

Lezione n.2
Peer-to-Peer Systems
and Applications
Capitolo 3

Sistemi P2P: Evoluzione
Laura Ricci

SCHEMA DELLA PRESENTAZIONE

1. Analisi del traffico P2P negli ultimi 5 anni
2. Le origini: ARPANET
3. NAPSTER: la applicazione 'killer'
4. GNUTELLA: una architettura completamente decentralizzata
5. Il successo dei sistemi P2P: motivazioni



3)



1)



2)

- 1) Freenet
- 2) Buzzpad
- 3) WuWu

[Most relevant P2P-Applications in the year 2001]

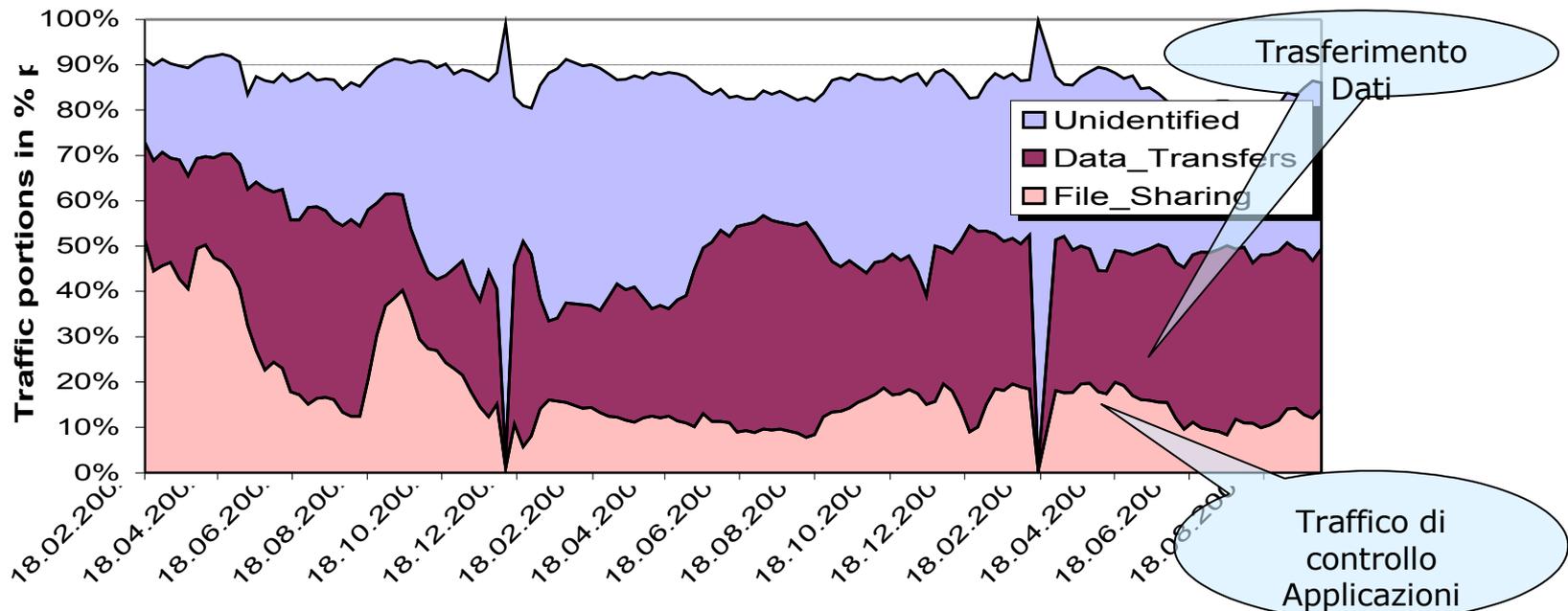
Analisi del traffico P2P

- **Maggio 1999**: presentazione di NAPSTER
- Analisi del traffico P2P, a partire dal 2002: <http://netflow.internet2.edu>
- Il traffico generato da applicazioni P2P rappresenta
 - il 30%-60% del traffico totale del backbone Abilene (periodo 2002-2004)
 - il 70% del traffico nella rete tedesca DFN
- LRZ (Munich Network Center) osserva una sostanziale uniformità tra Stati Uniti ed America

