Peer-to-peer

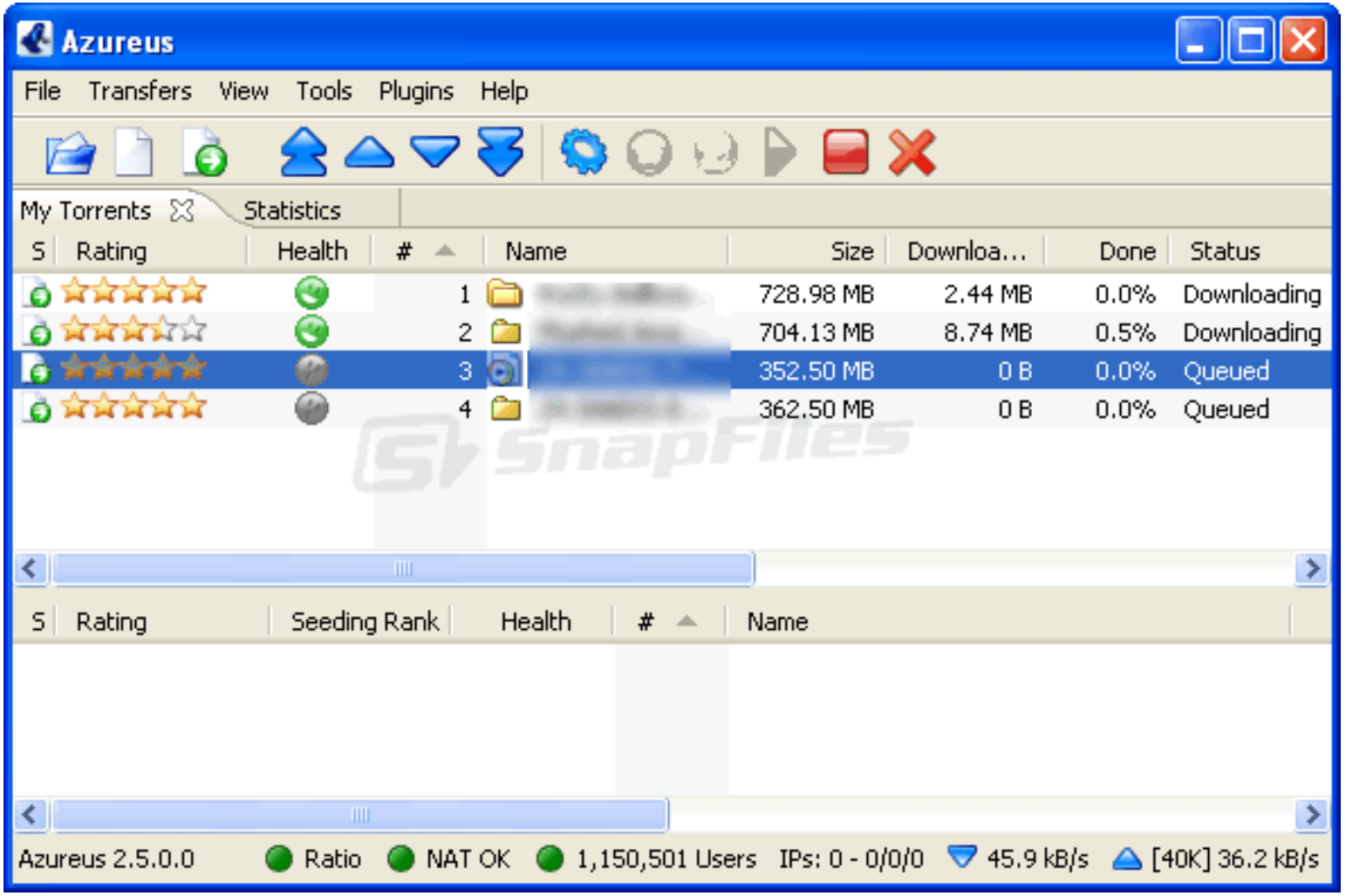
T-110.7100 Applications and Services in Internet, Fall 2009

