Survey on Energy Efficient Algorithms

MOHAMMED WASIM R D
BTECH IN COMPUTER SCIENCE,
PRESIDENCY UNIVERSITY,BANGALORE,INDIA

Abstract — ***Conservation of energy is one the main problem nowadays.Many industries are seeking for energy efficient products as the cost of energy to run their equipment has become one of the major concern.In terms of both cost and availability energy consumption is critical.As we can see when we use monitors, laptops they consume more energy while using and battery is reduced where we can also see it includes power management.To overcome these all problem energy efficient algorithms are introduced in various fields such as parallel computing,data mining,grid computing,cloud computing and so on.The purpose of the survey is to link all the energy efficient algorithms with respect to the data mining algorithms for handling large datasets.***

Index Terms — ***Energy efficient algorithm, cloud computing, data mining algorithm, grid computing, dataset evaluation***

I. INTRODUCTION

We know that energy has become one of the expensive resource which is true with respect many of the computing environments.Due the consumption of more energy power management is one of the critical problem that occurs in portable devices such as laptops,mobile where the capacity of the battery is reduced.Furthermore, electricity costs impose a substantial strain on the budget of many computing and data centers so that effective power management strategies are required .To solve this energy consumption and power management problem many algorithms have been introduced such as power down mechanisms,dynamic speed scaling and networking.

In power down mechanisms the main goal is if the system is ineffective for certain time then it can be transitioned to sleep mode.Main objective is to find transition schedules that minimizes the overall consumption of energy.

In dynamic speed scaling the problem is many processors run at irregular speed and frequency.If there is high speed then there is high performance and also the more energy consumption.The main objective is to utilize the high speed of processor and apply it to the low speed.

In networking the main goal is to figure out different data transmission and routing problems,to minimize the total energy consumed.

In these way many energy efficient algorithms are built for consumption of both energy and power management.

II.ALGORITHMS FOR ENERGY EFFICIENCY OF CLOUD COMPUTING

The requirement of energy efficiency for cloud computing is becoming higher these days.One of the main reason is rising of electricity consumption due to the utilization of virtualization and cloud computing techniques[1].More energy is consumed by the data center in large scale due to the rapid development of cloud computing [2].CO2 emission from data centers have led to the global warming and power consumption of data centers should also be reduced,hence Green Cloud Computing was introduced. [3].

An average data center power consuming is such a high that it serves as a power resource for around 20,000 homes which can be seen in Gartner report [4].At present cloud computing is providing different services such as Infrastructure as a service (IaaS), Platform as a service (PaaS) and Software as a service (SaaS),due to this there is large number of increasing demand for servers and many businesses have shifted from traditional to online.Some of the modern day major cloud providers are Amazon Web Services , Microsoft Azure and Google App Engine [5-7]. Due to this new data centers are coming into existence every year which consumes more power.Since data centers are consuming more power they will require cooling requirements which leads to more emission of CO2[8].To diminish the power consumption energy efficient cloud computing is required[9].

There are many energy efficient cloud computing algorithms. Some of them are discussed below.

Algorithm for exact virtual machine allocation:
The above algorithm comprises of a way known as Bin-Packing which extends its valid conditions as equalities and constraints in it [1].The power consumption says that Virtual machines are packed into servers.Limitization of server migration and capacity of servers with incuding of valid conditions is the main role of this algorithm [10].

Algorithm for exact virtual machine migration:
This algorithm describes the optimization of Virtual machines migration[1].To balance the load within the virtual machine capacity limits,there is need to fill the place of virtual machine optimally with the other virtual machine that are running on another servers. Migrating the Virtual machine from one server to other is that the servers that have less virtual machine working on it should be made active by optimally migrating to other server that will lead to saving of energy,which is the main objective of this algorithm.

Combination of allocation and migration algorithms:
The above algorithm describes the energy saving when both migration and allocation are merged[1]. The resources that are free after completion of jobs are used by this algorithms which are merged to optimally migrate virtual machine .

Minimizing the use of parallel machine by grouping virtual machine and parallel machine:
Minimization and grouping of virtual and parallel machine is done in this algorithm.Based on the need of execution

time and cores virtual machines are grouped in decreasing order.In the same way parallel machine are grouped according to the number of virtual machine running on them.

