International Journal of Science Engineering and Advance Technology
IJSEAT

Improve Query Response Time and Reduce CPU Cost In Web Search

[1]N.Devi,[2]P.Radhika Krupalini
[1]Final Master of Science in Computer Science, Ideal college of Arts and Sciences., Vidyuth Nagar,Kakinada,E.G.Dist.,A.P., India.
[2]Associate professor, Dept of Computer science, Ideal college of Arts and Sciences., Vidyuth Nagar,Kakinada,E.G.Dist.,A.P., India.

ABSTRACT:
We suggest the Predictive Energy Saving Online Scheduling Algorithm (PESOS) to choice the greatest suitable CPU frequency to procedure a query on a per-core basis. PESOS goal at procedure queries by their limits, and leverage high-level preparation info to decrease the CPU energy ingesting of a query dispensation node. PESOS base its result on inquiry efficacy predictors, guessing the meting out volume and meting out time of a query. We experimentally gauge PESOS upon the TREC ClueWeb09B collection and the MSN2006 query log. PESOS outpace also the best state-of-the-art entrant with a 20% oomph saving, while the player requires a fine restriction fine-tuning and it may invite in wild potential abuses.

KEYWORDS: Relocation, Query, global warming

1 INTRODUCTION

Web search engines are characteristically calm by thousands of these nodes, held in big datacenters which also comprise infrastructures for telecommunication, thermal cooling, fire suppression, power supply, etc at the same time, such many servers consume a important amount of energy, delaying the success of the search engines and rising ecological concerns. In fact, datacenters can consume tens of megawatts of electric power and the linked outflow can beat the innovative investment cost for a datacenter. Since of their energy consumption, datacenters are in control for the 14% of the ICT sector carbon dioxide emissions which are the focal cause of global warming. For this reason, governments are helping codes of deportment and best practices to condense the environmental power of datacenters. Then energy consumption has a key role on the lucrativeness and eco-friendly impact of Web search engines, civilizing their oomphproductivity is an chiefpart.

2LITERATURE SURVEY

2.1 Consuming Web search as a illustrative instance of this workload class, we first describe a manufacture Web search workload at cluster-wide scale. We deliver a fine-grain classification and representation the chance for power savings using low-power modes of each chief server component. Next, we advance and authenticate a presentation model to assess the impact of processor- and memory-based low-power modes on the exploration in expression distribution and deliberate the profit of current and likely low-power modes.

2.2we present PEGASUS, a feedback-based manager that meaning fully recovers the energy proportionality of WSC systems, as established by a real application in a Google search cluster. PEGASUS uses application latency statistics to energetically adjust server power running restrictions in a fine-grain manner, seriatim each server just fast adequate to happen global service-level inactivity aims. In large cluster experiments, PEGASUS decreases power ingesting by up to 20%.

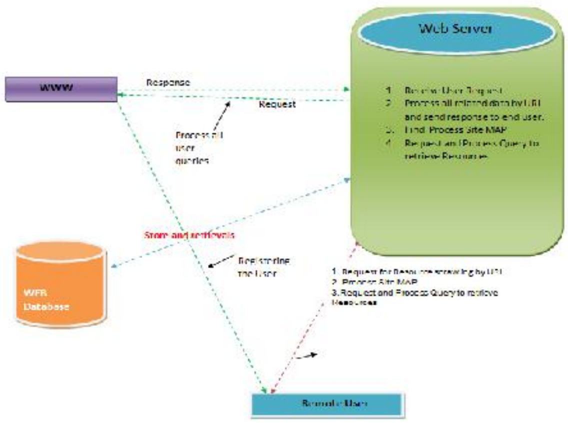
3 PROBLEM DEFINTION

Multi-site Web search engines, i.e., search engines unruffled by multiple and purely distant are datacenters. These studies suggest to use query forwarding, i.e., to change the query workload between datacenters. They target to minimalize the energy expenditure of the search engine. At the same time, the methods a guards that the remote sites can procedure forwarded queries deprived of exceeding their meting outvolume.

4 PROPOSED APPROACH

The scheme suggests the Predictive Energy Saving Online Scheduling algorithm (PESOS), which reflects the tail dormancy obligation of queries as a clear parameter. Via the DVFS technology, PESOS selects the most suitable CPU incidence to procedure a query on a per-core basis, so that the CPU energy ingesting is abridged while regarding obligatory tail latency. The algorithm bases its choice on query competence predictors rather than core use. Query efficiency predictors are methods to estimation the meting out time of a query before its meting out

5 SYSTEM ARCHITECTURE



6 PROPOSED METHODOLOGY

Request and Process Query to retrieve Resources

Relocationcovers at least the identifier of the text where the periodseems and its periodincidence, i.e., the number of incidences of the term in that specific document. The upturned index is typicallybeaten and kept in main memory to upsurge the presentation of the search engine.

