2009

TORRENTS – (BIT TORRENT)

Prof. Dr. Eduard Heindl – E- Business Technologies SS'09 BCM

SESHAGIRI VENKATACHALAM
HFU

TABLE OF CONTENTS

TORRENTS (BIT TORRENT)

This Term Paper on Torrents – (Bit Torrent) has been written by me, Seshagiri Venkatachalam after research and thought. Sources where information has been taken from, paraphrased or rephrased have been cited at the appropriate places.

- **SESHAGIRI VENKATACAHALM**
- **24’TH JUNE’09**

1. INTRODUCTION

The Internet isn't free. It just has an economy that makes no sense to capitalism.

- Brad Shapcott [1]

With almost a quarter of the world's populations already connected to the Internet and growth rates in Africa and the Middle East reaching almost 1000% over the past 8 years, the increases to on-line connectivity seem to be following a familiar pattern seen in other areas within the computing industry. [2]

One of the benefits of such acute penetration of the internet into the different cultures of the world has been the globalization and the collaborative economics that the internet has brought forth. The world today does not resemble the world of ten years ago. Businesses and business practices have changed all over the world and countries co-operate to produce and to distribute goods and services so that their individual economies prosper.Some of the factors that have brought about this change have been attributed to advancements in the field of computing (both software and hardware).

The computer hardware industry has also been characterised by exponential production volumes. Gordon Moore, the co-founder of Intel, in his famous observation in 1965 (made just four years after the first planar integrated circuit was discovered), predicted that the number of transistors on integrated circuits would double every few years. Indeed this prediction, thereafter called Moore's law, remains true today and Intel predicts that this will remain true at least until the end of this decade.Such acceleration in development has been made possible by the massive investment by companies who deal with comparatively short product life cycles. Each user now in this massive network has the CPU capability of more than 100 times that of an early 1990s.[3]

One of the key contributors to the globalised change has been the ability of the internet to provide a common platform for people to share, download and upload various contents on the internet.File sharing, downloading and uploading has been one of the key advances made in the last 2 decades and it has changed the very way the world functions.[4]

As any average internet user would know there are more than a few ways to download and upload content on the internet. This termpaper is targeted at the technological and the business model of torrents.

1.1 WHAT ARE TORRENTS?

Torrents are specialized files utilized in peer-to-peer (P2P) network environments. P2P is a network of personal computers that communicate with one another by running proprietary

---

[1] http://thinkexist.com/quotes/brad_shapcott/ 03'rd June'09
[2] http://www.internetworldstats.com/stats.htm 03'rd June'09
[3] **From P2P and Grids to Services on the Web – Sammes, A.J, Harrison, Andrew B, Taylor, Ian J. ISBN 978-1-8480-0123-7**
[4] http://wikisummaries.org/The_World_Is_Flat#Chapter_Two_.E2.80.93_The_Ten_Forces_That_Flattened_the_World 03'rd June'09

P2P software. One of the P2P software designed to utilize torrents was BitTorrent by Bram Cohen. Other torrent clients have followed.Torrents are distinguished by a unique transfer process. To compare how torrents download to standard files, let's first consider how normal files download off the Internet.

At any given website a user might click on a file to transfer it to his or her computer. Upon clicking on the file, the website's server starts sending the file to the visitor in discreet data packets. These packets travel various routes to reach the user's computer and are reconstructed upon receipt to complete the file transfer.

While this works fine for smaller files, it is cumbersome to transfer larger files this way. If the server is busy, download time can be very slow. Communication between your server and the computer can even crash, causing corruption in the transfer, or at best, delays.

Unlike downloads off the Web, torrents do not point to a single source on a P2P network when requesting files. Instead, torrents contain specific information that multiple computers in the network can read to send various parts of the requested file simultaneously and en masse. Torrents keep active track of which parts of the file are needed to complete the request. By downloading bits of the file from dozens, hundreds, or even thousands of sources, large files can download very quickly.[5]

In order to understand how torrents work, it is important to understand how P2P protocol works. This termpaper will focus on one of the applications or systems that use the P2P strategy namely the BitTorrent.

1.2 WHAT IS PEER TO PEER?

This section gives a brief background and history of the term peer to peer and describes how this definition has changed through the introduction of the World Wide Web and the change in focus on new Internet users.

Peer to peer was originally used to describe the communication of two peers and is analogous to a telephone conversation. A phone conversation involves two people (peers) of equal status, communication between point-to-point connections. Simply, this is what P2P is, a point-to-point connection between two equal participants. The Internet started as a peer-to-peer system. From the late 1960s until 1994, the Internet had one model of connectivity. Machines were assumed to be always switched on, always connected and assigned permanent IP addresses. [6] The original DNS system was designed for this environment, where a change in IP address was assumed to be abnormal and rare, and could take days to propagate through the system. However, another model began to emerge in the form of users connecting to the Internet from dial-up modems. This created a second class of connectivity because PCs would enter and leave the network frequently and unpredictably. Further, because ISPs began to run out of IP addresses, they began to assign IP addresses dynamically for each session, giving each PC a different, possibly masked, IP address. This transient nature and instability prevented PCs from being assigned permanent DNS entries, and therefore prevented most PC users from hosting any data or network-facing applications locally.

