NET0183 Networks and Communications

BitTorrent

All you want to know about *the* protocol

This assignment is all my own work. This work has not been submitted for assessment in any other context. I have not knowingly allowed others to copy my work.

Introduction [1]

What is BitTorrent protocol?

- peer-to-peer (P2P) file sharing protocol used for distributing large amounts of data
- one of the most common protocols for transferring large files
  - estimated to account up to 55 % of all Internet traffic (February 2009) [2]
- designed by programmer Bram Cohen (born in 1975)
- first implementation released on 2 July 2001
- maintained by Cohen's company BitTorrent, Inc.

Glossary [3, 4]

announce – the act of connecting to a tracker to update your status and to obtain information from it (eg. peer list)

client – the program that enables p2p file sharing via the BitTorrent protocol

hash – a unique "fingerprint" in the torrent file used in data verification

leecher – (negative) word for a downloading peer who does't contribute back

P2P (peer-to-peer) – file sharing technology where all nodes (computers) act as a client and a server, capable of both receiving and sending data

peer – a downloading user connected to the swarm with a BitTorrent client

piece – torrented files is divided into pieces which are distributed in a random fashion among peers

ratio –the amount of data uploaded / the amount of data downloaded

seeder – a peer with 100% of the data in the torrent contents

swarm – the collective group of peers (including seeds) that are connected by a common .torrent file

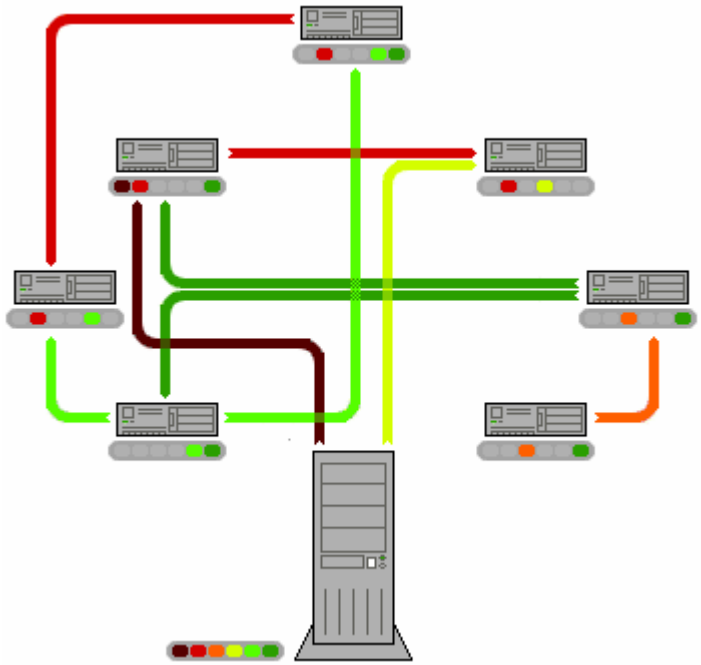
torrent file – contains metadata about the files and address of trackers

tracker – a server that keeps track of which seeds and peers are in the swarm

BitTorrent in Action [1]

How does BitTorrent work?

1. a peer creates a small (~50 kB) .torrent file
2. .torrent file is uploaded to the tracker
3. the initial .torrent file creator starts seeding
4. other peers download the .torrent file
   i. connect to the tracker to get the peer list
   ii. start downloading the pieces from the initial seeder and other peers



BitTorrent Protocol [1, 5]

How does BitTorrent differ from conventional downloading?

- many small data requests over different TCP connections to different machines
  - average packet size 587 kB [2]
- BitTorrent downloads file pieces in a random or in a "rarest-first" approach that ensures high availability
  - segmented downloading
  - does not allow streaming
- BitTorrent is not bound to specific port (typical ports 6881-6889)

Micro Transport Protocol (µTP) [6, 7, 8]

Torrent's switch to UDP

- developed within BitTorrent, Inc.
  - still work in progress
  - implemented in µTorrent 1.9 and newer
- µTP is transport protocol layered on top of UDP
  - must implement its own congestion control
  - packet loss control
  - own framing scheme similar to TCP (timestamps and SACK)
- fair protocol
  - tries to utilize the unused bandwidth fully
  - does not distrup other internet connections
  - yields to TCP

Adaptation [1, 9, 10]

Blizzard Entertainment

- »uses BitTorrent ("Blizzard Downloader") to distribute most content for World of Warcraft, including the game itself
- »distributes other Blizzard's game trailer (StarCraft II, Diablo III)

ASUS

- »offers BitTorrent download option besides regular HTTP for drivers and other software

Twitter

- »uses BitTorrent for server deployment
  - »Twitter servers act as a large BitTorrent swarm that will distribute file updates between them -> server-to-server technology

Network Problems [1, 11, 12]