Query efficiency predictors

Query efficiency predictors (QEPs) are methods that approximation the implementation time of a query before it is reallytreated. Theimplementation time of inquirieslicenses to recover the presentation of a search engine. Most QEPs feat the physiognomies of the query and the upturned index to pre-compute features to be brow beatenapproximate the query dispensation times.

Predicting processing volumes

If thoroughdispensation is done, it is likely to know a priori the number of scored postings, which is identical to the sum of the relocation, lists lengths of the query terms. But, when activetrimming is useful we do not know in early payment how much relocation will be scored, since slices of the relocation lists might be pranced. Then, we need a way to expect the number of keep countposition for a query.

7PREDICTIVE ENERGY SAVING ONLINE SCHEDULING ALGORITHM

INPUT:set of jobs ,time interval

STEP1:It works by analyzing each possible time interval *I* included in [*t*0, *t*1].

STEP2: it finds the *critical interval I_* that maximizes processor speed.

STEP3: It schedules the jobs in *JI_* using the *earliest deadline first* policy and processing speed.

STEP4: if not preempted, the jobs in *JI_* will terminate in time units since the beginning of their execution.

STEP5: Jobs in *JI_* are then removed from *J*. The interval *I_* as well is removed from [*t*0, *t*1],

STEP6:It repeatedly finds a new critical interval for the remaining jobs, until all jobs are eventually scheduled.

8RESULTS

Improve Query Response Time And Reduce CPU Cost In Web Search Index

THE HINDU

Improve Query Response Time And Reduce CPU Cost In Web Search Binary Details

Improve Query Response Time And Reduce CPU Cost In Web Search Pseudoinfo

EXTENSION WORK

Propose a framework to discover different user search goals for a query by clustering the proposed feedback sessions. Feedback sessions are constructed from user click-through logs and can efficiently reflect the information needs of users

9CONCLUSION

We experimentally appraised the recital of PESOS by means of the ClueWeb09B corpus and meting out queries from the MSN2006 log spread on two unlike dynamic pruning recovery strategies: Max Score and WAND. We linked the performance of PESOS with those of three baselines: perf, which always habits the thoroughgoing CPU core incidence, power, which throttles incidencesaffording to the core exploitations, and cons, which checksoccurrencesafording to the consumption of the interrogation servers.

10REFERENCES

[1] L. A. Barroso, J. Clidaras, and U. H¨olzle, The Datacenter as aComputer: An Introduction to the Design of Warehouse-ScaleMachines, 2nd ed. Morgan & Claypool Publishers, 2013.

[2] I. Arapakis, X. Bai, and B. B. Cambazoglu, “Impact of responselatency on user behavior in web search,” in Proc. SIGIR, 2014,pp. 103–112.

[3] U.S. Department of Energy, “Quick start guide to increasedata center energy efficiency,” 2009. [Online]. Available:http://goo.gl/ovDP26

[4] The Climate Group for the Global e-Sustainability Initiative,“Smart 2020: Enabling the low carbon economy in theinformation age,” 2008. [Online]. Available: http://goo.gl/w5gMXa

[5] European Commission – Joint Research Centre, “The EuropeanCode of Conduct for Energy Efficiency in Data Centre.” [Online].Available: http://goo.gl/wmqYLQ

[6] U.S. Department of Energy, “Best Practices Guide forEnergy-Efficient Data Center Design.” [Online]. Available:http://goo.gl/pikFFv

[7] D. C. Snowdon, S. Ruocco, and G. Heiser, “Power Managementand Dynamic Voltage Scaling: Myths and Facts,” in Proc. OfWorkshop on Power Aware Real-time Computing, 2005.

[8] The Linux Kernel Archives, “Intel P-State driver.”[Online].Available: https://goo.gl/w9JyBa

[9] D. Brodowski, “CPU frequency and voltage scaling code in theLinux kernel.” [Online]. Available: https://goo.gl/QSkft2

[10] C. Macdonald, N. Tonellotto, and I. Ounis, “Learning to predictresponse times for online query scheduling,” in Proc. SIGIR,2012, pp. 621–630.

[11] M. Jeon, S. Kim, S.-w. Hwang, Y. He, S. Elnikety, A. L. Cox,and S. Rixner, “Predictive parallelization: Taming tail latenciesin web search,” in Proc. SIGIR, 2014, pp. 253–262.

[12] Matteo Catena and Nicola Tonellotto, Energy-Efficient Query processing in Web Search Engines,2017

N. Devi is student of Computer Science Department at IDEAL COLLEGE OF ARTS AND SCIENCES, KAKINADA, AP, INDIA. Presently She is pursuing her final semester in Master of Science in Computer Science from Ideal college and She received her B.Sc(CS) from V.S. Lakshmi womens degree college,kakinada in the year 2016.

P. Radhika Krupalini is an Associate Professor of Computer Science Department at IDEAL COLLEGE OF ARTS AND SCIENCES, KAKINADA AP, INDIA. She passed M.Tech in Computer Science & Engineering from UCE, JNTU, Kakinada. She has Lectureship in Computer Science and Applications discipline and has an experience of 12 years of teaching.