A SURVEY ON BITTORRENT P2P PROTOCOL

[1]Shivanagouda S A, [2]Sathvik I K, [3]Samartha Shetty, [4]Jayasheela C S

[1]Student, **[2]**Student, **[3]**Student, **[4]**Assistant Professor
Department of Information Science & Engineering
Bangalore Institute of Technology/VTU/Bangalore/India

Abstract: Peer-to-Peer architecture is a decentralized platform whereby a number of individuals interact directly with each other, without intermediation by a third party. There are many Peer-to-Peer protocols like Gnutella, Kazaa, etc. BitTorrent belongs to the same architecture. Torrenting is the most popular form of peer-to-peer (P2P) file-sharing, and it requires torrent management software to connect to the BitTorrent network. Such software can be downloaded for free for a number of different devices. BitTorrent and a micro torrent client are similar but differ in the Transport layer. BitTorrent uses TCP for the transport layer whereas micro-torrent uses UDP as the transport layer. In this paper, we carry out the study on torrenting, what a torrent file contains, how BitTorrent client will help you download files by torrenting, and also have deep insight into security and safety issues related to the use of torrents.

I. INTRODUCTION

BitTorrent is a communication protocol based on the P2P model which enables users to distribute data in a decentralized manner. A user needs a BitTorrent client on their device to send or receive files. A BitTorrent client is a computer application that implements the BitTorrent protocol. Its advantage over traditional client-server download is that when multiple downloads of the same file happen at the same time, the downloaders upload to each other, making it possible for the file source to cater to many downloaders with only a small amount of growth in its load. In traditional client-server download, if many users are simultaneously downloading the same file from the same server, that server becomes clogged with requests and the bandwidth for that server is stretched. The result to which is that every person trying to download that file has to face very slow download speeds, or get denied eventually. Plus, the owner of that server could be dealing with enormous bandwidth costs. BitTorrent is designed to distribute the load.

II. PEER-TO-PEER NETWORKING: AN OVERVIEW

In the client-server model, a web browser is used to locate files on a server. In P2P, a software program specific to a protocol is used to locate files on other devices. They can be servers or normal computers collectively known as **peers** [1]**.** The process goes like this:

1. We run a P2P file-sharing software and send out requests for the file we need.
2. Our software program queries the devices to locate the file we need over the network. The file is divided into small pieces called **chunks.**
3. When the computer locates pieces of the file we need, the download starts.
4. If another device locates a file in our system, it can download our file too.
5. In order to prevent other systems from downloading files from our computer after we're finished downloading, we can disconnect immediately after we are done with our download.

III. BITTORRENT P2P PROTOCOL

BitTorrent is a P2P protocol that makes the distribution of large files easier, faster, and more efficient. The upload load is distributed between peers and there is a better utilization of download capacity. A large file is broken into chunks which would boost the concurrency. BitTorrent is not purely a P2P protocol since it involves a server called **Tracker** which we would discuss in further sections [2].

