

Bit Torrent

A brief introduction and working of Bit Torrent Protocol

Fateh *

Abstract

BitTorrent is a protocol for distributing files. It identifies content by URL and is designed to integrate into the internet. Its advantage over plain HTTP is that when multiple leechers start downloading a certain file at the same time, they become the seeders thus uploading the chunk of file to their peers. It supports sharing of file with a large number of peers with only a slight increase in load. With higher upload rate, peers can find better trading partners, and get the file faster.

The BitTorrent protocol can be used to reduce the server and network impact of distributing large files. The architecture of BitTorrent includes a torrent file, a centralized tracker and an associated group of peers. The centralized tracker provides different entities with an address list over available peers.

1 Introduction

Bit Torrent is a distributed peer-to-peer system. It is based on the notion of a torrent, which is a small file that contains metadata about a host, the tracker, that coordinates the file distribution and files that are shared. BitTorrent is completely free and open-source that is the reason of its success. But the way BitTorrent works also differs considerably from the other peer-to-peer protocols out there.

2 History

It was developed by Bram Cohen, a computer science graduate student at the University of Buffalo in 2001. In 2002, he presented it at a conference. His goal with this software was to give people a quick and simple way of distributing and swapping Linux software online. In 2004, pirate copies of movies and TV shows began dominating the BitTorrent traffic, and after that the growth has been explosive.

3 Protocol Architecture

As it is a distributed networking protocol, The BitTorrent architecture normally consists of the following entities:

- a torrent file
- a tracker
- Seeder (User who is uploading a file)
- Leecher/Peer (User who is downloading a file)

3.1 Torrent File

Torrent file is a meta-info file that contains the filename, size, hashing information and URL of the tracker. It acts like a table of contents that allows computers to find information through the use of a BitTorrent client.

3.2 Tracker

A tracker keeps the logs of the peers that are currently downloading a file, and help them find each other. It is not directly involved in the transfer of data and does not have a copy of file.

3.3 Seeder

It is a user that has the entire file. The seeder uploads the file and make it available for leechers or peers to download it by leaving its torrent link open.

*Email: fateh2018@namal.edu.pk

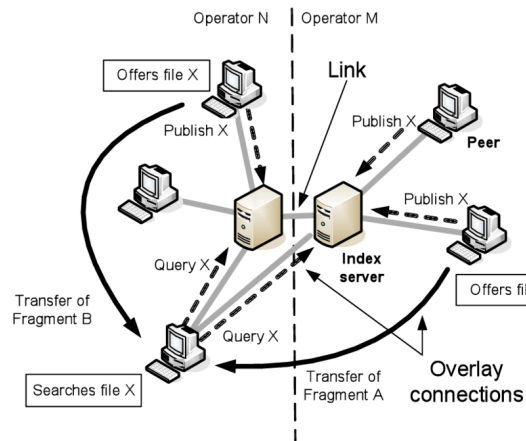


Figure 1. Hybrid P2P Networking

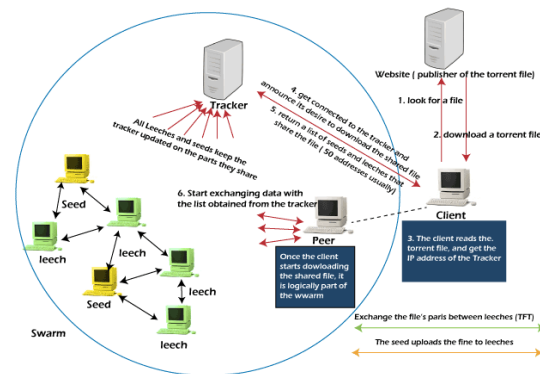


Figure 2. BitTorrent File Sharing

3.4 Leechers/Peers

Leechers are the ones who download the files from torrent links given by the seeders. A leecher can become a seeder if it has completed downloading and leaves its torrent link open.

BitTorrent in its original form matches the “hybrid” peer-to-peer concept. It’s all about the torrent file, the centralized tracker and the associated swarm of peers. The centralized tracker provides the different entities with an address list over the available peers. These peers will then contact each other to download pieces of the file from each other.

3.5 What is Hybrid Peer-to-Peer Networking?

A hybrid P2P network is one that has an index (table of contents) server containing information on the locations of resources at the center, and which uses the index server for search. A pure P2P network is one that directly executes the search process among hosts without a server between the hosts.

4 Working

At first, the meta-info file called torrent is created in order to publish a file using BitTorrent. It contains the filename, size, hashing information and URL of the tracker. Seeder uploads the torrent file on the torrent server. Seeder also provides the first seed to the tracker, that carries the list of peers. A peer joining the torrent has no chunks, but will accumulate over the time. A peer searches for the file from the search engine that has link with the torrent server, and retrieve the link to the torrent file. After getting the link, it contacts with the tracker. Tracker gives it the information of the seeder. Peer then contact the seeder for pieces of file. And then it trades the pieces with other peers. Each joining peer goes through the same process thus making a “swarm” of peers.

4.1 Tit for tat in BitTorrent

No centralized resource allocation exists in BitTorrent. Every peer is responsible for maximizing its download rate. A peer will, naturally, try to download from whoever they can. To decide which peers to upload to, a peer uses a variant of the “tit-for-tat” algorithm.

4.1.1 Choking and Unchoking

In order for a peer to peer network to be efficient, all peers have to contribute to the network. If a peer is just downloading a file and not uploading it, it is called free-rider. BitTorrent allows choking to prohibit large number of free-riders. A peer always unchoke a specific number of its peers (default number is 4). Unchoking is determined by the download rate. If your upload speed is high, more peers will allow you to download from them. Choking is a temporary refusal to upload to another peer. The time for deciding which peer to unchoke is 20 sec. Due to the use of TCP protocol, it's not desirable to choke and unchoke rapidly.

4.1.2 Optimistic Unchoking

BitTorrent also allows an additional unchoked peer, where the download rate criterion is not used. This is called optimistic unchoking. The reason for this is to see if there are any currently unused connections which might be better than the ones in use. Unchoke is shifted every 30 seconds to check for the optimistic peers. This is considered to be enough time for the upload to boost up to full speed and for the download to start and obtain full speed. If this new connection turns out to be better than one of the existing unchoked connections, it will replace it.

4.1.3 Snubbing

If a peers is choked by all its neighbouring peers then it will stop receiving data. To solve this problem we need to find new peers to trade, but optimistic unchoking checks unused connections. To help the download rate recover more rapid, BitTorrent introduces “snubbing”. If a client has not got anything from a particular peer for 60 seconds, it assumes that is has been “snubbed”. Following the mentality of “tit-for-tat” it retaliates and refuses to upload to that peer (except if it becomes an optimistic unchoke). It will then increase the number of optimistic unchokes in order to try to find new connections quicker.

5 References

- [1] <http://web.cs.ucla.edu/classes/cs217/05BitTorrent.pdf>
- [2] https://www.cs.helsinki.fi/webfm_send/1330
- [3] <https://lms.namal.edu.pk/mod/resource/view.php?id=2579>
- [4] <https://lms.namal.edu.pk/mod/resource/view.php?id=1226>
- [5] <http://www.dsc.ufcg.edu.br/~nazareno/bt/bt.htm>