Impatto delle traffico P2P nel backbone Abilene

- 90% del traffico = traffico di controllo P2P + trasferimento dati + traffico non identificato
- Inizio 2002: traffico P2P = 50% del traffico totale
- Fine 2004: traffico P2P ridotto al 15% del traffico totale

Nucleo della infrastruttura di Internet2.
Connette 190 università e centri di ricerca negli States



Fonte: <http://netflow.internet2.edu/weekly/>

Impatto delle traffico P2P nel backbone Abilene

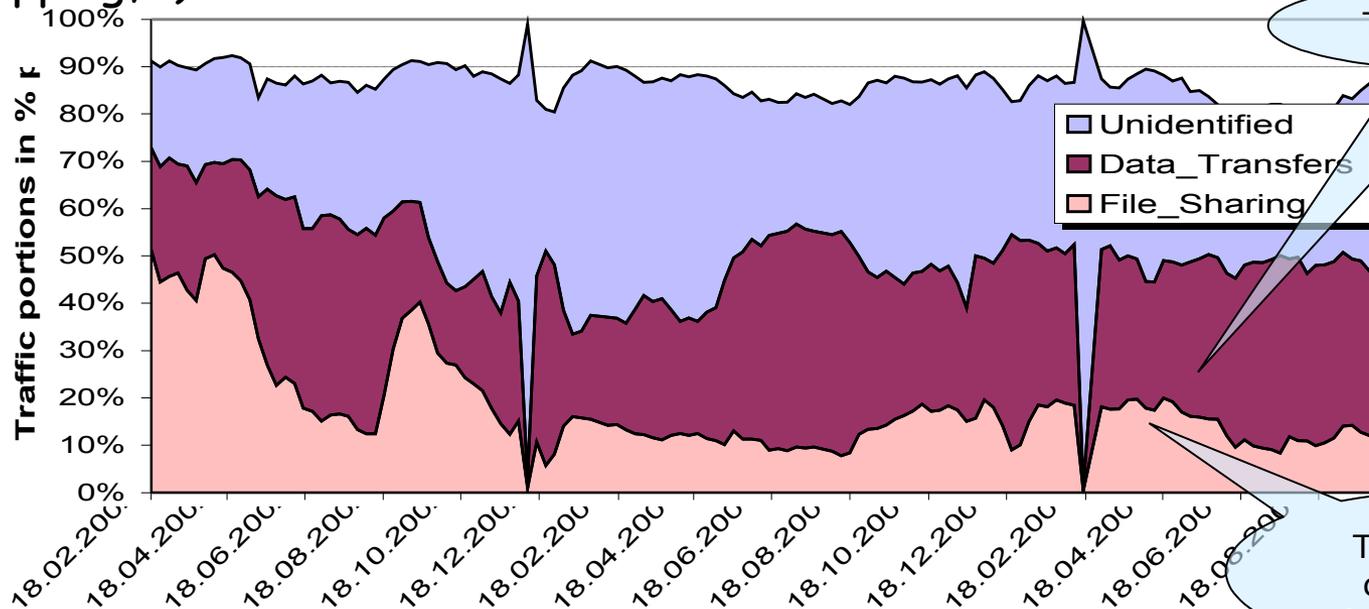
- Motivazione riduzione traffico P2P

Ottimizzazione dei protocolli P2P

Incremento di traffico non identificato + trasferimento dati

Si tende a "camuffare" il traffico P2P per evitarne

l'individuazione (esempio: utilizzo della porta 80, fenomeno del port hopping,...)



