

The innovation of BitTorrent file distribution protocol

Tong He

The efficiency of individuals' online activities is inextricably entwined with the connection speed of the Internet. Over the past few decades, the slow speed and a certain waste of source have been two persistent problems in the traditional transmission mode based on conventional protocols. Consequently, researchers have been devoted to creating a new protocol to overcome the limitations, and eventually BitTorrent file distribution protocol or BT for short, has emerged as a computer program in the field of information technology to considerably improve transfer efficiency online. This innovation has become necessary because it not only facilitates online activities by being embedded in various software, but also, as Silva Dias and Ricardo (2017) point out, allows a new peer-to-peer network that endows users with means to easily and efficiently swap files with each other. Some hackers, however, can also launch attacks to computers with the malicious use of this protocol. This essay will hence explain the innovation of BT protocol, describe its applications in various fields, evaluate both positive and negative effects on society and IT industry, and finally indicate a further research direction to improve the innovation.

In order to transfer files between two machines, there must be an underlying program operated before connection, consisting of a set of source codes and detailed data as to how the files will be transformed and encrypted into digital information, and how each side will send and receive it. This fundamental program, in telecommunications, is called 'protocol'.

BitTorrent, invented by Bram Cohen, a tech-savvy programmer, is a protocol based on peer-to-peer, that can be used to upload and download pieces of file to each other without the need for the big powerful machine that holds web pages and downloadable files called central server.

Once BitTorrent works, a small machine in our homes and office, called client which request a certain file, begins to connect a tracker, which is a transfer server that instantly reads and saves information about what file the client wants. Woodford (2017) gives a clearer explanation that the computer hosting the original file, in its entirety, called a seed and it then splits the file up into a great number of pieces to requesters; simultaneously the tracker finds all the online clients who have small pieces of the file and connect each other. Finally, at any given moment, each client is downloading some parts of the file from some of these peers and uploading other parts of the file to other peers. Unless closing the BT protocol, our computers download from others while offering pieces we have just already received.

In the past, people used conventional protocols, such as Hyper Text Transfer Protocol(HTTP) and Transmission Control Protocol(TCP), to share files online. Hoffman (2016) illustrates that the traditional transport method is to connect users' computers(client) with the web page's central server (hosting machine) and then download files directly from the server, which is firsthand and safe. Nevertheless, all upload cost will be placed on the hosting machine, which means the network is more likely to be paralyzed when a crowd of clients are transferring the same file simultaneously. As a result, it will experience very low uploads and downloads even with a fast Internet connection (Yādava & Singh, 2009). On the

contrary, with BT protocol, in order to receive files, users have to give them, which means the more clients who want to download the file are, and the faster the transmission is.

In a major study (Bhagwan, Moore, Savage, & Voelker, 2002), it was found that before BT was invented, many researchers had attempted to create new protocol to replace the old and indeed invented several small programs. However, these programs had not been previously deployed on a large scale as two major problems cannot be solved at that time. On one hand, Building and maintaining a peer-to-peer network incur a huge overhead; On the other hand, Real availability for use after test experiences tremendously high churn rates, because peers always cease uploading and close the window once they have downloaded what they want, thus hardly leading average concurrent users to a high level (Maymounkov & Mazières, 2002). Eventually, Bram Cohen solved these two problems well and thus invented BT protocol in 2003. He used a technique we refer to as 'rarest first', derived from the concepts of economic optimization. That is, when selecting which piece to continue, clients choose the fewest piece of their own peers, which guarantees more common are left for later. Therefore, the likelihood that a peer who now is offering upload will later have nothing interest to download is dramatically reduced. In addition, he utilized a series of algorithms to reliably saturate upload capacity. These are two good solutions to the second problems (Cohen, 2003). With optimization of BT protocol, it was also utilized in a variety of profitable projects, which solved the first problem.

BitTorrent, in recent years, not only has already been implemented, but also has been already widely deployed. Initially, BitTorrent has been regarded as the cheapest and most effective method of sending large files with the public. Ernesto (2010) reported that the UK

government used BitTorrent to release several large data sets detailing how the tax revenues was spent, as part of the UK Prime Minister's transparency initiative in 2010. Similarly, the National Aeronautics and Space Administration (NASA) in 2005 applied BitTorrent technology to the 'Visible Earth' program which enables online users to download images from the agency's catalog of satellite photography. These pictures are entirely massive, each of which even approach 3 GB in size. BitTorrent makes it possible for government and related organizations to be transparent about such issues or big data with the public. Furthermore, shuttling large files via BitTorrent saves significant bandwidth and thus hard cash.