1. TORRENTING: ARCHITECTURE OVERVIEW
Figure 1 shows the steps involved in the creation of a torrent and downloading files from a torrent file [2].
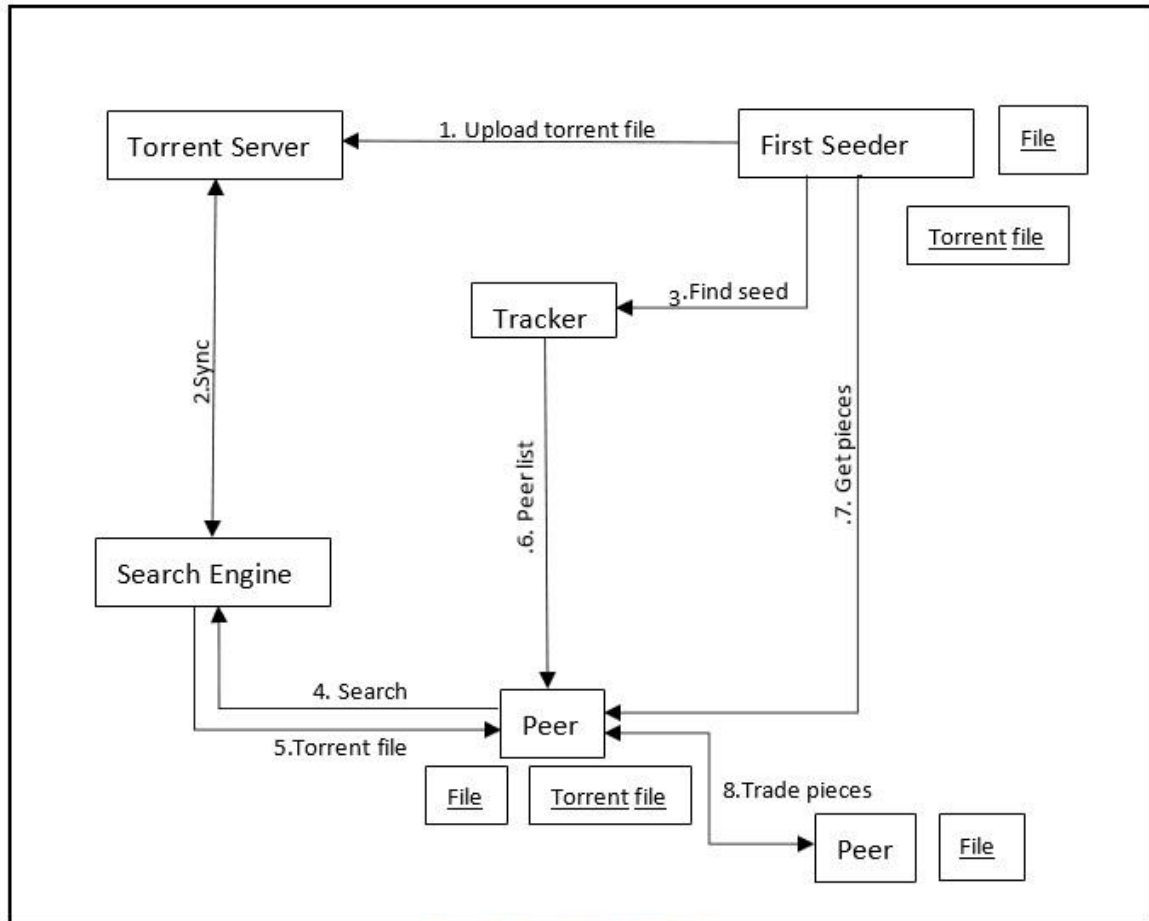

Figure 1: torrenting architecture

2. TERMINOLOGIES
1. **Pieces and Blocks:** Original file that is stored in the network is split into equal-sized pieces ranging from 256KB to 2MB. SHA1 of each piece is added in the .torrent file. Each piece is in turn split into **Blocks** [1]. This is illustrated in Figure 2.

    In one transfer, a block is transferred but a piece is served by a peer. Once a piece is downloaded by a peer, it informs the other peers in the swarm so that they can download it from each other.

Figure2: pieces and blocks

2. **Peer set:** Each peer maintains a list of peers it can send pieces and this is called a peer set.
3. **Active peer set** - A peer can only send data to a subset of its peer set, called an active peer set.
4. **Seeders and Leechers:** A peer can be a seeder or a leecher
    Leecher: when a peer is downloading
    Seeder: when a peer has all the pieces of the content
5. **Tracker:** A server that tracks all the information regarding peers, pieces, and file torrents.
6. **Torrent file and magnet torrent:** A torrent file is the session of transfer of single content to a set of peers. Each torrent is independent. This is like a metafile that contains information such as filename, file size, and information of pieces of the file that we'll be downloading. This file ends with the *.torrent* extension and can be found on the web. Figure 3 gives a sneak peek into a *.torrent* file.
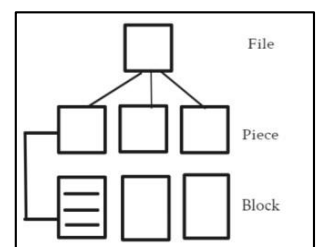
```
d8:announce34:udp://glotorrents.pw:6969/announce13:announcelistll34:udp://glotorrents.pw:6969/announceel40:udp://tracker4.piratux.com
```