[5] http://www.wisegeek.com/what-are-torrents.htm 03'rd June'09
[6] http://www.isoc.org/internet/history/brief.shtml#Origins 03'rd June'09

For a few years, treating PCs as clients worked well. Over time, though, as hardware and software improved, the unused resources that existed behind this veil of second-class connectivity started to look like something worth getting at.[7]

Within today's Internet, we rely on fixed IP addresses. When a user types an address into his/her Web browser (such as http://www.google.com/), the Web server address is translated into an IP address (e.g., 216.239.59.104) by a domain name server (DNS). The Internet protocol (IP) then makes a routing decision based on the IP address. If DNS is unavailable, then typing http://216.239.59.104/ into a browser would be equivalent since, at the time of writing, one of the Google Web sites is bound to this IP address. This is known as static or early binding. Early bindings form a simple architecture very similar to an address book on a mobile phone; e.g., the person's name is statically bound to his/her telephone number. This works in practice because typically people have long-term (early) bindings with their phone numbers and Web sites have long-term bindings with their IP addresses.

However, if a Web site changed its IP address several times a day, then this type of binding starts to become impractical. Within P2P networks this is the norm. Often devices do not have a fixed address as they are hidden behind Network Address Translation (NAT) systems and therefore need a late binding of their addresses with their network identifier.[8]

2. PEER TO PEER DISTRIBUTION SYSTEMS

The operation of any peer-to-peer content distribution system relies on a network of peer computers (nodes) and connections (edges) between them. This network is formed on top of- and independently from-the underlying physical computer (typically IP) network, and is thus referred to as an "overlay" network. The topology, structure and degree of centralization of the overlay network, and the routing and location mechanisms it employs for messages and content are crucial to the operation of the system, as they affect its fault tolerance, self-maintainability, adaptability to failures, performance, scalability, and security.[9]

Overlay networks can be distinguished in terms of their centralization and structure.Although in their purest form peer-to-peer overlay networks are supposed to be totally decentralized, in practice this is not always true, and systems with various degrees of centralization are encountered. Specifically, the following categories are identified.

2.1 PURELY CENTRALIZED ARCHITECTURES:

Napster (a P2P system) performed query resolution with the aid of centralized content registries The search mechanism of Napster client contracted a centralized registry and the registry responded with a list of peers that likely possessed the data object.[10]
The network structure of Napster has an Achilles Heel -- it is highly dependent on the static central server. If the central server is down, the network will collapse. This was shown by the actions of the recording industry, which forced the original Napster to be shutdown.[11]

---

[7] http://www.isoc.org/internet/history/brief.shtml#Origins 03'rd June'09
[8] **From P2P and Grids to Services on the Web – Sammes, A.J, Harrison, Andrew B, Taylor, Ian J. ISBN 978-1-8480-0123-7**
[9] http://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.html#p2p23 03'rd June'09
[10] http://books.google.com/books?id=NTF6jj_360UC&printsec=frontcover&dq=communication+network+analysis&hl=eN#PPA179,M1 03'rd June'09
[11] http://www.securityfocus.com/infocus/1843/2 03'rd June'09

2.2 DECENTRALIZED ARCHITECTURES:

Decentralized means a network structure with no dedicated central index servers. It is a trend for P2P evolution. These are of the following two types:

Purely Decentralized:

All nodes in the network perform exactly the same tasks, acting both as servers and clients, and there is no central coordination of their activities. The nodes of such networks are often termed "servents" (SERVers+clieENTS). E.g Edonkey [12]

Partially Decentralized or Partially Centralized:

The basis is the same as with purely decentralized systems. Some of the nodes, however, assume a more important role, acting as local central indexes for files shared by local peers. The way in which these supernodes are assigned their role by the network varies between different systems. It is important, however, to note that these supernodes do not constitute single points of failure for a peer-to-peer network, since they are dynamically assigned and if they fail the network will automatically take action to replace them with others. E.g Gnutella & KaZaA
The Napster case illustrates the vulnerability of a centralized network structure and greatly affects the subsequent P2P application. For legal, security, scalability, anonymity and some other reasons, more and more P2P applications nowadays work in a totally or partially decentralized network structure, or are moving in the direction. Major P2P file-sharing networks and protocols, such as Edonkey2k, FastTrack, Gnutella, Gnutella2, Overnet, Kad, all use this concept. And as a partially decentralized model, hybrid decentralized networks have won broad support from various P2P applications and are thus recognized as the most popular P2P network model. [13]

2.3 CONTENT DISTRIBUTION SYSTEM OF BIT TORRENT

Here it should be made clear that Bittorrent is not a general purpose P2P network although it is a popular P2P application. It still needs tracker servers; while the network structure of Bittorrent is partially decentralized.

In BitTorrent the groups of peers that are actively uploading and downloading a given data object are known as swarm.Peers search the web and download a *.torrent* file corresponding to the data object of that interest. The peer can then connect to a tracker that informs it of other peers currently participating in the swarm for that data object.

Decentralized versions of Bit Torrent employ protocols such as Kademlia.This protocol is similar to the one used by EDonkey and greatly reduces the need for tracker based architecture.[14]

3. THE P2P ENVIRONMENT

[12] http://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.html#CITEyangvldb01 03'rd June'09
[13] http://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.html#CITEyangvldb01 03'rd June'09
[14] http://books.google.com/books?id=NTF6jj_360UC&printsec=frontcover&dq=communication+network+analysis&hl=eN#PPA179,M1 03'rd June'09

In this section, a background is given into some of the technologies behind P2P networks, which helps set a more realistic P2P scene. The first section makes a brief excursion into switching technology for networks. The second section describes a particular subset of these that contains NAT systems.