Jukka K. Nurminen

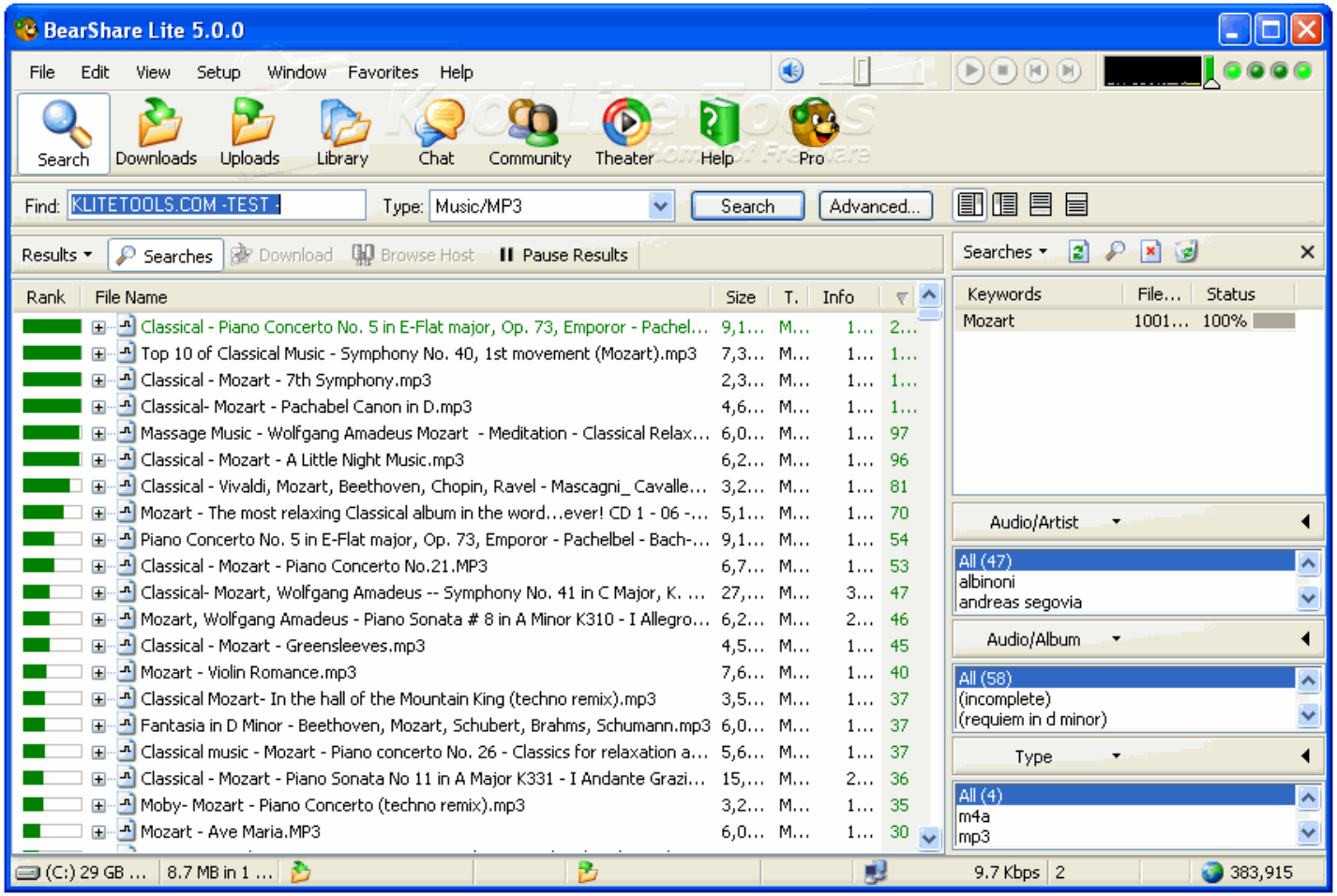
Schedule

Tue 15.9.2009 12-14	Introduction to P2P (example P2P systems, history of P2P, what is P2P)	Content delivery (BitTorrent and CoolStreaming)
Tue 22.9.2009 12-14	Unstructured content search (Napster, Gnutella, Kazaa)	Structured content search (DHT)
Tue 29.9.2009 12-14	Energy-efficiency & Mobile P2P	

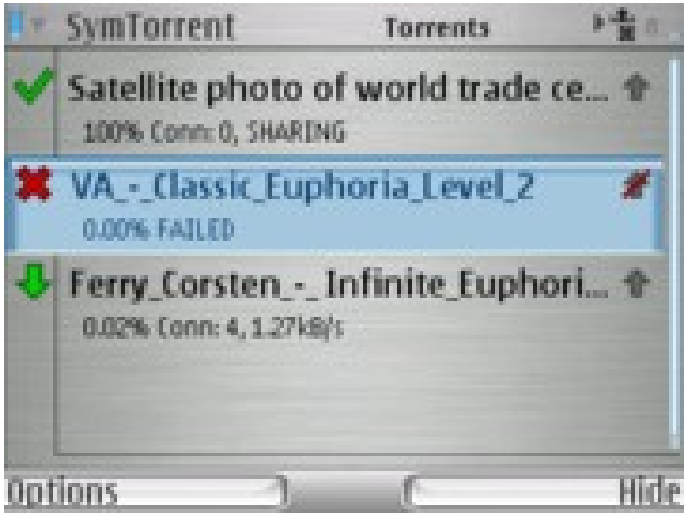
Azureus BitTorrent client



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
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

