

# Peer-to-Peer檔案交換軟體之使用於網路技術面所造成的影響及技術上可行之相關管理措施

清華大學  
資訊工程學系  
王建興

## Outline

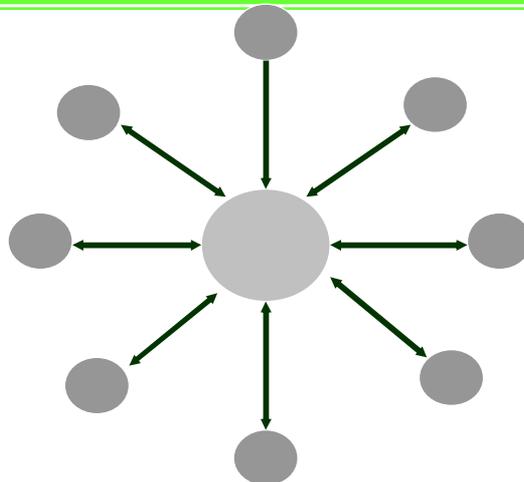
---

- P2P系統簡介
- P2P檔案交換系統所造成的問題
- 解決P2P檔案交換系統的技術策略
- 結論

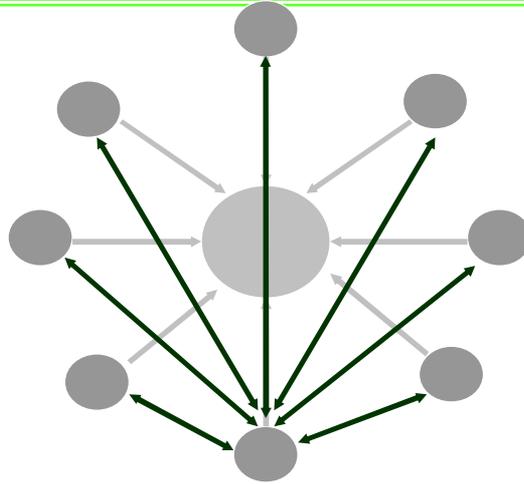
## 何謂Peer-to-Peer技術

- ❑ 點對點的對等式架構，簡稱P2P
- ❑ “Server”的角色淡化
- ❑ 集中式的Peer-to-Peer架構
  - Server仍舊存在，提供認證，授權，以及目錄服務
- ❑ 分散式的Peer-to-Peer架構
  - Peer被稱為Servent, 即 SERVer+cliENT