Apart from the direct use of BitTorrent for distribution mentioned above, at present, it is more prevalent to embed BitTorrent in download managers to swap files with each other effortlessly. Numerous download managers, such as Thunder, Flashget, Gui-Youget, and μ Torrent, can help users download diversiform digital data. For example, when an average BT user seeks to download a film, a video clip or other digital audio files, always with a large size and sometimes without source, he just runs the software to find the source easily from multiple BT users and downloads them far faster than using conventional software without BitTorrent.

Moreover, companies use BitTorrent for server deployment in their websites. Facebook, for instance, a for-profit corporation as well as a social networking service in America, has collaborated with BitTorrent inc since 2010. With the adoption of BitTorrent, Facebook can now push hundreds of megabytes of new code to all global servers in a few seconds, thus updating internal systems of tens of thousands of machines as quickly as possible. Tom

Cook, a programmer of Facebook's systems engineering group, delivered a speech in 2010 at the Velocity Conference titled 'A Day in the Life of Facebook Operations' where he discussed the efficiency of BitTorrent in server deployment. Cook (2010) mentioned that the daily code updates for Facebook was in a predicament until they discovered BitTorrent and he also contended that BitTorrent allow them to relieve loads of scaling concerns they had previously. Besides these huge brands, several universities also do the same. A Dutch university reported that after using BitTorrent, it liberated 20 of the 22 servers which were used to transmit update the workstations, saving time as well as overheads. The distributed nature of BitTorrent means a process that once took dozens of minutes, now finishes in less than a dozen seconds.

Inevitably though, BitTorrent is also maliciously used by perpetrators, which is showing a grim momentum in recently years. Lucian (2015) notes that BitTorrent can be abused by hackers to amplify denial-of-service attacks. A denial-of-service(DOS) attack is an internet-based in which the perpetrator repeatedly accesses internet service providers to overload them deliberately, thus causing the websites to load slowly or not at all. One of the best-known examples of this is MyDoom, the fastest spreading e-mail worm ever. On February 1, 2004, the website of the SCO Group was bombarded with a DOS attack amplified by BitTorrent by Myboom and in just the first hours, according to the security firm mi2g, Mydoom had caused \$800 million worth global economic damages (Wehner, 2007). Obviously, the major advantages to an attacker of using BitTorrent are that with its distributed nature, billions of unsuspecting BT users' computers can be abused by hackers to intensify attacks exponentially and unconsciously, and hence the behavior of each attack

machine can be stealthier, making it harder to track and shut down.

The emergence of BitTorrent may be the most remarkable twist on message passing among countless individuals in the so-called information era. It enables every Internet user to access more information and share files easily, efficiently, and freely. To be more specific, firstly, this new technology of dissemination assists in redistributing the cost of distribution, thus making offering a file to an unlimited number of users affordable. In addition, it drastically accelerates data download and avoids transmission redundancy of servers by allowing users to share their upload bandwidth with one another. By allowing a client to download pieces of a file simultaneously from various sources, BitTorrent reduces users' latency time as well as avoiding bottlenecks at a centralized server. Last but not least, it helps adopters save huge money on their bandwidth bills and resource consumption. These adopters could not afford to shuttle their files without employment of this technology and correspondingly reduced demands on the databases and bandwidth (Anastasi, Giannetti & Passarella, 2010).

Similarly, enterprises can yield money, from fledgling companies to big brands, for the profitability of BitTorrent. For one thing, BitTorrent Inc, founded by Bram Cohen, has been engaging in the ongoing development of BitTorrent and released a series of related products. To date, more than 170 million people use its products every month and BitTorrent protocol moves as much as 40% of the world's Internet traffic on a daily basis, which brings immeasurable economic benefits. For another, its corporate customers can prompt efficient server deployment mentioned above. These efficiencies reduce maintenance and downtime by a large margin, and hence minimize exposure to security vulnerabilities. Besides, BT file

delivery renders companies capable of conserving bandwidth and reduce networking costs.

With the proliferation of such inexpensive broadband connections, companies will save much more operating costs, even increase online traffic and fan acquisition without any charge.