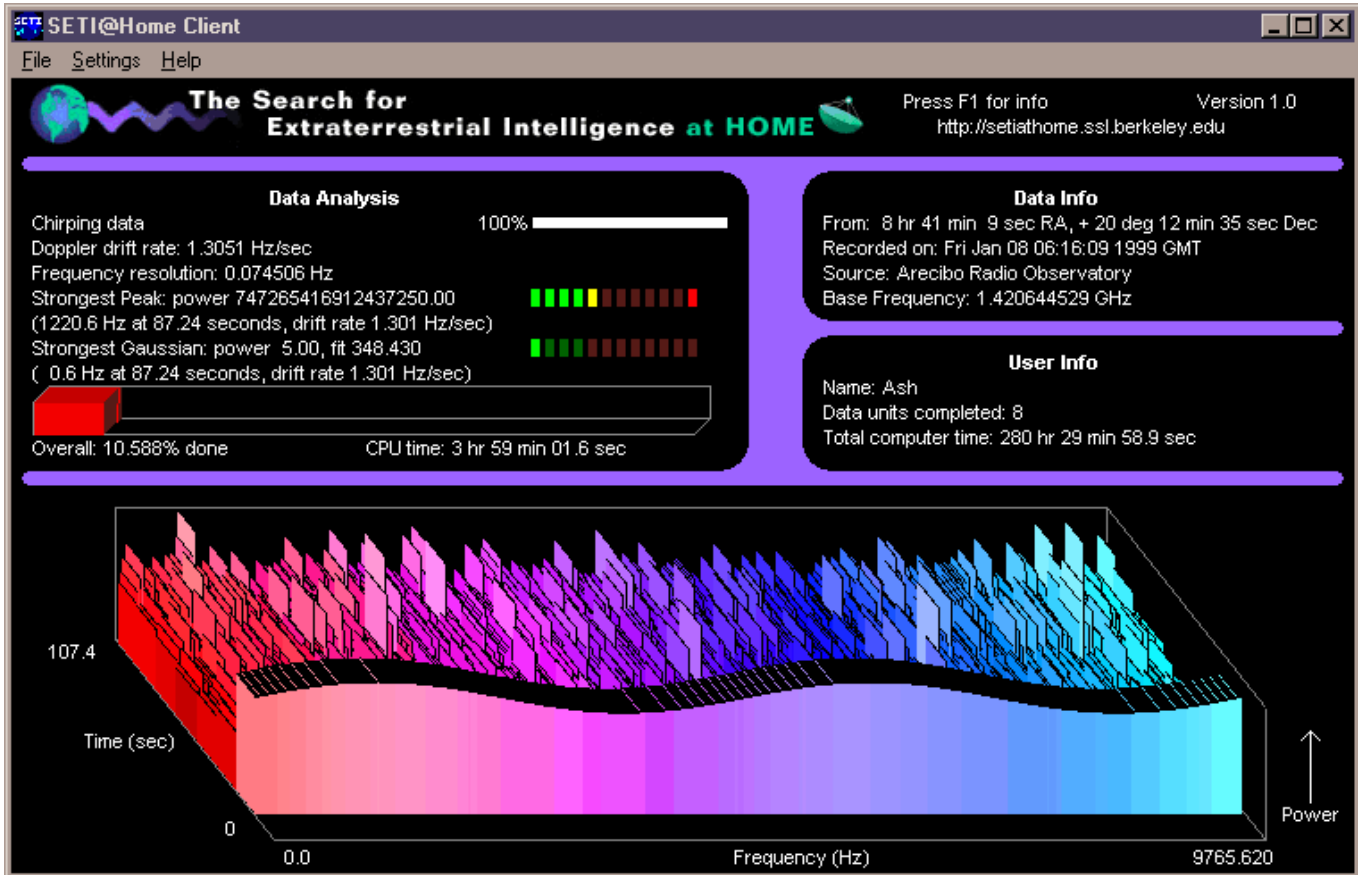


Skype

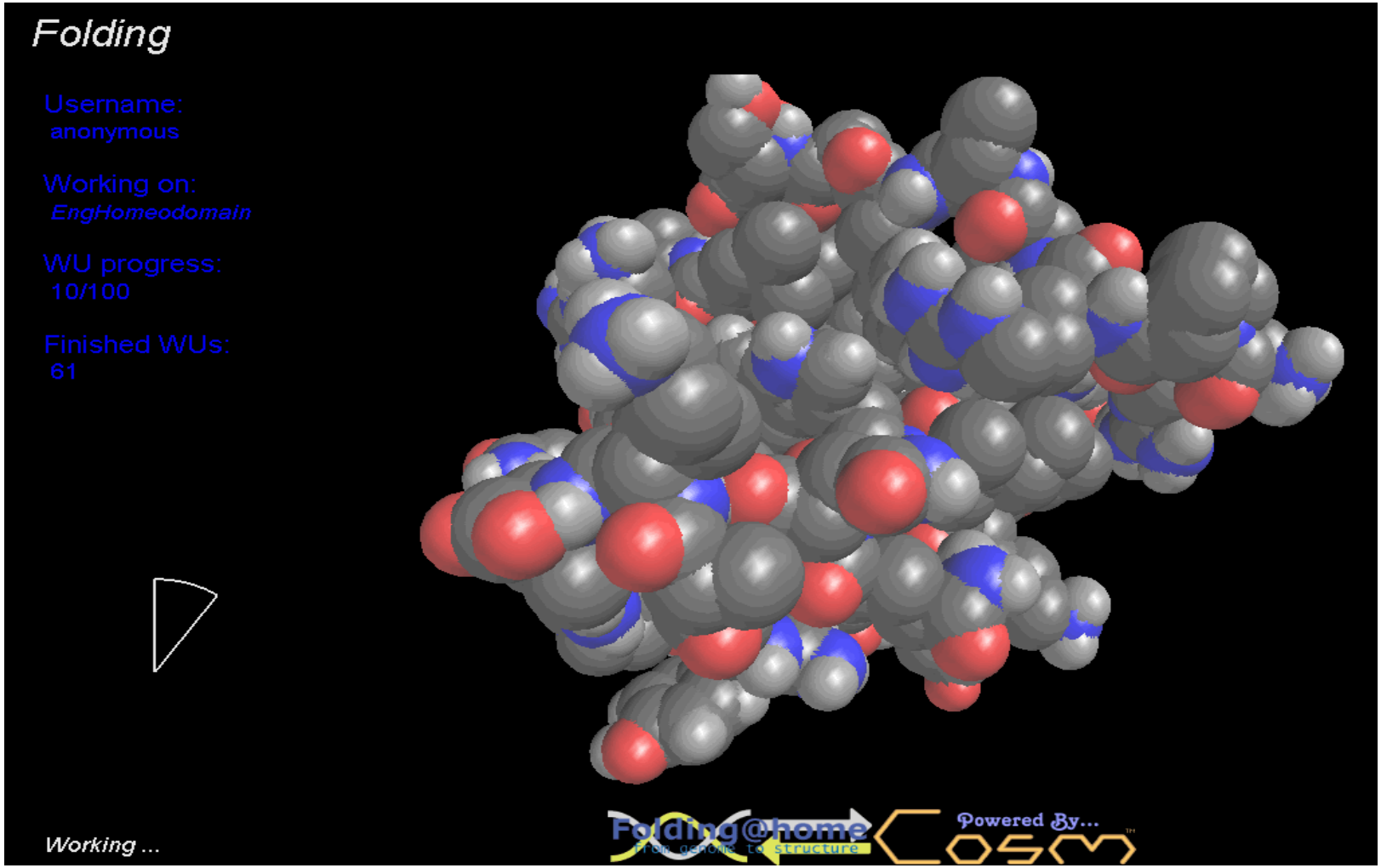
How skype works: http://arxiv.org/ftp/cs/papers/0412/0412017.pdf