Migration algorithm for energy awareness:

In this algorithm for the purpose of saving energy migration of virtual machine is done optimally.[12].This algorithm involves three steps and they are as follows:

(a) Selection of Victim:

During selection of victim to save energy[12] the servers should be switched off and those servers are founded in this. Servers below the threshold value as power of threshold are switched off.Afer the completion of the task by virtual machine the load on server decreases and whenever the load is less on their corresponding threshold value becomes less than power of threshold and migration of virtual machine to other servers is happened during this time so that the server victim is off and energy can be saved.[10]

(b) Selection of Target Server:

In this the it includes the selection the target servers that are needed for virtual machine to place from victim server. Based on first-fit allocation method target server is selected.[10]

(c)Switching on Server:

In switching on server the server which is in sleep mode is waken up.. When the threshold value of a server is above the Wake up Threshold then the server that are in switch off mode are woken up so that load from the Servers whose value of threshold is reached to Wake up Threshold is transfer to newly waken server.

III. ALGORITHMS FOR ENERGY EFFICIENCY OF GRID COMPUTING

In recent days grids [13] have been emerged as a global cyber-infrastructure in the field of e-Science and e-business applications.Middleware grid and various management tools such as Globus [14], UNICORE [15], Legion [16] and Gridbus [17] have been developed, so that it allows users to approach remote resources transparently over a secure scalable world-wide network. Service-oriented paradigm,the progress of grid computing [18,19] defines a new way of provisioning services based on various computing models in which grid contains each resources and consumer can bargain the usage and service quality.

To manage and process large datasets many research areas such as bio informatics,physics,astronomy are adopting grids. In return of it which grid supports complex experiments with large resources such as computational devices, data, applications, to manage the workflow within the grid environment [20]. Automation resources define workflow of grids whereas data and various files that are passed between the thing that takes part in activity to achieve main goal [21].

The workflow system of grid have the capacity for creating,managing and executing different grid applications with the higher efficiency.Since the workflow system of grid is highly effective,diversified and distribution in nature it is very difficult to solve the problem in grid environment. One of the scheduling algorithm for grid workflow is Directed Acyclic Graph which provides various resources during execution and it also satisfies need of the user. These diversified resources belong to set of problems known as NP-Hard problems [22].Though it is difficult to solve the problem in grid environment many heuristic and meta heuristics algorithms related to workflow environment have been introduced for fulfilment of grid environment.

Scheduling algorithms	Algoritm	Type of scheduling	Year of publication	Objective
Swarm algorithm(Brain storm) [23]	Multiobjective brain storm algorithm	Meta Heuristic	2014	It reduces execution time and economic cost
Game Theory Approach[24]	Genetic and Random Algorithm	Meta Heruristic or Heruristic	2013	It reduces Time and cost
Ant Colony with intelligent water drop[25]	Ant Colony algorithm and Intelligent water drop algorithm	Meta Heuristic	2013	It reduces makespan, Improve system utilization and load balancing
Deadline satisfactory degree of workflow[26]	Deadline satisfaction enhanced scheduling algorithm	Heuristic	2011	It reduces total execution time
Task with maximum completion time assign to fastest machine[27]	Efficient algorithm on hetrogenous computing system	Heuristic	2011	Improved makespan, Improve system utilization and load balancing
Budget Constraints [28]	Reliability enhanced grid workflow scheduling algorithm with budget constraint	QoS based Heuristic	2011	Better reliability and adaptability.
HGreen Heuristic[29]	HGreen Heuristic algorithm and energy efficient model	Heuristic	2011	It reduces power consumption
A-star algorithm[30]	Data aware workflow scheduling algorithm based on a-star	Heuristic	2011	It reduces turnaround time

IV. ALGORITHMS FOR ENERGY EFFICIENCY OF PARALLEL COMPUTING

[1] Divide and conquer algorithm:

Given a problem it splits the main problem into subproblems which is the easiest way to solve original problem and again it solves the subproblems and combines the solution with the subproblem to generate solution for the original problem[31].It plays a vital role in improvement of modularity of program and has usually lead to easy and efficient algorithms.For sequential algorithm designers,it is one of the best tool.

[2] Randomization algorithm

It is one of the parallel algorithm which uses random numbers to verify that various processors can make selection with more feasibility and good decisions.Some of the levels of the randomness are as follows,

a] Sampling randomness[31]

Given set of elements a sample is being selected.Solve the problem based on that sample and using solution for sample,guide the solution for the original set.This is the function of sampling.It represents the barrier of intervals.It is used in fields such as geometry,graph etc.