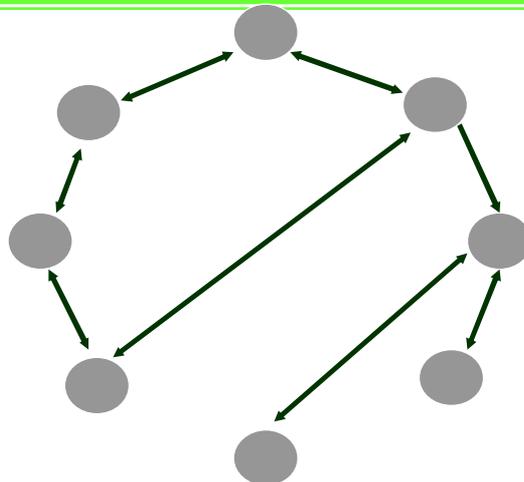
## Client/Server



## Centralized Peer-to-Peer



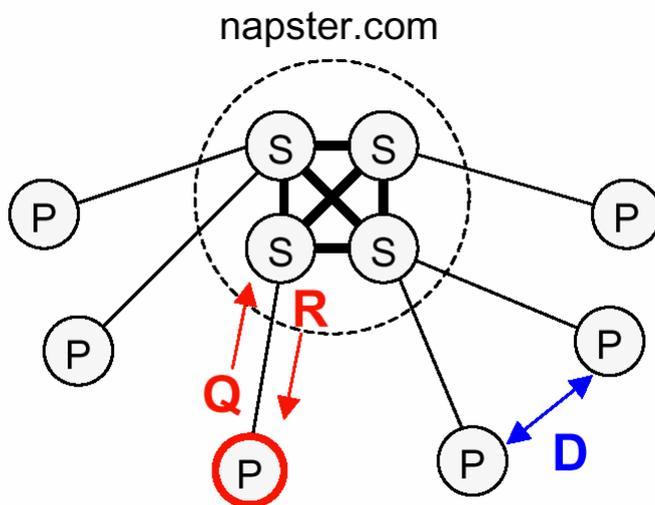
## Decentralized Peer-to-Peer



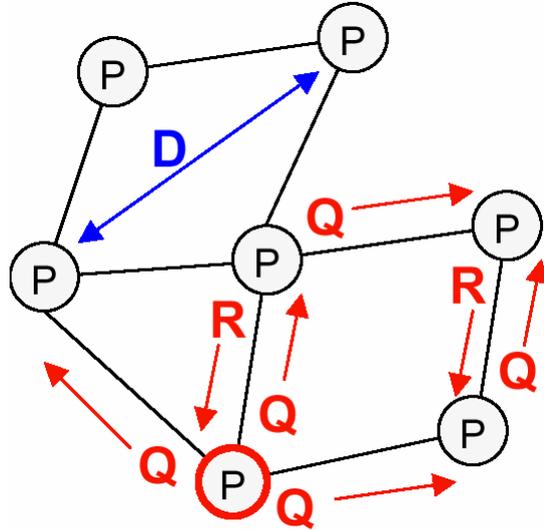
## 常見的P2P應用

- Instance Messaging
  - ICQ, MSN Messenger
- 分散式計算
  - [Seti@home](#)
- 檔案交換
  - Napster
  - Gnutella
  - KaZaA
  - eDonkey

## 集中式的P2P檔案交換架構



## 分散式的P2P檔案交換架構



## P2P架構的優點

- ❑ Large Scale Distributed Computing
  - Computation power
  - Storage
- ❑ 容錯
- ❑ 平衡負載
  - Computation Power
  - Storage
  - Network Bandwidth
- ❑ Power-to-People

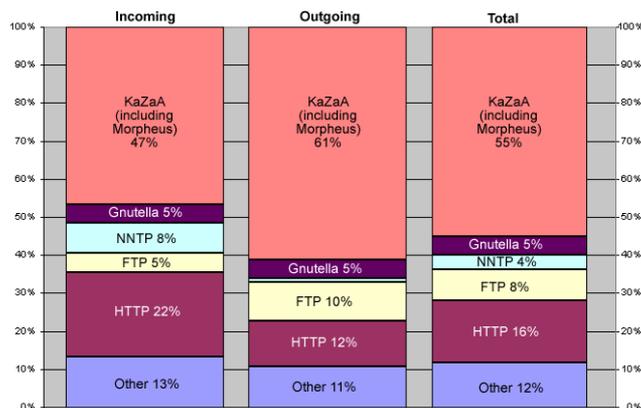
## P2P檔案交換系統所造成的衝擊 (1/2)

- ❑ P2P檔案交換Client快速的普及
- ❑ 根據www.slyck.com的報導 (2003/2/4)

Client名稱	使用者數目
FastTrack	4,114,120
iMesh	1,375,199
eDonkey	569,097
DirectConnect	136,552
Blubster	97,128
Gnutella	92,678
Cvnet	91,750

## P2P檔案交換系統所造成的衝擊 (2/2)

- ❑ P2P檔案交換系統取代HTTP，成為Internet最主要的流量
- ❑ 根據cornell.edu (2002/3)，P2P佔據60%的流量



## P2P檔案交換對網路生態的改變

---

- P2P改變網路內容供需機制
  - Web is not enough
  - Power-to-People
- 資料的複製散佈與取得，速度更快，成本更低廉，甚至免費！
- 許多P2P系統的匿名與極端分散的本質，對現有法律來說是一大挑戰

## P2P網路的運作原理

---

- Peer的搜尋
- 檔案搜尋
- 檔案上傳與下載
- Power Server與Super-node

## Peer與檔案的搜尋

---

- Peer的搜尋
  - 透過大量類似Ping的訊息，來搜尋存在的Peer或提供內容的節點
- 檔案的搜尋
  - 利用大量（可能達上百個）同時的連線，來搜尋欲下載的檔案