Home routers lock up

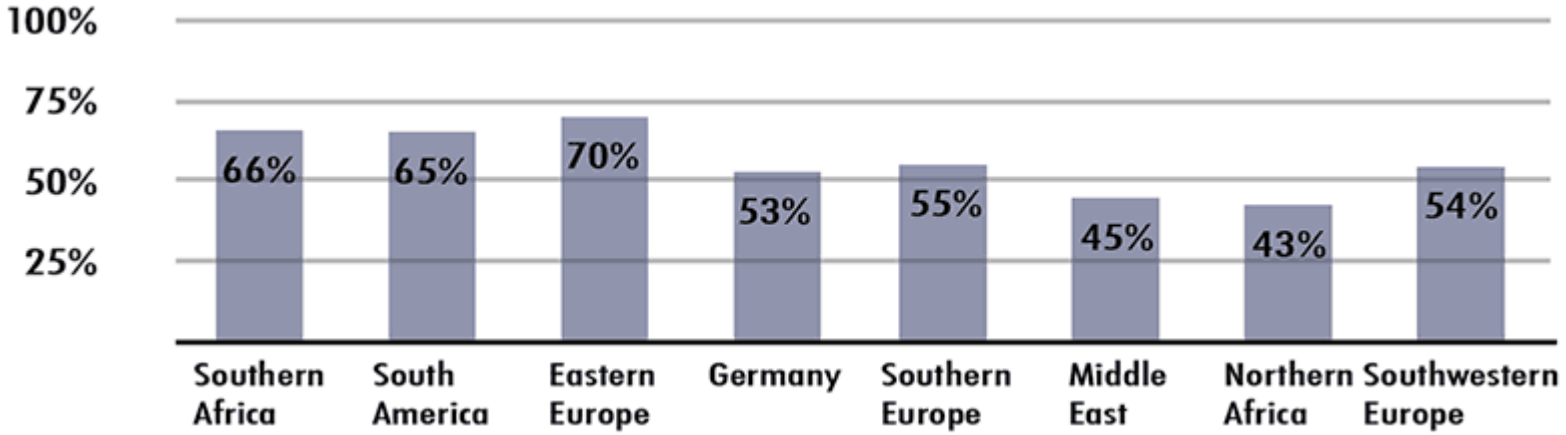
- NAT routers must maintain tables of source and destination IP addresses and ports
  - Cheap home routers are limited to about 2000 table entries
  - BitTorrent frequently contacts 300-500 servers per second rapidly filling the NAT tables

BitTorrent protocol overhead

- BitTorrent traffic analysis in large network in China (2008)
  - overhead comprises up to 80 % of total BitTorrent traffic
  - DHT (distributed hash table) is the major source of overhead

Network Impact [2]

Relative P2P traffic volume



Most popular P2P protocols

Protocol	Southern Africa	South America	Eastern Europe	Northern Africa	Germany	Southern Europe	Middle East	SW Europe
All P2P	65,77%	65,21%	69,95%	42,51%	52,79%	55,12%	44,77%	54,46%
BitTorrent	48,34%	30,02%	80,83%	74,51%	70,77%	48,94%	78,85%	58,20%
DirectConnect	0,01%	0,00%	17,87%	0,08%	0,85%	0,00%	0,12%	0,30%
eDonkey	2,48%	25,99%	1,16%	7,70%	24,22%	47,17%	15,37%	35,99%

Conclusions

- BitTorrent protocol is fairly new
  - still being developed + new applications & uses
  - massive traffic overhead -> room for improvement
- cheap way to distribute large files
  - does not need expensive servers or a big centralized bandwidth
  - slowly, commercial companies begin to realise its advantages too
- account for big portion of total Internet traffic
  - some ISPs try to throttle BitTorrent traffic
  - Internet is not neutral, some ISPs favour other traffic over BitTorrent
- controversial protocol
  - distribution of copyrighted material quite easy

Appendix: References

1) http://en.wikipedia.org/wiki/BitTorrent_(protocol)
2) Schulze H, Mochalski K. Internet Study 2008/2009 [pdf on the Internet]. Leipzig, Germany: ipoque; 2009 [cited 2010 Apr 12]. Available from: http://www.ipoque.com/resources/internet-studies/internet-study-2008_2009
3) http://www.bittorrent.com/btusers/guides/bittorrent-user-manual/glossary
4) http://en.wikipedia.org/wiki/Terminology_of_BitTorrent
5) Cohen B. The BitTorrent Protocol Specification [html on the Internet]. San Fransisco, CA: BitTorrent, Inc.; 2008 [cited 2010 Apr 12]. Available from: http://www.bittorrent.org/beps/bep_0003.html
6) http://en.wikipedia.org/wiki/UDP_Torrent_Protocol
7) Nordberd A. uTorrent transport protocol [html on the Internet]. San Fransisco, CA: BitTorrent, Inc.; 2010 [cited 2010 Apr 12]. Available from: http://bittorrent.org/beps/bep_0029.html
8) Van Beijum I. µTorrent's switch to UDP and why the sky isn't falling. [html on the Internet]. arsTechnica; 2008]cited 2010 Apr 12]. Available from: http://arstechnica.com/old/content/2008/12/utorrents-switch-to-udp-and-why-the-sky-isnt-falling.ars
9) http://torrentfreak.com/twitter-uses-bittorrent-for-server-deployment-100210/

Appendix: References

10) http://torrentfreak.com/asus-uses-bittorrent-to-boost-downloads-090720/
11) Ellis L. BitTorrent's Swarms Have a Deadly Bite On Broadband Nets. Multichannel News; 2006 [cited 2010 Apr 12]. Available from: http://www.multichannel.com/article/87516-BitTorrent_s_Swarms_Have_a_Deadly_Bite_On_Broadband_Nets.php
12) Qi J, Zhang H & Ji Z. Analyzing BitTorrent Traffic Across Large Network. International Conference on Cyberworlds 2008: IEEE Computer Society; 2008.