

Energy Efficiency in Web Search Engines

Matteo Catena

Gran Sasso Science Institute National Research Council of Italy
67100 L'Aquila, Italy 56124 Pisa, Italy
matteo.catena@gssi.infn.it

Today, Web search is a frequent action in the everyday life of many people. To perform it on a large scale, Web companies need energy-hungry data center, which raise environmental and economical challenges. For these reasons, Green Information Retrieval promotes energy and energy-cost awareness in contemporary Web search engines. In this document, we propose to further the research on Green Information Retrieval, which is still at its early stage. Moreover, we illustrate our first results in evaluating and improving the energy efficiency of search servers.

Keywords: Web search, energy efficiency, hardware power management

1. INTRODUCTION

Web search engines continuously crawl and index large amount of web pages, which have to be promptly retrieved in response to user queries. To do so, Web search companies – e.g., Google, Yahoo!, Microsoft, Yandex, Baidu, etc. – need computer systems with large computational power and data storage capabilities. Such systems are reported to be composed by thousands of computers organized in clusters (Barroso et al. 2003), which can efficiently handle big quantities of data. These companies started building large data centers to house such computer clusters. A data center hosts large computer systems together with the associated infrastructures, such as: telecommunications, power supplying, thermal cooling, fire suppression, etc. While data centers enable large-scale search, they also raise environmental and economical issues. The ICT sector has been reported to be responsible for roughly 2% of global carbon emissions in 2007, with general purpose data centers accounting for 14% of the ICT footprint (GeSI 2008). Moreover, power and cooling cost for 15,000 commodity servers could exceed 280,000 \$/month in 2003 (Barroso et al. 2003).

For such reasons, improving data center energy efficiency has become an attractive and active research area. Nevertheless, little literature exists about energy efficiency in search engines data centers. Chowdhury is the first to explicitly write about Green Information Retrieval and to propose a research agenda for evaluating and reducing energy consumption in search services (Chowdhury 2012). In line with this agenda, we want to evaluate the

energy expenditure due to the different components of a search engine. Also, we aim to investigate on possible energy saving strategies at the software (e.g., which algorithms are used to implement a component) and software architectural level (e.g., how the components are combined to form a green Information Retrieval system).

2. FIRST RESULTS

In this section, we briefly describe our initial findings on both evaluating and improving the energy efficiency of a Web search engine. At this stage, our work focuses on single search servers.

2.1. Query energy consumption

Query energy consumption is the energy consumed by a search server to solve a single query. Such information is important, since electric energy is expensive and commercial Web search engines have to keep a low cost-per-query to be profitable (Barroso et al. 2003). Moreover, recent works try to reduce search engines expenses and carbon footprint by taking into direct account their energy consumption. For instance, energy cost has been recently considered for devising energy-saving caching mechanism (Sazoglu et al. 2013). Then, precise measurements of query energy consumptions would be beneficial for such approaches.

It is possible to experimentally show that query energy consumption is linear in the query processing time. Details can be found in (Catena and Tonelotto 2015), where we experiment using the TREC ClueWeb09 corpus and MSN 2006 query log

to measure the energy consumption of a search server. Results reinforce the importance of efficiency improvements in Information Retrieval. More specifically, the carbon footprint of search engines can be lowered by reducing query response times without demanding additional, energy consuming, hardware. Therefore, low latencies are necessary not only to achieve user satisfaction, but also to tackle the economical and environmental costs of data centers.

2.2. Load-sensitive CPU Power Management

Typically, the energy consumption of a server is dominated by its CPU. Dynamic Frequency Scaling (DFS) technologies trade performance for reduced energy consumptions, by throttling CPU frequency (Snowdon et al. 2005). When running at low frequencies, processors absorb less power but also have lower performance than processors running at full speed. Operating systems (OS) have mechanisms that can exploit DFS to achieve energy savings. For instance, OS-level frequency governors throttle the server CPU speed accordingly to its utilization. However, the OS misses domain-specific information about the search engine application and the incoming queries.

We advocate that a more refined CPU power management is possible, knowing the search server utilization and load. In (Catena et al. 2015) we propose search engine-specific frequency governors, that manage the processor speed from within the search server application. These governors increase the CPU speed whenever the search server is struggling with processing incoming queries. Similarly, CPU speed is decreased when the search server is easily processing the arriving requests.

We conduct extensive experimentation upon the TREC ClueWeb09 corpus and the MSN 2006 query stream, to evaluate the benefits and drawbacks of our approach compared to standard OS-level frequency governors. Results show that our solution can absorb ~24% less power than a system which operates at maximum CPU frequency, with only a limited detriment in query processing quality. When compared to more energy efficient OS configurations, we find that our governors can still save at least 7% in power absorption. Such energy savings are important for data centers, as reduced processor frequencies reduces heat output and thermal cooling cost. Greater energy savings can be obtained by allowing more substantial degradation in query processing quality.

3. CONCLUSIONS AND FUTURE WORK

In this work, we illustrate our first results in evaluating and improving the energy efficiency of Web search engines. Future work will continue in this direction,

for instance in evaluating the energy expenditure of other search engine components (e.g., query expansion/reformulation, machine-learned document re-ordering, snippets generation, etc.). We also believe there is still space to further improve the CPUs power management in Web search engines data centers.

Up to this point, our work focuses on single search servers but we also wish to evaluate and improve energy efficiency at the intra-data center level, i.e., on search server clusters. Similarly, we would like to reduce energy consumption at the inter-data center level, i.e., on geographically distributed data centers owned by the same search company.

Many aspects besides hardware power management could be worth exploring. For example, one could try to understand if it is possible to trade search engines effectiveness (recall, MAP, NDCG, etc.) for energy efficiency; or what is the relationship between energy savings and corpus size; etc.

Finally, we must observe that search results are increasingly often consumed from mobile platforms (such as smartphones, tablets, etc.) with limited battery life. For this reason, future work should also promote energy efficient interactions between mobile clients and Web search engines.

4. ACKNOWLEDGMENT

The author thanks ACM SIGIR for the awarded scholarship for participating in ESSIR 2015 and the FDIA Symposium.

REFERENCES

- Barroso, L.A., Dean, J., and Hoelzle, U. (2003) Web Search for a Planet: The Google Cluster Architecture. *IEEE Micro*, 23(2), 22–28.
- Catena, M., Macdonald, C., and Tonellotto N. (2015) Load-sensitive CPU Power Management for Web Search Engines. To appear in: *Proc. of SIGIR*. ACM.
- Catena, M., and Tonellotto, N. (2015) A Study on Query Energy Consumption in Web Search Engines. In: *Proc. of IIR*. CEUR-WS, <http://ceur-ws.org/Vol-1404/>.
- Chowdhury, G. (2012) An agenda for green information retrieval research. *Information Processing & Management*, 48(6), 1067–1077.
- Sazoglu, F.B., Cambazoglu, B.B., Ozcan, R., Altinogovde, I.S., and Ulusoy, O. (2013) A Financial Cost Metric for Result Caching. In: *Proc. of SIGIR*. ACM, 873–876.
- Snowdon, D.C, Ruocco, S., and Heiser, G. (2005) Power Management and Dynamic Voltage Scaling: Myths and Facts. In: *Proc. of PARC Workshop*.
- The Climate Group for GeSI (2008) *Smart 2020: Enabling the low carbon economy in the information age*. Available from: <http://goo.gl/413vun> (16 June 2015).