## 檔案上傳與下載

---

- 透過WAN與Internet來傳輸大型的檔案
- 同時從不同的Peer下載檔案的不同部份
- 應用程式基於擷取效率，會盡可能的佔據頻寬，以求快速傳輸完檔案
- 極具侵略性：P2P會排擠其他共用相同頻寬的應用程式
- 流量是上下對稱的：所有節點都可能下載與上傳檔案

## Power Server與Super Node

---

- ❑ 許多P2P應用程式都會試著找出所謂的Power Server
  - 頻寬充足
  - 儲存空間充足
- ❑ Power Server會成為許多其他Peer上傳或下載檔案的目標
- ❑ 可能會需要服務超過20,000同時的連線
- ❑ 資源愈多，對外提供的服務愈多

## 範例：Gnutella Network

---

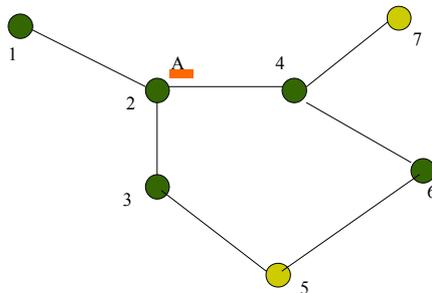
- ❑ 規模極大：超過500K個節點，100TB資料量，10M個檔案
- ❑ 自我組織的網路
- ❑ 開放性架構，有著簡單且彈性的通訊協定

## Gnutella通訊協定簡介

- Gnutella是個P2P檔案交換應用程式，建構於overlay network上
  - 節點通訊採用TCP連線
  - 訊息以廣播方式送出

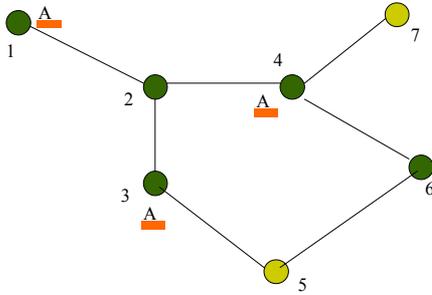
Protocol	Broadcast (Flooding)	Back-propagated	Node to node
Membership	PING	PONG	
Query	QUERY	QUERY HIT	
File download			GET, PUSH

## Gnutella的搜尋機制



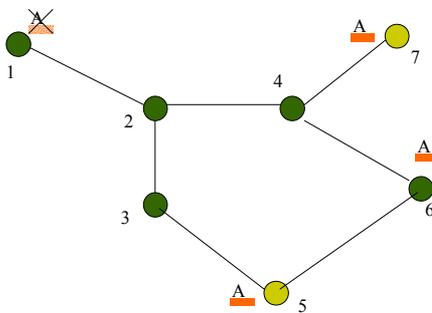
- 節點2想要搜尋檔案A

## Gnutella的搜尋機制



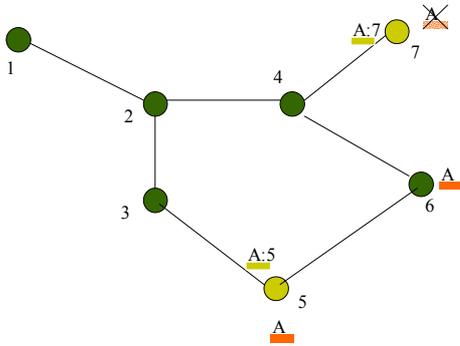
- 節點2想要搜尋檔案A
- 將訊息發送給其鄰點

## Gnutella的搜尋機制



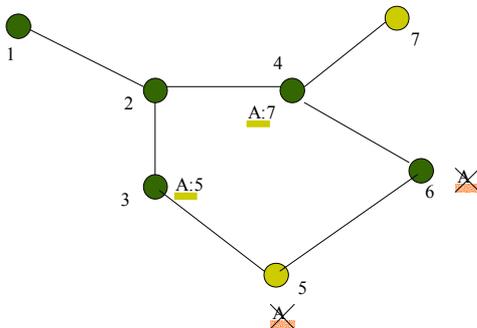
- 節點2想要搜尋檔案A
- 將訊息發送給其鄰點
- 鄰點轉遞該搜尋訊息

## Gnutella的搜尋機制



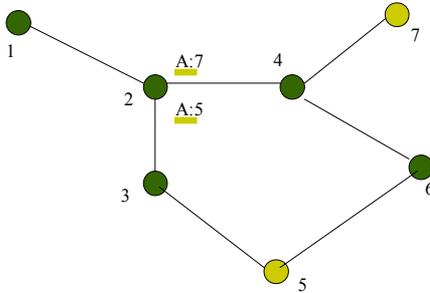
- 節點2想要搜尋檔案A
- 將訊息發送給其鄰點
- 鄰點轉遞該搜尋訊息
- 含有檔案A的節點回傳

