A novel approach to energy efficient content distribution with BitTorrent

Matthew Forshaw and Nigel Thomas

School of Computing Science
Newcastle University, Newcastle upon Tyne, UK
{m.j.forshaw,nigel.thomas}@ncl.ac.uk

Abstract. The energy efficiency aspects of IT infrastructure and communications systems are facing increased scrutiny, and a broad range of compelling financial, social, political and legislative factors is emerging. In this paper, energy efficiency considerations are addressed in the context of BitTorrent. We provide mechanisms to facilitate energy efficiency and energy proportionality, and propose an energy-efficient content distribution system employing these mechanisms to minimise energy consumption and reduce cost.

Keywords: Energy efficiency, Peer to Peer (P2P), Content distribution, BitTorrent.

1 Introduction

Energy costs now dominate IT infrastructure total cost of ownership (TCO), with data centre operators predicted to spend more on energy than hardware infrastructure in the next five years. With western european data centre power consumption estimated at 56 TWh/year in 2007 and projected to double by 2020 [4], the need to improve energy efficiency of IT operations is imperative. The issue is compounded by social and political factors and strict environmental legislation governing organisations.

BitTorrent [7] is a peer-to-peer (P2P) file sharing protocol, accounting for approximately 17.9% [15] of overall Internet bandwidth use. Contrary to traditional client-server approaches, BitTorrent relies less on the distributor's centralised infrastructure and bandwidth, offering a scalable content distribution solution with reduced provider-side power consumption and cost. This scalability makes BitTorrent particularly resilient to *flash crowds* [10], vast numbers of users accessing content simultaneously, a behaviour often observed for new and popular content.

In this paper we introduce provider-side mechanisms to promote energy-efficient and energy-proportional operation of a BitTorrent based content distribution system. Our approach is complementary to the proxy scheme proposed in [1], and alleviates the need for centralised peer control as imposed in [2] and [5]. In this research we consider situations where such centralised control cannot be guaranteed, and present mechanisms which do not require alterations to client logic. These relaxed conditions make our approach more broadly applicable as well as simplifying deployment.

2 Related Work

Early research considering BitTorrent energy efficiency focused primarily on file sharing using devices with limited battery and computational power [11].

Anastasi et al. [1] propose a scheme allowing multiple peers within a typical LAN environment to delegate the task of downloading to a designated proxy server which takes part in the BitTorrent protocol on their behalf. Meanwhile these peers "behind" the proxy can be switched off without interrupting the download. Upon completion of the download, the requested files are transferred back to the peers.

Blackburn and Christensen [5] introduce a wake-up semantic to the BitTorrent protocol, allowing peers to sleep while remaining active in the system. Centralised control is assumed whereby these peers may be sent a packet and woken up remotely.

Andrew et al [2] propose a system to balance the power consumption of servers and peers involved in a peer-to-peer download. This approach assumes centralised control over all peers, enabling these peers to be powered on and off to maximise the download rate of a subset of awake peers.

3 BitTorrent

When a downloader (*peer*) initiates a download via BitTorrent, they first obtain a *torrent file*, a file containing metadata for the requested content. This metadata includes an endpoint to a BitTorrent tracker node. The *tracker* is essential to the operation of any BitTorrent system. The tracker maintains records of all peers uploading or downloading particular content (known collectively as the *swarm*), and coordinates content distribution and enables peer discovery. This component must remain online at all times in order for newly arriving peers to be able to connect.

Once the peer has established a connection with the tracker, the tracker responds with a peer list containing the details of a random subset of the other peers transferring the requested content. The peer may then connect to, and obtain content from, these peers. Additionally, the peer may elect to obtain up-to-date peer lists from the tracker periodically according to an *announce interval* specified by the tracker.

Files in BitTorrent are split into multiple *pieces*, allowing peers to share pieces of the file they hold while obtaining the pieces they require. BitTorrent peers' ability to download and upload simultaneously benefits performance and makes BitTorrent significantly more scalable than client-server file distribution approaches.

BitTorrent peers may belong to one of two states; *leeching* or *seeding*. Peers actively downloading in the system but who do not currently hold a full copy of the file are referred to as *leechers*. Once a peer has obtained all the pieces of their download, they may either depart from the system or remain active as a *seed*. Seeds remain active participants in the system, altruistically sharing upload bandwidth to distribute content to other peers.

