



能快速處理大量文物影像之數位化儲存備  
份機制-以國立故宮博物院實作經驗為例

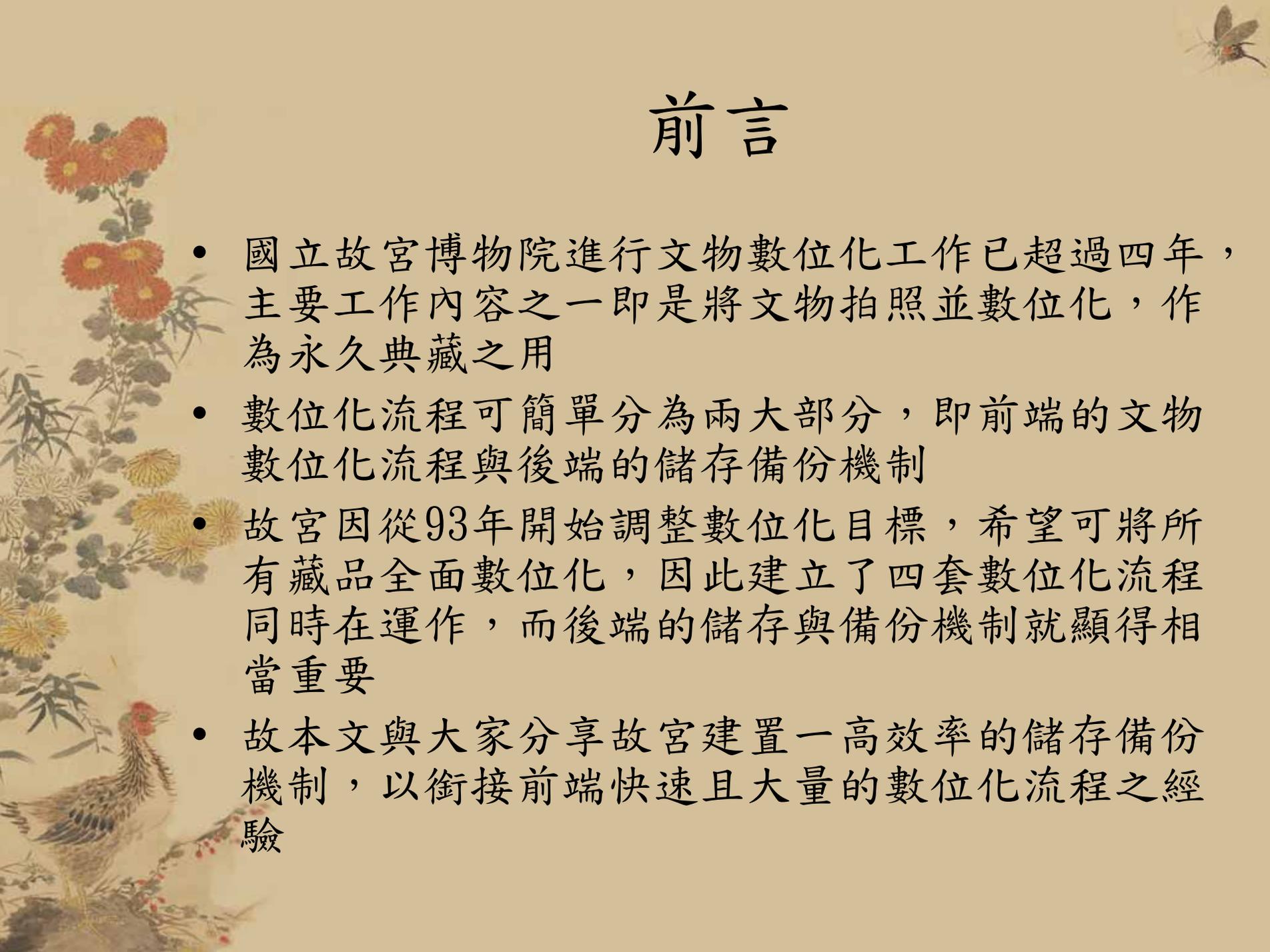
國立故宮博物院  
資訊中心



# 報告項次

1. 前言
2. 故宮初期文物影像數位化儲存備份機制
3. 大量文物數位化之瓶頸
4. 可快速且大量處理之文物影像數位化儲存備份機制
5. 結論及未來發展





# 前言

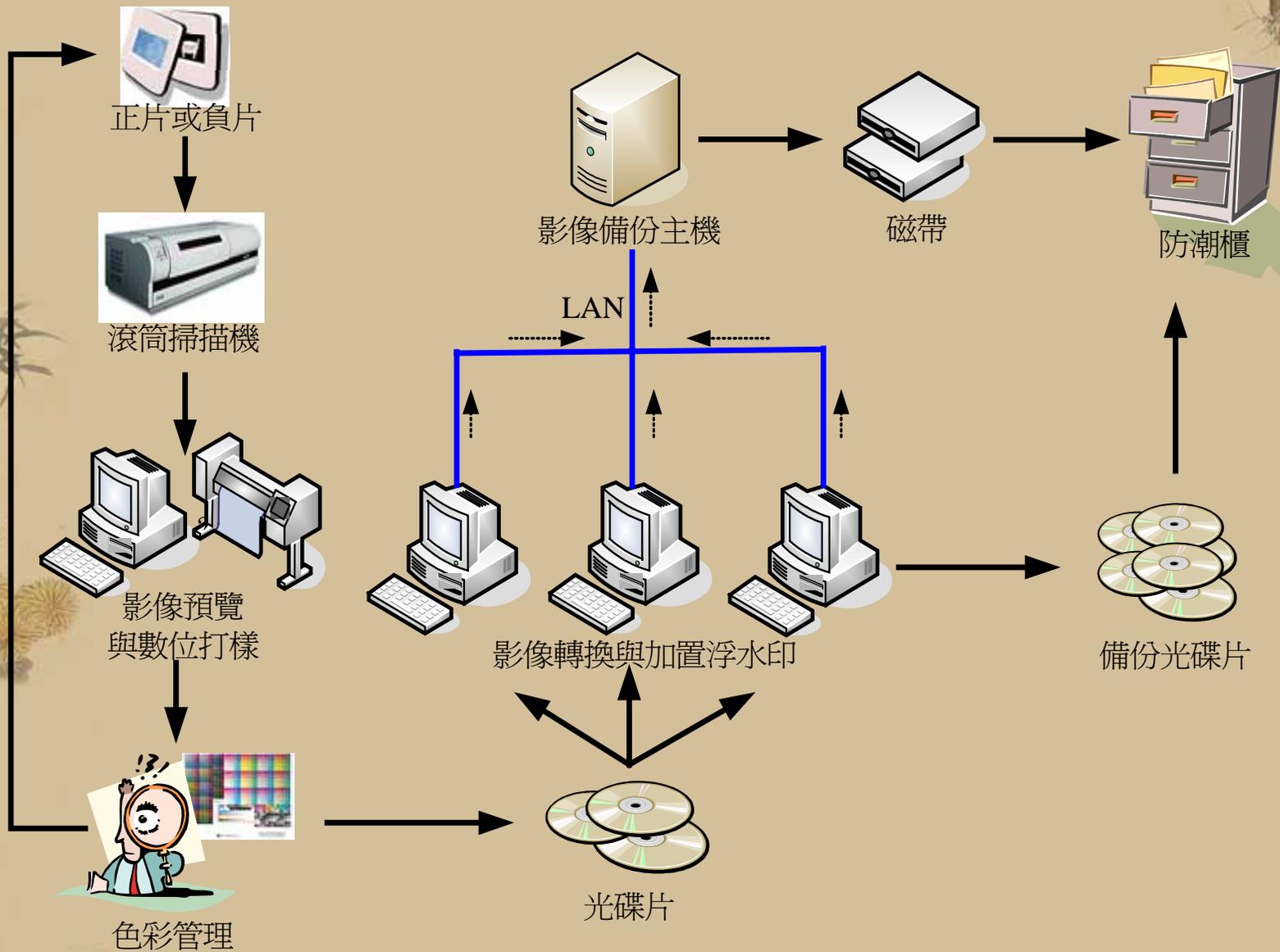
- 國立故宮博物院進行文物數位化工作已超過四年，主要工作內容之一即是將文物拍照並數位化，作為永久典藏之用
- 數位化流程可簡單分為兩大部分，即前端的文物數位化流程與後端的儲存備份機制
- 故宮因從93年開始調整數位化目標，希望可將所有藏品全面數位化，因此建立了四套數位化流程同時在運作，而後端的儲存與備份機制就顯得相當重要
- 故本文與大家分享故宮建置一高效率的儲存備份機制，以銜接前端快速且大量的數位化流程之經驗



