I J C T A, 9(37) 2016, pp. 145-149

A Comprehensive View of Energy Efficient Mr Scheduling Algorithms in Hadoop Clusters of Data Centers

M. Ramla* and M. Ramesh**

Abstract: We live in an era of unprecedented flow of data. A data center is a dedicated space with ICT infrastructure set up by companies to support their business. Most of the large scale data intensive applications run in data centers as Map Reduce jobs using the open source implementation Hadoop. There are numerous Hadoop scheduling algorithms with different objectives that reduce the make span of the MR jobs. However, reducing the make span is not the only target, as energy footprints is the paramount concern in datacenters. Along with the energy management policies used in data centers, there are also significant optimization opportunities within the MR framework agreeing with the Service Level Agreement (SLA). To realize this, we make a detailed survey of some of the proposed techniques to improve the MR cluster energy efficiency. These heuristic optimization techniques provide considerable energy savings with tradeoffs between performance and availability. We also organize the studied techniques based on the taxonomy and conclude with possible directions for future research.

Index Terms: Hadoop, MR, Energy-aware, data centers

1. INTRODUCTION

Hadoop is a popular open source implementation of map reduce. Map Reduce is a programming framework for data intensive computing on large scale distributed systems [1]. Nowadays Map reduce is used in many business and educational data centers [2]. Data centers contain Hadoop MR clusters to process both batch and interactive Big data jobs. The report from Environment Protection Agency states that the second highest cost in data centers after labour cost is the electricity cost [7]. The increase in the size of data centers has led to increase in energy consumption. Trends indicate that the data centers consume upto 100 billion KWh per year in US [3]. Researchers are turning their attention towards Green Data centers. Energy savings for data centers and computational grids is the need of the hour.

Simple energy management policies like redirection and server shutdown techniques will prove good only for workloads which is not data intensive [9]. Also scaling down of nodes will cause unavailability of data. It is therefore necessary to involve the underlying file distribution and programming frameworks for energy efficiency. Applications must be scheduled by taking into consideration the power and performance characteristics of cluster while ensuring the service level agreement [7]. This will greatly reduce the energy costs. This paper is a comprehensive study of various energy aware techniques proposed for Map Reduce.

2. BACKGROUND

Hadoop is Java based programming model for large data set processing in distributed environment sponsored by Apache Software foundation [8]. It provides much needed robustness and scalability options to a distributed system. It is fault tolerant and can be deployed on low cost hardware. The storage system is not physically separated from processing system. MR is the programming model for processing large data sets and HDFS is used to stream those large data sets. Hadoop Distributed File System is Hadoop's

* Department of Computer Applications, SRM University, Chennai, India. *Email: ramla.m@ktr.srmuniv.ac.in*
** Department of Computer Science, SRM University, Chennai, India. *Email: ramesh.ma@ktr.srmuniv.ac.in*

implementation designed to hold a large amount of data and provide access to data to many clients across network. It comprises of two nodes- Data node for storing data and name node (master node) for monitoring data nodes [8]. Scheduling decisions are taken by Master nodes called Job trackers and the slave nodes called Task Trackers execute the tasks. HDFS is resilence, fault tolerant and it also minimizes disk seeks. The scheduler which is the center piece of Hadoop decides the scheduling of tasks to reduce the makespan of the job. But minimizing the energy consumption in data centers is also a critical concern.

