Aalto University

P2P Networks-General

Zhonghong Ou
Post-doc resercher
Data Communications Software (DCS) Lab,
Department of Computer Science and Engineering,
Aalto University

With slides partially from Prof. Jukka K. Nurminen, Aalto University

Schedule

Tue 13.9.2010 14-16	Introduction to P2P (example P2P systems, concepts)	Content delivery (BitTorrent and CoolStreaming)
Tue 20.9.2010 14-16	Unstructured content search (Napster, Gnutella, Kazaa)	Structured content search (DHT)
Tue 27.9.2010 14-16	Energy-efficiency & Mobile P2P	

Introduction to P2P

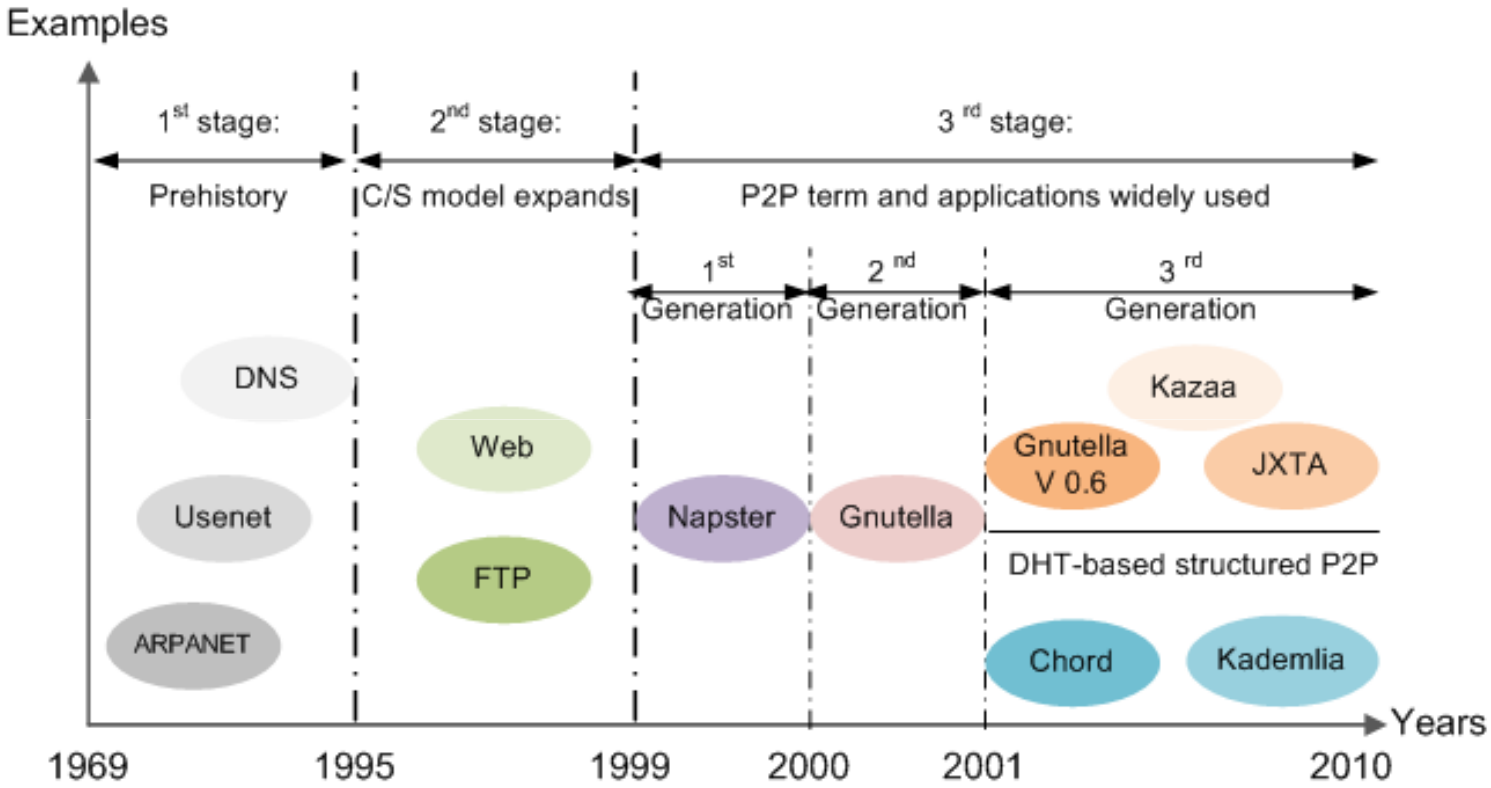
- History of P2P networks
- Definition of P2P
- Example systems
- Why P2P is successful
- Multiple view points
  1. A set of widely used applications
  2. Interesting set of technologies
  3. Increasingly finding legal use to save server costs

History of P2P Networks

Humans are born equal

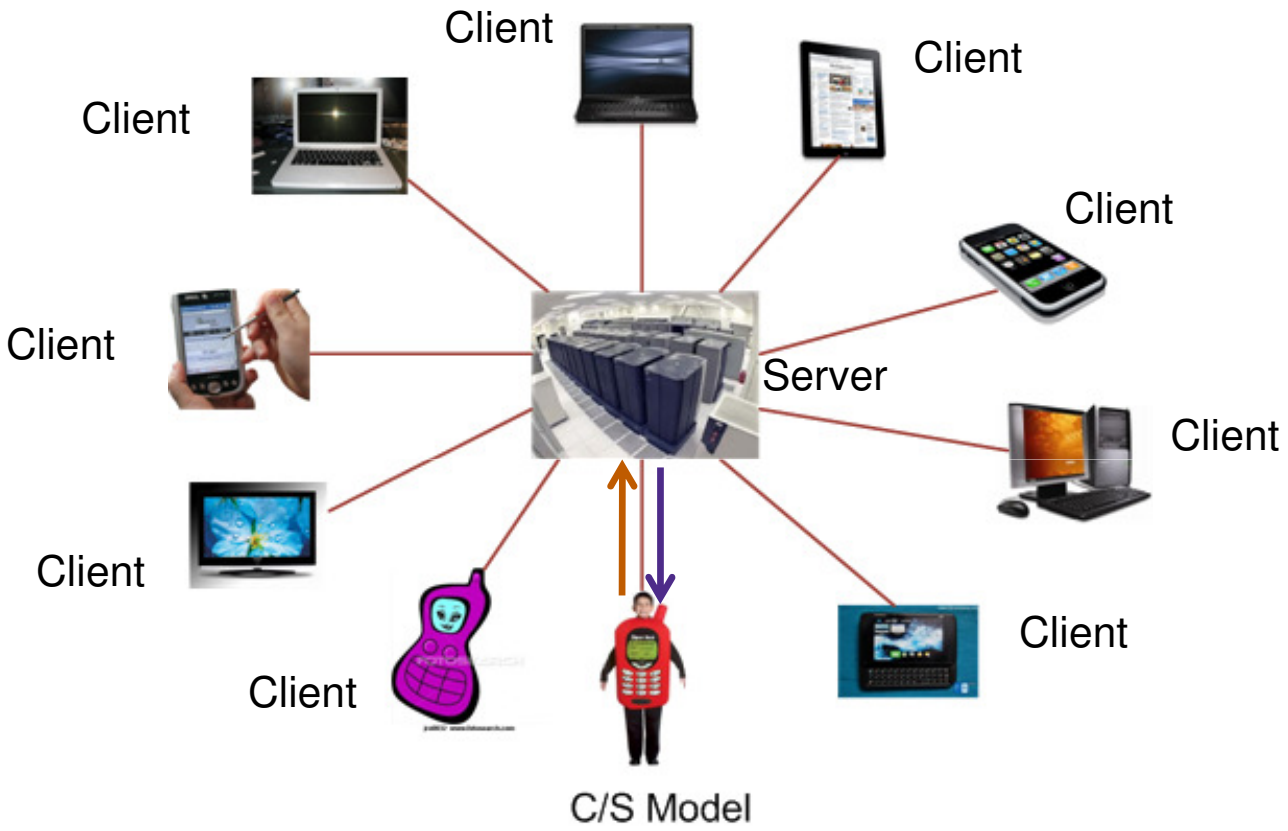
Source: http://lordofdesign.com/ancient-people-of-china-psd/

History of P2P Networks (Cont.d)



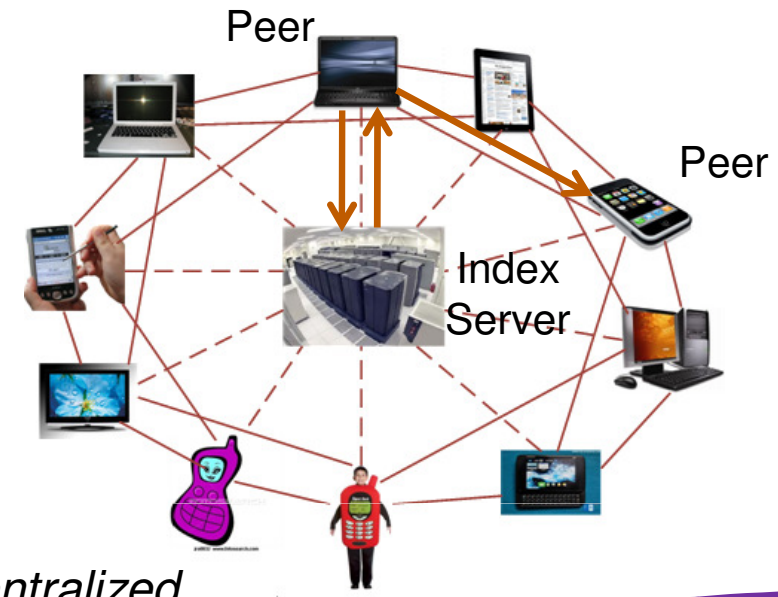
1. (1969–1995) Prehistory: P2P thought burgeoning;
2. (1995–1999) Internet explosion: P2P concept retrogressive;
3. (1999–?) P2P term widely used: P2P-based applications blossom.