# 故宮初期之文物影像數位化儲存備份機制

- 故宮初期數位化目標為國寶或重要古物等級以上之文物，故預期的量不多，且大多已拍攝過照片，留下不少正片與負片，所以初期數位化採掃描之方式
- 故宮初期數位化的進度是每年2000張正負片左右，每張會產生約90MB左右之影像檔，因此初期的儲存備份機制每星期只需處理約 3.75GB 即可





圖一 故宮初期之文物影像數位化儲存備份機制



# 大量文物數位化之瓶頸

- 93年起，故宮調整數位化目標，朝向全面數位化而努力，總數高達六十多萬件
- 前端數位化流程除了增加正、負片的掃描量外，亦陸續增加了三套數位攝影設備，大幅提昇前端數位化流程之速度
- 數位攝影系統每星期共可拍攝約340張影像檔，每張影像檔約會產生260MB之檔案大小，每星期將產生88.4GB之檔案，再加上正負片掃描數量提升至每年約8000張，每星期也約有15G之數位化檔案產出，因此每星期共約有103.4G，數位化檔案量是初期掃描流程的27倍之多





# 大量文物數位化之瓶頸

- 因增加了27倍之多的檔案量，造成了儲存備份的瓶頸，分析原因有以下幾項：
  1. 儲存空間不足
  2. 網路速度慢且佔用內部網路頻寬
  3. 安全疑慮
  4. 無法自動備份光碟