SETI@home (setiathome.berkeley.edu)

- Currently the largest distributed computing effort with over 3 million users
- SETI@home is a scientific experiment that uses Internet-connected computers in the Search for Extraterrestrial Intelligence (SETI). You can participate by running a free program that downloads and analyzes radio telescope data.

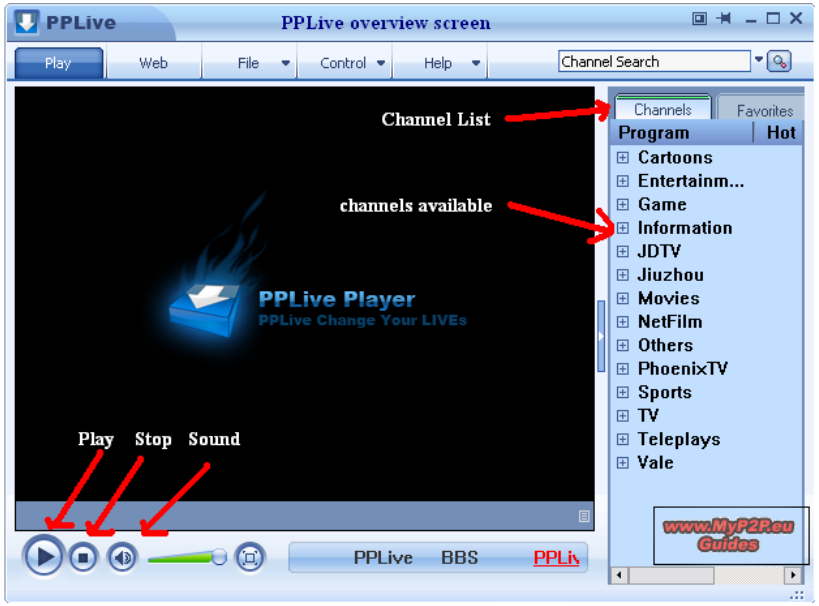


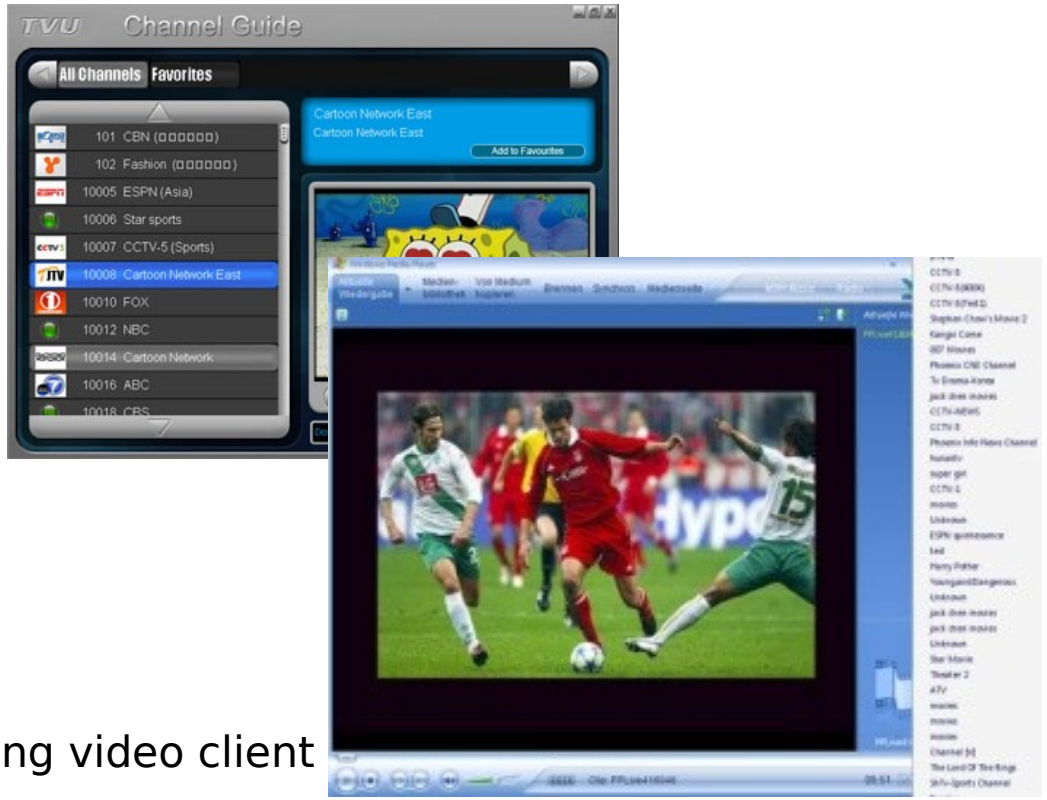
Folding@home (http://folding.stanford.edu/)



PPLive, TVU, …

“PPLive is a P2P television network software that famous all over the world. It has the largest number of users, the most extensive coverage in internet.”