Nucleo della infrastruttura di Internet2. Connette 190 università e centri di ricerca negli States

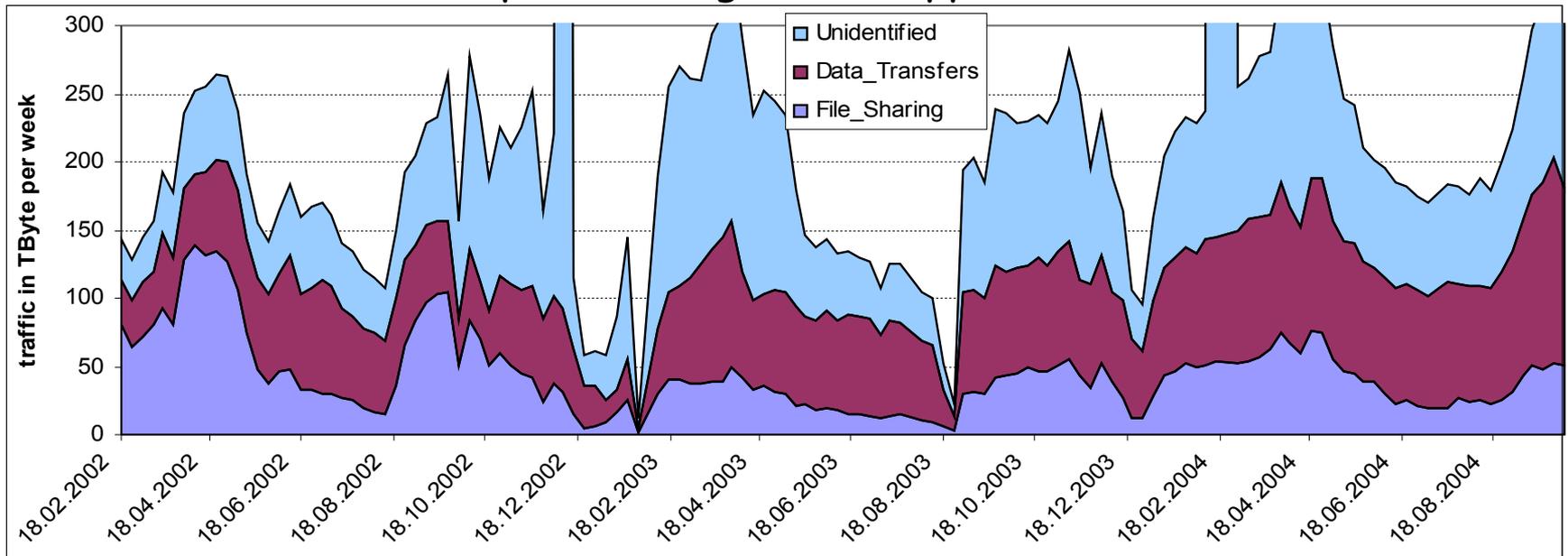
Trasferimento Dati

Traffico di controllo Applicazioni

Fonte: <http://netflow.internet2.edu/weekly/>

Impatto del traffico P2P nel backbone Abilene

- Il traffico P2P (solo quello di controllo) risulta alto (~50 TByte per ogni settimana) costituisce una parte costante del traffico totale
- I picchi possono essere attribuiti a:
 - Chiusura di porte (firewalls, NATs)
 - Condanne giudiziarie (Napster,...)
- Il trasferimento dati è in parte collegato alle applicazioni P2P



Fonte: <http://netflow.internet2.edu/weekly/>

Un esempio : il client LimeWire

The screenshot shows the LimeWire client interface. The main window title is "LimeWire: Enabling Open Information Sharing". The menu bar includes "File", "View", "Navigation", "Resources", "Tools", and "Help". Below the menu bar are icons for "Search", "Monitor", "Connections", "Library", and "LimeShop". The "Lime" logo is in the top right corner.