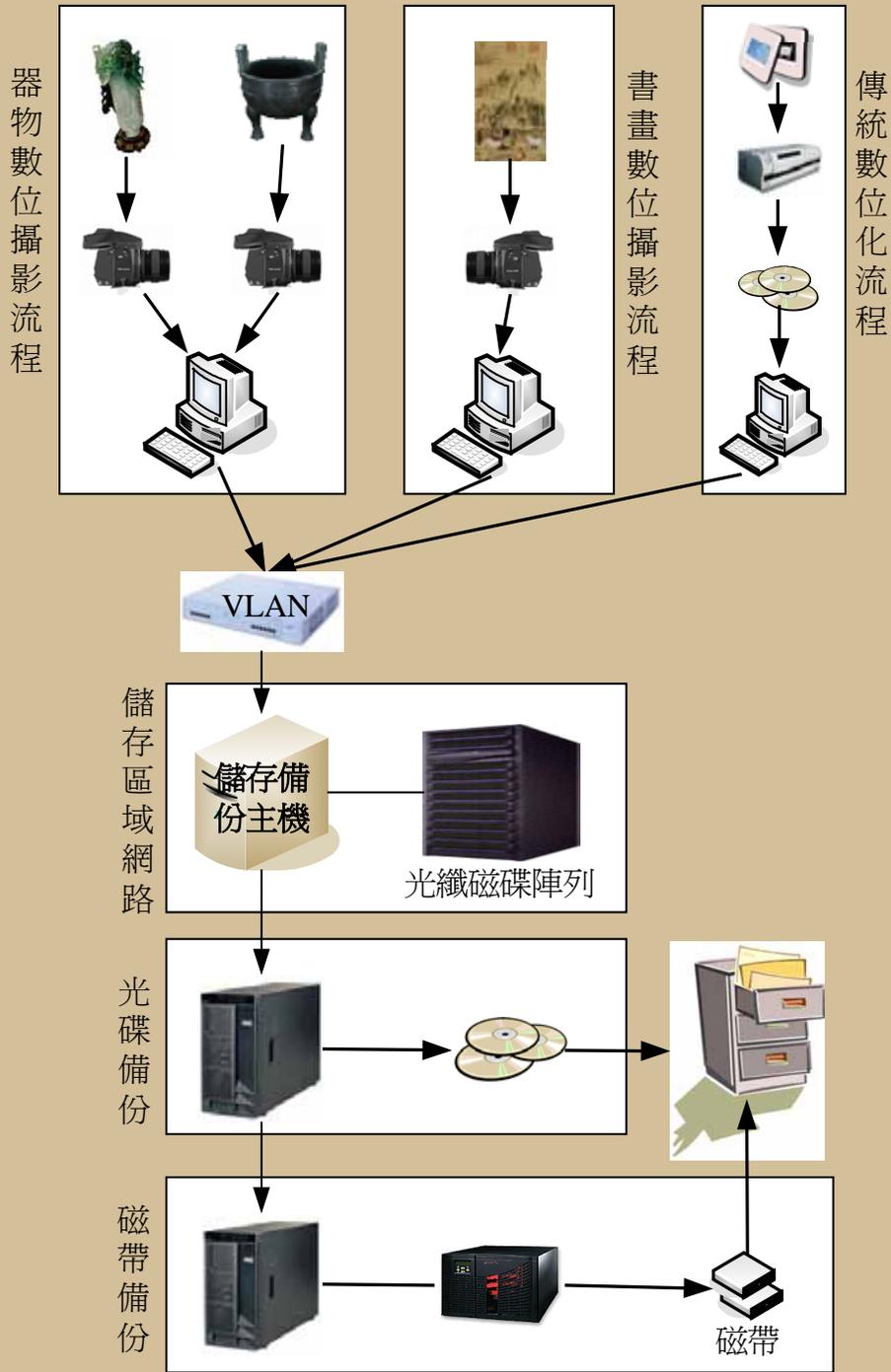
# 大量文物數位化之瓶頸

- 為了解決上述之問題，故宮利用以下三項改善方法：
  1. 建置高速虛擬區域網路(VLAN)
  2. 建置儲存區域網路(SAN)
  3. 採用光碟櫃備份光碟





圖二 故宮現行儲存備份機制





# 大量文物數位化之瓶頸

- 但現行之機制只能勉強應付三套數位化流程（兩套數位攝影與傳統數位化流程），如果要應付四套數位化流程同時上線，則無法負荷。分析了其原因，故宮現行儲存備份機制仍有以下幾項問題：
  1. 網路頻寬不足與易產生碰撞
  2. 移動大量數位影像檔
  3. 儲存空間不敷使用