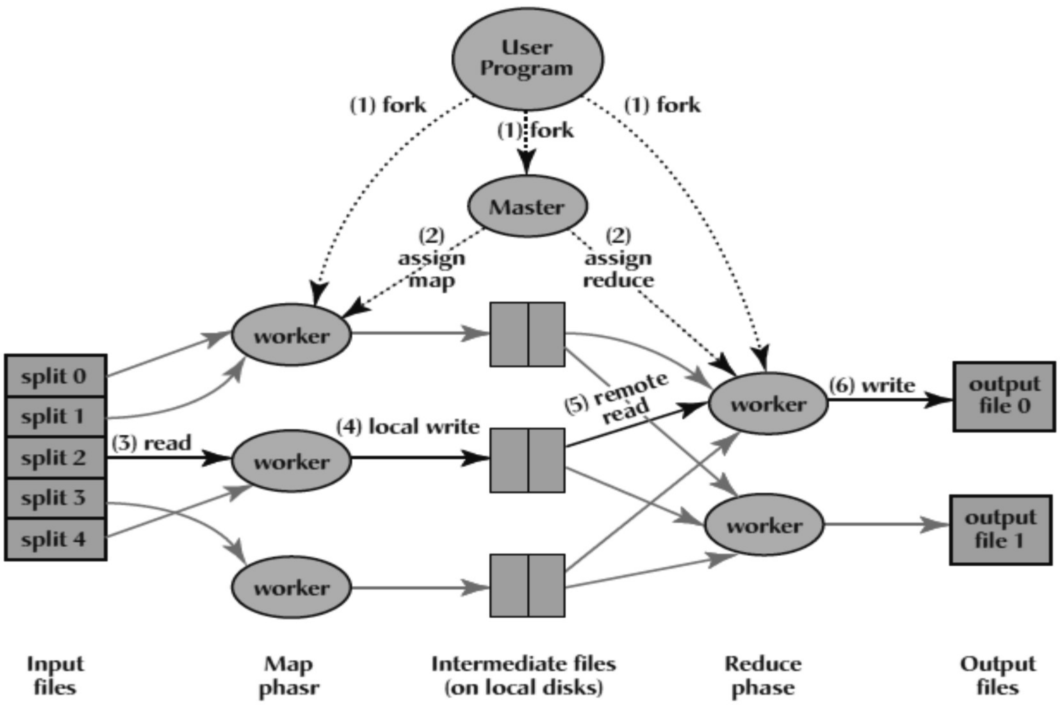

Figure 1:Map Reduce Archietecture

3. CLASSIFICATION

More intelligent MR techniques are required for improving the MR clusters energy efficiency. The techniques can be broadly classified as:

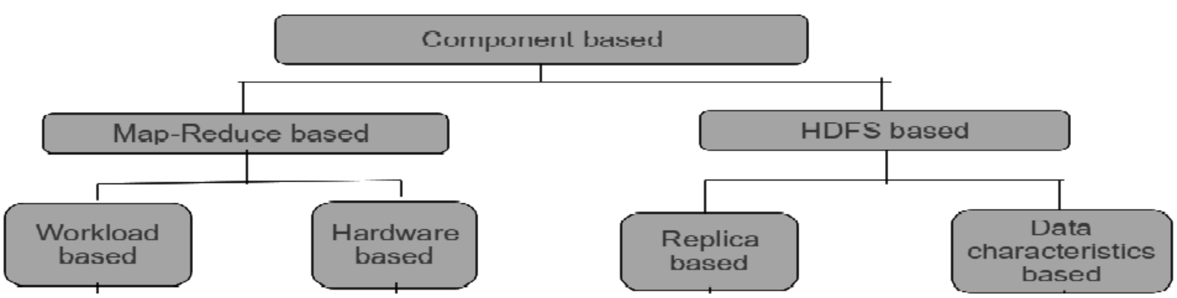

Figure 2: Classification of Energy Efficient Techniques

A. MR programming model modification techniques

1. *Workload Energy Aware Scheduling:* The central theme of this scheduling technique is to run jobs on all nodes and power down all nodes when no work [9]. The time taken to transit from one power state to another would have significant impact on energy consumption. Therefore trade-off between performance and energy consumption is necessary. A mathematical model in terms of constraint optimization problem is formulated and used. A workload prediction model can also be used to channel this technique.

2. *Hardware energy aware scheduling:* This technique is based on two heuristic techniques and it intelligently places the jobs on its corresponding energy efficient machines. The heuristics that was derived was, IO bound workloads have better energy efficiency on low power nodes and CPU bound workloads achieve better efficiency on high power nodes [9]. The map tasks are CPU intensive and reduce task is IO intensive.

