Large Scale Analysis of BitTorrent Proxy for Green Internet File Sharing

Sena Cebeci[1]*, Oznur Ozkasap[1], and Giuseppe Anastasi[2]

[1]*Department of Computer Engineering, Koc University, Istanbul, Turkey*
[2]*Department of Information Engineering, University of Pisa, Pisa, Italy*
{senacebeci,oozkasap}@ku.edu.tr,giuseppe.anastasi@iet.unipi.it

Abstract. Addressing energy efficiency in P2P services has the potential to make the Internet greener since they comprise a major source of the Internet data traffic. In this paper, we consider approaches for green Internet file sharing. We develop simulation models of proxy-based energy efficient BitTorrent as well as legacy BitTorrent on PeerSim P2P simulator, and explore their characteristics on large-scale scenarios. Our aim is first to identify the operating conditions where proxy-based BitTorrent outperforms the legacy protocol in terms of overall energy efficiency, and then to develop optimizations for the proxy-based approach.

Keywords: BitTorrent, Energy Efficiency, Green Internet

1 Introduction

With energy efficiency becoming a high-priority concern in today's world, design and development of energy-aware ICT services attract increasing attention. Recent studies have shown that the internet related energy consumption represents a significant part of the overall energy consumption of ICT because Internet traffic continues to increase rapidly [1]. Moreover, P2P traffic due to file sharing represents a large fraction of the Internet traffic. Thus, P2P services and protocols largely contribute to the Internet related energy consumption. Therefore, it becomes more and more important that the P2P protocols need to consider energy efficient approaches to support green communication and content sharing.

Among the P2P file sharing protocols, BitTorrent is the highly popular protocol since BitTorrent P2P traffic constitute more than 45-78% of all P2P Internet traffic, and roughly 27-55% of all the Internet traffic depending on geographical location [2]. Solutions for green file sharing approaches are broken down into three categories and interpreted according to the power management technique they used [3]. These categories are adaptive link rate, different power management levels and proxy-based. In adaptive link rate and power management levels energy efficiency is dependent on the hardware performance of the NIC. These techniques are not good candidates for internet file sharing, thus we have chosen

* Corresponding Author

proxy-based solution. In proxy-based approach file download is fully performed by the proxy, by switching off the requested user's PC in course of the download.

A proxy-based BitTorrent called Energy Efficient BitTorrent (EE-BT) has been proposed in [4], its comparison with the legacy BitTorrent protocol has been carried out using an experimental testbed with limited number of peers. In this work, we intend to compare the performance of these approaches in large scale scenarios with a large number of peers, and we can control protocol parameters. With the analysis, we aim to find out under which operating conditions EE-BT outperforms the legacy protocol in terms of overall energy efficiency. Goals of our study include (1) Developing models of EE-BT and legacy BitTorrent on PeerSim P2P simulator [5], (2) Performing simulations of EE-BT and legacy BitTorrent on large scale scenarios to investigate several parameters, (3) Developing extensions/optimizations to EE-BT such as investigating the optimal number of peers a proxy can serve without degrading the performance, and (4) Analyzing the effect of increasing the number of proxies on the overlay.

2 BitTorrent Architecture Types

In this section we briefly describe the Legacy BitTorrent protocol [7] and Energy Efficient BitTorrent protocol [4].

2.1 Legacy BitTorrent

BitTorrent is a file distribution system essentially developed to distribute files with large size. According to the protocol terminology, nodes of the overlay are called peers, and the set of peers involved in the distribution of a file is referred to as torrent or swarm. Each peer downloads the desired file, in chunks, from a multitude of other peers instead of fetching it from a single server (as in the conventional client-server model). While downloading missing chunks, peers upload to other peers in the same torrent the chunks they have already obtained. For each torrent there is a tracker, i.e., a node that constantly tracks which peers are involved in the torrent. A peer that wants to join a torrent must register with the tracker and, then, it must periodically inform the tracker that it is still in the torrent.