.

b] Symmetric Breaking randomness[32]

Given a problem select a graph which contains large independent vertices in parallel I.e vertices are said to be independent if no two vertices are neighbours.Each vertex should decide with all other vertices to join the set or not.The condition is like this,if one vertex joins the set then all the other vertices should not join the set.The choice is difficult but if each vertex has same number of neighbours it can be dealt by using randomness so that the symmetry between the vertices is broken.

c] Load Balancing randomness [31]

Separation of large number of data items intoaccumulation of evenly sized subset.Randomly assign each element to subset.When the average size of subset is at leat logarithmic in size of the original set,this technique works best.

[3] Parallel pointer techniques algorithm:

Parallel techniques cannot be easily translated by sequential ways of managing trees,graph and lists.Many techniques such as linked list traverse,postorderin visiting the nodes of tree, depth-first traversal of a graph appear to be intrinsically subsequent which are being replaced by the parallel techniques with same power.

Some of the parallel pointer techniques are mentioned above:

a] Pointer jumping parallel technique [33]

This is one of the oldest techniques involved in parallel pointer techniques.This can be used with either tree or list. During the process the each step involved in pointer jumping , each node in parallel is replaced by its parent or one of its successor.Labelling each node of an n-node tree (or list) with respect to labelling of last root sis to use pointer jumping.

b] Euler Tour parallel technique [34]

It is directed graph path where every edge is traversed exactly once. Two oppositely directed edges are replaced in case of undirected graph in euler tour which is followed by the perimeter of the tree visiting each edge twice, once on the way down and once on the way up.By this we can determine the size of each subtree. Parallel depth that is independent of the depth of the tree is the technique used here.It can be used to replace depth first traversal.

c] Graph contraction parallel tehnique[35,36]

In the above parallel pointer technique the size of the graph is reduced to maintain its original structure.After performing graph contraction the solution obtained is kept to form the final solution of the contraction. It can showed as,first contracting the graph by grouping some of the vertices with neighboring vertices, then finding the connected components of the contracted graph, and finally delete the contraction operation.One of the technique for contracting tree is known as tree contracting.

d] Ear decomposition parallel technique [37,38]

It can be defined as separation of its edges into an ordered collection of paths. First path is known as the cycle and the remaining are called ears. If we found an ear decomposition of a graph , it is not difficult to decide if two edges lie on a common cycle. The above information can be used in algorithms for determining biconnectivity, triconnectivity,planarity. The above algorithm can be found in linear work and logarithmic depth, independent of the structure of the graph. This technique can also be used to solve depth first search.

V. ALGORITHMS FOR HANDLING DATASETS

The below mentioned are the different algorithms used in data mining for the purpose of handling large datasets.

Rtree algorithm used for

a] Indexing multi-dimensional information.
b] Handling geospatial coordinates.
c] Implementation of virtual maps.
d] Handling game data.

Nearest neighbour search technique :

Finding the point in a given set that is closest to a given point I.e Closeness.

Decison tree technique:

Decision tree algorithm falls under the category of supervised learning. They can be used to solve both regression and classification problems.

Decision tree uses the tree representation to solve the problem in which each leaf node corresponds to a class label and attributes are represented on the internal node of the tree.

Genetic algorithm:

Genetic Algorithms are adaptive heuristic search algorithms that belong to the larger part of evolutionary algorithms.They are commonly used to generate

high-quality solutions for optimization problems and search problems.

AUTHOR NAME	TECHNIQUE	CHARACTERISTIC	SEARCH TIME
N. Beckmann, H. P.Kriegal, R. Schneider, B.Seeger[39]	R-Tree and R*_Tree technique	It has performance bottleneck	$O(3^D)$
S. Arya, D. Mount, O. Netanyahu, R. Silverman, A.Wu[40]	Nearest Neighbour Search technique	It is expensive when object is in High dimensional space	It grows exponentially with the size of searching space. O(dnlogn)
Lawrence O Hall, Nitesh Chawla, Kevin W Bowyer [41]	Decision Tree Learning technique	It is reasonably fast and accurate	It is Less time consuming
Zhiwei Fu, Fannie Mae[42]	Decision Tree C4.5 technique	Practice local greedy search throughout dataset	It is Less time consuming
D.V.Patil, R.S.Bichkar[43]	GA Tree(Decision Tree + Genetic Algorithm) technique	Improvement, in classification, performance, and reduction in size of tree,with no loss in classification accuracy	Improved Performance and Problems like slow memory execution can be reduced
Yen-Ling Lu, Chin-Shyurng Fahan[44]	Hierarchical Neural Network technique	It requires High accuracy rate of recognizing data; have high classification accuracy	It is Less time consuming improved performance