¤ù9·âe»²}‰ÀÖÞåoR'ïÉ‹

Ôü,,@A6îéfÂËëLw³ÕØñžÖ¤dÖ|ÜµSZxH¨k¤Þápè

6Ö´Ä|Ùº^ –é¿$æµûë ê—ï˜%*?•Ñzê
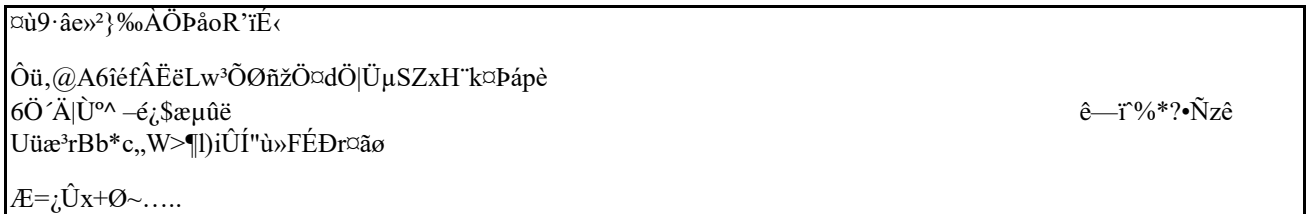
Uüæ³rBb*c,,W>¶l)iÛÍ"ù»FÉÐr¤ãø

Æ=¿Ûx+Ø~…..

Figure 3: contents of a .torrent

This is the torrent file of the movie Southpaw. Since torrent uses a different encoding format – **Bencoding**, you're seeing weird text in the file [3]. But at the very beginning, we can observe the announced statements, this is critical information. Announce or a "Tracker Announce", is a request sent to a tracker. A request is sent, a connection to the tracker is established, information related to peers having the same file is exchanged, then the connection is closed [3]. Another way to get these details is by using the magnet torrent link. On the P2P networks, every file is hashed using the SHA-1 algorithm for finding the details about its peers and pieces from **Kademlia**, a distributed database. So, if you have the hash, you can get the details by generating a magnet link from the hash as shown in Figure 4:

What follows the hash in the magnet link is the info about the trackers that have been added automatically. You can add these or you can let the torrent client do that for you.
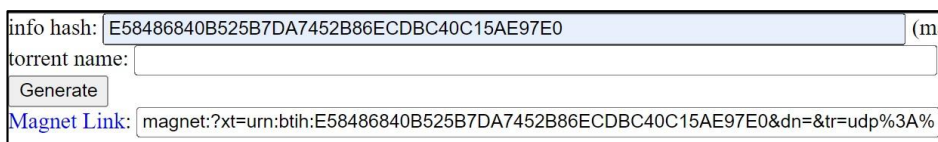
info hash:	E58486840B525B7DA7452B86ECDBC40C15AE97E0	(ma
torrent name:		
Generate		
Magnet Link:	magnet:?xt=urn:btih:E58486840B525B7DA7452B86ECDBC40C15AE97E0&dn=&tr=udp%3A%	

Figure 4: generating magnet torrent from the hash

3. PIECE SELECTION ALGORITHMS

Since a file is split into a number of pieces and in turn blocks, a piece selection strategy is very important to ensure that the swarm has all the required parts of the file. In case a certain piece of the file is missing from the swarm then the download will never finish for all the peers in the swarm which is not desirable. To ensure that this does not happen, the piece selection algorithm is implemented which tries to ensure that all the pieces availability by replicating the pieces as quickly as possible.

1. **Random first piece -** The first piece to be downloaded is chosen at random. This way a peer can receive and start quickly uploading [2].

2. **Rarest First –** Sometimes, the original seeder is removed from the swarm due to cost reasons. As the name itself suggests, the next piece to select for downloading is the rarest piece in the swarm. This way, this rarest piece can be spread in the swarm quickly and risks of losing a piece are avoided [2].

3. **Strict policy -** Since pieces are again subdivided into blocks. When a 'block' is requested, the remaining blocks of the same piece are requested before any other piece. With this, pieces can be assembled quickly and will be available for other peers [2].