The main area displays a list of files under the "Any Type" filter. The list has columns for "Quality", "#", "Name", "Type", "Size", "Speed", "Chat", and "Bitrate". The files listed include various classical and modern music tracks, such as "Moonlight Sonata (Beethoven)", "Beethoven - Fuer Elise", and "Mozart - Pachabel Canon in D (1)".

At the bottom of the main list are buttons for "Download", "Chat...", and "Browse Host", along with a "Grouped" checkbox. Below this is a "Downloads" section with a table showing the progress of several files:

Name	Size	Status	Chat	Progress	Speed	Time	Vendor/Ver...
Classical Remix - Beethoven - Moonlight ...	5,288 KB	Downloading from 2 hosts	●	25%	41KB/s	1:36	Morpheus ...
Fantasy Impromptu (Chopin).mid	37.2 KB	Coud Not Download; Awaiting Sourc...		0%			
Mozart - Pachabel Canon in D (1).mp3	4,525 KB	Downloading from 216.187.120.29	●	75%	0KB/s		Morpheus ...
Alla Turca (Mozart).mid	29.2 KB	Coud Not Download; Awaiting Sourc...		0%			
Tacotta in D Minor (Bach).mid	23.1 KB	Coud Not Download; Awaiting Sourc...		0%			
Fugue in D Minor (Bach).mid	22.2 KB	Coud Not Download; Awaiting Sourc...		0%			

At the bottom of the downloads section are buttons for "Kill Download", "Resume", "Launch", "Chat...", and "Clear Inactive".

At the very bottom of the window, there is a status bar showing "Sharing 3 files", a link "Click here to support open source and open protocols.", and a "LimeWire Media Player" control bar.

● Fine anni sessanta: La rete ARPANET

Scopo: condividere documenti e risorse tra centri di ricerca americani.
Ogni host fornisce servizi e documenti e beneficia dei servizi offerti dagli altri hosts

⇒ rete strutturata secondo la filosofia P2P.

MA non può essere considerata una vera rete P2P perchè:

- la rete virtuale in gran parte coincide con quella fisica (nessun overlay)
- L'organizzazione della rete viene stabilita da un organo centralizzato (manca la capacità di auto-organizzazione)
- Applicazioni FTP e Telnet, strutturate secondo il modello client/server (non definisce meccanismi decentralizzati di ricerca)

● 1979: Sviluppo del protocollo UseNet

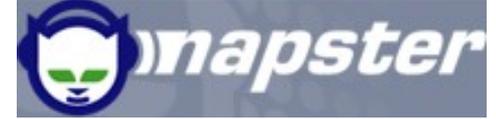
Applicazione basata su newsgroups

Capacità di inserire e rimuovere gruppi di newsgroup

Ma... l'applicazione rimane organizzata secondo il modello client server

Napster: un'applicazione 'killer'

- **Maggio 1999:** Prima generazione di sistemi P2P



Napster: un'applicazione 'dirompente' (disruptive)

Gli utenti non sono più solo consumatori di contenuti (download di file da servers), ma

essi stessi mettono a disposizione risorse per la comunità

Gli utenti definiscono una **rete virtuale**, indipendente dalla struttura della rete fisica

Basata su connessioni TCP ed UDP tra gli utenti

Non esiste una autorità centrale di controllo

Esiste un server centralizzato che memorizza riferimenti alle informazioni condivise dagli utenti

- **Dicembre 1999:** RIAA (Recording Industry Association of America) intraprende un'azione legale contro Napster Inc.

Obiettivo : chiudere l'attività del server centralizzato di Napster,

Napster: una applicazione 'killer'

- **Febbraio 2001:** 2.79 miliardi di file al mese scambiati mediante Napster
- **Luglio 2001:** condanna di Napster Inc.

Napster Inc. è costretta ad interrompere l'attività del server centralizzato

la rete virtuale costituita dagli utenti Napster cessa di esistere

MA...: in questa data sono già disponibili un buon numero di nuove applicazioni con funzionalità simili a quelle di Napster (Gnutella, Kazaa/FastTrack, Freenet...)