REFERENCES

[1] Chaima Ghribi, Makhlof Hadiji and Djamal Zeghlache, "Energy efficient VM scheduling for cloud data centers: exact allocation and migration algorithms", in 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid) pp.671-678, may 2013

[2] Ye, D. Huang, X. Jiang, H. Chen, and S. Wu, "Virtual machine based energy-efficient data center architecture for cloud computing: a performance perspective", in Proceedings of the IEEE/ACM International Conference on Green Computing and Communications, pp. 171 – 178, Dec. 2010

[3] Anubha jain, Manoj Mishra, Sateesh Kumar Peddoju and Nitin jain, "Energy efficient comuting-green cloud computing", International Conference on Energy Efficient Technologies for Sustainability (ICEETS), pp. 978-982, April 2013.

[4] Gartner report, financial times, 2007.

[5]AmazonElasticComputeCloud,"' http://aws.amazon.com'ec2

[6]WindowsAzurePlatform http://www.microsoft.com.azure/ (March 17,2010)

[7] Google App Engine, http://appengine.google.com(March 17, 2010)

[8] Ching-Chi Lin,Pangfeng Liu and Jan-Jan Wu, " Energy-aware virtual machine dynamic provision and scheduling for cloud computing", IEEE International Conference on Cloud Computing, pp. 736-737, july 2011.

[9] R.Yamini, "Power management in cloud comuting using green algorithm", IEEE-International Conference On Advances In Engineering,Science And Management(ICAESM), pp. 128-133, March 2012.

[10] Anuj Prasher , Rajkumari Bhatia "A Review On Energy Efficient Cloud Computing Algorithms Research Scholar UIET Panjab University, Chandigarh, India ,Assistant Professor, UIET, Panjab university, Chandigarh, India.4, April 2014

[11] Cesar O. Diaz, Harold Castro, Mario Villamizar, Johnatan E. Pecero and Pascal Bouvry, " Energy-aware VM allocation on an opportunistic cloud infrastructure", in 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid), pp. 663-670, may 2013.

[12] Mohammad H.AL Shyaeji and M.D.Samrajesh, "An energy-aware virtual machine algorithm", International Conference on Advances in Computing and Communication(ICACC), pp. 242-246, Aug. 2012

[13] Foster, I., Kesselman, C. (eds.): The Grid: Blueprint for a Future Computing Infrastructure. Morgan Kaufmann Publishers, USA (1999)
[14] Foster, I., Kesselman, C.: Globus: A Metacomputing Infrastructure Toolkit. International Journal of Supercomputer Applications 11(2), 115 – 128 (1997)
[15] Almond, J., Snelling, D.: UNICORE: Uniform Access to Supercomputing as an Element of Electronic Commerce. Future Generation Computer Systems 15, 539 – 548 (1999)
[16] Grimshaw, A., Wulf, W.: The Legion vision of a worldwide virtual computer. Communications of the ACM 40(1), 39 – 45 (1997)
[17] Buyya, R., Venugopal, S.: The Gridbus Toolkit for Service Oriented Grid and UtilityComputing: An overviewand StatusReport. In:The 1st IEEEInternational Workshop on Grid Economics and Business Models, GECON 2004, Seoul, Korea, April 23 (2004)
[18] Benkner, S., et al.: VGE - A Service-Oriented Grid Environment for On-Demand Supercomputing. In: The 5th IEEE/ACM International Workshop on Grid Computing (Grid 2004), Pittsburgh, PA, USA (November 2004)
[19] Foster, I., et al.: The Physiology of the Grid, Open Grid Service Infrastructure WG. In: Global Grid Forum (2002)
[20] Mayer, A., et al.: Workflow Expression: Comparison of Spatial and Temporal Approaches.In:Work fl owinGrid SystemsWorkshop,GGF-10, Berlin,March 9(2004)
[21] Hollinsworth, D.: The Workflow Reference Model, Workflow Management Coalition, TC00-1003 (1994)
[22] J. Yu, R. Buyya, and K. Ramamohanarao. "Workflow scheduling algorithms for grid computing." Metaheuristics For Scheduling in Distributed Computing Environments, Springer Berlin Heidelberg, pp. 173-214, 2008.
[23] M. Arsuaga-Rios, and M. A. Vega-Rodriguez. "Cost optimization based on brain storming for grid scheduling." in IEEE 4th International Conference on Innovative Computing Technology, pp. 31-36, 2014.
[24] M. Yaghoobi, A. Fanian, H. Khajemohammadi, and T. Aaron Gulliver. "A non-cooperative game theory approach to optimize workflow scheduling in grid computing." in 2013 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, pp. 108-113, 2013.
[25] P. Mathiyalagan, S. N. Sivanandam, and K. S. Saranya. "Hybridization of modified ant colony optimization and intelligent water drops algorithm for job scheduling in computational grid." ICTACT Journal on Soft Computing, vol. 4, no. 1, 2013.
[26] X. Li, Z. Hu, and C. Yan. "A deadline satisfaction enhanced workflow scheduling algorithm." in 19th Euromicro International Conference on Parallel, Distributed and Network-Based Processing, pp. 55-61, 2011.
[27] S. Bansal, and C. Hota. "Efficient algorithm on heterogeneous computing system." in International Conference on Recent Trends in Information Systems, pp. 57-61, 2011.
[28] C. Yan, Z. Hu, X. Li, Z. Hu, and P. Xiao. "Reliability enhanced grid workflow scheduling algorithm with budget constraint." in IEEE International Conference on Information and Automation, pp. 629-633, 2011.
[29] F. Coutinho, L. Alfredo V. de Carvalho, and R. Santana. "A workflow scheduling algorithm for optimizing energy-efficient grid resources usage." in IEEE 9th International Conference on Dependable, Autonomic and Secure Computing, pp. 642-649, 2011.
[30] D. Yin, and T. Kosar. "A data-aware workflow scheduling algorithm for heterogeneous distributed systems." in International Conference on High Performance Computing and Simulation, pp. 114-120,2011