4 System Models and Objectives

In our model we represent peer power consumption as manufacturer specified *nameplate* power consumption figures. Selecting readily available power consumption values provides sufficient accuracy for our system to make valuable energy savings while minimising the overhead associated with collecting the information. We also maintain details of the download and upload capacity of individual peers. These may be figures obtained out of band or taken from real-time observations of the running system.

We model a *seed pool* as a group of servers under centralised control, heterogeneous in terms of power consumption and upload capacity. The upload capacity of these servers is assumed to be considerably greater than that of typical peers. Membership is assumed to be dynamic, with servers arriving to and departing from the pool periodically. Where members of the seed pool may be considered internal architecture across one or more data centre facilities, we may assume physical access for detailed in-situ power profiling. Multiple linear regression models calibrated for each resource will provide accurate estimates based on real-time resource utilisation measurements, including CPU, RAM and disk activity. Software agents instrumenting each machine communicate this utilisation data to the tracker.

Our model considers tracker and seed instances to belong to one of two distinct states; *sleep* or *active*. An active resource is fully powered up and is able to execute operations and serve requests from the system. A resource may be placed in a sleep state, where the machine is no longer able to serve requests but consumes significantly less power. While asleep, system state is stored in memory allowing the machine to transition into an active state quickly. We model the time taken to transition between these two states, during which the resources consume power but are unable to contribute to the system.

Content distribution networks are typically large shared infrastructures, distributed across multiple data centre facilities nationally or globally. Hence, it is imperative that our system model adequately represents the differences between data centre facilities and global variation in the cost and cleanliness of their power sources. Facility modeling includes the Power Usage Effectiveness (PUE) rating, a metric representing the proportion of facility overheads (for example, power, cooling and lighting infrastructure) in terms of the power consumption of the IT equipment. We account for variations in the price and ecological impact of energy supply in our model, representing these in pence and kg CO_2 per kWh respectively.

We consider modeling of network devices outside the data centre facility as beyond of the scope for this research. Peer-to-peer approaches have greater total bandwidth requirements than client-server approaches due to peers communicating with one another. The impact of this communication overhead on power consumption is difficult to assess. Despite significant recent improvements in energy-efficiency of hardware [16], typical network hardware is found to be energy-disproportional [13]. This power characteristic results in a narrow dynamic power range, limiting the potential impact of variable traffic workload on power consumption. Furthermore, these network devices must remain online at all times and are outside of the administrative control of content providers. Existing research has compared client-server and peer-to-peer approaches, finding peer-to-peer to demonstrate greater network-related power consumption but lower overall power consumption in a communication-intensive scenario such as file distribution [14].

It is unrealistic for an organisation to minimise its power consumption without first considering the trade-offs between energy efficiency, cost and reliability. In an inter-organisational scenario such as software patch distribution in an office environment or large-scale deployment across a cluster, stakeholders of the system will most likely be concerned with minimising the aggregate energy consumption and cost of a system. Conversely, in situations where peers are external to the organisation (e.g. video on demand or public content distribution), stakeholders are likely to prioritise provider-side energy efficiency and cost over those of the peers. Our approach must remain flexible in order to satisfy the various optimisation goals of the stakeholder.

5 Approach

5.1 Energy Proportional Tracker Migration

Energy Proportional Tracker Migration leverages heterogeneous hardware to promote energy proportionality of the tracker component. During periods of low utilisation the tracker will reside on a computationally constrained but energy-efficient machine, autonomically migrating to a more performant (but more costly in terms of power) server during periods of increased load. This will minimise the load-independent component of our system's overall power consumption and achieve near energy proportional operation.

Existing research has demonstrated the ability to compose a number of non energy-proportional servers, combining power saving mechanisms to deliver an energy-proportional aggregate system [17] [12]. We acknowledge the heterogeneous nature of typical real-world data centres (often caused by machine failures, and upgrades, etc) [9] and contribute mechanisms which specifically leverage hardware heterogeneity to achieve aggregate energy proportionality.

5.2 Elastic Capacity Provisioning

In Elastic Capacity Provisioning, we propose a variation of typical BitTorrent use, whereby a content distributor operates a pool of specialised seeds. It is the role of these seeds to share content to other peers, ensuring satisfactory levels of performance, energy consumption and cost. This pool is said to be elastic because instances are provisioned dynamically in response to real-time service demand. We consider the heterogeneous nature of this pool of specialised seeds when periodically recalculating and provisioning the minimum active set of seed resources to achieve desired performance, cost and energy optimisations.