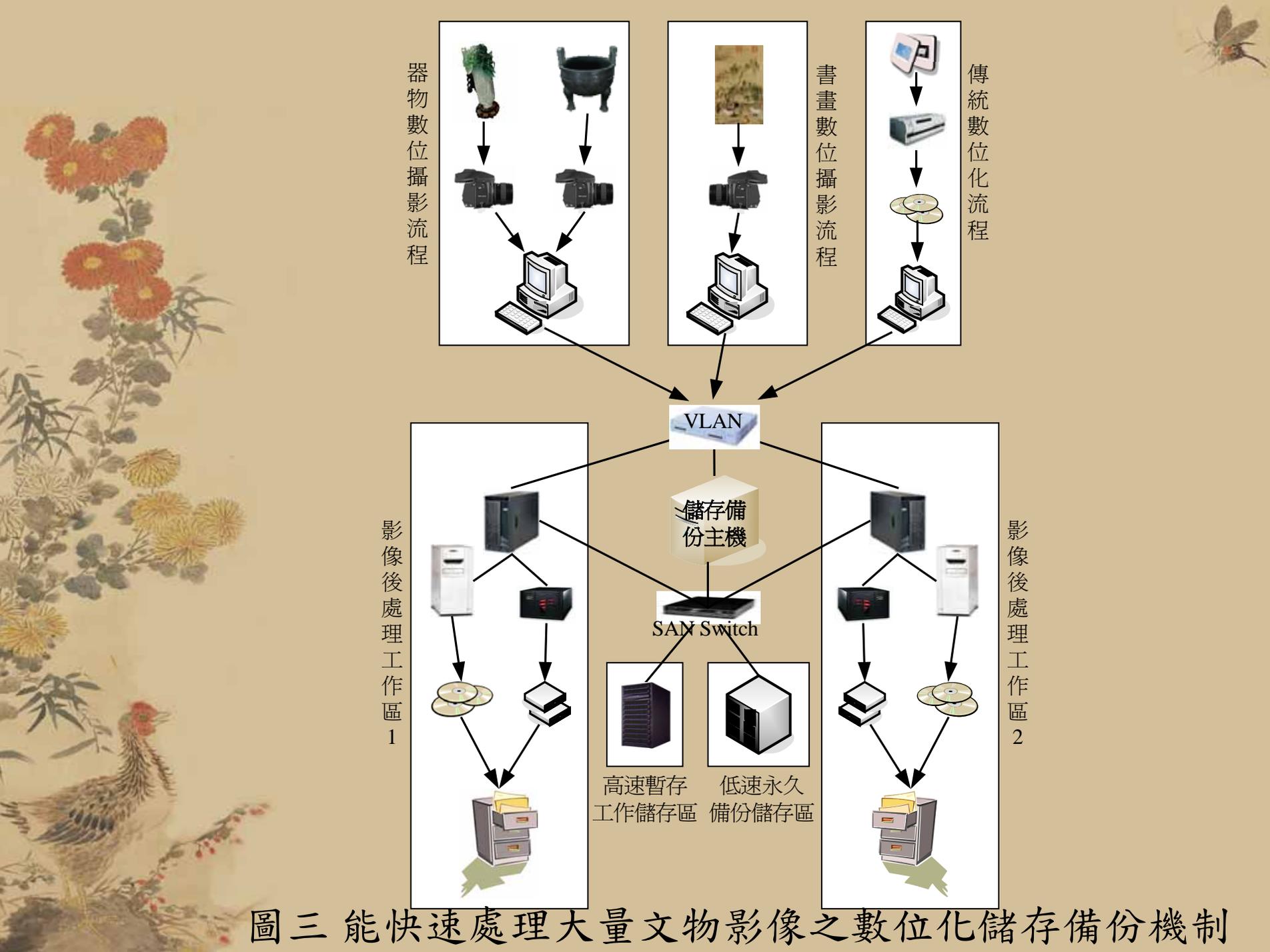




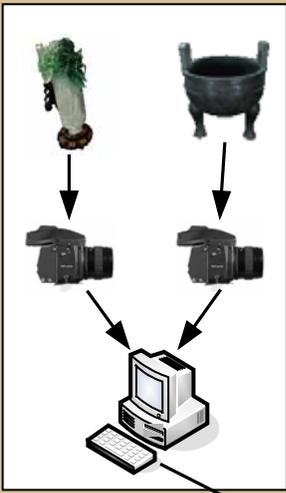
# 能快速處理大量文物影像數位化儲存備份機制

- 因前述之問題，改善方法如下：
  1. 建置暫存工作儲存區
  2. 建置永久備份儲存區
  3. 結合磁帶備份與光碟備份
  4. 建置兩個影像後處理工作區