[31]Parallel Algorithms,Guy E. Blelloch and Bruce M. Maggs School of Computer Science Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15213
[32] M. Luby. A simple parallel algorithm for the maximal independent set problem. SIAM Journalon Computing, 15(4):1036 – 1053, Nov. 1986.
[33] J. C. Wyllie. The complexity of parallel computations. Technical Report TR-79-387, Depart
ment of Computer Science, Cornell University, Ithaca, NY, Aug. 1979
[34] R. E. Tarjan and U. Vishkin. An efficient parallel biconnectivity algorithm. SIAM Journal of
Computing, 14(4):862–874, Nov. 1985.
[35] G. L. Miller and J. H. Reif. Parallel tree contraction part 1: Fundamentals. In S. Micali,
editor, Randomness and Computation. Volume 5 of Advances in Computing Research, pages
47–72. JAI Press, Greenwich, CT, 1989.
[36] G. L. Miller and J. H. Reif. Parallel tree contraction part 2: Further applications. SIAM
Journal of Computing, 20(6):1128–1147, Dec. 1991.
[37] Y. Maon, B. Schieber, and U. Vishkin. Parallel ear decomposition search (EDS) and st
numbering in graphs. Theoretical Comput. Sci., 47:277–298, 1986
[38] G. L. Miller and V. Ramachandran. A new graph triconnectivity algorithm and its paralleliza
tion. Combinatorica, 12(1):53–76, 1992.
[39] N. Beckmann, H. -P. Kriegal, R. Schneider, and B. Seeger, "The R*-Tree: An Efficient and Robust Access Method for Points and Rectangles," Proc. ACM SIGMOD, May 1990
[40] S. Arya, D. Mount, N. Netanyahu, R. Silverman, A. Wu, "An Optimal Algorithm for Approximate Nearest Neighbor Searching in Fixed Dimensions, " Proc. Fifth Symp. Discrete Algorithm (SODA), 1994, pp. 573-582.
[41] Lawrence 0. Hall, Nitesh Chawla , Kevin W. Bowyer, "Decision Tree Learning on Very Large Data Sets", IEEE, Oct 1998
[42] Zhiwei Fu, Fannie Mae, "A Computational Study of Using Genetic Algorithms to Develop Intelligent Decision Trees", Proceedings of the 2001 IEEE congress on evolutionary computation, 2001.
[43] Mr. D. V. Patil, Prof. Dr. R. S. Bichkar, "A Hybrid Evolutionary Approach To Construct Optimal Decision Trees with Large Data Sets", IEEE, 2006 [
[44] Yen-ling Lu, chin-shyurng fahn, "Hierarchical Artificial Neural Networks For Recognizing High Similar Large Data Sets. ", Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, August 2007 [