Architecture of C/S Model



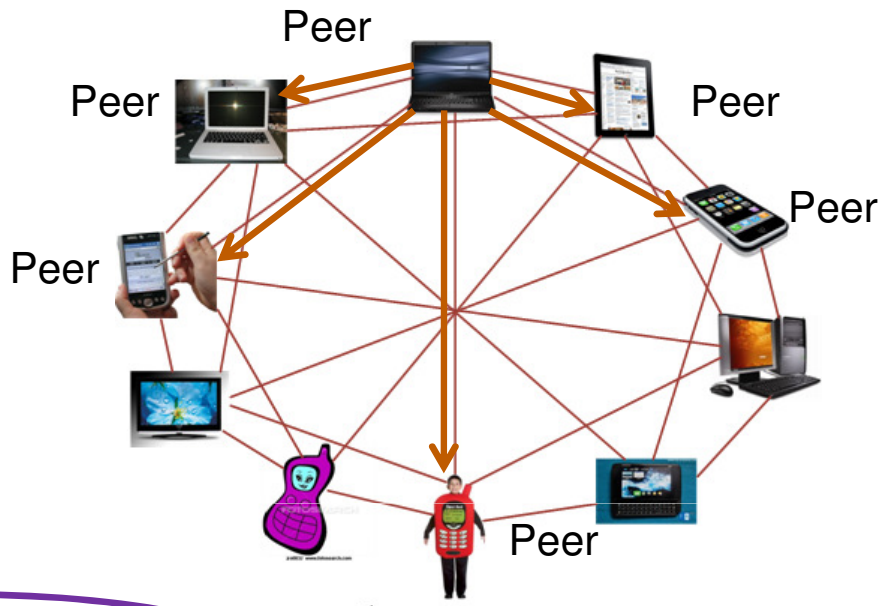
Every client contacts to the centralized server to get the desired content