Le nuove applicazioni sono progettate in modo da evitare la definizione di un server centralizzato, anche per evitare di incorrere in nuove sanzioni penali



- **Marzo 2000:** Nullsoft (Gene Khan) propone Gnutella, un progetto open-source
- Idee base: Rinunciare al server centralizzato.
 - Motivazione giuridica:* rendere più complessa l'individuazione del traffico P2P
 - Motivazioni tecniche:* eliminazione del collo di bottiglia, maggior affidabilità del sistema nel suo complesso
- Funzionalità
 - Ad ogni peer è affidato, oltre che al compito di server (server+client) per il trasferimento dati, anche il compito di effettuare il routing delle queries

Gnutella

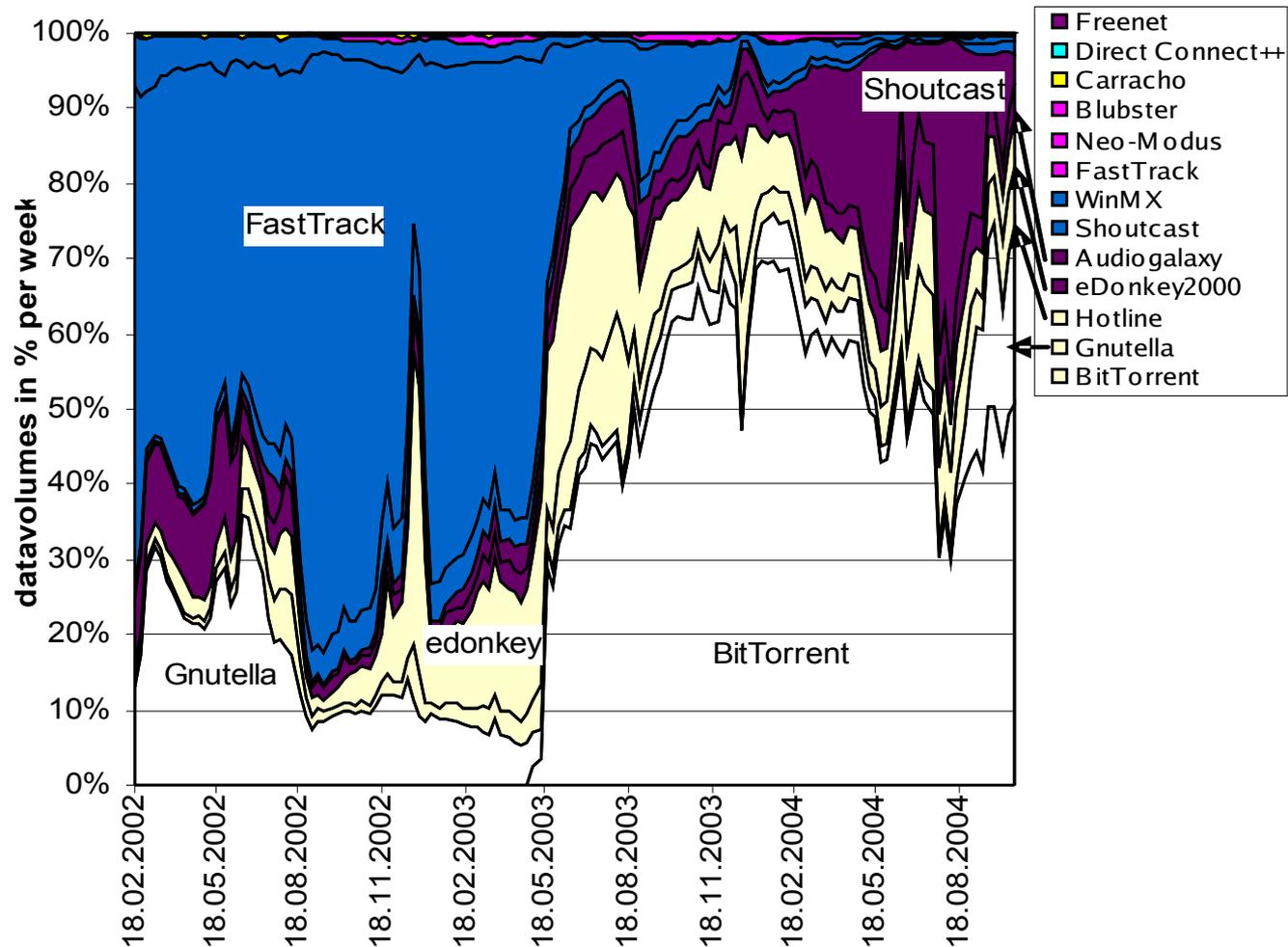
- Ottobre 2000: P2P di seconda generazione. Introduzione di livelli di routing gerarchico
 - Gnutella:** introduzione del concetto di Reflector/Superpeer. Aumento significativo della scalabilità
- Diversi protocolli P2P simili vengono proposti immediatamente dopo Gnutella:
 - Audiogalaxy
 - FastTrack/KaZaA
 - iMesh
 - Freenet

