLI Table。由於 RLS 系統是屬於階層式的架構，從 RLI Server 上面可以查到我們所要搜尋的 LFN 對應是儲存在哪個 LRC Server 之上，然後再到該 LRC Server 查詢 LFN 與 PFN 之間的對應關係。同時 RLS 的查詢也支援萬用字元可以根據我們輸入的 pattern 做部分符合的搜尋因此使用的彈性可以變大。

因此在 Broker 的運作與 Resource 挑選時，首先必須知道那些站台是扮演 RLI 的身份，然後再透過 globus-job-run 的指令，對所有的 RLI server 做查詢，找出所有跟檢索條件相符的 PFN，以達到對所有格網系統中機器查詢的目的。

5. 結論與未來發展

數位典藏格網系統可以整合各種不同的典藏應用，提供一個方便共通的平台供使用者使用。本論文探討了建構數位典藏格網系統的應用技術，技術方面已日趨成熟，目前系統也已在 Taiwan Unigrid 系統下架設完全，在 Linux 運作正常，目前正積極在開發改善

Web Portal 上的功能服務，最終目的要發展出一套接近使用者需求，且易於操作的典藏系統。

承如第三章談及的二個階段，我們完成了第一階段系統建置的工作，爾後，我們將陸續將格網上的研究與技術應用到數位典藏格網系統之中，如下列事項：

- (1) 一致性：當檔案進行複製，而被散落的儲存在不同的站台之中，該如何保證這些檔案在使用的過程中仍可確保檔案的正確性與一致性。
- (2) 平行下載：對於龐大的影音檔案等，採用平行下載的方式取得檔案或進行複製是個有效且省時的方法。
- (3) 檔案分割：對於影音之類可被切割儲存的檔案，使用者可針對不同的需求，選擇真正有用的檔案區塊儲存，以節省系統儲存空間。
- (4) 常駐程式：當系統中不斷的有檔案被進行複製，則該有個常駐程式自動判斷，刪除較不被常用到站台上重複的檔案。

目前，格網技術與數位典藏的技術正不斷的被研究當中，數位典藏可說是格網系統中一個可行的應用，若能將之推廣，相信典藏的技術又將引領出一股潮流。未來我們也將繼續從事這方面的研究與技術的研發，期望可以建立出一個系統更完善，服務更周到，使用更方便，內容更豐富的數位典藏格網系統。

6. 誌謝

- [1] 行政院國家科學委員會補助專題研究計畫
計畫編號: NSC 94-2422-H-259-001
- [2] 國家高速網路與計算中心補助之研究計畫
計畫編號: 94A0227SB
- [3] 國立東華大學資訊工程研究所—無線網路暨網際網路實驗室。

7. 參考文獻

- [1] <http://dipeco.sinica.edu.tw/main/main1.html>
- [2] 中央研究院珍藏歷史文物數位典藏計劃，網址：<http://saturn.ihp.sinica.edu.tw/~dahcr/index.html>
- [3] 數位典藏國家型科技計劃，網址：<http://www.ndap.org.tw/>
- [4] <http://d1m.ntu.edu.tw/d1m/>
- [5] <http://www.globus.org/>, The Globus Project.
- [6] Enabling Applications for Grid Computing with Globus, ibm.com/redbooks, 2002.
- [7] <http://purl.oclc/dc>
- [8] <http://lcweb.loc.gov/ead/>

- [9] <http://www.fgdc.gov/>
- [10] http://www.access.gpo.gov/su_docs/gils/index.html
- [11] 王祥安、范記文、何建明，「影音數位典藏系統—以台灣社會人文電子影音數位博物館為例」
- [12] 王祥安、陳貴青、林彥君、李道明，「影音數位典藏系統之技術研究—以電子影音創意加值應用計劃為例」
- [13] Ruay-Shiung Chang, Chih-Min Wang and Po-Hung Chen, "Replica Selection on Co-Allocation Data Grids," Proceedings of Second International Symposium on Parallel and Distributed Processing and Applications (ISPA'2004), Hong Kong, China, December 2004. (NCHC-KING_010200). Book Chapters : Lecture Note on Computer Science, vol. 3358, pp. 584-593. (SCI)
- [14] Ruay-Shiung Chang, Chih-Min Wang and Po-Hung Chen, "Fragmented Replica in Data Grids," Proceedings of The First Workshop on Grid Technologies and Applications (WoGTA 2004), HsinChu, Taiwan, December 2004. (NCHC-KING_010200)
- [15] Ruay-Shiung Chang, Chih-Min Wang and Po-Hung Chen, "Fragmented Replica Selection and Retrieval," Proceeding of the International Conference on Grid Computing and Applications (GCA'05), Las Vegas, USA, June 2005.(NCHC_94A0227SB, NSC_94-2422-H-259-001)