Apparently, BitTorrent can be the idealist method to transfer large files to billions of locations in a short period of time, which has promoted technological revolution and has a profound influence on society. The hazards to network security and copyright, however, still exist owing to its openness and lack of supervision.

Although the protocol itself and related platforms eschew piracy, perpetrators, on one hand, can employ its open-source file-sharing nature to facilitate the unauthorized dissemination and reproduction of copyrighted material, which is deemed illegal. Law enforcement and prosecutorial agencies are bending themselves to address this avenue of copyright infringement. Information can be propagated uninhibited over numerous users without permission and paying, which is guilty of copyright infringement (Karunaratne, 2012). On the other hand, malwares inserted with BiTorrent lure tons of unsuspecting peers to download files that will damage their computers. They launch DOS attacks to paralyze the whole Internet, thus exerting devastating influence on victims.

In conclusion, the innovation of BiTorrent file distribution protocol makes a great contribution to the whole society and facilitate daily life. It is noteworthy that there is certainly nothing inherently wrong with BitTorrent file distribution protocol. The problem is how the protocol is used, legally or maliciously. Therefore, for relevant department, there is abundant room for

further progress in strengthening supervision of piracy and pursuing legal action against malicious use; for average users, they are bound to filter legitimate content sources. From my own perspective, it is worthwhile devoting much effort to the study in information security in my field such as firewalls, attack mitigation devices, to solve these problems ultimately.

References

- Anastasi, G., Giannetti, I., & Passarella, A. (2010). A BitTorrent proxy for Green Internet file sharing: Design and experimental evaluation. *Computer Communications*, 33(7), 794-802.
- Bhagwan, R., Moore, D., Savage, S., & Voelker, M. G. (2002). Replication strategies for highly available peer-to-peer storage systems. In A. Schiper, A. Shvartsman, H. Weatherspoon, & B. Zhao (Eds.), *Future Directions in Distributed Computing* (pp. 153-158). New York, America: Springer.
- Castro, M., Druschel, P., Kermarrec, A., Nandi, A., Rowstron, A., & Singh, A. (2003). SplitStream: high-bandwidth multicast in cooperative environments. *Operating Systems Review*, 37(5), 292-313. doi:10.1145/1165389.945474.
- Cohen, B. (2003). Incentives build robustness in BitTorrent. *Workshop on Economics of Peer-to-Peer systems*, 6(1), 68-72.
- Cook, T. (2010, June 24). *A Day in the Life of Facebook Operations* [Audio podcast]. Retrieved from https://www.youtube.com/watch?v=T-Xr_PJdNmQ
- Ernesto. V. (2008). Dutch University uses BitTorrent to update workstations. Retrieved from <https://torrentfreak.com/university-uses-utorrent-080306/>
- Ernesto, V. (2010). UK government uses BitTorrent to share public spending data. Retrieved from <https://torrentfreak.com/uk-government-uses-bittorrent-to-share-public-spending-data-100604/>.
- Hoffman, C. (2016). How does BitTorrent work. Retrieved from <https://www.howtogeek.com/141257/htg-explains-how-does-bittorrent-work/>.
- Karunaratne, S. B. (2012). The Case Against Combating BitTorrent Piracy Through Mass John Doe Copyright Infringement Lawsuits. *Michigan Law Review*, 111(2), 283-309.
- Maymounkov P., & Mazières, D. (2002) Kademlia: A Peer-to-Peer information system based on the XOR Metric. In P. Druschel, F. Kaashoek, A. Rowstron (Eds.), *Peer-to-Peer Systems*. (pp.53-65). Heidelberg, Germany: Springer.
- Silva, P. D., Dias, J., & Ricardo, M. (2017). Mistrustful P2P: Deterministic privacy-preserving P2P file sharing model to hide user content interests in untrusted peer-to-peer networks. *Computer Networks*, 120(1), 87-104. doi:10.1016/j.comnet.2017.04.005

Yādava, S. C., & Singh, S. K. (2009). *An introduction to client/server computing*. New Delhi: New Age International.

Wehner, S. (2007). Analyzing worms and network traffic using compression. *Journal of Computer Security*, 15(3), 303-320.

Woodford, C. (2017). How does BitTorrent work? Retrieved from <http://www.explainthatstuff.com/howbittorrentworks.html>.