2.2 Energy Efficient BitTorrent (EE-BT)

EE-BT is a proxy-based version of BitTorrent where download requests of peers are served by a BitTorrent proxy. Peers that are not directly involved in the torrent are called passive peers. They delegate the task to the associated proxy, which downloads the file on behalf of them. The proxy participates to the BitTorrent overlay network, just like any other regular peer, and takes care of the overall process. Therefore, the user's PC can be switched off during the download phase. The file will be transferred from the proxy to the user's PC later, when the user reconnects. Figure 1 shows the actions performed by the various

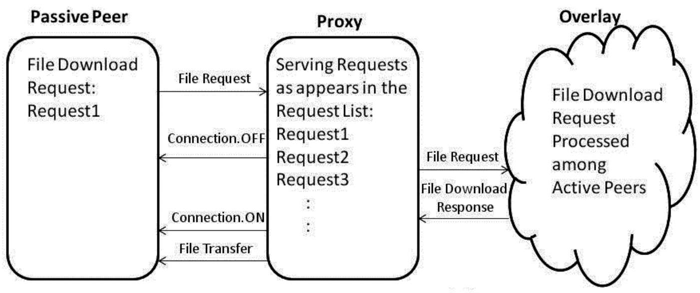

Fig. 1. EE-BT Protocol Sequence Diagram

actors of EE-BT. Upon receiving a request from the user, the software running on the user's PC contacts the proxy and requests the desired file. The proxy acknowledges the request with file request message. If the requested file is already available in the local cache of the proxy, it is immediately transferred to the user. Otherwise, the proxy starts downloading the requested file from the BitTorrent overlay network, acting as a regular peer and following the legacy BitTorrent protocol. The user's PC can be switched off just after receiving the acknowledgement from the proxy. Later, when the user reconnects, he/she can check the status of the download process at the proxy. If the file is completely available, it can be transferred from the proxy to the user's PC.
Other potential tasks that might be performed by the proxy can be listed as follows:

1. Proxy connects to other proxies in the network in order to serve the requests of the other peers belonging to different proxies.
2. Proxy is capable of serving a request of a peer which is not necessarily a neighbor of itself.

3 Analysis and Experimental Results

We have developed simulation models for Legacy and EE BT protocols. The simulation environment for these protocols is PeerSim. Peersim provides a dynamic environment to develop and test any kind of P2P protocol with high scalability (scales up to 1 million nodes), easy to configure and simple component-based architecture.

Scenario Description: We designed scenarios to measure average download time, average number of pieces uploaded/downloaded by peers and energy savings. In our system, network size is configurable and distribution of the seeders and leechers in the swarm has default values, equal to 20% seeders and 80% leechers, respectively. We consider the distribution of a single file in each experiment. Initially, all the peers in the system have file pieces randomly distributed, and all the peers in the network request the same file. The simulation terminates after all the peers downloaded the file successfully. These peers form the topology in the peerlist generated by the tracker and all the peers in the list are presumed to be neighbours of each other. We used three different file sizes 100 MB, 500 MB and 1 GB for our scenarios and file pieces are 256 KB each. Upload and download data rates per peer are selected randomly from four different options, 640 Kb/s, 1 Mb/s, 2 Mb/s and 4 Mb/s, in order to use in the transmission time of a file. To set the delays, Peersim Transport Package is used. This package includes Uniform Random Transport protocol that provides an environment for reliable message delivery between peers with random delays. Default values of minDelay (minimum delay of messaging) is set to 50 ms and maxDelay (maximum delay of messaging) is set to 200 ms. Default values used in the experiments are given in Table 1.

To explore the protocol behaviours in the network, analysis on average download time, average number of pieces uploaded/ downloaded and energy savings carried out. The analyzed metrics for a peer are:

Average Download Time: The duration (in milliseconds) to download the torrent file by a generic peer in the network.

Average Number of Pieces Uploaded/Downloaded : This metric measures the uploaded and downloaded pieces of the file by peer during the simulation, thus the file download percentage can be easily kept track of.

Energy Saving: We have calculated performance metric considering energy savings of the Legacy BT and EE BT protocol relative to each other. Without losing in generality, we assumed that all PCs and proxy have the same power consumption as in [4] and [6]. Under this assumption, it can be shown that the energy saving achieved when using EE BT is given by [6]; (1) where with respect to Legacy BT.

The formula of the energy savings introduced by EE BT with respect to Legacy BT, for single user i, is represented by:

$$S^i_{usr}(\mathrm{N}) = 1 - \frac{\sum_{i=1}^{N}(D_0(i) + 2\sum_{i=1}^{N}(D_t(i))}{\sum_{i=1}^{N} D_L(i)} \qquad (1)$$

N: Number of BitTorrent peers in the network

D_0(i): Time taken by the proxy to download the file (requested by peer i) from the BitTorrent overlay.