## Gnutella的搜尋機制



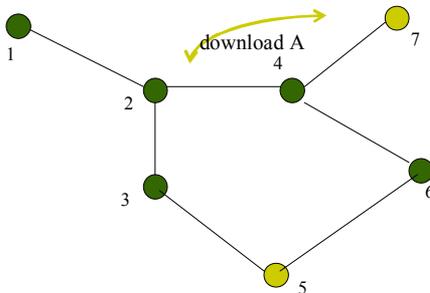
- 節點2想要搜尋檔案A
- 將訊息發送給其鄰點
- 鄰點轉遞該搜尋訊息
- 含有檔案A的節點回傳
- 回應訊息沿原路回傳

## Gnutella的搜尋機制



- 節點2想要搜尋檔案A
- 將訊息發送給其鄰點
- 鄰點轉遞該搜尋訊息
- 含有檔案A的節點回傳
- 回應訊息沿原路回傳

## Gnutella的搜尋機制



- 節點2想要搜尋檔案A
- 將訊息發送給其鄰點
- 鄰點轉遞該搜尋訊息
- 含有檔案A的節點回傳
- 回應訊息沿原路回傳
- 自節點7下載檔案

## P2P常見下載技術：多線擷取檔案

---

- ❑ 常用於大檔案下載
- ❑ 產生多個連線，分別取得大檔案的不同部份
- ❑ 平行化處理，盡可能吞蝕頻寬，以求快速下載檔案

## P2P檔案交換系統在管理上的問題

---

- ❑ 檔案著作權問題
- ❑ 網路流量管理
- ❑ 網路安全管理

## P2P檔案交換系統的流量統計資訊

---

- ❑ 研究指出：P2P流量約佔ISP網路的50-70%
- ❑ P2P佔掉95%上傳流量
- ❑ 75%的寬頻使用者透過P2P網路來分享資訊

## P2P檔案交換系統在流量管理的特性

---

- ❑ 使用時間
- ❑ Flow數目
- ❑ 頻寬使用量
- ❑ 受網路延遲的影響程度

## 使用時間的比較

---

- P2P中的普通Peer
  - 持續性使用
  - 下載與上傳不需使用者於其旁照料
- P2P中的Super Node
  - 持續性使用
  - 額外多出許多其他Peer的流量
- Web網站瀏覽
  - 間歇性使用
  - 受使用者使用時間限制

## Flow數目的比較

---

- P2P中的普通Peer
  - 同時可能會有上百個flow
  - 用於搜尋與下載
- P2P中的Super Node
  - 同時可能會有上千個flow
  - 用於搜尋或是被搜尋
- Web網站瀏覽
  - 同時不超過十個flow

## 頻寬使用量的比較

---

- P2P中的普通Peer
  - 極耗頻寬
  - 餵多少吃多少
- P2P中的Super Node
  - 比普通的Peer更嚴重
  - 供給愈多消失愈多
- Web網站瀏覽
  - 頻寬耗費量中等
  - 間歇性流量

## 受網路延遲影響的比較

---

- P2P中的普通Peer
  - 不受影響
- P2P中的Super Node
  - 不受影響
- Web網站瀏覽
  - 使用者的等候耐心有限

## P2P檔案交換軟體是吃頻寬的怪獸

---

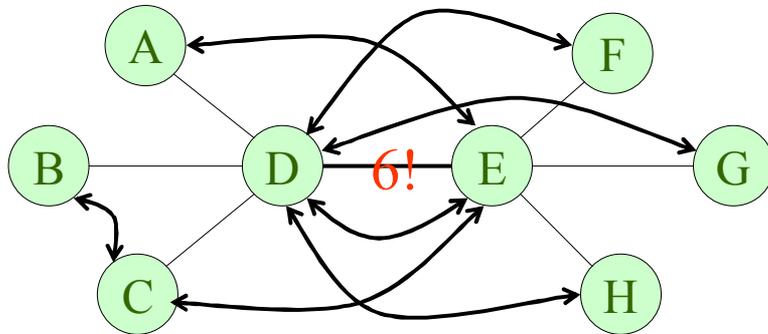
- ❑ 為了有效率的交換檔案，Peer會建立大量的連線，持續性吞蝕大量的頻寬
- ❑ 網路上的其他應用程式皆會受到排擠