3.1 HUBS, SWITCHES, BRIDGES, ACCESS POINTS AND ROUTERS

This section gives a brief overview of the various devices used to partition a network, which gives the context for the following two sections on NAT and firewalls often employed within a P2P network. Briey, the critical distinction between these devices is the level or layer at which they operate within the International Standard Organization's Open System Interconnect (ISO/OSI) model, which defines seven network layers.

(i) **A hub** is a repeater that works at the physical (lowest) layer of OSI. A hub takes data that comes into a port and sends it to the other ports in the hub. It does not perform any filtering or redirection of data. You can think of a hub as a kind of Internet chat room. Everyone who joins a particular chat is seen by everyone else. If there are too many people trying to chat, things get bogged down.

(ii) **Switches and Bridges** are pretty similar. Both operate at the Data Link layer (just above Physical) and both can filter data so that only the appropriate segment or host receives a transmission. Both filter packets based on the physical address (i.e., Media Access Control (MAC) address) of the sender/receiver although newer switches sometimes include the capabilities of a router and can forward data based on IP address (operating at the network layer), referred to as IP switches. In general, bridges are used to extend the distance capabilities of the network while minimizing overall traffic, and switches are used primarily for their filtering capabilities to create multiple, smaller virtual local area networks (LANs) out of one large LAN for easier management/administration (V-LANs).[15]

(iii) **Routers** work at the Network layer of OSI (above Data Link) and operate on the IP address. Like switches and bridges, they filter by only forwarding packets destined for remote networks thus minimizing traffic, but are significantly more complex than any other networking device; thus they require much more maintenance and administration. The home networker typically uses a DSL or cable modem router that joins the home's LAN to the wide area network (WAN) of the Internet. By maintaining configuration information in a routing table" routers also have the ability to filter traffic, either incoming or outgoing, based on the IP addresses of senders and receivers. Most routers allow the home networker to update the routing table from a Web browser interface. DSL and cable modem routers typically combine the functions of a router with those of a switch in a single unit.

(iv) **A firewall** is a system designed to prevent unauthorized access to or from a private network. All messages entering or leaving the computer system pass through the firewall, which examines each message and blocks those that do not meet the specified security criteria. Specifically, firewalls are implemented by blocking certain ports, thereby disabling certain types of services that operate on those ports. Some firewalls permit only email traffic, thereby protecting the network against any attacks other than attacks against the email service. Other firewalls provide less strict protections, and block services that are known to be problematic.

---

[15] **From P2P and Grids to Services on the Web – Sammes, A.J, Harrison, Andrew B, Taylor, Ian J. ISBN 978-1-8480-0123-7**

Generally, firewalls are configured to protect against unauthenticated interactive logins from the outside world. This, more than anything, helps prevent unauthorized users from logging into machines on your network.A firewall blocks traffic to and from specified ports.[16]

3.2 NAT SYSTEMS

For a computer to communicate with other computers and Web servers on the Internet, it must have an IP address. An IP address is a unique 32-bit number that identifies the location of your computer on a network. There are, in theory, 232 (4,294,967,296) unique addresses but the actual number available is much smaller (somewhere between 3.2 and 3.3 billion). With the explosion of the Internet and the increase in home networks and business networks, the number of available IP addresses is simply not enough. An obvious solution is being developed and is called IPv6, but it may take several years to deploy because it requires modification of the entire infrastructure of the Internet. A network address translation system allows a single device, such as a router, to act as an agent between the Internet (public network) and a local (private) network. This means that only a single, unique IP address is required to represent an entire group of computers. The internal network is usually a LAN, commonly referred to as the stub domain. A stub domain is a LAN that uses IP addresses internally. Any internal computers that use unregistered IP addresses must use NAT to communicate with the rest of the world. There are two types of NAT translation, static or dynamic, which are illustrated in Fig.1. & Fig. 2.

Static – One to One:

Static NAT involves mapping an unregistered IP address to a registered IP address on a one-to-one basis. Particularly useful when a device needs to be accessible from outside the network (i.e., in static NAT), the computer with the IP address of 192.168.0.0 will always translate to 131.251.45.110 (see upper part of Fig.1).

STATIC ONE TO ONE

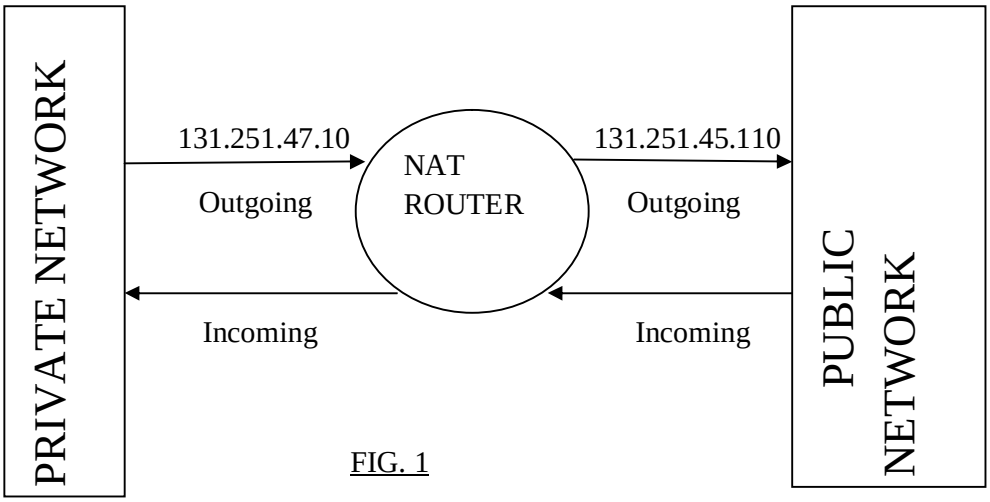

FIG. 1

DYNAMIC MANY TO ONE:

[16] **From P2P to Web Services and Grids : Peers in a Client/Server World by Ian J. Taylor ISBN 978-1-8462-8074-0**

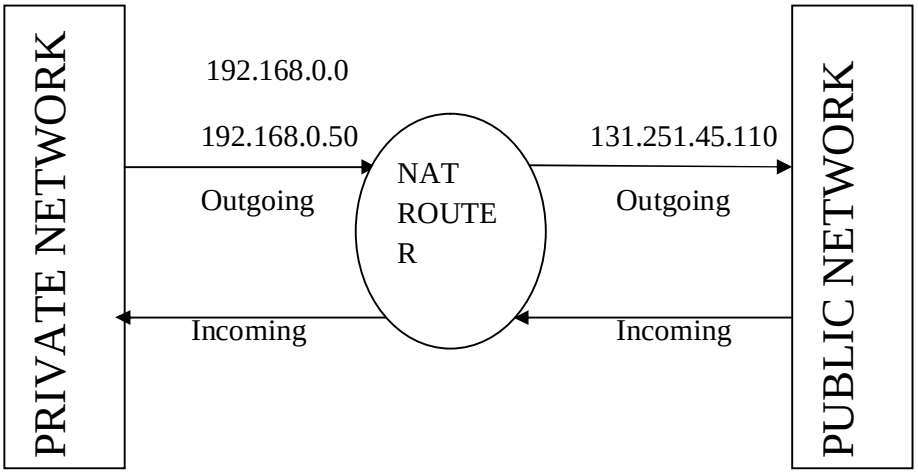

FIG: 2 *A NAT system can be allocate dynamic address or translate from fixed stub domain address to outside ones.*

Dynamic – Many to one:

Dynamic NAT, on the other hand, maps an unregistered IP address to a registered IP address from a group of local dynamically allocatable IP addresses, i.e., the stub domain computers will be allocated an address from a specified range of addresses, e.g., 192.168.0.0 to 192.168.0.50, in Figure 2 and will translate these to 131.251.45.110 for the outside world. In this circumstance, it is easy to see why NAT systems are problematic since you could have potentially hundreds of stub domain computers masquerading as one external IP address. [17]

3.3 P2P APPLICATIONS

No universally accepted definition of peer to peer (P2P) exists. It is used to cover systems with vastly different architectures, and various taxonomies have been used to classify P2P systems according to how decentralized they are, the nature of the overlay network, the kinds of data that are shared between peers and the target audience. Although the term P2P implies communication between equals, as discussed in the previous section, most systems employ unequal client/server-type relationships. Flagship examples of systems and applications that fall under the P2P category include Gnutella, Skype, Napster, Bit Torrent, Freenet, KaZaA, Instant messaging and Skype. [18]

4. BIT TORRENT

You get so tired of having your work die. I just wanted to make something that people would actually use. – BRAM COHEN [19]

[17] **From P2P and Grids to Services on the Web – Sammes, A.J, Harrison, Andrew B, Taylor, Ian J. ISBN 978-1-8480-0123-7**
[18] **From P2P and Grids to Services on the Web – Sammes, A.J, Harrison, Andrew B, Taylor, Ian J. ISBN 978-1-8480-0123-7**
[19] http://www.brainyquote.com/quotes/quotes/b/bramcohen219752.html 03‘rd June’09

The above quote is from Bram Cohen, BitTorrent's author, in an interview with Wired in 2005.The first version of the BitTorrent protocol was presented in the first CodeCon conference in San Francisco in February 2002 and subsequently became one of the most popular Internet file sharing protocols. In essence, BitTorrent introduced two key concepts that were novel to its file-sharing competitors at the time. First, rather than providing a search protocol itself, it was designed to integrate seamlessly with the Web and made files (torrents) available via Web pages, which could be searched for using standard Web search tools. Second, it enabled so-called file swarming; that is, once a peer starts downloading that file, it also makes whatever portion of the file that is downloaded immediately available for sharing. The fileswarming process is enabled through the use of a tracker, which is an HTTP- based server used to dynamically synchronise and update the peers as they are downloading as to the locations and availability of pieces of the file in question on the network. [20]

Reuters on the 11'Th Nov'04 reported that Bit Torrent accounted for 35% of all internet traffic and bandwidth. [21] The result of the file swarming techniques made BitTorrent an extremely attractive tool for sharing files because it allowed users to download files to the maximum of their download capability of their broadband connection by enabling simultaneous downloads of pieces of the same file from multiple users. This is significant because typically a broadband connection has a far lower upload bandwidth than a download one (the upload bandwidth can be typically ten times slower than the download). This means that being able to connect to, say, ten peers, will balance this mismatch and enable the full potential of your Internet link, which results in files being downloaded several times faster than other file sharing systems on the Internet at that time. The BitTorrent protocol therefore has had a massive impact on file sharing applications and similar schemes have been adopted by competitors since. Further, its use has far outgrown the illicit file-sharing arena and nowadays the BitTorrent protocol or similar techniques are used in a multitude of different applications in science and business and it has even being integrated into hardware devices. In this sense, the protocol has grown up and is now taken very seriously throughout the Internet community.[22]