D_t(i): Transfer time of the file from EE BT to user

D_L(i): Download time of the torrent file with Legacy BT

Table 1. Default Parameter Values in the Simulation Setting

Parameter	Value
Network Size	1000
Piece Size	256 KB
Seeder Distribution	20%
Leecher Distribution	80%
File Size	100 MB
Upload/Download Rate	640-4096 Kb/s
Message MinDelay	50 ms
Message MaxDelay	200 ms

Figure 2 (a) indicates the average download time of a torrent file having, 100 MB, 500 MB and 1 GB file sizes in legacy BT and EE BT for samples (1 to 4) of generic peers for different network sizes 10, 100, 1000, 10000, respectively. In this case, it is assumed that according to the network size all the peers either uses Legacy or Proxy-based solution for the file download. We observe that the average download times for small file sizes (i.e.,100 MB) are close to each other for both protocols. For instance, for a sample from 10 peer network size (i.e., sample 1), average download times of legacy BT and EE BT are 1.812 min (0.0302 hour) and 1.548 min (0.0258 hour), respectively. On the other hand, for large file sizes (i.e.,1 GB) the differences between these two protocols become more clear, average download rate of legacy BT is 29.8 hours whereas EE BT is 24.9 hours.

Figure 2 (b) shows the average number of uploaded and downloaded pieces during the simulation. For the legacy BT, average upload/ download rates exhibits some small differences. For example, for different samples the range of pieces differ fro 184.33 to 140.10 for 100 MB file. Furthermore, this rate decreases while the number of peers increases. In other words, protocol message traffic (choke messages) affects the performance in a negative manner. On the other side, for the EE BT protocol shows average upload/download rates of pieces do not differ significantly from each other.

We have calculated Energy Savings Percentages for 100 MB and 500 MB file sizes, with upload rate 1 Mb/s. D_t(i) values for file sizes 100 MB and 500 MB are 0.22 h and 1.11 h, respectively. Average download time samples for different network sizes (i.e. Sample 1 :10 peers, Sample 2: 100 peers, etc.) are shown in Table 2. Energy Savings of a single user calculations in terms of percentages based on equation 1 are represented in Table 3.

We also consider another (hybrid) scenario when the peers in the overlay are defined as passive and active. Peers connected to the proxy for file download request are identifed as passive. The remaining peers in the system are called active. In Figure 3 (a), we investigate how average download time is affected by two protocols if the percentage of passive peers in the network increases. The results show EE BT average download time is less than Legacy BT. For example, for a file size 500 MB EE BT download time is 13.75% better. Furthermore, for 1 GB file size EE BT average download time is 16.25% lower. As shown in Figure

3 (a), for both protocols, the average download time slightly decreases when the percentage of passive peers increases.

In Figure 3 (b), the increment on passive peers in the network causes relatively large decreases on average number of pieces uploaded/downloaded per peer. The significant reduction of number of pieces uploaded/downloaded is the reason of passive peers do not upload but download from the proxy.

Table 2. Average Download Times (Single User)

Samples	D_L (100 MB)	D_{EEBT}(100 MB)	D_L(500 MB)	D_{EEBT}(500 MB)
1	0.0302 h	0.0258 h	0.1632 h	0.1123 h
2	0.0310 h	0.0260 h	0.1680 h	0.1118 h
3	0.0313 h	0.0252 h	0.1711 h	0.1112 h
4	0.0318 h	0.0247 h	0.1746 h	0.1108 h

Table 3. Single User Energy Savings

Samples	100 MB File	500 MB File
1	15.42%	14.29%
2	15.03%	13.87%
3	14.86%	13.62%
4	14.61%	13.34%

4 Future Work and Conclusions

In this work, we have developed Legacy BitTorrent and EE BitTorrent implementations on Peersim P2P simulation environment. We have evaluated the simulation results in large scale scenarios (i.e. with large number of peers) using three performance metrics. The first metric represents the average download time taken by the two protocols and the other one investigates average number of pieces uploaded/downloaded per peer in the run time. Lastly, user energy savings comparison measured by performance metric. We have measured the energy saving introduced by EE BitTorrent with respect to legacy BitTorrent. Our simulation results have revealed that proxy based EE BT decreases the average download time of the torrent file by 16.25%.

For future work, we plan to extend our simulation model increasing the number of proxies in the network, and in order to explore how it affects the system in terms of energy efficiency. Additionally, we will investigate optimizations and extension methods to EE BitTorrent to examine parameters such as optimal number of peers a proxy can serve without degrading the performance.