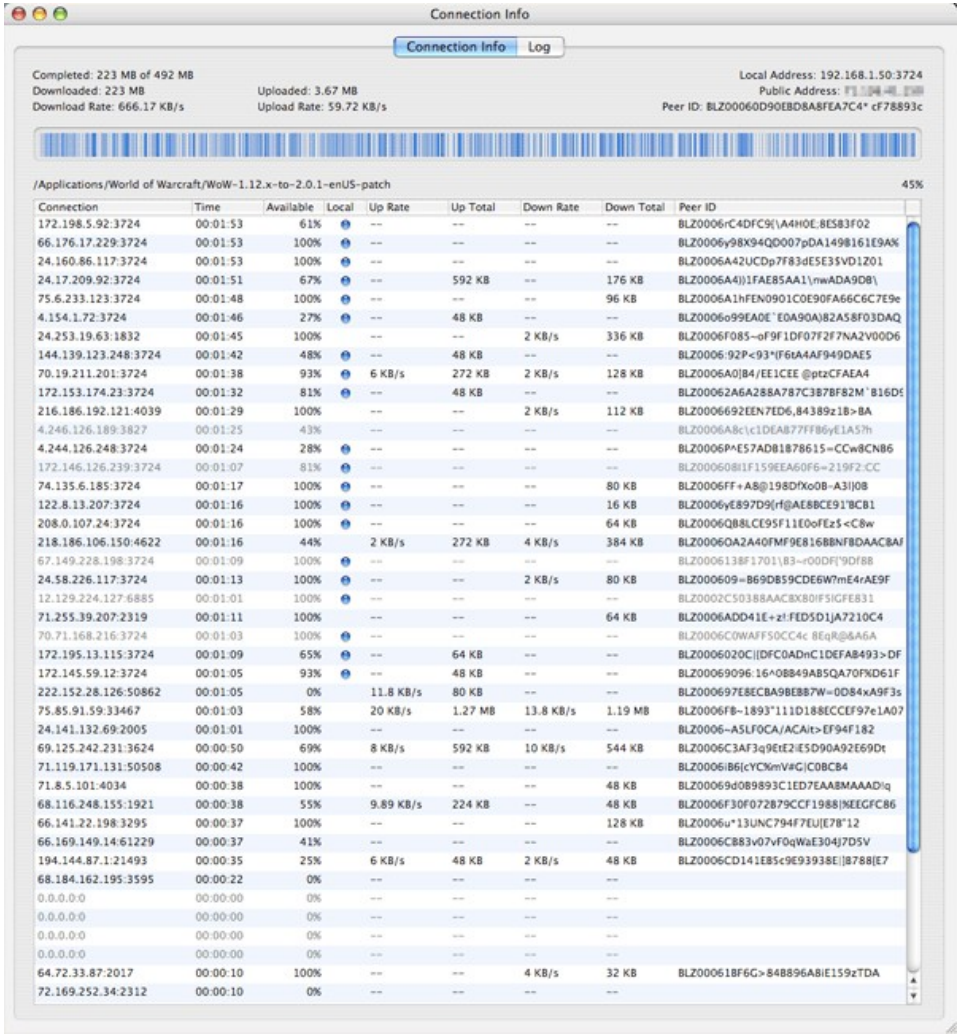


PPLive

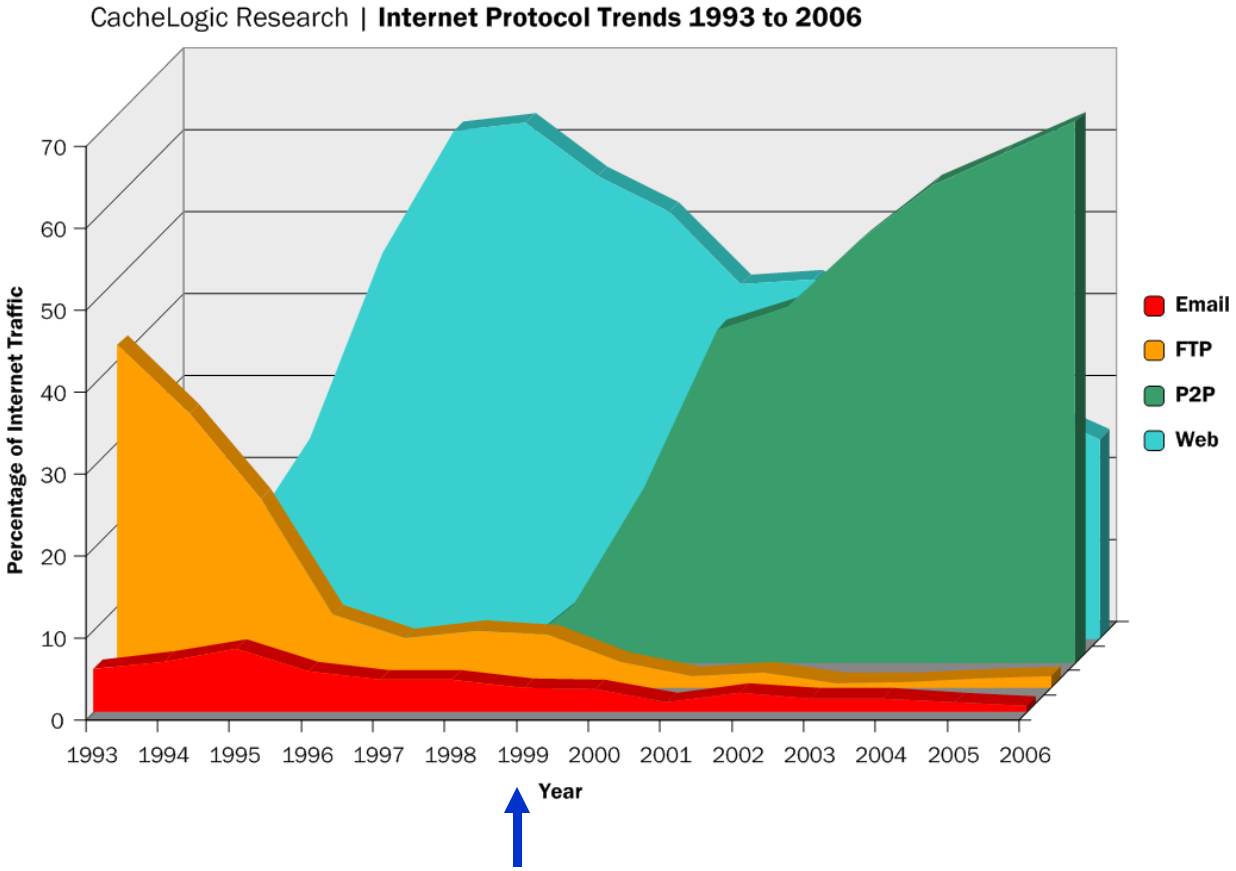
- 100 million downloads of its P2P streaming video client
- 24 million users per month
- access to 900 or so live TV channels
- 200 individual advertisers this year alone

