The Journal of Systems and Software 87 (2014) 38–47

Contents lists available at ScienceDirect

The Journal of Systems and Software

journal homepage: www.elsevier.com/locate/jss

A three-phase energy-saving strategy for cloud storage systems

Saiqin Long, Yuelong Zhao*, Wei Chen

School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China

ARTICLE INFO

Article history:
Received 22 April 2013
Received in revised form 17 August 2013
Accepted 17 August 2013
Available online 27 August 2013

Keywords:
Energy saving
Cloud storage systems
Cost efficient

ABSTRACT

In the running process of cloud data center, the idle data nodes will generate a large amount of unnecessary energy consumption. Furthermore, the resource misallocation will also cause a great waste of energy. This paper proposes a three-phase energy-saving strategy named TPES in order to save energy and operational costs for cloud suppliers. The three phases are replica management based on variable replication factor, cluster reconfiguration according to the optimal total costs and state transition based on observed and predicted workloads. These three phases save energy for the system at different levels which enhance the adaptability of our strategy. We evaluate our strategy using the expanded CloudSim toolkit and the results show that the proposed strategy achieves better energy reduction under different conditions in comparison with the existing schemes.

1. Introduction

Energy consumption of cloud data centers is a major problem for cloud suppliers.

It has been reported that energy consumption of Google data centers is equivalent to the total energy consumption of a small city (Buyya et al., 2009b). If the trend goes on, the energy costs consumed by a server in its life cycle will be more than the hardware costs of it (Barroso, 2005). Additionally, according to the reports of International Data Corporation (IDC), the number of global servers (virtual and physical) and the number of files in data centers will grow by 10 times and 75 times respectively in next 10 years, while up to 10% of the data will be maintained in the cloud (Gantz and Reinsel, 2011). These lead to the fact that with the rapid growth of the scale of data centers, the power consumption and costs become increasingly prominent issues. Especially in the case of lack of energy and significant growth of energy costs, we must consider the problem of energy efficiency to save energy and operational costs for cloud suppliers.

Numerous energy saving techniques for cloud data centers have been studied. They focus on improving the energy efficiency of data centers in all aspects—distribution of power, data center cooling, energy consumption of IT components (Guerra et al., 2010) and resource allocation and job scheduling (Zikos and Karatza, 2011; Galizia and Quarati, 2012; Mavromoustakis and Karatza, 2012; Wang et al., 2012a). Among these different aspects, reducing the energy consumption of IT components has played an important role since improving its energy efficiency also reduces the required capacity of the cooling and power distribution systems. Of all the IT components in a data center, storage is the next largest consumer of energy after processors and cooling systems. Recent work has shown that the storage system accounts for up to 37–40% of the energy consumption of all IT components (Schulz, 2007). This percentage of energy consumption by storage systems will only continue to increase, as data intensive applications demand fast and reliable access to on-line data resources (Kim and Rotem, 2011).

A rich body of existing work (Gurumurthi et al., 2003a,b; Irani et al., 2005; Zhu et al., 2005; Weddle et al., 2007; Xie, 2008) has already investigated the energy efficiency of storage systems. Generally, these algorithms use the idea of spinning down the data nodes from the high energy consumption mode (working mode) into a lower energy mode (standby mode) after they experience a period of inactivity. The common feature of them is that they only apply the spin-down techniques to individual data node. The energy saving approach we considered in this paper is different from existing approaches, as it is based not only on handling energy saving in a group of data nodes, but also controlling data nodes individually. In addition, our approach exploits data replication which is used in many datacenters due to fault tolerance and load balancing. We optimize the default replication management in Hadoop Distributed File System (HDFS) (Shvachko et al., 2010).

Three important problems must be solved in order to achieve the energy saving:

- What is the minimal number of replicas in the system to meet reasonable requirement for availability is an important issue to be thoroughly investigated. With the number of replicas increasing, the system maintenance cost will significantly

* Corresponding author. Tel.: +86 15219834211.
E-mail addresses: longsaiqin@163.com (S. Long), zhaolab@126.com, 282293237@qq.com (Y. Zhao), chw521@163.com (W. Chen).

0164-1212/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jss.2013.08.018

increase, and too many replicas may not increase availability, but bring unnecessary spending instead (Sun et al., 2012).

- How many data nodes should be kept in an active state to make the operational costs of cloud suppliers reach the optimal value is another important issue for further research. Keeping more data nodes active may improve system task execution rates by shortening waiting time of users, but consumes more energy instead.
- The last important issue is to determine when to transit the data nodes to a lower power state. One typical approach for this is to use a set of thresholds for a load metric.