4.1 BIT TORRENT INC:

BitTorrent, Inc. was founded by Bram Cohen and Ashwin Navin in 2004. [23]They secured a funding of $28.75 million from Accel Partners and DCM (formerly Doll Capital Management) to pursue the development of a software platform based around BitTorrent that aims to help software providers and media companies to download high-resolution files faster for real-time on-line content distribution in 2007. In a Wired article in late 2007, they were listed as one of the top 10 startups worth watching in 2008. The company initially employed 60 people. At the time of writing, the company had 27 partner logos listed on their partner page and their current CEO Doug Walker stated in an interview in early 2008 that the company was talking to 55 partners from the media industry, and was in active discussions about how BitTorrent could help them deliver media content and games over the Internet. These companies include

---

[20] **From P2P to Web Services and Grids : Peers in a Client/Server World by Ian J. Taylor BITTORRENT ISBN 978-1-8462-8074-0**

[21] http://www.brainyquote.com/quotes/quotes/b/bramcohen219752.html 03'rd jUNE'09

[22] **From P2P to Web Services and Grids : Peers in a Client/Server World by Ian J. Taylor BITTORRENT ISBN 978-1-8462-8074-0**

[23] http://kepler.ss.ca.gov/corpdata/ShowAllList?QueryCorpNumber=C2676150 03'rs June'09

20th Century Fox, Comedy Central, Lionsgate Films, MTV, Paramount, Spike TV and Warner Brothers. The company boasts an installed base of over 160 million clients worldwide.[24]

BitTorrent has three lines of business:

- BitTorrent DNA for enabling Websites to add BitTorrent technology to their current content delivery infrastructure.
- BitTorrent Device Partners, which is a program designed to help hardware and software companies to embed BitTorrent through the use of their Software Development Kit (SDK).
- Torrent Entertainment Network (TEN) which is an online destination for downloadable and ad-supported streaming entertainment content powered by BitTorrent DNA.[25]

The network provides fast, on-demand access to the most comprehensive catalog of licensed content on the Internet, featuring thousands of movies, TV shows, music titles and games. Cohen authored both the peer-to-peer (P2P) BitTorrent protocol and the first file sharing program that used the protocol, also called BitTorrent. He developed BitTorrent in 2001 and around the same time, he co-founded the CodeCon conference with his roommate Len Sassaman. CodeCon is a low-cost conference (i.e., less than $100) aimed at hackers and technology enthusiasts, with a focus on presentations from developers with working code, in contrast to other conferences, which they suggested focused more on companies with products to sell. The first CodeCon conference was on 24 February 2002 and BitTorrent was one of the several presentations. It continues to this day to claim the “most famous presentation” title for this conference series.

4.2 THE BIT TORRENT PROTOCOL

BitTorrent is simply a protocol for distributing files. It was designed to be based around the Web with content (file to be downloaded) being identified by URLs and accessed via standard Web pages. However, it difiers from a typical http scenario, because when a user clicks on a BitTorrent a conventional client-server download of the content between the browser and the Web server does not take place. Rather, when the link is clicked, a MIME- type mapping (from a .torrent file) takes place that results in a BitTorrent client-side application starting and joining the BitTorrent network. Thereafter, multiple downloads of the same file happen concurrently by breaking the file into pieces and by exchanging these pieces between users using TCP until everyone has the complete file. A tracker is used to update the network as more pieces are downloaded, which results in continuous updating of the state of a group of users download- ing a particular file (a swarm), which results in massive increases in download speeds for files that are popular in the network. Therefore, the more people trading a file, the more options there are for obtaining its pieces. So, unlike the Web or P2P systems like Napster, popularity of a specific file does not slow down the process; rather, it allows for more simultaneous downloads and therefore makes for much faster downloads. BitTorrent trackers also keep track of how much you contribute to hosting files for the group and it further expedites a user's download speed if that peer shares files to the swarm group, by implementing policies in such a way that the more you share, the faster you can download (i.e., the more hosts the tracker notifies you of). This ensures a more even ratio of file sharing

[24] http://www.bittorrent.com/pressreleases?csrc=splash 03’rd June‘09
[25] http://www.bittorrent.com/ 03’rd June‘09

versus file downloading which helps to limit so-called leaches or free riders and therefore helps balance the network as a whole.[26]

The BitTorrent protocol is an open specification that can be found in full on the BitTorrent Web site and is updated periodically in order to keep various BitTorrent applications compatible. This section provides a summary of the protocol and describes the interactions between the various BitTorrent components. We begin with a description of the terminology and an overview of the entities in a typical BitTorrent application to specify the interactions, before describing the low-level messages in more detail.[27]

4.3 TERMINOLOGIES AND ENTITIES IN BIT TORRENT:

In this section, we use the following terms that are defined in the BitTorrent protocol:
Torrent: a metadata file containing the information about a file to be shared on the BitTorrent network (files ending in .torrent).
Peer: a participant in the network. Peers are defined as elsewhere in this paper, being both consumers and providers of files on the distributed network.
Seed: the peer that has a complete copy of the file (who probably created the torrent).
Swarm: peers who are connected (interested) in a particular file.
Tracker: server responsible for keeping track of the people in a swarm.
Piece: is a piece of a file in Bittorrent that is distributed around the network. Pieces typically have a length which is a power of 2 and very much depends on the size of the files that are to be distributed on the network. Common sizes in BitTorrent applications are 256K, 512K or 1MB.[28]