Source: iResearch, August, 2008

WoW Distribution of patches and software



Rise of P2P

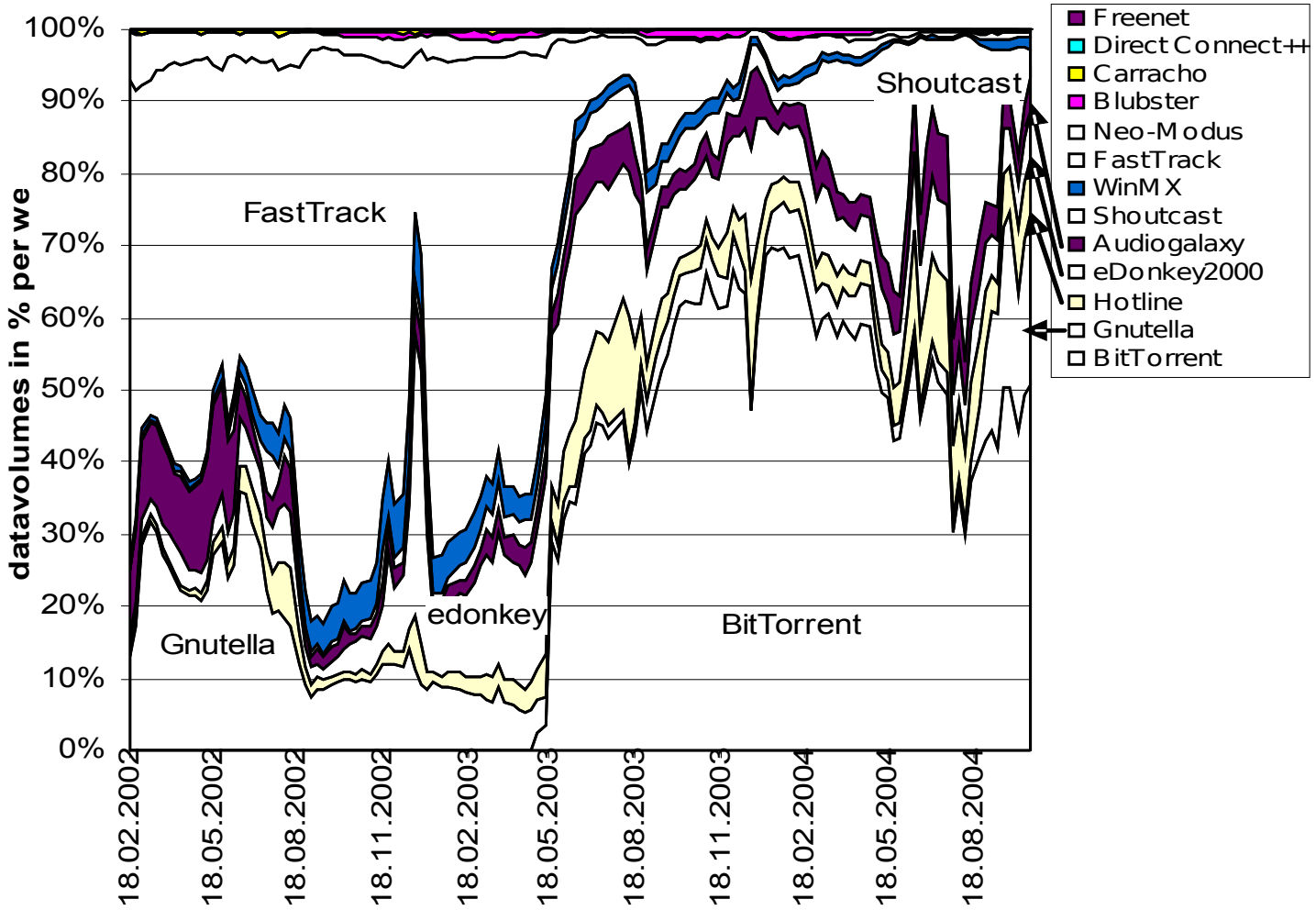


- 1999: *Napster*, first widely used p2p-application

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

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

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

Evolution of P2P via example applications

P2P Protocols:

- 1999: Napster
- 2000: Gnutella, eDonkey
- 2001: Kazaa
- 2002: eMule, BitTorrent
- 2003: Skype
- 2004: PPLive
- 2005: TVKoo, TVAnts, PPStream, SopCast…
- 2006: WoW distribution via P2P
- 2007: Joost, Vuze
- Next: Gaming, mobile P2P

Application type:

File Download

Streaming

Telephony

Video-on-Demand

Gaming

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