Architecture of P2P Networks



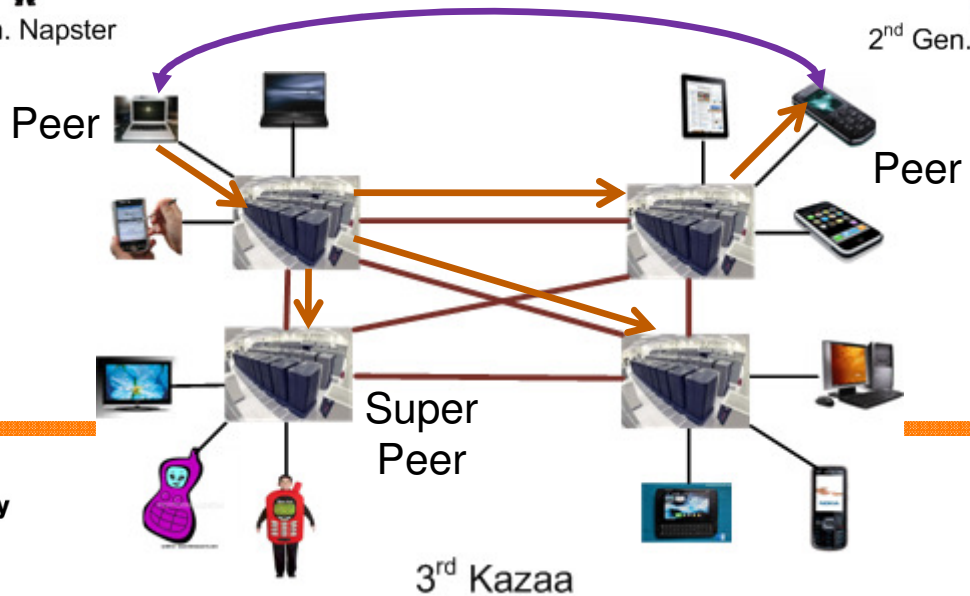
Centralized Index Server

1st Gen. Napster



2nd Gen. Gnutella

Flooding-based

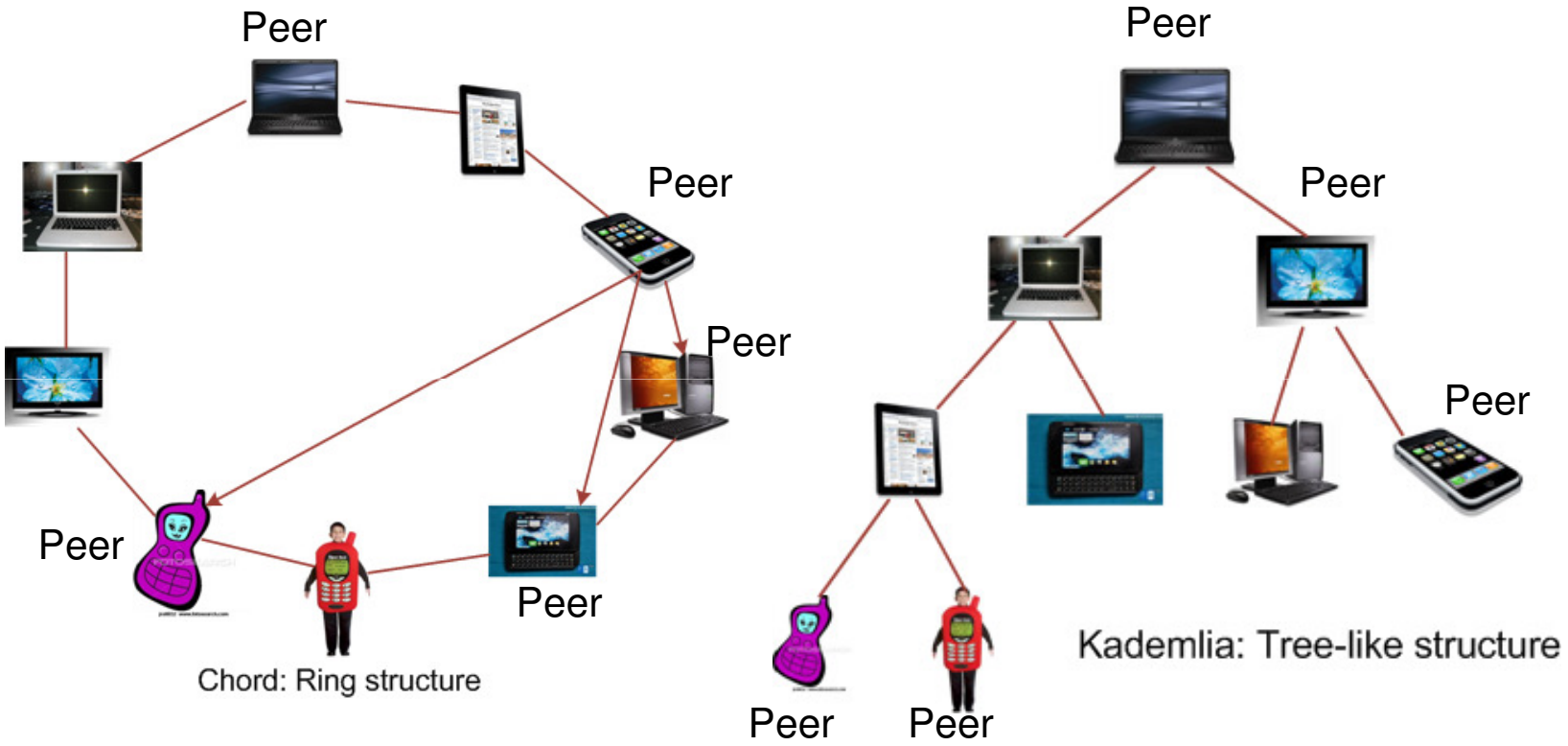


3rd Kazaa

Hierarchy-based

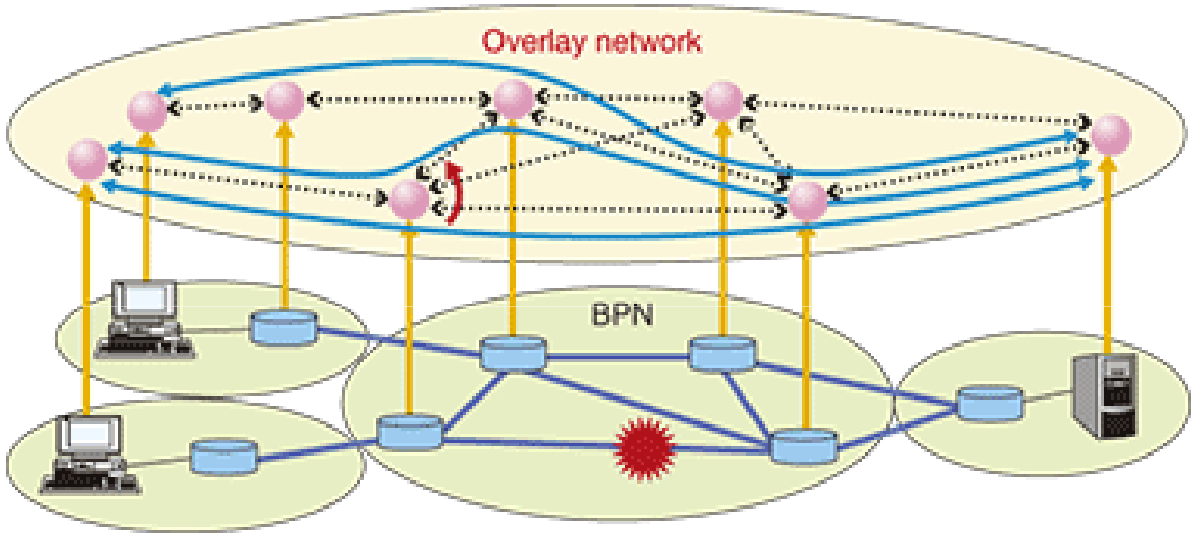
Aalto University

Related Concepts: Structured P2P Networks



Structured P2P networks organize the participants of the whole overlay network according to certain DHT algorithm, which makes the overlay network look like some structure, e.g. a ring, a tree, thus inspires the name "structured".

Related Concepts: Overlay Network



Source: NTT

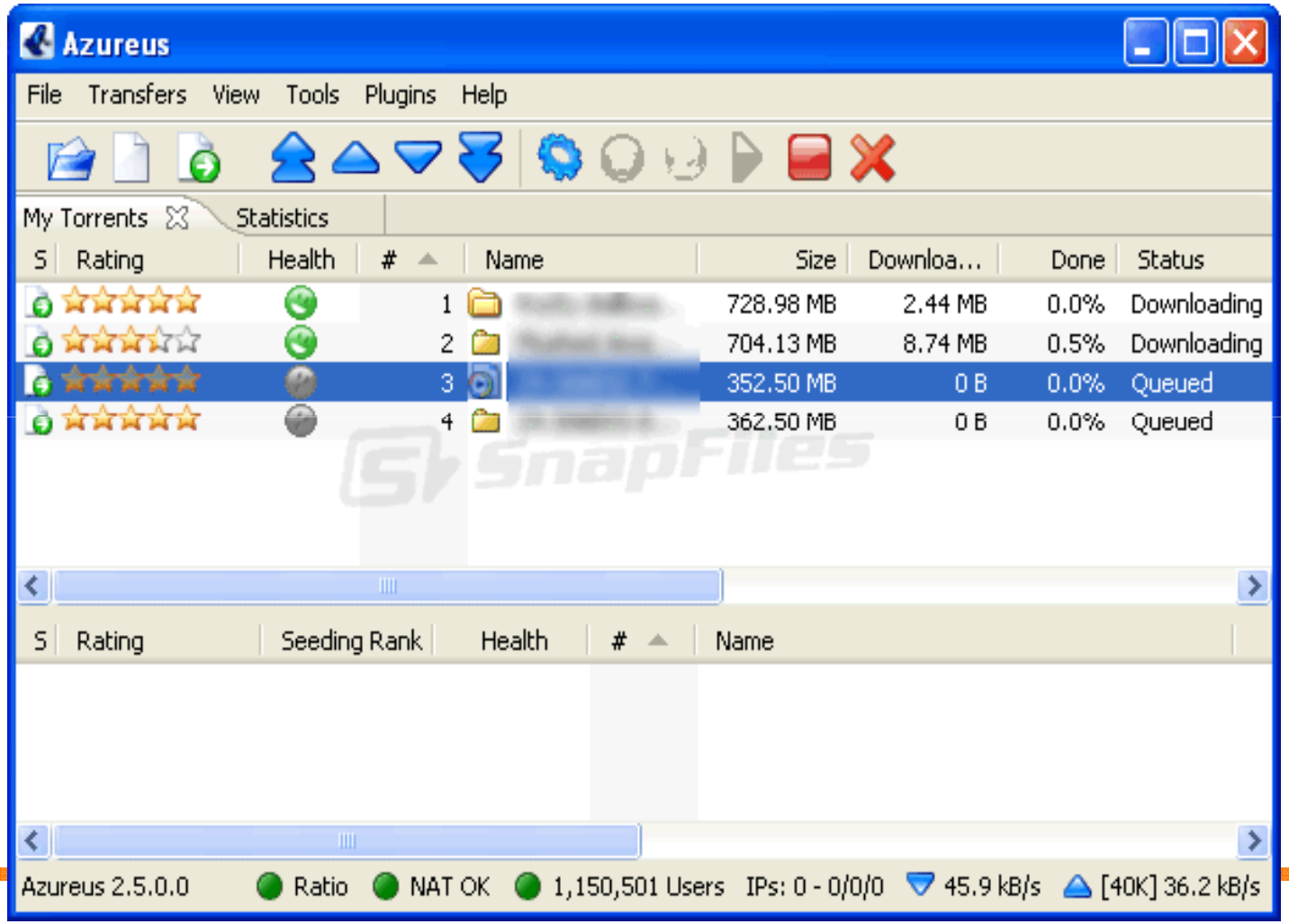
One hop at an overlay network can be mapped to multiple hops at a physical network

Definition of Peer-to-peer (or P2P)

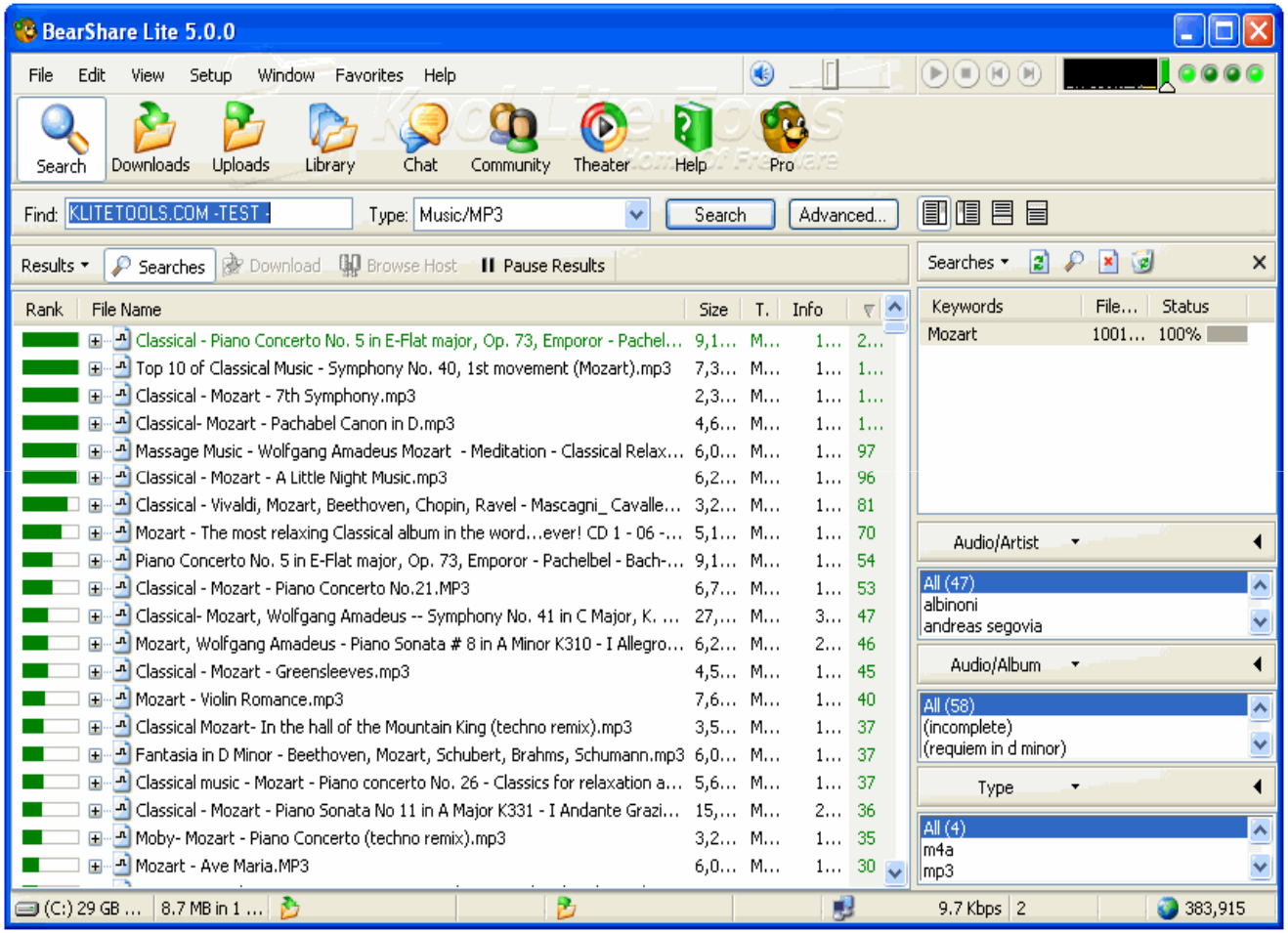
- A peer-to-peer (or P2P) computer network is a network that relies primarily on the ***computing power*** and ***bandwidth*** of the ***participants*** in the network rather than concentrating it in a relatively small number of servers.
- A pure peer-to-peer network does not have the notion of clients or servers, but only equal peer nodes that simultaneously function as both "clients" and "servers" to the other nodes on the network, so-called “servent”.
- This model of network arrangement differs from the client-server model where communication is usually to and from a central server.