Traditionally, BitTorrent seeds abide by strategy where seeder upload capacity is allocated proportionally to those peers with higher download rates, optimistic that those peers may themselves become seeds more quickly and serve other peers. We propose a hybrid scheme whereby upload bandwidth is allocated on a combination of observed download rates and peer energy inefficiency. Peers who are particularly energy-inefficient relative to the rest of the swarm will be provided with a larger proportion of the seeder's upload capacity. Enabling these peers to complete their download and leave the system more quickly reduces their power consumption. In situations where upload capacity is limited among members of the swarm, and such actions threaten the overall health of the swarm, the traditional strategy is observed to prevent starvation.

5.3 Peer Connectivity Shaping

Peer Connectivity Shaping augments the peer lists returned by the tracker, giving some peers preferential treatment by providing them with the details of a larger peer set, or of peers with greater available upload bandwidth. This promotes greater connectivity between the peer and the swarm, lowering the peer's download time and consequently reduces its energy consumption.

Once a peer list has been received, a client typically selects a random subset of peers with which to connect to in the first instance. Peers are unaware of the upload capacity of the peers when they select which peers to connect to, so it is important when a peer requests its initial peer list that the list comprises a smaller proportion of peers with slow upload rates. Subsequent peer lists may include a wider range of peer upload capabilities, as BitTorrent's "tit-for-tat" mechanism will favour peers with higher upload rates and ensure the peer receives fair download rates. In the case of a particularly energy-inefficient peer, it may be more beneficial to provide small peer lists to increase download performance at the expense of increasing the peer's connectivity with the swarm.

The interval between a peer's requests to the tracker may also be optimised to improve performance and lower energy consumption and cost. In highly dynamic systems where peers and seeds are arriving and departing frequently, it may be preferable to lower the interval between peer requests in order for them to remain responsive to the changing state of the system. Increased requests to the tracker will place the tracker under greater load so there is a subtle trade-off between increasing performance for peers without incurring greater power consumption.

The impact of these approaches should be both equitable and proportional, such that energy-efficient peers are not penalised excessively in terms of download performance, and be beneficial to the swarm as a whole. Decisions made by the system are informed by comprehensive measures of system performance collected by the tracker, and are subject to the optimisation goals of the policy currently being enforced by the service provider, and the current state of the system.

6 Experimentation

To evaluate the efficacy of our approach we have developed a simulation environment based on TorrentSim [3]. Our simulation environment extends the underlying simulation framework to allow dynamic provisioning of nodes during execution, adds power consumption modelling support, and augmented BitTorrent tracker and seed components implementing our proposed energy efficiency approaches. A monitoring component periodically collects real-time power consumption and performance metrics for offline analysis. Simulation studies are carried out for scenarios with varying levels of swarm heterogeneity and bandwidth availability. We also compare our approach with traditional client-server and naive BitTorrent approaches.

A lightweight implementation of our energy-efficient BitTorrent system is written in Python, with system models and operational data stored in a MySQL database. The implementation will be tested locally using a virtualised testing environment [8], providing greater control over the conditions under which the implementation is evaluated. Large scale testing is to be carried out on PlanetLab [6].

7 Preliminary Findings

In the initial evaluation of the Energy Proportional Tracker Migration approach we consider two normalised tracker workload traces shown in Figures 1 and 2. Workload traces WL_1 and WL_2 represent tracker requests during the arrival and service of 100 and 200 peers respectively. In each case three seeds are active in the system, and all peers depart from the system upon completing their download.

Workload WL_1 is characterised by larger peer inter-arrival times and greater availability, resulting in smaller mean peer service time. Conversely, in WL_2 peer inter-arrival times are much smaller and peer download rates are constrained by limited availability and greater competition for available upload capacity. The request rate at a given period is largely dependant on the number of peers and seeds active in the system. Observed increases in request rate over time indicate the arrival of new peers, while decreases signify peers' completion and subsequent departure from the system.

The efficacy of our provisioning approach is evaluated for two groups of servers. The first group is homogeneous in terms of both performance and power consumption, while the second comprises servers from two heterogeneous classes of server. In Figure 3 we present relative energy savings for our approach when

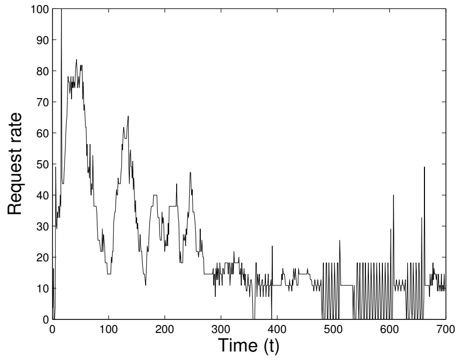

Fig. 1. Tracker workload trace WL_1.