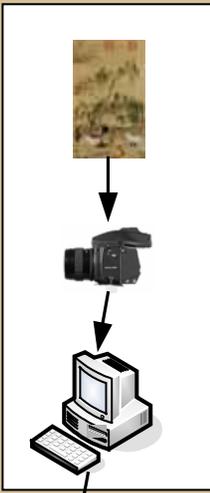




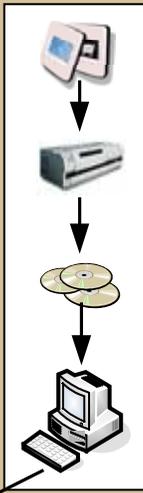
器物數位攝影流程



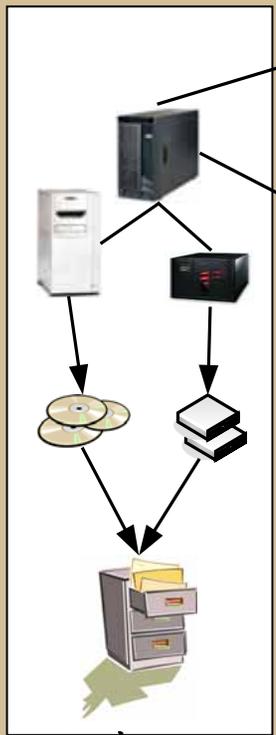
書畫數位攝影流程



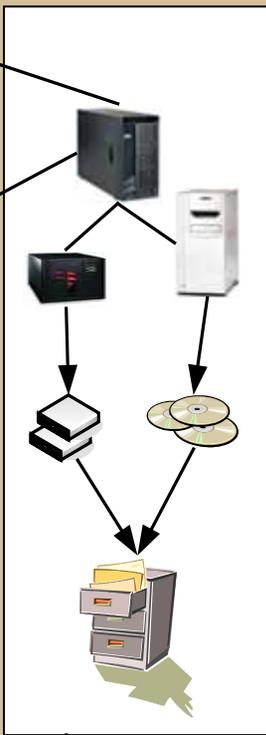
傳統數位化流程



影像後處理工作區 1



影像後處理工作區 2



VLAN

儲存備份主機

SAN Switch

高速暫存  
工作儲存區

低速永久  
備份儲存區

圖三 能快速處理大量文物影像之數位化儲存備份機制



## 能快速處理大量文物影像數位化儲存備份機制

- 此備份機制即能處理每星期約100G以上的數位檔案，處理的量是故宮最初儲存備份機制的27倍之多
- 改善後的文物影像數位化儲存備份機制，搭配前端快速的數位化流程，可使整體文物影像數位化的效率大幅提升



# 結論及未來發展

- 此儲存備份機制，主要是利用高速且獨立之網路環境，讓大量的文物影像可以快速的傳送到後端儲存設備中，藉由高速的光纖磁碟陣列與兩組影像後處理工作區，來處理影像轉換、加置浮水印與備份等工作。並利用光碟機與磁帶機同時備份成光碟與磁帶，避免多次的移動大量影像檔，以大幅縮短資料傳輸、處理與備份的時間，實現可快速處理大量的文物數位化影像，大幅的加快整體數位化效率



## 結論及未來發展

- 此文物影像數位化儲存備份機制雖大大的加快了儲存備份的速度，但仍有一瓶頸無法突破—即是影像檔的檢查工作
- 因影像檔備份成光碟或磁帶之前與之後，皆需人工檢視，檢查影像檔都正常，才能確保備份的影像檔無誤，典藏才有意義





# 結論及未來發展

- 但每張影像檔約為70MB左右，檢查如此多與大之影像檔，需耗費相當多的人力與時間，這不但是原本數位化流程的瓶頸之一，也是本儲存備份機制的瓶頸
- 因此未來將繼續尋求此問題的解決之道，避免太多的人力負擔，使得整體數位化流程能更為快速、有效率





報告完畢

謝謝

