Modelling BitTorrent with PEPA

Ashok Argent-Katwala*

Abstract

In this talk I will present some initial work modelling BitTorrent, a very popular peer-to-peer application, using PEPA.

1 About BitTorrent

BitTorrent [1] is a popular peer-to-peer application for transferring large files to a wide audience, using the peers to share the load of distribution.

Anyone can publish content using BitTorrent by creating a small file with metadata about the content (the `.torrent` file) and publishing that, typically on the Web. The `.torrent` includes the size and name of the file or files in the torrent, the size of each *piece*, an SHA-1 hash of each piece and the address of a *tracker* that co-ordinates all the peers. They then *seed* the file by starting a BitTorrent client on some machine that has the full file.

When other peers start downloading the file, they receive from the tracker a random selection of peers currently downloading or serving this torrent. They first ask those peers for pieces at random, and after they have their first complete piece they try to download the rarest pieces first. When a peer finishes downloading a piece, they check it is valid using hashes stored in the `.torrent` file and inform the peers they are connected to that they have it.

Once a peer has the complete file, and until its owner shuts down the client that peer becomes another seed. This means that once there are enough peers with enough of the file still running the original publisher can remove the initial seed without harming the distribution of the file. However, the distribution will stall if there are pieces that no active peer has.

2 Models

The models to be presented differ substantially from the one presented by Qiu and Srikant in [2], where a fluid model is used to study a network of

*Department of Computing, Imperial College London. `a.argent-katwala@doc.ic.ac.uk`

identical peers at steady-state. Here we hope to capture more of the variety between peers, since even only accounting for home users there is a wide variety in the speed of connections people will be using.

There are peculiar challenges for peer-to-peer network applications, where we have large number of agents, which we would like to be identical. This will be discussed in further detail at PASTA.

3 Future work

- Verifying the model with real-life measurements.
- Studying the transient behaviour of the system – which is essential since in reality many torrents may never reach a steady-state.
- Modifying the protocol or communication systems to provide some anonymity, firstly for initial seeders but also for the other peer participants.
- Model the system in the presence of malicious peers, who just try to consume resources of other peers.

References

[1] Bram Cohen, *BitTorrent* `http://bitconjurer.org/BitTorrent/`

[2] Dongyu Qiu and R. Srikant, *Modeling and Performance Analysis of BitTorrent-Like Peer-to-Peer Networks*, to appear in the Proceedings of ACM Sigcomm, August 2004