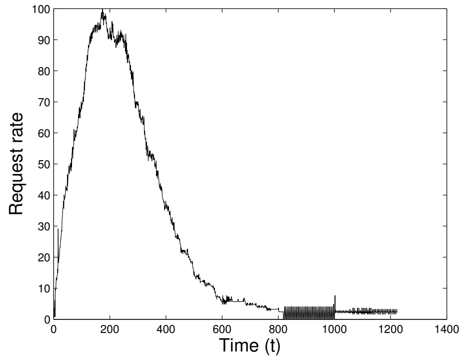

Fig. 2. Tracker workload trace WL_2.

compared to a group of servers right-sized to satisfy the peak request rate observed over the duration of the traces. In each case we find increasing the number of servers is beneficial in reducing energy consumption, allowing for finer grained provisioning of resources to satisfy the offered workload.

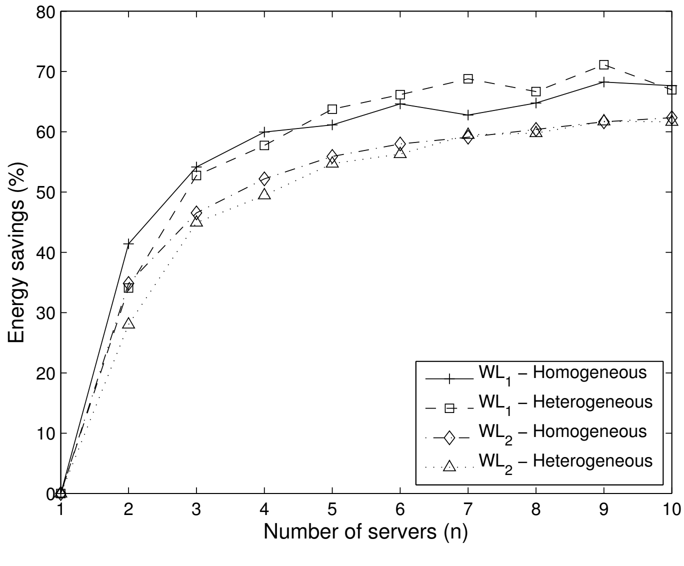

Fig. 3. Comparison of energy savings for two workload traces with homogeneous and heterogeneous groups of servers of size n.

Our initial results demonstrate the potential for considerable energy savings and reduction in the load-independent portion of aggregate power consumption with negligible SLA violations. Further detailed analysis is ongoing and will form the basis of future publications.

8 Conclusions and Further Work

The potential trust and security implications of the system proposed in this paper are of great interest. We acknowledge complexities in guaranteeing the veracity of a peer's self-reported power consumption profile in inter-organisational settings, and design the system ensuring that any free riding permitted by our mechanisms is not detrimental to the overall performance or energy efficiency of the system.

A common challenge in peer-to-peer systems is accountability [18]. The usage data collected by our energy-efficient tracker not only informs the behaviour of our system, but also allows fine-grained attribution of utility, energy-efficiency and cost. The application of this accountability information in a class of energy-aware incentive mechanisms for BitTorrent will be addressed in our ongoing research.

Peer Exchange (PEX) [19] is an extension to the BitTorrent protocol enabling decentralised peer discovery in BitTorrent swarms. In PEX, peers periodically communicate directly among themselves, sharing details of the peers with whom they are connected. PEX has been shown to be beneficial to performance. However, as peers are not equipped with global knowledge of the system this approach cannot easily be made energy-aware. We will look to investigate approaches to an energy-aware PEX-like system without impacting upon overall performance.

This paper considers the use of BitTorrent as a content distribution mechanism in a single management domain. An interesting area of future research is to extend our approach to facilitate energy-efficient use of BitTorrent in a federated network of interconnected content distribution networks. Such a federated approach would allow organisations to share resources, further reducing the need to over-provision to meet peak demand. Service Level Agreements (SLAs) between these organisations may be enforced on a combination of utility, cost and energy efficiency. Audit and accountability information may be used to facilitate billing for service between organisations. Of particular interest is the ability to reconcile the conflicting optimisation goals of multiple service providers on shared infrastructure, and energy-aware incentive mechanisms in a federated context.

