

The bittorrent p2p file-sharing system: measurements and analysis

Johan Pouwelse, Paweł Garbacki, Dick Epema, Henk Sips

2005

Measurement

- Supernova (Slovanian based), closed in 2004, legal threats. Ordinary tracker, used mirroring system (mirror sites)
- Detailed trace gathered over 8 months (2003 June – 2004 March), maximum of the trace: 54845 peers informations in a two weeks period
- Measurement software:
 - Global: access to supernova tracker informations
 - Local: 100 peers (distributed supercomputers), who gather information from their neighbor
- Goal: understanding and modeling P2P systems

Analysis

- Overall system activity – number of downloads per time
 - Notice daily cycle: minimum and maximum at same time each day (max at 23:00 GMT, min at 11:00)
 - Strongly influence by the availability of the trackers (Supernova architecture was vulnerable, DoS attacks)
- Availability (peers) - examined a three months window (one minute resolution), 90155 peers:
 - Almost 60% of peers were not behind firewalls
 - Free riding: only 17% had an uptime longer than one hour after finished downloading (3.1% 10 hours) → Long uptimes are rare (incentives to seed)

Analysis

- Integrity
 - Meta-data: hosted a mirror of Supernova, corrupted meta-data → cause integrity and privacy risks
 - Content: actively tried pollute the system in decentralized way, failed because of the moderators:
 - Moderation very effective on remove fake and corrupted files (20 moderator can manage 8000 submitters with 0.008 omission)
 - Centralized, hard to distribute
- Flashcrowds effect – sudden popularity of a new content:
 - One seeder, a lot of leechers → time gap is long until the first downloader who finishes → could be a problem
 - Influence by the tracker's availability and the only seeder's properties (e.g. is it behind a firewall ?)
 - After that, typically highly scalable
 - Arrival process not Poisson process

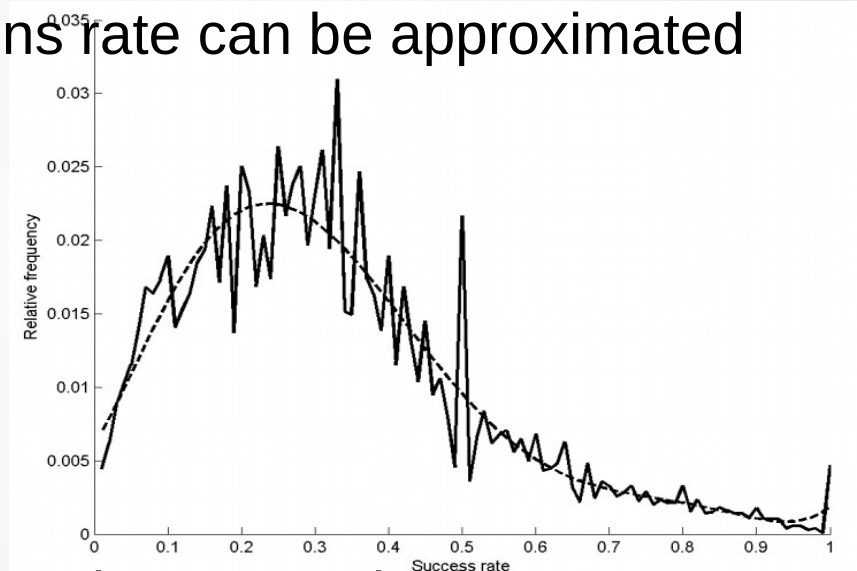
Analysis and conclusions

- Download performance
 - Outdated bandwidth measurement (avg 240 kbps)
 - Power-law relation between avg download speed and number of downloaders at that speed
 - Number of seeders after 10 days is not an accurate predictor for the content lifetime (one seeder should be enough)
- Conclusions
 - BitTorrent itself not have incentives to seed but trackers that do so. What can be incentives to seed ? (this course)
 - Supernova has high integrity (centralized moderators) but low system availability, decentralized solutions of tracker has low integrity, but high availability (Tribler)

Another related measurement study:

- Analysis of Mobile BitTorrent Client Behavior - Kristóf Csorba, Péter Ekler, András Bori, Hassan Charaf (2013):

- Distribution of successful connections rate can be approximated by a polynomial of 6th degree:



- Defining and understanding smartphone churn over the internet: a measurement study - Árpád Berta, Vilmos Bilicki, and Márk Jelasity (2014)
 - Also measured the same daily activities
 - Smart phones: more than 95% of peers behind firewall