Source: www.wikipedia.org

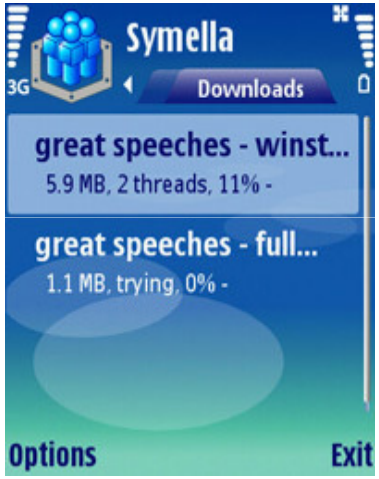
Azureus BitTorrent client



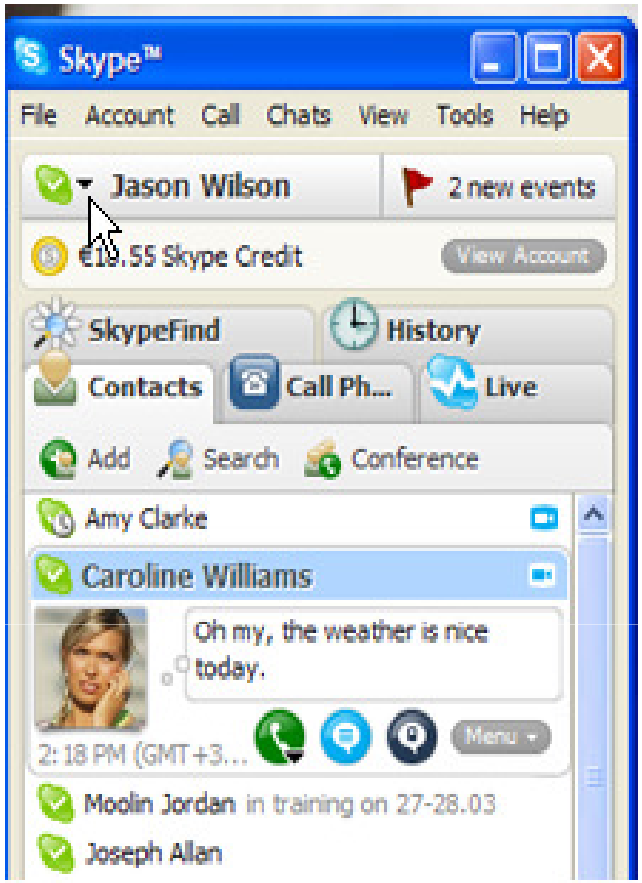
BearShare


BearShare Lite 5.0.0
File Edit View Setup Window Favorites Help
Search Downloads Uploads Library Chat Community Theater Help Pro
Find: KLITETOOLS.COM -TEST -
Type: Music/MP3
Search
Advanced...
Results
Searches
Download
Browse Host
Pause Results
Rank File Name Size T. Info
Classical - Piano Concerto No. 5 in E-Flat major, Op. 73, Emporor - Pachel... 9,1... M... 1... 2...
Top 10 of Classical Music - Symphony No. 40, 1st movement (Mozart).mp3 7,3... M... 1... 1...
Classical - Mozart - 7th Symphony.mp3 2,3... M... 1... 1...
Classical- Mozart - Pachabel Canon in D.mp3 4,6... M... 1... 1...
Massage Music - Wolfgang Amadeus Mozart - Meditation - Classical Relax... 6,0... M... 1... 97
Classical - Mozart - A Little Night Music.mp3 6,2... M... 1... 96
Classical - Vivaldi, Mozart, Beethoven, Chopin, Ravel - Mascagni_ Cavalle... 3,2... M... 1... 81
Mozart - The most relaxing Classical album in the word...ever! CD 1 - 06 -... 5,1... M... 1... 70
Piano Concerto No. 5 in E-Flat major, Op. 73, Emporor - Pachelbel - Bach-... 9,1... M... 1... 54
Classical - Mozart - Piano Concerto No.21.MP3 6,7... M... 1... 53
Classical- Mozart, Wolfgang Amadeus -- Symphony No. 41 in C Major, K. ... 27,... M... 3... 47
Mozart, Wolfgang Amadeus - Piano Sonata # 8 in A Minor K310 - I Allegro... 6,2... M... 2... 46
Classical - Mozart - Greensleeves.mp3 4,5... M... 1... 45
Mozart - Violin Romance.mp3 7,6... M... 1... 40
Classical Mozart- In the hall of the Mountain King (techno remix).mp3 3,5... M... 1... 37
Fantasia in D Minor - Beethoven, Mozart, Schubert, Brahms, Schumann.mp3 6,0... M... 1... 37
Classical music - Mozart - Piano concerto No. 26 - Classics for relaxation a... 5,6... M... 1... 37
Classical - Mozart - Piano Sonata No 11 in A Major K331 - I Andante Grazi... 15,... M... 2... 36
Moby- Mozart - Piano Concerto (techno remix).mp3 3,2... M... 1... 35
Mozart - Ave Maria.MP3 6,0... M... 1... 30
Searches
Keywords File... Status
Mozart 1001... 100%
Audio/Artist
All (47)
albinoni
andreas segovia
Audio/Album
All (58)
(incomplete)
(requiem in d minor)
Type
All (4)
m4a
mp3
(C:) 29 GB ...
8.7 MB in 1 ...
9.7 Kbps
2
383,915