BitTorrent was designed to be interfaced seamlessly with a user's Web browser. A user launches the BitTorrent application by simply clicking on a .torrent hyperlink within a standard Web browser. A BitTorrent application therefore generally has the following components:
1**. An `original' downloader**, referred to as a seed.
2**. An ordinary Web server.**
3. **An end user's Web browser.** A user clicks on a hyperlink to a .torrent file containing the `metainfo', which results in the launching of Bit Torrent Client Application.
4. **An end user's BitTorrent client application** helps initiate downloading files. There are typically ideally many end users for a single file. E.g Bitcomet, Bittornado, µTorrent, ZipTorrent and others. Most of these could be downloaded for free and declare they are adware and spyware free.
5. **A BitTorrent tracker** for synchronising the downloading process.

A BitTorrent tracker keeps track of all of the users that form the swarm for a particular file. As users join or leave the network, the tracker maintains a list of peers who are currently registered as interested in a particular file. Trackers receive information from each peer in the swarm containing their contact information and statistics on how much they have uploaded or downloaded so far. This helps the tracker to make decisions about the priority of the peers, which results in deciding the number of seeds they are given for the downloading of the file.

---

[26] http://www.bittorrent.com/btusers/guides/bittorrent-user-manual/chapter-01-introduction-bittorrent-mainline/about-bittorrent-m 03'rd June'09
[27] **From P2P to Web Services and Grids : Peers in a Client/Server World by Ian J. Taylor BITTORRENT ISBN 978-1-8462-8074-0**

[28] http://dessent.net/btfaq/#where 03'rd June'09

The tracker returns an update to each BitTorrent node in the group as new pieces are downloaded within the network. BitTorrent nodes make further connections to other BitTorrent nodes after simultaneous downloads of the different pieces from the seed result in the tracker notifying peers that new pieces are available in the network.[29]
The communication from the peer to the tracker is an HTTP GET request, which contains the following information:
IP: IP address or IP name for the peer.
Port: port number for the peer tries port 6881 and if that port is taken, it tries 6882, then 6883, etc. and gives up after 6889.
Uploaded: total amount uploaded so far
Downloaded: total amount downloaded so far.
Left: number of bytes this peer still has to download.
Event: optional key which maps to started, completed, or stopped or it is empty.
The tracker then responds with either a failure reason or a bencoded dictionary containing the list of peers that the requesting peer can connect to in a swarm. If a tracker response has a key failure reason, then it maps to a human-readable string which explains why the query failed, and no other keys are required. Otherwise, it must have two keys:
Interval is the number of seconds the downloader should wait between requests.
Peers are a list of dictionaries corresponding to the peers available to this peer in the swarm. Each peer has:
Peer ID: is the peer's self-selected identifier.
IP: the IP address (or DNS name) for the peer.
Port: the port that the peer can be contacted on.
A peer distributing a file simply treats it as a number of identically sized pieces. The peer creates a SHA-1 hash for each piece and records it in the torrent file. When a peer downloads a particular piece, the hash of that piece is computed locally and then compared to the one specified in the .torrent file to check for consistency (data integrity). If they match, then the peer assumes that the data has not been corrupted during transmission or by the other peer. Once all pieces of the file have been downloaded, the peer has a complete copy and can choose to leave the group or swarm for that particular file by exiting their BitTorrent application. As more pieces are downloaded, there are more options for the various pieces and therefore downloading increases in speed. Setting the piece size is key to the success of the protocol. A piece size too small will cause too much latency for opening the TCP connections whereas a piece size too large may reduce the eficiency of the protocol because less pieces (and therefore locations on the BitTorrent network) are available for simultaneous download.[30]

4.4 LEGAL AND SOCIAL ISSUES

BitTorrent trackers have been subjected to raids and shutdowns due to claims of copyright infringement. BitTorrent metafiles do not store copyrighted data, so it has been claimed that BitTorrent trackers, which only store and track the metafiles, must therefore be legal even if sharing the data in question would be considered a violation of copyright.In December 2004, the Finnish police raided a major BitTorrent site, Finreactor. The case is before the courts, and 32 people, in September 2006, mostly administrators and moderators, are facing charges. Software and media companies are seeking damages worth 3.5 million euros (about 5.4 million USD) in total. Two defendants were acquitted by reason of being underage at the time, but they are being held liable for legal fees and compensation for illegal distribution ranging up to 60,000 euros. The court set their fine at 10% of the retail price of products distributed.