## P2P檔案交換軟體在流量上的特性

---

- ❑ 忽略實際的網路拓樸
- ❑ 使用者的流量不再是非對稱
  - 違反Service Provider的規劃
- ❑ 使用者的流量不再是間歇性的使用，而是持續性長期的佔用
- ❑ 使用者透過網路擷取超大型檔案
- ❑ 處處皆Server
  - 違反傳統上對Internet流量的設計與預期

## Overlay Network：忽略實際網路拓樸



## P2P失控的流量

- ❑ 引發頻寬必須升級的需求
- ❑ 正常使用者的使用權利反而遭受剝奪
- ❑ 非傳統的架構使得網路規劃更形困難
- ❑ 非標準公開化的通訊協定形成安全上的漏洞

## 控制P2P檔案交換軟體的困難點

---

- 非標準化的應用程式通訊協定
  - 非標準，非公開
- 協定隱藏
  - 隱藏在HTTP之類的通訊協定中，以規避偵測與防堵
- 多連線
  - 產生多個連線只是為了取得單一個檔案
- Port Hopping
  - 使用非註冊，隨機，非固定的TCP/UDP port

## 解決P2P流量問題

---

- 駝鳥法
- 增加頻寬
- 期待P2P軟體開發者進行流量最佳化
- Port Blocking
- Bandwidth Quota
- Quality of Service