4. **Endgame mode -** When the pieces are requested from peers, sometimes a peer with a low upload speed is selected. This can be amplified during the end of the download as the download reaches completion due to the tendency for the remaining pieces to be downloaded from peers with saturated connections. This is avoided by sending a request to all peers for the piece instead of waiting for a single peer. The overhead for this is negligible as the endgame mode is relatively short, but cannot be used for all pieces as that would cause considerable overhead [2].

4. THE CHOKE ALGORITHM: THE POWER OF BITTORRENT

Since BitTorrent works on a P2P network, there is no central resource allocation unit. Then, how to ensure maximum download speed and let no peer abuse the network?

The **choke** algorithm was introduced to guarantee that everybody gets a fair share and reciprocation. This algorithm is a variant of tit-for-tat.

Free riders are peers that download from the network but never upload. They should be penalized. Hence, our criteria to choose a peer to send our content to cannot be simple, it should be based on *reciprocation*.

Choking and Interested:
From [2], Choking is the refusal to upload. This becomes necessary to reduce TCP congestion and prevent abuse and starvation. Choking and unchoking aren't perpetual but periodic. An interested peer is someone who wants a piece that you have. At the very beginning, every peer is by default choked by you. You'll unchoke a peer if that peer has a piece that you need. Any peer will upload to peers, who give the best download rate. This prohibits free-riders and encourages peers to let others download. The choke

algorithm works differently in different cases as when a peer is – leecher, seeder. The mechanism is huge and we're not going in-depth into it now.

IV. VoD STREAMING ON BITTORRENT

The video-streaming technology has been evolving over the years. Since from initial HTTP streaming and DASH (Dynamic Adaptive Streaming Over HTTP), it has seen tremendous changes and scalability solutions. The CDN (Content Delivery Network) overcame the scalability problem at the cost of increasing bandwidth as the number of clients increased. CDNs are clusters of servers covering some part of geographical locations to reduce the load on the main server. Content replication enhances robustness so that CDNs can maintain reliable service in case of failures. Today's well-known notable appeal of video-on-demand (VoD) streaming service has at first posed a huge challenge for cellular network operators: the increase in video traffic greatly outpaced the improvement in the cellular network capacity [8]. This makes the server-based video streaming solution more expensive.

BitTorrent reduces bandwidth load on ISP by utilizing the peer's download bandwidth and incentives for cooperation, which are two attractive characteristics of VoD. In order to meet the requirements of VoD systems, many works have been brought by focusing on two essential factors of BitTorrent protocol: the piece selection and the peer selection algorithms [8].

BitCover: Enhanced BitTorrent for interactive VoD streaming over 5G and Wi-Fi Direct

The authors of [7] have proposed a new algorithm for VoD Streaming using BitTorrent protocol over 5G and Wi-Fi Direct. 5G is the 5th generation technology standard for broadband-cellular networks. Wi-Fi Direct is a Wi-Fi standard for a peer-to-peer wireless connection that allows devices to connect without an intermediary wireless access point, router, or Internet connection.

In BitCover, the decision for using 5G or Wi-Fi-Direct links is a function of the physical distance between the two communicating devices (i.e., the network peers that are exchanging video pieces). If in case the distance is beyond the Wi-Fi-Direct radio range, the piece delivery occurs through 5G using the mobile-backhaul network; otherwise, the endpoints deploy a Wi-Fi Direct link only. This is done in such a manner that the pivotal features of both 5G and Wi-Fi-Direct communication technologies may be jointly leveraged [7].

V. EXPLOITING AND STEALING FROM BITTORRENT

You may have come across this advice not to do torrenting or you'll be put into jail or made to pay a fine. Is it really dangerous to do torrenting? Well, it depends on the file you're trying to download. If you're downloading any media files, and someone has copyrights on them, it's a copyright violation and the owner may sue you for that. But the government wouldn't bother itself with all this until and unless a case is filed. To be on the safer side, a VPN can be employed while torrenting. VPN provides a virtual IP address hiding the original information. In cases of downloading some Linux software that is free, it's legal to download such files. There is also a chance that the file being downloaded is infected with malware/virus. So, it's important to download files in torrents from trusted uploaders.