The basic idea of this technique is formulated as follows : The task tracker calculates the metrics for energy efficiency and sends them to Job tracker along with heart beats. The Job tracker schedules the map and reduce tasks identifying the best match for the energy efficiency. If the match is not found, fairness and data locality is used to schedule the tasks.

B. HDFS Cluster modification techniques

These techniques capitalize on opportunity provided by data volume, distribution and redundancy across nodes in HDFS. Nodes are partitioned into zones and the zones are periodically disabled to save energy [9]. This idea of zoning will create the adverse effect of data unavailability which hampers the performance. Various data redistribution and zoning strategies have been proposed.

1. *Replica binning based data placement & Zoning:* In this technique, subset of nodes called covering subset is defined. The basic policy being, at least one copy of each data should be present on one of the node of covering subset node. This makes the data available till the nodes outside the covering are disabled. About 10 to 30 % of cluster size is the cover size. HDFS was changed in such a way that one copy of data stored in the node which created, second copy on a node in covering subset and third copy on a node which is not in the same rack [9].
2. *Temporal binning based data placement & Zoning:* This technique focus on the temporal quality of data. Data caching can be performed to save the most recently used data and flush off the least recently used one. Then data is classified into hot/active and cold/passive zones based on the temporal quality. Cold zones can be put to sleep and hot zones can be used to serve most of the request. GreenHDFS uses this strategy. The cold zone data are not replicated and they follow file migration policy, server power conservation policy and file reversal policy [1].

4. ENERGY AWARE SCHEDULING ALGORITHMS

1. *All-in-Strategy (AIS):* Proposed by Willis et al. is a workload energy aware All-in-strategy. AIS uses batching for low utilization period, batches the jobs and powers on all nodes, performs the jobs and power down the nodes when all jobs are completed. AIS will not work well for time sensitive interactive jobs [9].
2. *Berkeley Energy Efficient Map Reduce (BEEMR):* This energy efficient Map Reduce manager segregates the interactive and batch workloads into separate sub clusters [9]. The cluster is split into 2 zones a small interactive zones and a larger batch zone. Interactive zone is always in full power mode and batch zones switches between full and low power modes.
3. *Covering Subset of Nodes:* HDFS do not rely on data protection mechanism such as RAID, instead replicas are used to provide data availability. This multiple copies of data makes it energy inefficient. Leverich and Kozyrakis proposed an energy aware technique to define a covering subset of nodes that includes atleast one copy of each data [7]. These covering nodes will provide data availability when the nodes outside the cover are disabled.
4. *Seggroup Based Policy:* Nedeljko et al. proposed an approach to carefully place the replicas in the distributed nodes to allow some nodes to be put in low power state. A seggroup metadata is used to store the replica details. The replica policy of HDFS is modified and a new state called sleep is added for data nodes [9]. If the data block is not available and the machine is in sleep mode then replication is not triggered. Thus the speculative task executions are avoided which saves energy costs.

5. *GreenHDFS:* Kaushik et al. proposed GreenHDFS technique to partition servers as hot and cold zones based on their performance, cost and power characteristics [1]. Hot zones are always powered on and cold zones are mostly idling and thus conserved energy. Very aggressive power management policies are used in cold zones and the nodes are powered on only on demand.
6. *Dynamic HDFS:* Nitesh et al. proposed a strategy where the cluster is started with minimum number of nodes and when the workload increases, more number of nodes are added [9]. Every time the cluster is scaled up or scaled down, the data is redistributed in rack aware manner.
7. *E-Ant:* Proposed by Dazhao Cheng et al., E-Ant aims to minimize the overall energy cost in heterogenous Hadoop Cluster without sacrificing the performance. It employs the Ant Colony optimization method to assign the tasks to the nodes based on the feedback of the energy consumption reported by each task trackers [4]. Energy profiling is done at each task tracker nodes.
8. *Dynamic Voltage and Frequency Scaling:* Thomas Wirtz et al. proposed how adjusting the processor frequency based on the workload computational needs can improve the energy efficiency of the MR clusters. The key of DVFS scheduling is to identify the workload phases and adapt the processor frequency to match each phase [6].
9. *Snake Like Data Placement (SLDP):* SLDP adopts the idea of dividing the data nodes into several virtual storage tiers [5]. Then it determines the replication factor of each data block based on hotness and places the data blocks in each VST circuitously