[29] http://dessent.net/btfaq/#where 03'rd June'09
[30] http://dessent.net/btfaq/#where 03'rd June'09

The Pirate Bay torrent website, formed by a Swedish anti-copyright group, is notorious for the "legal" section of its website in which letters and replies on the subject of alleged copyright infringements are publicly displayed. On May 29, 2007, a federal judge ordered TorrentSpy, a torrent website, to begin monitoring its users' activities and to submit these logs to the Motion Picture Association of America. TorrentSpy's attorney, Ira Rothken, has stated that TorrentSpy would likely turn off access to U.S. users before it started monitoring anyone, since such monitoring is in violation of TorrentSpy's own privacy policy. HBO, in an effort to combat the distribution of its programming on BitTorrent networks, has sent cease and desist letters to the Internet Service Providers of BitTorrent users. Many users have reported receiving letters from their ISP's that threatened to cut off their internet service if the alleged infringement continues. HBO, unlike the RIAA, has not been reported to have filed suit against anyone for sharing files as of April 2007. On the other hand, in 2005 HBO began "poisoning" torrents of its show Rome, by providing bad chunks of data to clients.[31]

Copyrighted Material being uploaded is a major problem on the internet and methods are being sought to put this practice to an end. There are two major differences between BitTorrent and many other peer-to-peer file-trading systems, which advocates suggest make it less useful to those sharing copyrighted material without authorization. First, BitTorrent itself does not offer a search facility to find files by name. A user must find the initial torrent file by other means, such as a web search. Second, BitTorrent makes no attempt to conceal the host ultimately responsible for facilitating the sharing: a person who wishes to make a file available must run a tracker on a specific host or hosts and distribute the tracker address (es) in the .torrent file. Because it is possible to operate a tracker on a server that is located in a jurisdiction where the copyright holder cannot take legal action, the protocol does offer some vulnerability that other protocols lack. It is far easier to request that the server's ISP shut down the site than it is to find and identify every user sharing a file on a peer-to-peer network. However, with the use of a distributed hash table (DHT), trackers are no longer required, though often used for client software that does not support DHT to connect to the stream.[32]

4.5 BIT TORRENT'S BUSINESS MODEL

BitTorrent's creator, Bram Cohen, has been attempting to shift attention onto the legitimate uses to which his technology can be put. Since 2005, Cohen has been heavily focused on harnessing BitTorrent to facilitate the legal distribution of films and TV programmes. Cohen hopes that the availability of quick, cheap and easy film downloads will do its part in reducing the level of illegal downloading.

The most recent demonstration of Cohen's intentions came in the form of the release of BitTorrent's new service in October 2007, called Delivery Network Accelerator (DNA). Media firms are increasingly making content available via the Internet to consumers in an attempt to counter illegal downloads. DNA will allow any company to tap into the existing network of around 150 million users of the BitTorrent protocol and is being marketed as a supplement for client-server content delivery networks, increasing the speed and reducing the cost of downloads by utilising the existing BitTorrent network.

As part of its legal download project, BitTorrent has launched The BitTorrent Entertainment Network, which offers downloads of films and television shows from some of the big names in the world of film, including MGM, Paramount, Fox and Warner Brothers. The fact that BitTorrent Inc has signed agreements with such major film studios speaks volumes for the film industry's reformed attitude to the Internet as a vehicle for distribution of its creative works.

---

[31] http://en.wikipedia.org/wiki/Legal_issues_with_BitTorrent#cite_note-TPBlegal-0 03'rd June'09

[32] http://news.bbc.co.uk/2/hi/technology/4463372.stm 03'rd June'09

Whilst the proliferation of illegal downloads has demonstrated the ease with which BitTorrent technology can enable films to be downloaded, the aim now is to harness that reputation and use it to drive the sale of legal film downloads with the consent of the film studios. Different download options are available, with both daily rentals and permanent downloads on offer at competitive prices (typically US$2.99 to $3.99 for a 24 hour rental). On the subject of price sensitivity, Ashwin Nawin, president and co-founder of BitTorrent, has been quoted as saying: *"We're really hammering the studios to say "Go easy on this audience". We need to give them a price that feels like good value relative to what they were getting for free"*.

The move towards offering legitimate methods of online delivery has also been embraced by BBC Worldwide, the commercial arm of the BBC, which has signed a deal with technology firm Azureus. Azureus uses BitTorrent technology to distribute BBC TV programmes legally across the Internet. In the past, any BBC programmes that were found on peer-to-peer file sharing networks were illegal copies.

Tellingly, Bitorrent Inc's new 'DNA' service is reportedly aimed at males between the ages of 15 and 35, perhaps the Internet users most likely to illegally download films. In many ways, one can see this as the equivalent of legal music download services such as Apple's iTunes. ITunes has been a success, but only time will tell whether BitTorrent Entertainment Network can match it. The music industry has to date remained fairly reticent about Internet distribution services as a whole, but it too may look to exploit BitTorrent in future, especially if the film studios are able to make a success of charged-for BitTorrent downloads.

A final point to note is that another layer of protection has been added to legal BitTorrent downloads, in the form of digital rights management. The BitTorrent video store has incorporated Windows Media DRM, so BitTorrent's content will only play back using Windows Media Player. Whether this will discourage potential customers from using the legitimate service remains to be seen. DRM has proved to be highly unpopular with consumers in the past, with iTunes ultimately bowing to pressure and negotiating with rightholders so that it can now offer DRM-free versions of some tracks, albeit at a slightly higher price than the standard download rate. Indeed, BitTorrent Inc acknowledges that DRM is unpopular and has openly stated that it hopes to be able to offer DRM-free download in due course.[33]

Bit Torrent's foray into venture capital markets has only helped it in the long run. At the time of writing, the total funding on Bit Torrent was US$ 40.8 millions. DCM, ACCEL Parteners and DAG Ventures have provided Bit Torrent money in funding in a bid for commercial respectability.[34] [35] [36] [37]

Bit Torrent plans to move from being recognized as popular software for piracy to popular software that can help combat piracy by working in tandem with the same firms that are fighting it. Bit Torrent accounts for between 20% - 35% traffic on the internet. It has over 10 million simultaneous users as on 2006. Bit torrent plans to slowly wean its users away from illegal distribution of copyright content.It has two plans. One is to introduce a club model wherein a peer gets paid for sharing a file with another introduced peer who will pay for it and so on and so forth. Another is to introduce the compulsory advertisement strategy where a user has to click on an advertisement to proceed to the download. [38] Firms could also save money by opting to distribute their copyright content like video and Music over Bit Torrent which needs no central server compared to the conventional model of hosting files on a central server which could cost them money.