Symbian S60 versions: Symella and SymTorrent

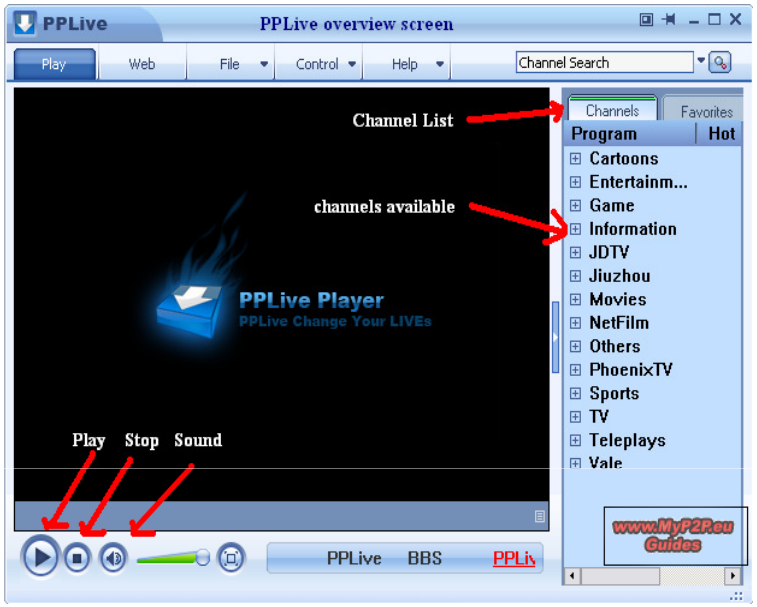


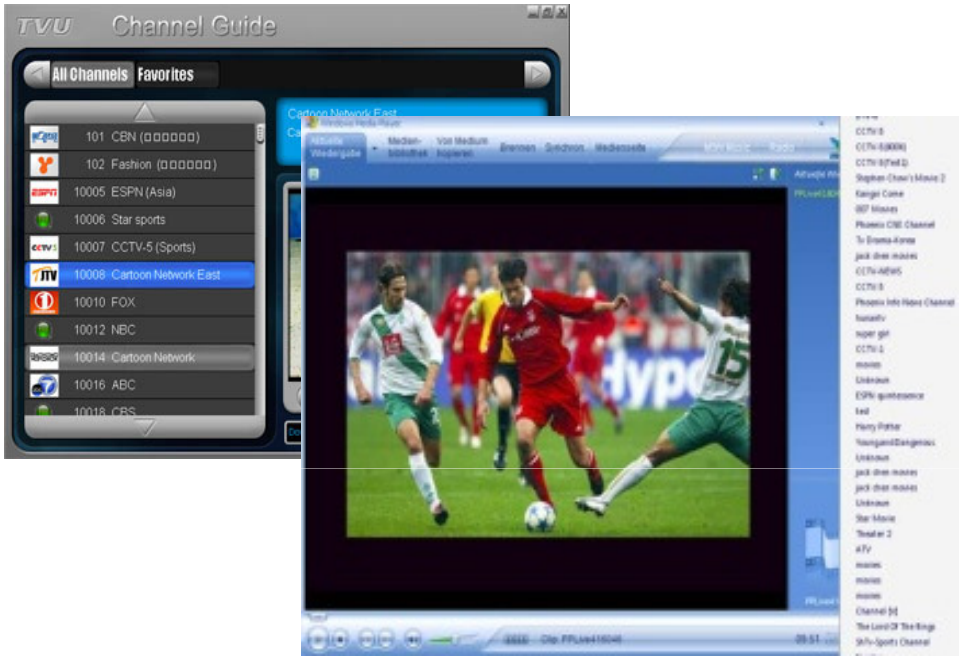
Skype



How skype works: http://arxiv.org/ftp/cs/papers/0412/0412017.pdf

PPLive, PPS, TVU, …





PPLive

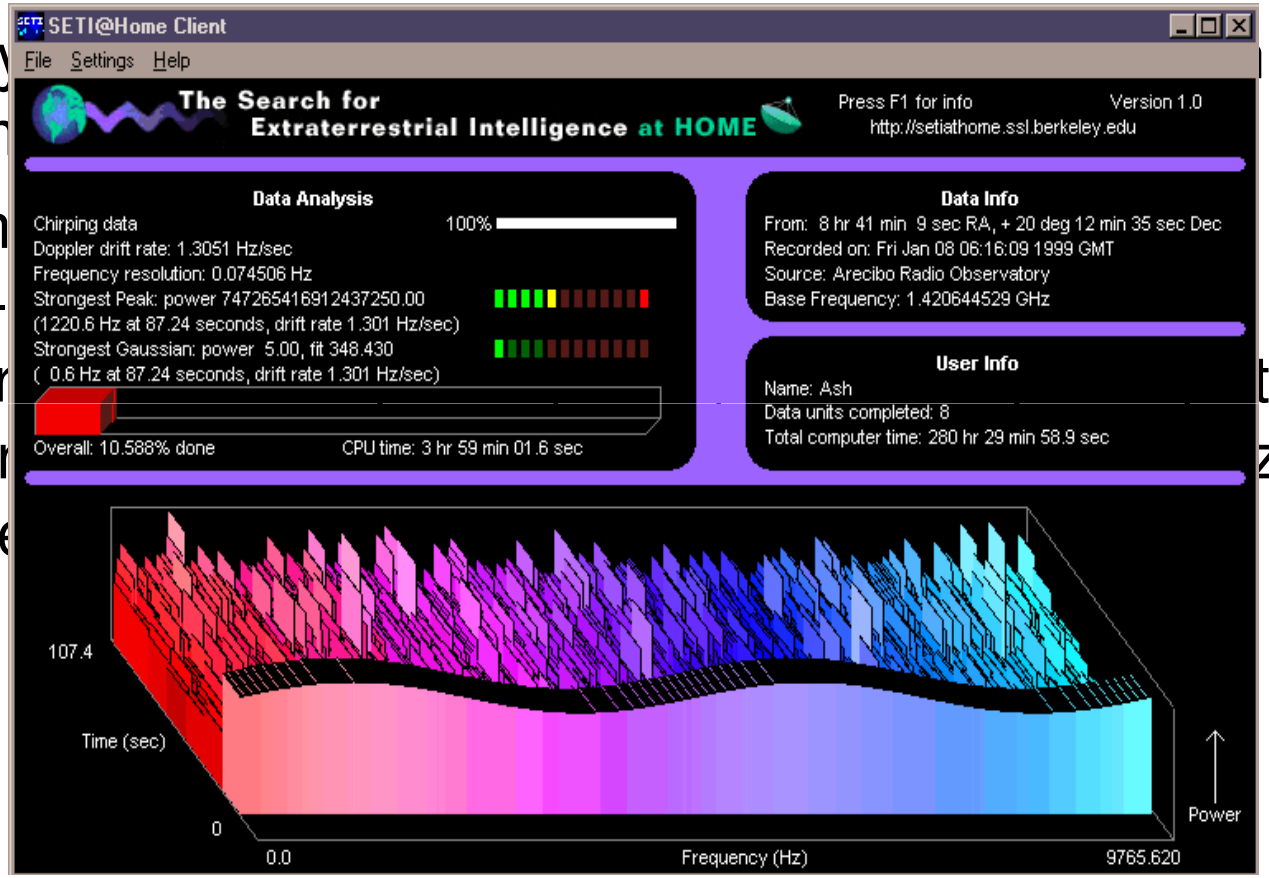
- Founded in 2004, the first online video service provider in China.
- The largest aggregator of China TV programs with over 120 TV stations, thousands of TV shows and programs.
- Has more than 200 million user installations and its active monthly user base (as of Dec 2010) is 104 million, i.e, PPLive has a 43% penetration of Chinese internet users.
- Average viewing time per person per day has reach over 2 hours and 30 minutes.

Source: http://www.synacast.com/en/

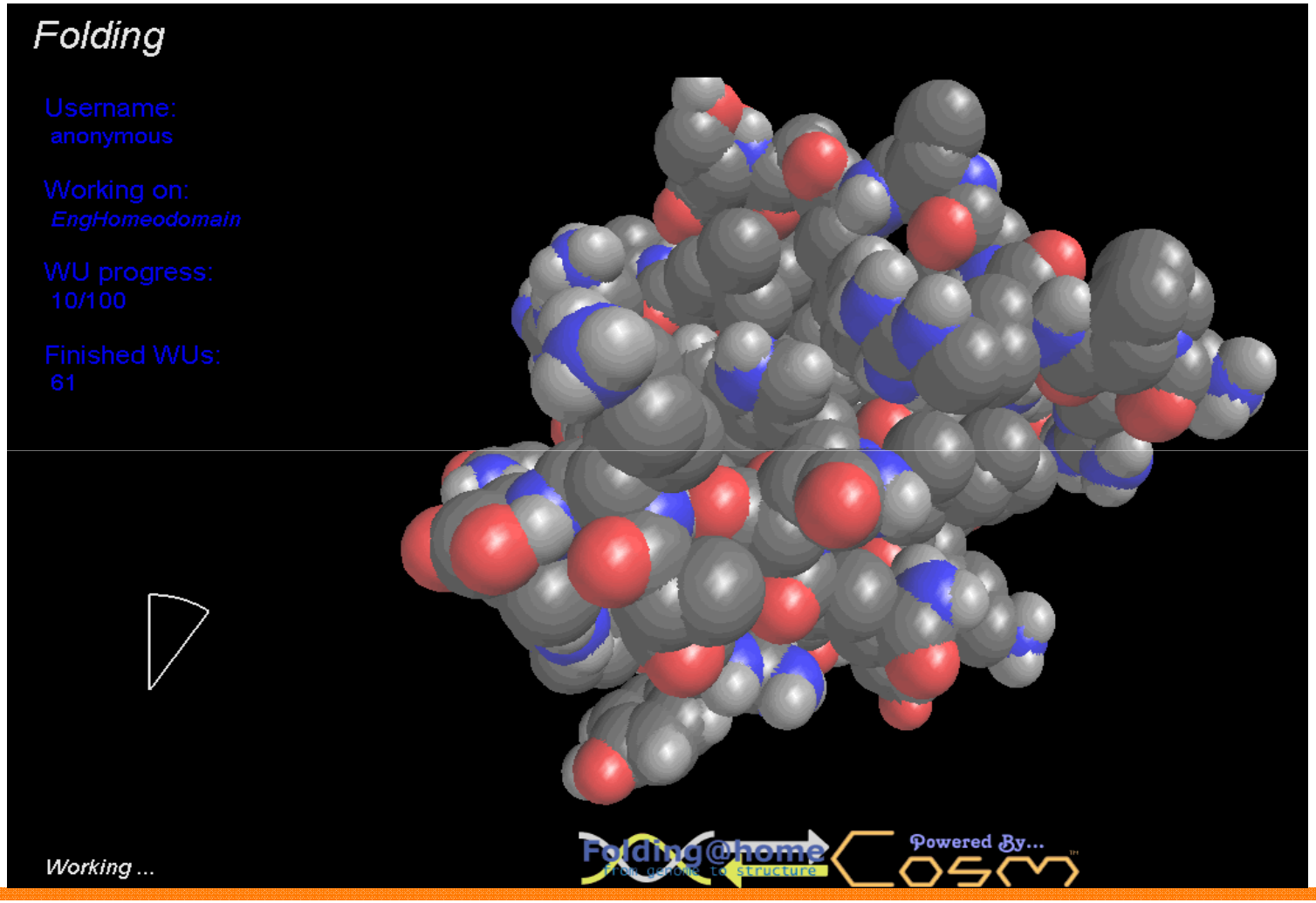
http://cool.pptv.com/

SETI@home (setiathome.berkeley.edu)