5. COMPARISION OF ENERGY EFFICIENT TECHNIQUES

Table 1
Comparison of Energy Efficiency Techniques for Map Reduce for varied parameters

S.No	*Techniques*	*Response Time*	*Energy Efficiency*	*Data Availability*	*Scalability*	*Description*	*Category*
1	AIS	Moderately degraded due to batching	Improved	None	Scalable	Suited only for production jobs	Workload Based MR modification
2	BEEMR	High degradation for batch jobs	40-50%	Yes	Partial	Suited for interactive and production jobs. Overhead on batch jobs was higher.	Workload Based MR modification
3	Peak Power, Hetero Nodes	Improved	27%	None	-----	Based on the energy metric reported by the task trackers	Hardware based MR modification
4	SegGroup Based Policy	30%	--	---	---	Identification of nodes for power state transition is a problem	Replica Binning Based Cluster Modification
5	GreenHDFS	26%	Greatly improved	Partial		File migration, Server power conservation, file reversal policies are used. But a Frequent power transition is a overhead. Time factor to be considered in file migration policy	Temporal Binning Based Cluster modification
6	Dynamic HDFS Cluster	33 – 54%	--	--	--	Scaling up and scaling down the cluster based on the workload. Redistribution of data in rack-aware manner is overhead.	HDFS cluster modification

6. CONCLUSION

In this paper, the various data placement strategies and job scheduling techniques to improve the Hadoop MR Energy efficiency has been discussed. The algorithms are compared based on varied parametric aspects and tabulated. Based on the study made, we would embark on exploring more optimization techniques for energy efficient data centers. Our future work would focus on the challenging aspect of reducing the speculative execution of tasks (Straggler tasks) which result in resource and energy wastage.

References

1. Thomas Wirtz, Rong Ge, "*Improving MR energy efficiency for computation intensive workloads*".
2. Lena Mashayekhy, Mahyar Movahed Nejad, Daniel Grosu, Dajan Lu, Weisong Shi, "*Energy aware scheduling of Map Reduce jobs*", 2014 IEEE International Congress on Big Data.
3. Krish K.R, M. Safdar Iqbal, M.Mustafa Rafique, Ali R.Butt, "*Towards energy awareness in Hadoop*", 2014 Fourth International Workshop on Network Aware Data management, IEEE Computer Society.
4. Dazhoa Cheng, Palden Lama, Changjun Jiang, Xiaobo Zhou, "*Towards energy efficiency in heterogenous Hadoop clusters by adaptive task assignment*," 2015, IEEE-35th International Conference on Distributed Computing Systems, IEEE Computer Society.
5. Runqun Xiong, Junzhou Luo, Fang Dong, "*SLDP : A novel data placement strategy for large scale heterogeneous Hadoop cluster*", 2014 ,Second International Conference on Advanced Cloud and Big Data.
6. Nidhi Tiwari, Umesh Bellur, Maria Indrawan, Dr.Santonu Sarkar, "*Improving the Energy Efficiency of MR Framework*", PhD Forum.
7. Lena Mashayekhy, Daniel Grosu, Quan Zhang, Weisong Shi, "*Energy aware scheduling of Map Reduce jobs for Big Data Applications*", 2014, Vol. 25, IEEE Transactions on parallel and distributed systems.
8. Seyed Reza Pakize, "*A Comprehensive view of Hadoop MR Scheduling Algorithms*", Internaltional journal of computer networks and communications security, Vol. 2, No. 9, Sep 2014.
9. Nidhi Tiwari, "*Scheduling and Energy Efficiency Improvement Techniques for Hadoop Map-reduce: State of Art and Directions for Future Research*", PhD Seminar report.