## 駝鳥法

---

- 問題依舊不變
- 正常使用者使用權利遭受剝奪
- 頻寬持續被佔用

## 增加頻寬

---

- P2P是吃頻寬的怪獸
- 餵多少吃多少

## 期待P2P軟體開發者進行流量最佳化

---

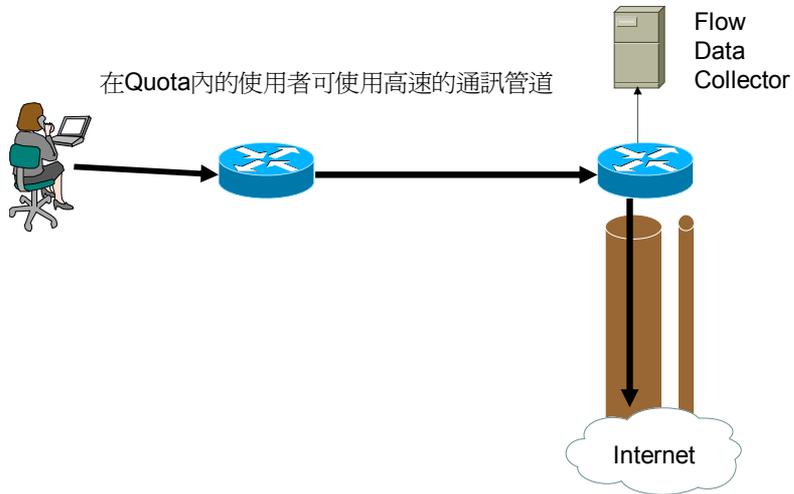
- ❑ 需要時間：開發者還有優先權更高的事要做
- ❑ 需要開發者的自制

## Port Blocking

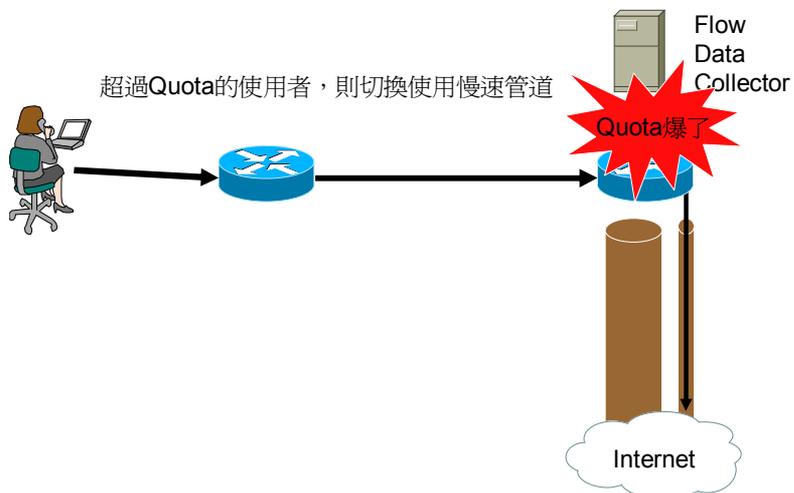
---

- ❑ 並非所有的P2P應用程式都能發揮作用
- ❑ 某些P2P應用程式會使用HTTP tunneling，或隱藏在某些公開的通訊協定之內
- ❑ 某些P2P應用程式會使用Port Hopping的技術，所以不會使用固定的埠號

## Bandwidth Quota (1/2)



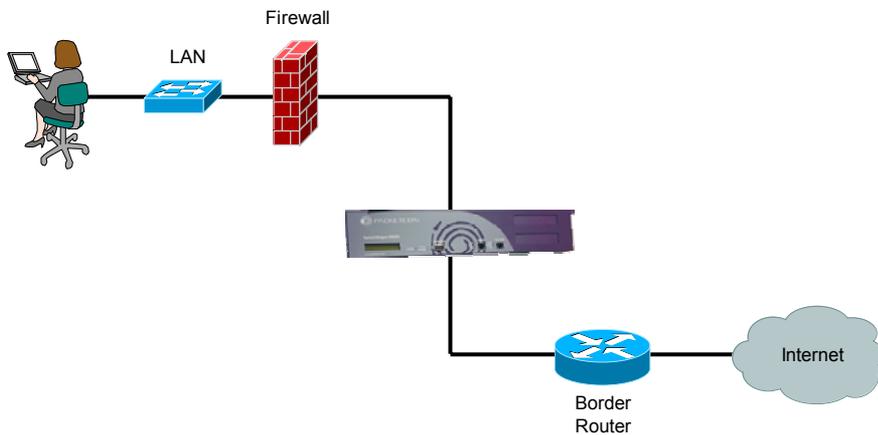
## Bandwidth Quota (2/2)



# Quality of Service

- 透過外部設備來管理流量
  - By Application
  - By Session
  - By User
- 建立Policy來管理頻寬的使用
- 對流量做分類
- 對流量做調整

## 透過外部設備來管理流量



## 流量管理機制

---

- Traffic Classification
- Traffic Policy Enforcement
- Traffic Shaping
- Reporting

## Traffic Classification (1/2)

---

- Layer 7的Application Signature
- 依應用程式再細分
  - Oracle：依資料庫instance再細分
  - VoIP：依協定或CODEC再細分
- HTTP的細分
  - URL wildcard
  - MIME Type
  - 瀏覽器類型

## Traffic Classification (2/2)

---

- Layer 4
  - UDP/TCP特定埠號
  - UDP/TCP埠號範圍
  - UDP/TCP埠號列表
- Layer 3
  - 特定IP位址
  - IP位址範圍
  - Subnet
  - 主機列表
- MAC Address
- 流量方向(進或出)

## Traffic Shaping

---

- Per application minimum
- Per application maximum
- Per session minimum
- Per session maximum
- Dynamic per-user minimum & maximum
- TCP & UDP rate control

## P2P檔案交換軟體的安全性問題

---

- 防火牆阻擋不住P2P應用軟體
- P2P檔案交換軟體允許使用者自由交換檔案
  - 可能取得未經授權的軟體
  - 可能破壞安全管理法規
- P2P檔案交換軟體可能會防礙內部重要系統的運作
- P2P檔案交換軟體允許使用者輕易盜走機密或重要資料
- P2P檔案交換軟體可能有設計上的瑕疵，而遭有心人士利用
- P2P檔案交換軟體本身可能就是特洛伊木馬

## 結論

---

- P2P檔案交換軟體已經成為網路上不可忽視的力量，而且佔掉極大比例的頻寬
- 針對P2P檔案交換軟體必須進行有效的管理，否則其
  - 所導致的巨大流量將嚴重衝擊正常的網路使用
  - 所衍生的安全性議題將造成危害