---

[33] http://links.cecollect.com/409/2517/BitTorrent.pdf 03'RD jUNE'09
[34] http://arstechnica.com/old/content/2005/09/5363.ars 03'rd June'09
[35] http://www.businessweek.com/technology/content/may2006/tc20060508_693082.htm 03'rd June'09
[36] http://torrentfreak.com/bittorrent-inc-receives-20-million-in-funding/ 03'rd June'09
[37] http://www.crunchbase.com/company/bittorrent 03'rd June'09
[38] http://www.businessweek.com/technology/content/may2006/tc20060508_693082.htm 03'rd June'09

5. CONCLUSION

Although the term P2P, in many peoples' minds, is linked with distributing copyrighted material illegally, it has in fact much more to offer. P2P file-sharing applications have addressed a number of important issues when dealing with large-scale connectivity of transient devices. There are a number of practical real-world applications for such a technology, both on the Internet and on wireless networks, e.g., for mobile sensors applications, and in many different kinds of scientific and social experiments.

P2P could provide more useful and robust solutions over current technologies in many different situations. For example, current search engine solutions centralize the knowledge and their resources. This is an inherent limitation. Google, for example, relies on a central database that is updated daily by scouring the Internet for new information. Simply due to the massive size of this database (more than 1.6 billion entries) not every entry gets updated every day, and as a result, information can often be out of date. Further, it is impractical (from a cost perspective) that such solutions will be scalable for the future Internet.

For example, even though Google, at the time of writing, runs a cluster of 10,000 machines to provide its service, it only searches a subset of available Web pages (about 1.3 x 108) to create its database. Furthermore, the world produces two exabytes (2 x 1018 bytes) each year but only publishes about 300 terabytes (3 x 1012 bytes) i.e. for every megabyte of information produced, one byte gets published. Therefore, finding useful information in real-time is becoming increasingly difficult. Corporations could also provide specialized information available that current search engines cannot reach. Further, if the user's server disconnected from the network then the search service would also become unavailable and therefore users searching would not receive results for unavailable resources as they do at present. This solution outlines an extreme P2P solution, but in practice some combinational technique could prove very effective.[39]

The legal distribution of films via the Internet using BitTorrent may help to reduce piracy, but the film industry will also not want this to impact on its other revenues, for example by reducing the number of DVDs sold. In many ways, BitTorrent's Entertainment Network is approaching the problem of unauthorised film downloads in a manner reminiscent of Apple's offering to the music industry with iTunes. By offering daily rentals of videos for relatively small payments - or permanent downloads of films for slightly more - the aim is to gradually wean downloaders away from unauthorised copies of films and on to legal downloads.

In this way, BitTorrent is seeking to open up a new world full of opportunities for content owners to generate revenue. The technology which was once seen as the enemy of the film studios is now being used to try and benefit them - only time will tell whether that transformation will be effective. The very nature of the Internet means that whilst it offers unparalled opportunities to deliver content to paying customers, it is also practically impossible to prevent all illegal distribution of content. It may be that for now, rightholders

[39] **From P2P to Web Services and Grids : Peers in a Client/Server World by Ian J. Taylor BITTORRENT ISBN 978-1-8462-8074-0**

are content to focus on growing the legitimate market by persuasion rather than legal threats.[40]

REFERENCES:

1. http://thinkexist.com/quotes/brad_shapcott/

2. http://www.internetworldstats.com/stats.htm

3. **From P2P and Grids to Services on the Web – Sammes, A.J, Harrison, Andrew B, Taylor, Ian J. ISBN 978-1-8480-0123-7**

4. http://wikisummaries.org/The_World_Is_Flat#Chapter_Two_.E2.80.93_The_Ten_Forces_That_Flattened_the_World

5. http://www.wisegeek.com/what-are-torrents.htm

6. http://www.isoc.org/internet/history/brief.shtml#Origins

7. http://books.google.com/books?id=NTF6jj_360UC&printsec=frontcover&dq=communication+network+analysis&hl=eN#PPA179,M1

8. http://www.securityfocus.com/infocus/1843/2

9.http://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.html#CITEyangvldb01

10. http://www.brainyquote.com/quotes/quotes/b/bramcohen219752.html

11. **From P2P to Web Services and Grids : Peers in a Client/Server World by Ian J. aylor BITTORRENT ISBN 978-1-8462-8074-0**

12. http://www.bittorrent.com/

13. http://dessent.net/btfaq/#where

14. http://news.bbc.co.uk/2/hi/technology/4463372.stm

15. http://www.businessweek.com/technology/content/may2006/tc20060508_693082.htm

16. http://links.cecollect.com/409/2517/BitTorrent.pdf

17. http://www.crunchbase.com/company/bittorrent

18. http://en.wikipedia.org/wiki/Legal_issues_with_BitTorrent#cite_note-TPBlegal-0

[40] http://links.cecollect.com/409/2517/BitTorrent.pdf 03'RD jUNE'09