Definition of Peer-to-peer (or P2P)

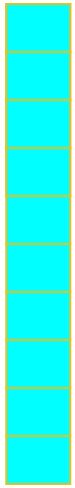
- A peer-to-peer (or P2P) computer network is a network that relies primarily on the computing power and bandwidth of the participants in the network rather than concentrating it in a relatively small number of servers.
- A pure peer-to-peer network does not have the notion of clients or servers, but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

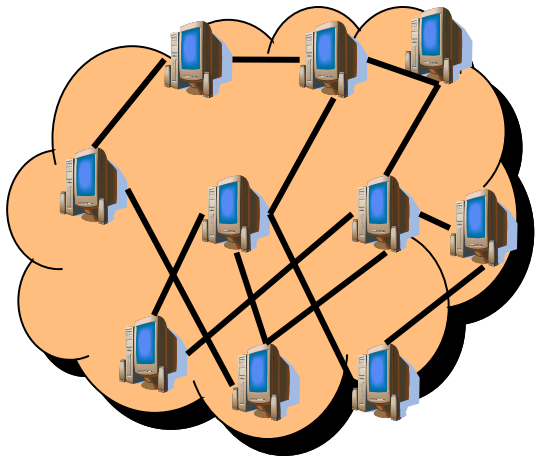
Taken from the wikipedia free encyclopedia - www.wikipedia.org

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

capacity





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
  - ...

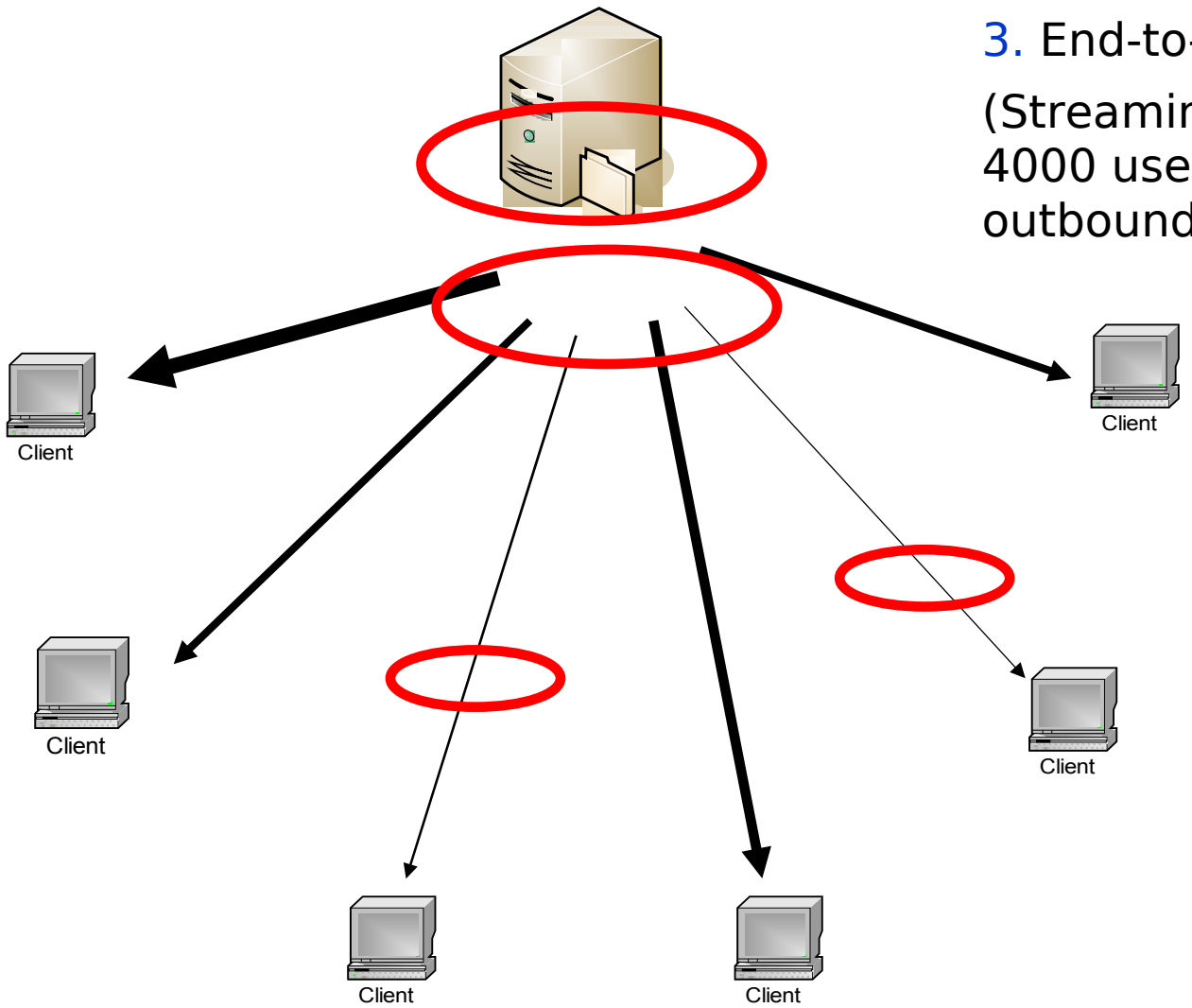
- High availability – Content accessible most of time