VI. LEGAL ISSUES WITH BITTORRENT

YTS.MX and PIRATE-BAY are the most prominent peer-to-peer release groups known for distributing large numbers of movies as free downloads through BitTorrent. YTS.MX also known as YIFY was shut down in 2015 by the MPAA (Motion Picture Association of America); however, numerous websites imitating YIFY/YTS brand still receive a significant amount of traffic [1].

The Pirate Bay torrent website, formed by a Swedish group, is under the "legal threats" section of its website in which letters and replies on the subject of alleged copyright infringements are publicly displayed [1]. In 2006, their servers were raided by Swedish police on allegations by the MPAA of copyright infringement. The site was back online in less than 72 hours, and returned to Sweden, accompanied by public and media backlash against the government's actions [4].

VII. APPLICATIONS

i. **µ- Torrent:** Once a popular independently developed app, uTorrent is now owned and operated by BitTorrent the company (not to be confused with the protocol). Despite a continued emphasis on keeping the application small, fast, and light, uTorrent is now loaded with features, including a personal favourite, and built-in remote control [6].

ii. **Facebook and Twitter use BitTorrent internally:** Moving a 1.5GB binary blob to countless servers is a non-trivial technical challenge. After exploring several solutions, Facebook came up with the idea of using BitTorrent, the popular peer-to-peer file-sharing protocol. BitTorrent is very good at propagating large files over a large number of different servers [6].

iii. **Game uploads and downloads:** Blizzard Entertainment uses its own BitTorrent client to download World of Warcraft, Star-craft II, and Diablo III. When you purchase one of these games and download it, you're actually just downloading a BitTorrent client that will do the rest of the work. When an update is available, the BitTorrent client built into the game's launcher automatically downloads it for you [6].

iv. **Linux ISOs:** Whether you're downloading the latest release of Ubuntu, Fedora, Debian, or any of the other best Linux distributions, there's a good chance you're getting it via BitTorrent. These distributions offer themselves for free to everyone and they're often 1 GB or larger [6].

v. **The Internet Archive:** The Internet Archive is a non-profit organization that preserves content and makes it downloadable on the Internet. It recommends people use BitTorrent to download its content, as it's the fastest method and allows the non-profit organization to save on bandwidth costs [6].

VIII. CONCLUSION

In just a short span of time, BitTorrent became the most used protocol in peer-to-peer networking. It is lightweight, easy to use, has fast speed download, and provides fairness among peers. Originally the protocol involved a central tracker in order to let peers find other peers, i.e. a hybrid peer-to-peer protocol. New versions of the protocol have removed this central tracker and distributed this function to the peers themselves. This improvement has made BitTorrent a pure peer-to-peer protocol [6].

The protocol is highly self-configuring, where the peers themselves get the seeders, and leecher's information from the tracker server, which may be centralized or distributed. The protocol has found a niche as a preferred method for the decentralized distribution of large files. BitTorrent can also leverage the costs of distributing shows and movies, making broadcasting possible for almost every Internet user. This can have tremendous effects on large networks, especially content providers. The industry needs to change and adopt the new technology before they are overrun.

IX. REFERENCES

[01] A Survey on BitTorrent File Sharing Protocol – 2019 Gavin Koli, Chetna Achar

[02] https://arpitbhayani.me/bittorrent-internals/179

[03] http://www.bittorrent.org/beps/bep_0003.html

[04] https://en.wikipedia.org/wiki/Legal_issues_with_BitTorrent

[05] A Review on Torrent and Torrent Poisoning over the Internet – 2018 Pooja Balhara

[06] https://www.makeuseof.com/tag/8-legal-uses-for-bittorrent-youd-be-surprised/

[07] BitCover: Enhanced BitTorrent for interactive VoD Streaming over 5G and Wi-Fi-Direct – 2023 Vladimir Rocha, Carlo K. da S. Rodrigues

[08] Peer-To-Peer Live Streaming and Video-On-Demand Design Issues and Its Challenges – 2015

[09] https://en.wikipedia.org/ wiki/Wi-Fi-Direct

[10] https://web.cs.ucla.edu/classes/cs217/05BitTorrent.pdf