Gnutella ed altri protocolli simili: the story goes on...



- **Agosto 2001**
 - Gli utenti si adattano rapidamente alla chiusura di Napster
 - 3.05 billioni di files al mese scambiati mediante la rete Gnutella
- **2001:** Terza generazione di sistemi P2P
 - Prime proposte di sistemi di terza P2P di terza generazione: le reti P2P strutturate
 - Caratteristica base: Utilizzo di algoritmi di routing basati su Distributed Hash Tables (DHTs)
- **Agosto 2002**
 - La quantità di dati scambiati nella rete KaZaA diminuisce, la causa è da attribuirsi al gran numero di file corrotti (problemi nell'algoritmo di hashing)
 - Edonkey and Gnutella riacquistano popolarità
- **Maggio 2003**
 - Introduzione di Bittorrent
 - Rapidamente diventa l'applicazione che genera la maggior quantità di traffico P2P. La ragione può essere attribuita al fatto che il traffico gerato è individuabile più facilmente
- **Metà 2003**
 - Utilizzo di P2P per applicazioni diverse dal file sharing
 - Sviluppo di Skype e Voice over P2P
- **Oggi....**
 - Gli sforzi maggiori sono orientati ad aumentare la affidabilità delle reti P2P, ad utilizzare le reti P2P su reti mobili, etc.
 - Ebay acquisisce a metà del 2005 Skype per utilizzarlo per la comunicazione tra i partecipanti all'asta e i venditori.

Sviluppo delle applicazioni P2P



Porzioni di traffico di diverse applicazioni e protocolli P2P rilevate sul Backbone Abilene nel periodo da 18.02.2002 fino a 18.10.2004

Data source: <http://netflow.internet2.edu/weekly/>

Fattori che hanno favorito lo sviluppo dei sistemi P2P

Sviluppo della tecnologia:

- **1992:**
Dimensione media di un hard disk: ~0.3Gbyte
Potenza media di un personal computer : ~ 100MHz
- **2002:**
Dimensione media di un hard disk: 100 Gbyte
- **2004:**
Potenza media di un processore : ~ 3GHz ⇒
I personal computers raggiungono capacità di calcolo paragonabili a quelle possedute dai server negli anni 90.s

Sviluppo delle reti di comunicazione:

- Primi anni 90: utenti privati iniziano a connettersi ad Internet mediante modem a 56k.
- **1997/1998**
Prime connessioni a banda larga per utenti privati
Modem a 10Mbps
- **1999**
Introduzione delle connessioni DSL ed ADSL
Connessioni possibili fino a 8.5Mbps mediante connessione telefonica
La deregulation nel mercato della telefonia mostra I primi effetti, portando alla riduzione delle tariffe

⇒

Maggiore banda a minor prezzo