Client/Server: Poor scalability

3 bottlenecks

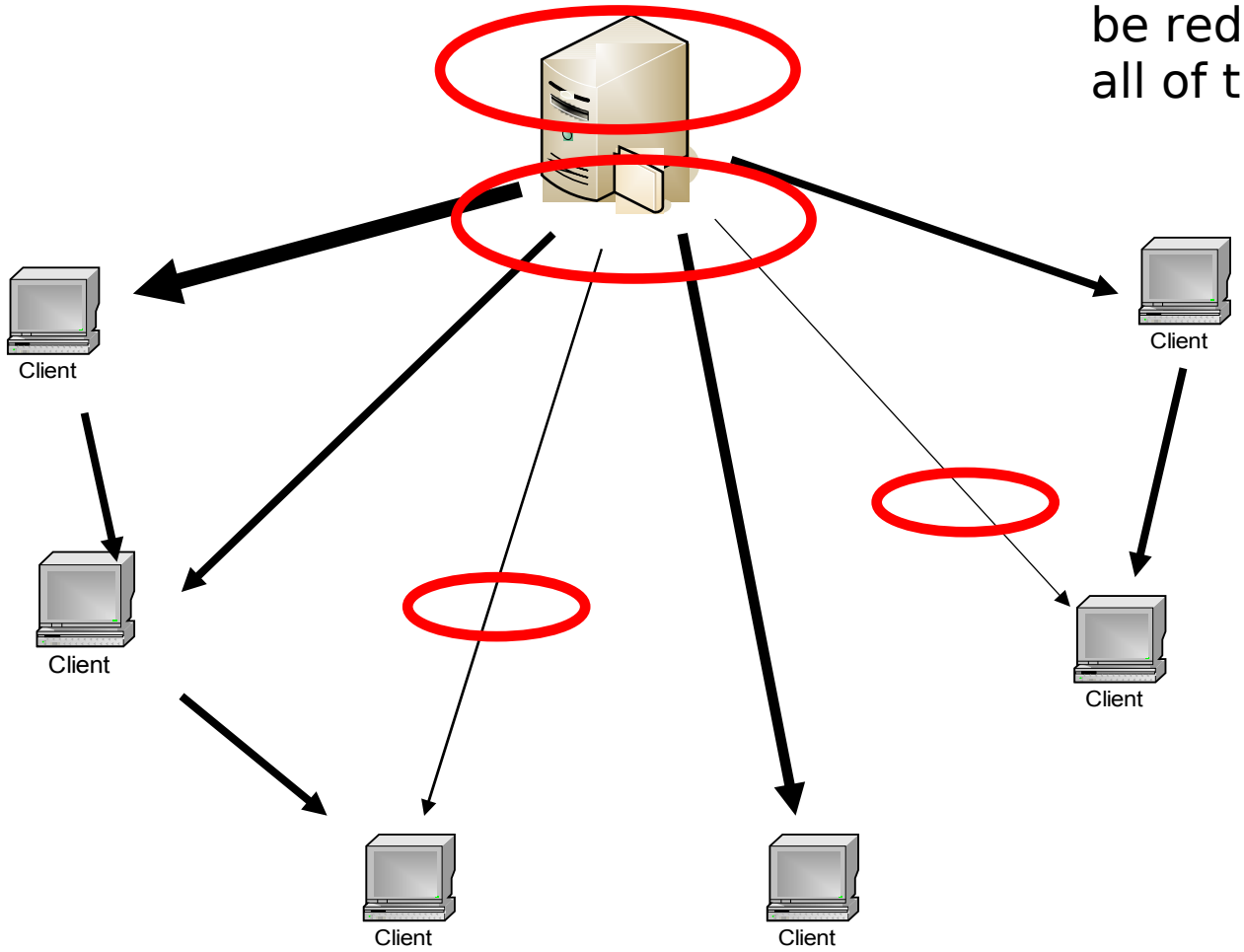
1. Server load
2. Edge capacity
3. End-to-end bandwidth

(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).

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)