- Currently
  over 3 m
- SETI@h
  Internet-
  Extraterr
  by runnir
  radio tele

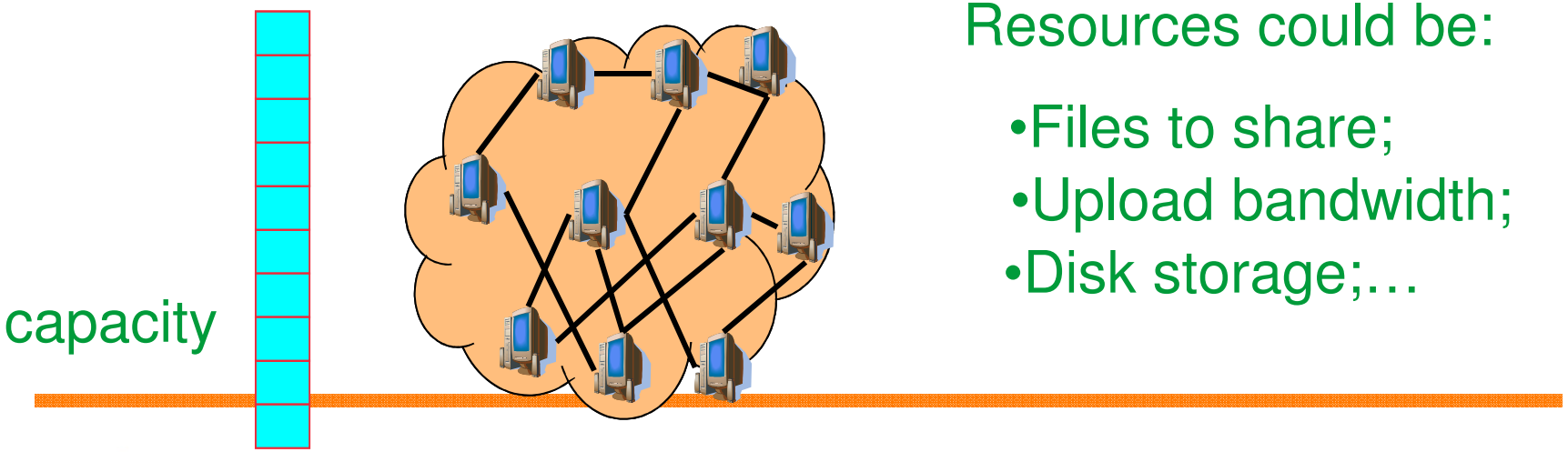


Folding@home (http://folding.stanford.edu/)



Why is P2P so successful?

- Scalable – It's all about sharing resources
  - No need to provision servers or bandwidth
  - Each user brings its own resource
  - E.g. resistant to flash crowds
    - flash crowd = a crowd of users all arriving at the same time



Resources could be:

- Files to share;
- Upload bandwidth;
- Disk storage;…

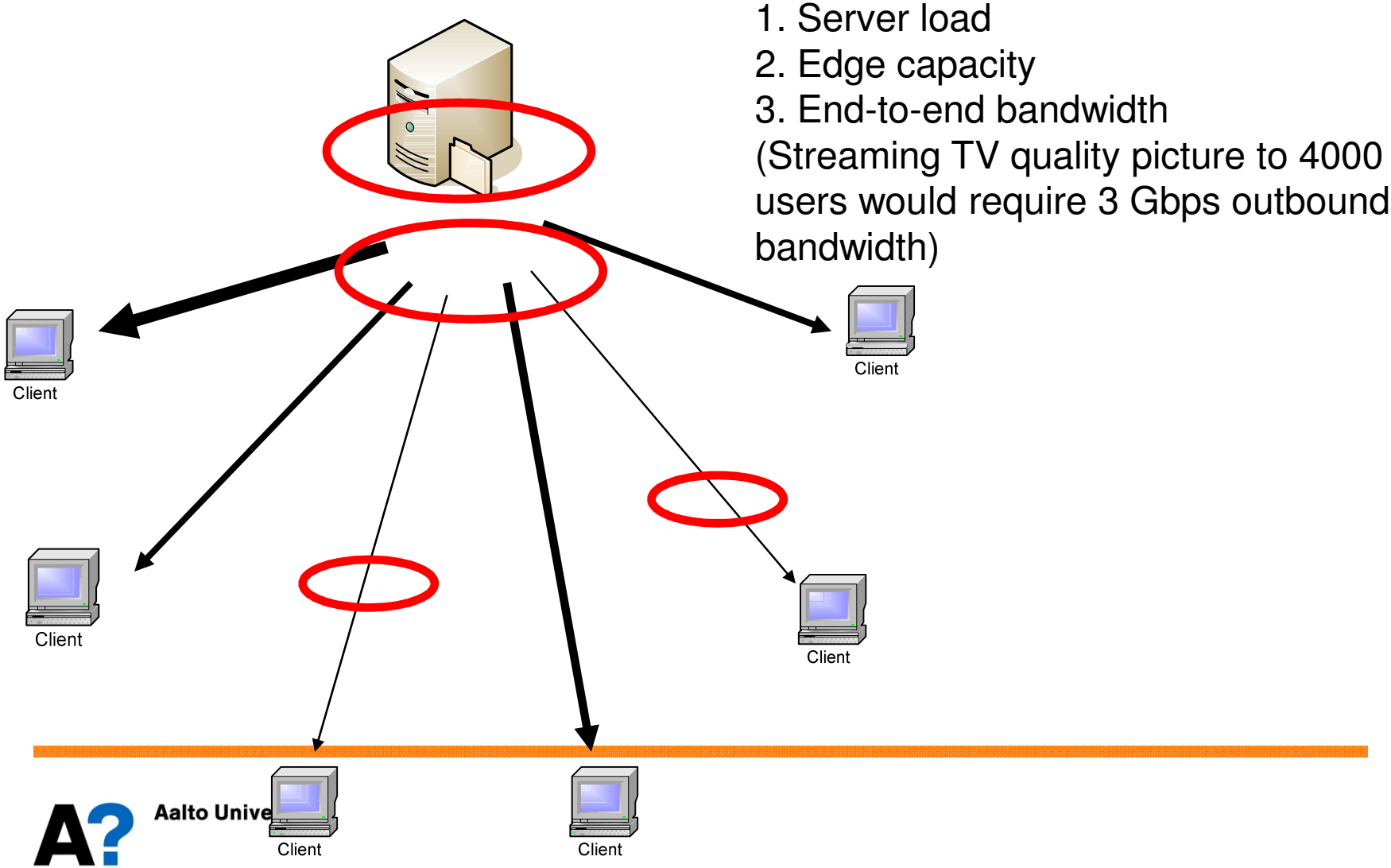
Why is P2P so successful? (cont'd)

- Cheap - No infrastructure needed
- Everybody can bring its own content (at no cost)
  - Homemade content
  - Ethnic content
  - Illegal content
  - But also *legal* content
  - …
- High availability – Content accessible most of time

Client/Server: Poor scalability

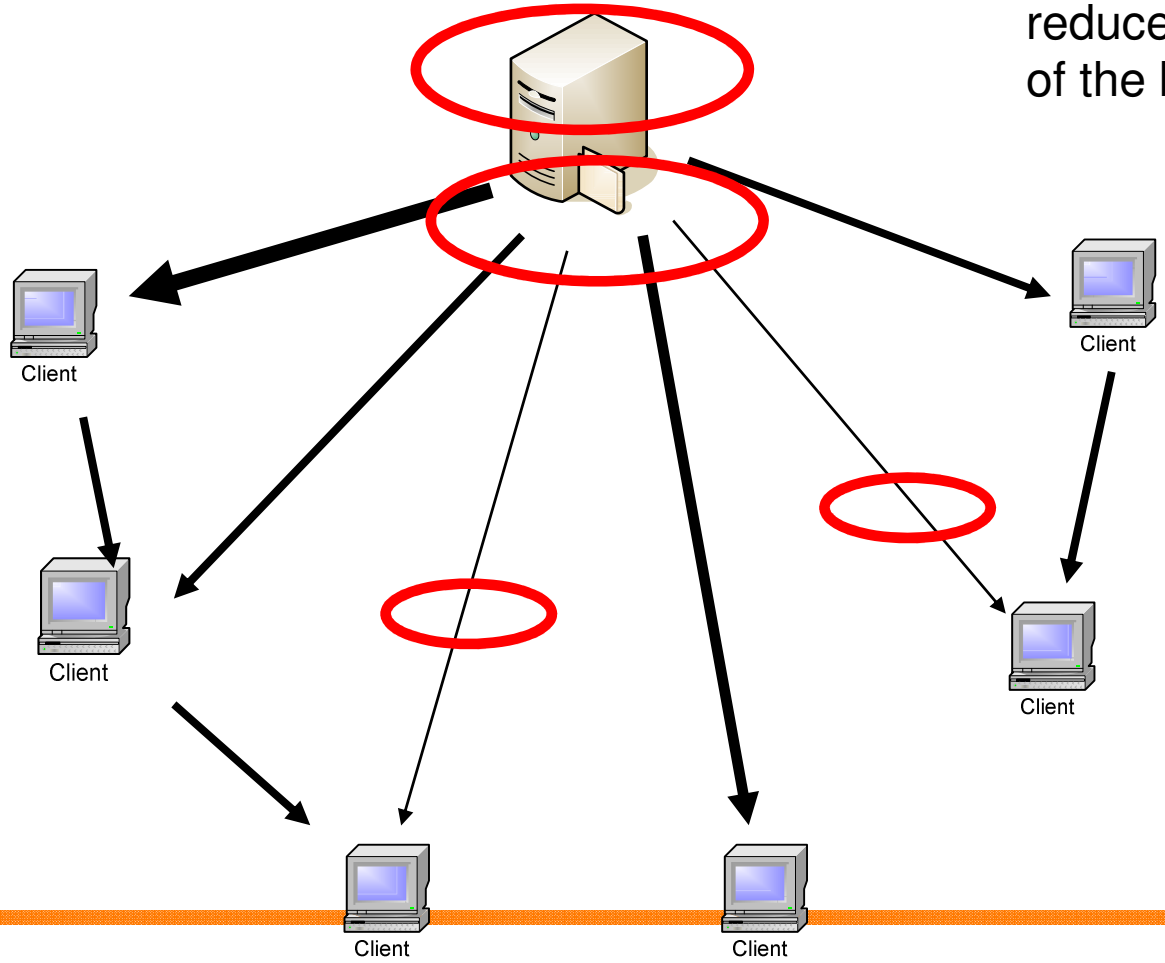
Three bottlenecks:
1. Server load
2. Edge capacity
3. End-to-end bandwidth