References

1. Anastasi, G., Giannetti, I., Passarella, A.: A bittorrent proxy for green internet file sharing: Design and experimental evaluation. Comput. Commun. 33, 794–802 (2010)
2. Andrew, L.L.H., Sucevic, A., Nguyen, T.T.T.: Balancing peer and server energy consumption in large peer-to-peer file distribution systems. In: Proc. IEEE GreenCom (2011)

3. Barcellos, M., Mansilha, R., Brasileiro, F.: Torrentlab: investigating bittorrent through simulation and live experiments. In: Computers and Communications, 2008. ISCC 2008. IEEE Symposium on. pp. 507 –512 (july 2008)
4. Bertoldi, P., Atanasiu, B.: Electricity consumption and efficiency trends in european union. Renewable Energy Status Rep (2009)
5. Blackburn, J., Christensen, K.: A simulation study of a new green bittorrent. In: Communications Workshops, 2009. ICC Workshops 2009. IEEE International Conference on. pp. 1–6 (2009)
6. Chun, B., Culler, D., Roscoe, T., Bavier, A., Peterson, L., Wawrzoniak, M., Bowman, M.: Planetlab: an overlay testbed for broad-coverage services. SIGCOMM Comput. Commun. Rev. 33, 3–12 (2003)
7. Cohen, B.: Incentives build robustness in bittorrent (2003)
8. Deaconescu, R., Milescu, G., Aurelian, B., Rughiniş, R., Ţăpuş, N.: A virtualized infrastructure for automated bittorrent performance testing and evaluation. Internation Journal on Advances in Systems and Measurements 2(2&3), 236–247 (2009)
9. Heath, T., Diniz, B., Carrera, E.V., Meira, Jr., W., Bianchini, R.: Energy conservation in heterogeneous server clusters. In: Proceedings of the tenth ACM SIGPLAN symposium on Principles and practice of parallel programming. pp. 186–195. PPoPP '05, ACM, New York, NY, USA (2005)
10. Izal, M., Uroy-Keller, G., Biersack, E., Felber, P.A., Hamra, A.A., Garces-Erice, L.: Dissecting bittorrent: Five months in torrent's lifetime. pp. 1–11 (2004)
11. Kelenyi, I., Ludanyi, A., Nurminen, J.: Energy-efficient bittorrent downloads to mobile phones through memory-limited proxies. In: Consumer Communications and Networking Conference (CCNC), 2011 IEEE. pp. 715–719 (2011)
12. Krioukov, A., Mohan, P., Alspaugh, S., Keys, L., Culler, D., Katz, R.H.: Napsac: design and implementation of a power-proportional web cluster. In: Proceedings of the first ACM SIGCOMM workshop on Green networking. pp. 15–22. Green Networking '10, ACM, New York, NY, USA (2010)
13. Mahadevan, P., Sharma, P., Banerjee, S., Ranganathan, P.: A power benchmarking framework for network devices. In: Proceedings of the 8th International IFIP-TC 6 Networking Conference. pp. 795–808. NETWORKING '09, Springer-Verlag, Berlin, Heidelberg (2009)
14. Nedevschi, S., Ratnasamy, S., Padhye, J.: Hot data centers vs. cool peers. In: Proceedings of the 2008 conference on Power aware computing and systems. pp. 8–8. HotPower'08, USENIX Association, Berkeley, CA, USA (2008)
15. Price, D.: An estimate of infringing use of the internet. Tech. rep., Envisional Ltd (2011)
16. Ryckbosch, F., Polfliet, S., Eeckhout, L.: Trends in server energy proportionality. Computer 44(9), 69 –72 (2011)
17. Tolia, N., Wang, Z., Marwah, M., Bash, C., Ranganathan, P., Zhu, X.: Delivering energy proportionality with non energy-proportional systems: optimizing the ensemble. In: Proceedings of the 2008 conference on Power aware computing and systems. pp. 2–2. HotPower'08, USENIX Association, Berkeley, CA, USA (2008)
18. Turcan, E., Graham, R.L.: Getting the most from accountability in P2P. In: 1st International Conference on Peer-to-Peer Computing (P2P'01). pp. 95–96 (2001)
19. Wu, D., Dhungel, P., Hei, X., Zhang, C., Ross, K.: Understanding peer exchange in bittorrent systems. In: Peer-to-Peer Computing (P2P), 2010 IEEE Tenth International Conference on. pp. 1–8 (2010)