Therefore, in this paper, we address the above three issues by designing a corresponding three-phase energy-saving strategy (TPES) for cloud storage systems. In the first phase, through the dynamic replica management algorithm driven by the users' availability requirements, the system only maintains as few replicas as possible to avoid energy waste caused by allocating excess resources to users. In the second phase, our TPES dynamically reconfigures the cluster according to the current workloads and the costs of the whole system so as to achieve energy efficiency at the data center level. In the third phase, our algorithm dynamically switches the state of data nodes through load prediction to enable energy saving at the single data node level. These three phases achieve the purpose of energy saving at different levels of the system in a step-by-step way which enhances the adaptability of our strategy. Our strategy overcomes the limitation of the existing energy-saving strategies in cluster systems which merely scale clusters or change the number of replicas unilaterally. It optimizes the cluster not only at the data center level but also the single data node level.

The main contributions of this paper are as follows:

- A mathematical model is formulated to describe the HDFS.
- A mathematical model formulated to describe the relationship between file availability and replication factor and a replica management algorithm based on the minimal replication factor is proposed.
- A cost-based energy-saving algorithm used to determine the optimal number of data nodes in the running state according to the access rate of requests and its distribution is presented.
- An energy-saving algorithm based on the state transition of the single data node is described. It makes predictions based on the historical information.
- Evaluation results are shown to demonstrate the energy efficiency of our proposed strategy.

The rest of the paper is organized as follows. In Section 2, we discuss the related work. In Section 3, we formulate the problem and present our three-phase energy-saving strategy. In Section 4, we evaluate performance of our strategy. Finally, conclusions and future work are given in Section 5.

2. Related work

A significant amount of research work in the field of energy saving techniques for cloud storage systems has been done. Existing energy-saving techniques can be roughly divided into the following categories.

The first class of techniques focuses on the hardware components. It reduces the energy consumption of the system by reducing the energy consumption of the hardware devices which are used to constitute the cloud systems. This class of techniques can be divided into two types (Wang et al., 2012b). One focuses on computer components constituting the cloud systems. This type of technique reduces the energy consumption of a single node or even the entire system through the use of new architecture or hardware technology. A disk structure with a multi-phase rotational speed which changes with workloads change is proposed by Gurumurthi et al. (2003a,b). In addition, Hamilton (Hamilton, 2009) proposes a low-power server rack structure based on an Athlon processor. The other is focused on the data centers. In order to achieve the goal of energy reduction for the system, it uses the low-power, low-performance hardware devices to replace the high-power, high-performance devices to build the data centers. A research team in Korea (Kim et al., 2010) applies low-power components to build a Hadoop platform. Although the performance is slightly lower, the results show that this scheme makes 113 times less energy consumption. These techniques are easy to manage, but they need to update the system architecture and hardware devices which make them very inflexible. In addition, large quantities of hardware replacement will cause excessive hardware costs for those already deployed applications.

The next set of techniques uses intelligent data placement and/or continuous data migration, resulting in large time-scale to adapt. In some studies, energy is saved mainly through migration of virtual machines and optimization of data layout (Hu et al., 2008; Liu et al., 2009; Milenkovic et al., 2009). They aggregate data into fewer storage devices whenever performance requirements permit, or place/move data into storage devices that best fit its performance requirements. Beloglazov and Buyya (2012) in the CLOUDS laboratory propose an adaptive heuristic dynamic virtual machine consolidation algorithm based on the historical records of resource utilization of the virtual machine. Xie (2008) presents a striping-based energy-aware (SEA) strategy which can be integrated into data placement in RAID-structured storage systems. He first separates disks in a disk array into two zones: the hot disk zone and the cold disk zone. After that, he places popular data onto disks in the hot disk zone and assigns unpopular data onto disks in the cold disk zone. Beyond that, diversion of newly written data to enable spinning down disks for longer periods, coupled with opportunistic movement of data to storage devices when they become active, is used to reduce energy consumption by Narayanan et al. (2008).

Finally, another class of techniques benefits from the availability of hardware-based active and inactive low energy modes in disks. Multi-speed disks have been adopted by some traditional energy-saving schemes such as Multi-speed (Carrera et al., 2003), DRPM (Ghandeharizadeh et al., 1995), PDC (Pinheiro and Bianchini, 2004), and Hibernator (Zhu et al., 2005). Typically, these techniques dynamically switch disks from one-speed mode to another to better serve disk access requests and save energy. The problem of subset not coverage when turning down some data nodes is solved by Harnik et al. (2009) in the IBM lab through adding an auxiliary data node. Leverich and Kozyrakis (2010) modify Hadoop and make it possible to dynamically scale up or down according to the cluster configuration at the data center level. However, they ignore the data availability requirements of users while scaling the cluster. Maheshwari et al. (2012) in India also study the Hadoop cluster for energy saving. Their proposed strategy can dynamically reconfigure the cluster based on the current workloads and turn nodes on or off when the average cluster utilization rise above or fall below the thresholds specified by the administrator, respectively.

In contrast to the discussed studies, we propose a three-phase energy-saving strategy for cloud storage systems which is inspired by the work of Maheshwari et al. (2012). The main difference between their work and ours is as follows: (1) they only shift gears across multiple data nodes, while we achieve it not only across data nodes but also for an individual data node, and we also consider performance guarantee and cost constraints. (2) Their gear-shift mechanism takes a reactive approach based on utilization with thresholds, while we use a prediction technique based on