(Streaming TV quality picture to 4000 users would require 3 Gbps outbound bandwidth)



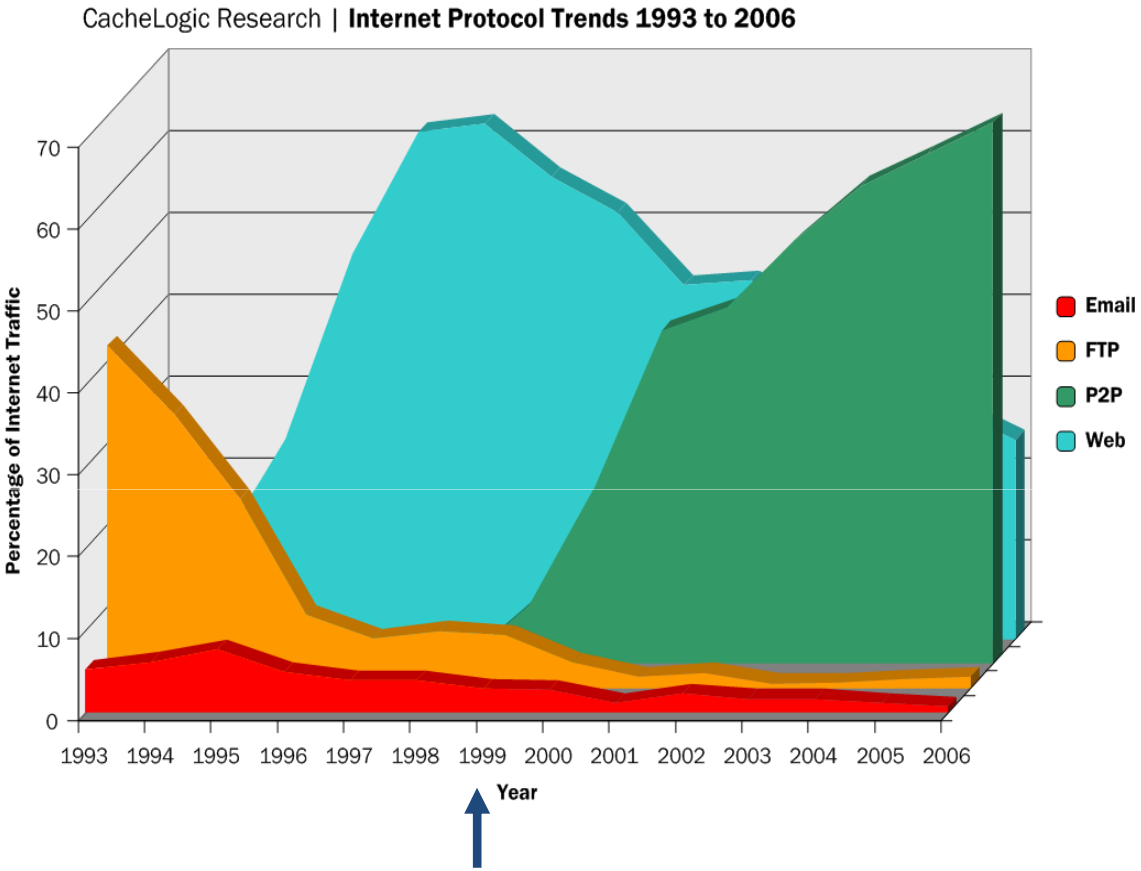
Collaborative Communications

Through cooperation, data transfer from the server can be reduced. Releases some or all of the bottlenecks.



"The server workload is reduced by 41% even when users share only videos while they are watching. When users share videos for one day, the server workload reduces by a tremendous 98.7%, compared to a client-server approach."
Cha, M., Kwak, H., Rodriguez, P., Ahn, Y., and Moon, S. 2007. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In Proceedings of the 7th ACM SIGCOMM Conference on internet Measurement (San Diego, California, USA, October 24 - 26, 2007).

Rise of P2P



P2P represented ~65% of Internet traffic at end 2006, CacheLogic 2007

- P2P data currently represents 44.0% of all consumer traffic over the Internet and 33.6% in North America. Much of this data is audio and video files (over 70%).
- P2P Traffic to Grow Almost 400% over the Next 5 Years
- legitimate P2P traffic is expected to grow 10 times as fast as illicit P2P traffic Multimedia Intelligence, November 2008

- 1999: *Napster*, first widely used p2p-application

Driving Forces Behind Peer-to-Peer

Development of the terminal capabilities:

- 1992:
  - Average hard disk size: ~0.3Gbyte
  - Average processing power (clock frequency) of personal computers: ~ 100MHz
- 2002-04:
  - Average processing power (clock frequency) of personal computers: ~ 3GHz
  - Average hard disk size: 100 Gbyte
  - → Personal computers have capabilities comparable to servers in the 1990s
- 2007: Nokia N95 mobile phone
  - ARM9 ~1GHz clock frequency
  - Up to 2GB external Micro SD

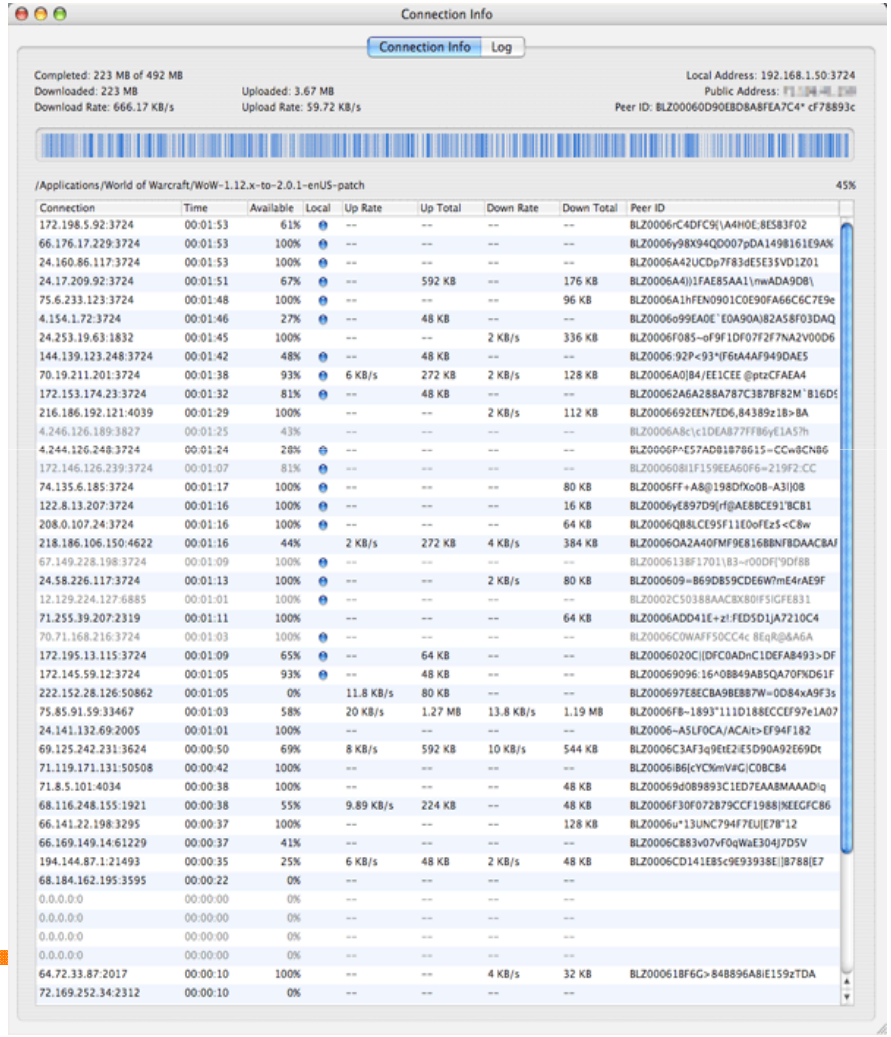
Development of the communication networks:

- Early 1990s: private users start to connect to the Internet via 56kbps modem connections
- 1999
  - Introduction of DSL and ADSL connections
  - Data rates of up to 8.5Mbps via common telephone connections become available
  - The deregulation of the telephone market shows first effects with significantly reduced tariffs, due to increased competition on the last mile
  - → bandwidth is plentiful and cheap!
- 2007 Nokia N95
  - HSDPA 1.8 Mbps

From Piracy to Business

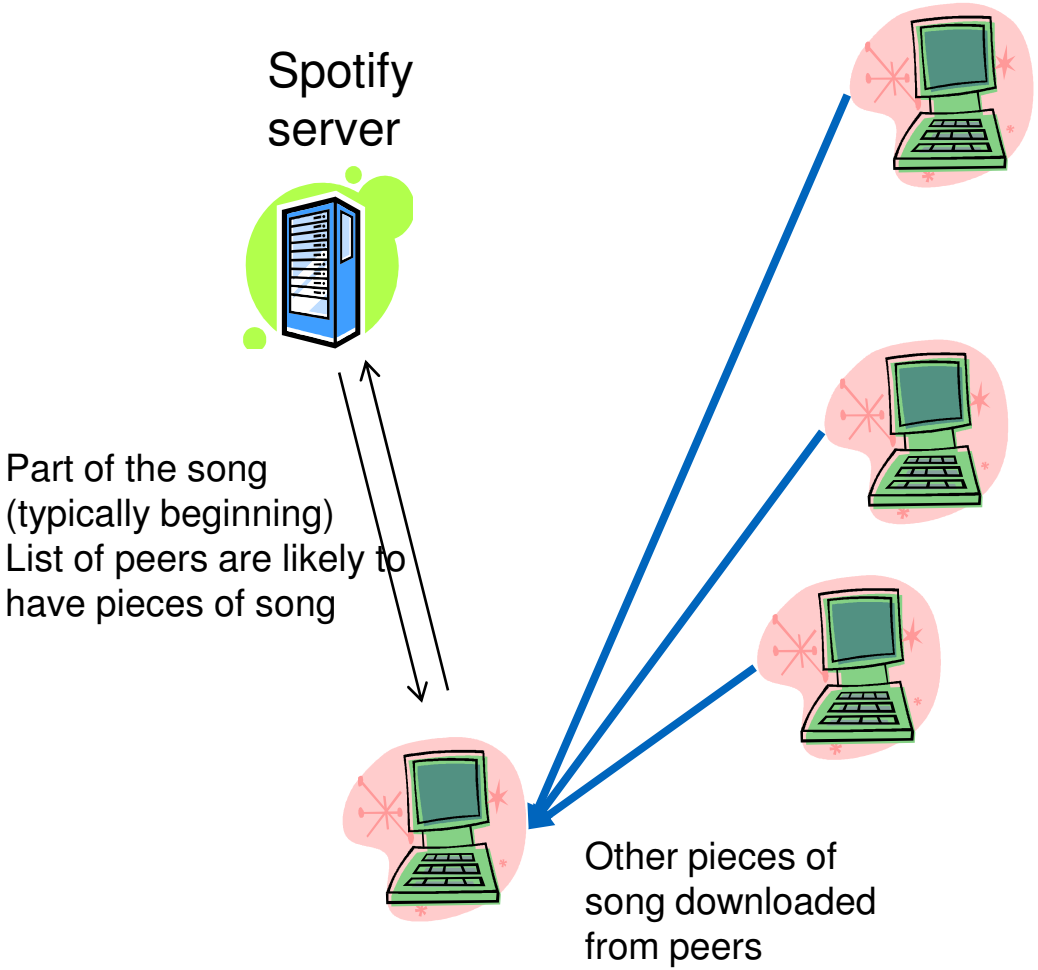
- **The share of P2P networks of Internet traffic is on decline**
  - 2007 about 40% of all traffic
  - 2009 about 18% of all traffic
  - Lähde: Arbor Atlas, 2009
- **P2P technologies increasingly used as a service platform**
  - Skype calls
  - Word-of-Warcraft updates
  - Spotify music streaming
- **According to theoretical analysis YouTube could save 40-98% of network capacity with P2P technology** (Cha et al., 2007)
  - This is a lot since it is estimated that YouTube data transfer costs are even $1.000.000 per day (Credit Suisse, 2009)

WoW Distribution of patches and software



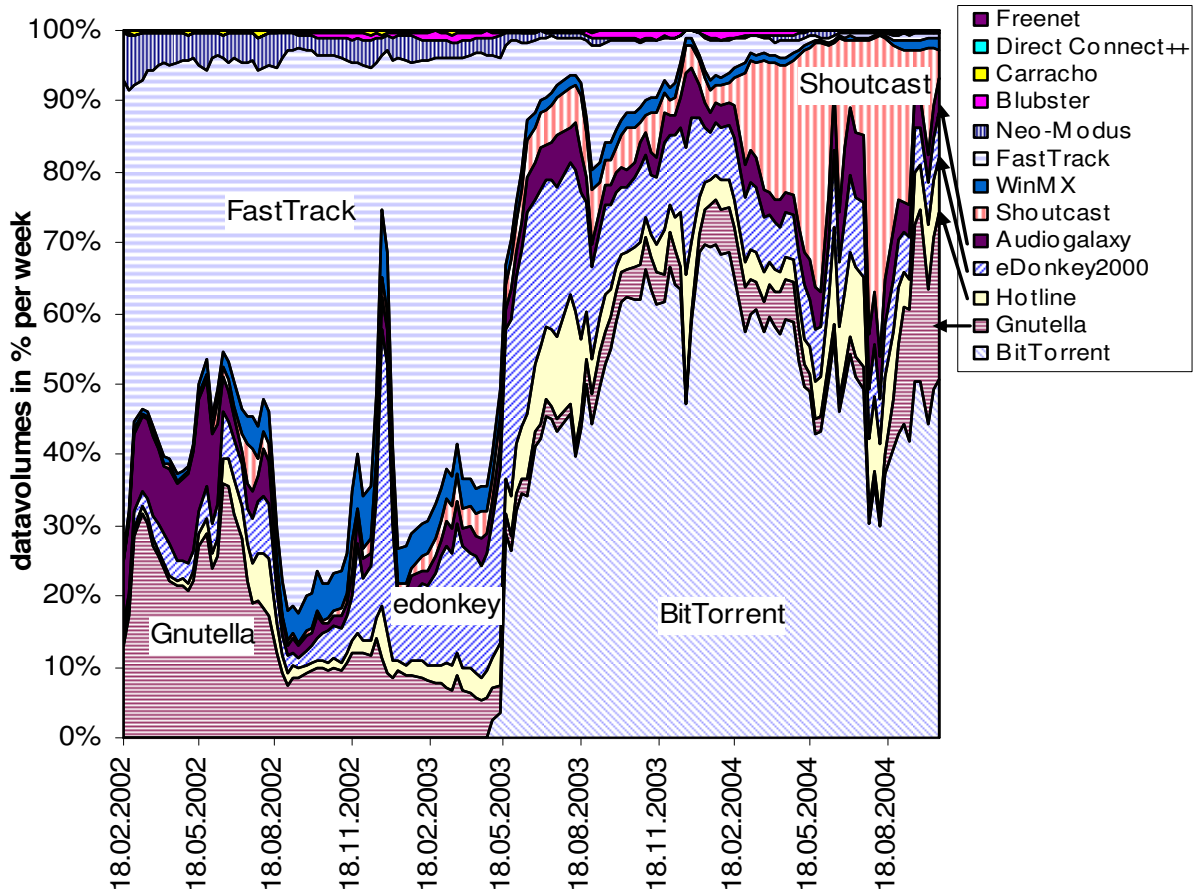
Aalto University

Spotify, 2008



- Music streaming, encrypted content
- Commercial, legal
- Proprietary protocol
- P2P reduces the load of Spotify servers

Development of P2P Applications



Traffic portions of the different P2P applications and protocols from the traffic measured per week in the Abilene backbone from 18.02.2002 until 18.010.2004

Aalto University

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

Some P2P research topics

Science of networks

Self-organization

Business models

P2P middleware

Applications

Mobile use

Copyrights & legal

Trust and reputation

Anonymity

Security

Incentives

Content search (unstructured, structured)

Content distribution (downloading, streaming)

Contact Information

- Course web page:
  - https://noppa.aalto.fi/noppa/kurssi/t-110.5150/etusivu
- Contact email:
  - zhonghong.ou@aalto.fi
- Office hour:
  - Fri 10-11 room A 109
- Qustions & Suggestions?