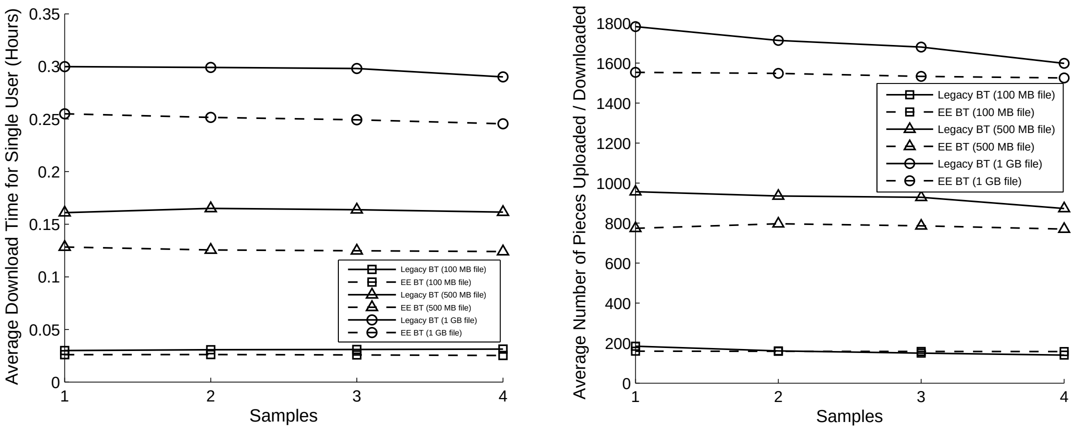

Fig. 2. Legacy BT vs EE BT (a) Average download time for single user (b) Average number of pieces downloaded/uploaded for single user

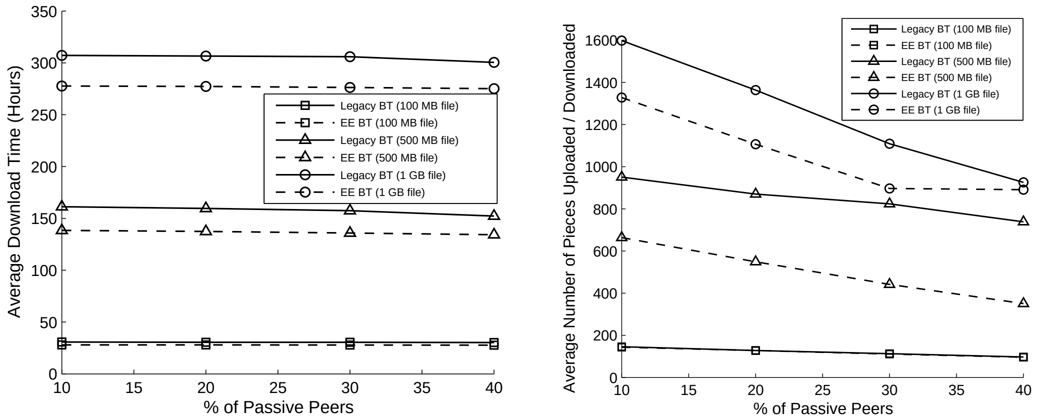

Fig. 3. Passive node percentage in Legacy BT and EE BT (a) Average download time per peer (b) Average number of pieces downloaded/uploaded per peer. The total number of peer is constant and equal to 1000.

Acknowledgments

Our work was partially supported by the COST (European Cooperation in Science and Technology) framework, under Action IC0804: Energy efficiency in large scale distributed systems, and by TUBITAK (The Scientific and Technical Research Council of Turkey) under Grant 109M761.

References

1. Y. Audzevich, A. Moore, and A. Rice, R. Sohan, S. Timotheou, J. Crowcroft, S. Akoush, A. Hopper, A. Wonfor, H. Wang, R. Penty, I. White, X. Dong, T. El-Gorashi, J. Elmirghani: Intelligent Energy Aware Networks. In: Handbook of Energy-Aware and Green Computing. Chapman and Hall/CRC (2012)
2. Y. J. Lee, J.-H. Jeong, H. Y. Kim, and C. H. Lee, Energy-saving set top box enhancement in bittorrent networks, in Network Operations and Management Symposium (NOMS), 2010 IEEE, pp. 809812, (2010)
3. G. Anastasi, M. Conti, A. Passarella," Power management in mobile and pervasive computing systems," Algorithms and Protocols for Wireless and Mobile Networks, (2005)
4. G. Anastasi, I. Giannetti, A. Passarella, A BitTorrent proxy for Green Internet file sharing: Design and experimental evaluation, Computer Communications, Volume 33, Issue 7, pp. 794-802, (2010)
5. The PeerSim Simulator, http://peersim.sf.net, 2013.
6. I. Giannetti, G. Anastasi, M. Conti, "Energy-efficient P2P file sharing for residential BitTorrent users," Computers and Communications (ISCC), 2012 IEEE Symposium pp.524-529, (2012)
7. L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang, "Measurements, analysis, and modeling of BitTorrent-like systems," 5th ACM SIGCOMM conference on Internet Measurement (IMC '05). USENIX Association, Berkeley, CA, USA, (2005)