Applied Soft Computing xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Applied Soft Computing

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

Reducing energy usage in drive storage clusters through intelligent allocation of incoming commands

E. Smart[a,*], D. Brown[a], K. Toumi Borges[b], N. Granger-Brown[b]

[a] Institute of Industrial Research, Unit 1 St Andrews Court, University of Portsmouth, Hampshire, PO1 2PR, United Kingdom
[b] Seagate Technology, Langstone Road, Havant, Hampshire, PO9 1SA, United Kingdom

ARTICLE INFO

Article history:
Received 10 April 2015
Received in revised form
24 September 2016
Accepted 4 October 2016
Available online xxx

Keywords:
Decision support systems
Disk drives
Data storage systems
Pareto optimization
Optimal scheduling

ABSTRACT

Although significant research has been undertaken to reduce high level energy consumption in a data centre, there has been very little focus on reducing storage drive energy consumption via the intelligent allocation of workload commands at the file system level. This paper presents a method for optimising drive energy consumption within a custom built storage cluster containing multiple drives, using multi-objective goal attainment optimization. Significantly, the model developed was based on actual power consumption values (from current/voltage sensors on the drives themselves), which is rare in this field.

The results showed that command energy savings of up to 87% (17% overall energy) could be made by optimising the allocation of incoming commands for execution to drives within a storage cluster for different workloads. More significantly, the transparency of the method meant that it showed exactly how such savings could be made and on which drives. It also highlighted that whilst it is well known that solid state drives use less energy than traditional hard disk drives, the difference is not consistent for different sizes of data transfers. It is far larger for small data transfers (less than or equal to 4 kB) and our algorithm utilised this.

Significantly, it highlights how much larger energy savings can be made through using the optimisation results to show which drives can be safely put into a low power state without affecting storage cluster performance.

1. Introduction

Data centres are a key part of the world's economic infrastructure and vary in size from server farms that support the likes of Google and Facebook to the server rooms that support small-to-medium corporations. This heavy reliance on data has made data centres one of the biggest consumers of electricity (estimated at 91 billion kilowatt hours in America alone in 2013) [1]. However, according to the same report, as much as 50% of this is wasted. As such, energy efficiency in data centres is a key topic for the research community.

The topic of energy reduction within a data centre is well researched with significant activity in the following areas; power usage within a disk drive, reducing disk drive power states within a storage cluster and intelligent data assignment. However, nearly all of this research focusses on high level data flow within sets of storage clusters.

Our research focuses on the problem from a new angle and analyses energy management at a much lower level. It looks at modifying the way in which read/write commands are managed within a storage system in order to minimise the energy used in this process. The algorithm proposed is intended to operate at a level between the operating system and the storage devices and would be embedded within the kernel. The algorithm would then take commands that come from the upper layers and redirect them to the most appropriate drives. To conceptually illustrate the algorithm in action, suppose that the file storage system is blank. If a given application wants to store data that would be rarely retrieved (read), the algorithm would redirect that data to drives with the least write energy costs. However, if the data were to be retrieved (read) often, it would be stored (written) on those drives with the lowest read energy costs (even if they had higher write energy costs than other drives). This would occur until all of the storage devices were full, thus creating a storage system with an overall low optimal energy consumption profile.

The following areas are out of scope of this piece of research; algorithm implementation, where it should sit within the storage

* Corresponding author.
E-mail addresses: edward.smart@port.ac.uk (E. Smart), david.j.brown@port.ac.uk (D. Brown), kais.toumi@seagate.com (K. Toumi Borges), nick.grangerbrown@seagate.com (N. Granger-Brown).

http://dx.doi.org/10.1016/j.asoc.2016.10.005
1568-4946/

system and also what happens to data once it has been initially written (e.g. should it be replicated onto other drives).

We propose the use of the multi-objective goal attainment algorithm to optimally assign individual commands to drives in order to achieve minimal command energy usage at the storage cluster level. In storage clusters, it is not unreasonable to see multiple copies of regularly accessed files stored on different drives (e.g. popular YouTube videos). This means that any energy management algorithm could access a file by reading different sections of it from different drives. To the best of the author's knowledge, there are no other methods that operate so deep within the file system at such a low level. For this reason, it was not meaningful to compare the proposed method with others in the literature review (see Section 2).

A custom built storage cluster was fitted with two types of drive: solid state drive (SSD) and hard disk drive (HDD). Their energy usage was computed directly using current and voltage sensors to ensure accurate values. This is a significant advantage over previous research (e.g. Chou et al. [2]) as usually storage clusters are simulated numerically, which can lead to inaccuracies if drive background activity is not taken into consideration.

The paper is organised as follows. Section 2 contains the literature review. Section 3 describes the multi-objective goal attainment method. Section 4 explains the experimental setup. Section 5 describes the method. Section 6 details the results and Section 7 lists the conclusions.

2. Literature review

In this paper, we assess the available literature for energy saving methods, hardware and software based, for disk storage clusters as well as for optimisation methods.

2.1. Storage cluster energy saving

The main methods used to perform energy saving focus on general dynamic power techniques and general performance enhancing techniques that additionally improve power usage.

Modern drives have multiple power states such as 'active' (data transfer ready), 'idle' (waiting for data transfer) and 'standby' (some electronics turned off and disk platters stopped) [3]. For more details on drive power states, see Section 4.4. Significant power savings can be made by putting HDDs into standby mode but if a drive receives a new I/O request, it takes time and energy to spin up, which often has too much of an impact on performance in certain applications. The duty cycle rating is defined as the number of spin ups and spin downs a disk can withstand before the failure risk on spin up exceeds 50%. For server based drives, this figure is much less than for laptop drives so energy saving measures should not spin up or spin down a drive indiscriminately [4].

A basic approach is to use Traditional Power Management (TPM) methods, which are usually threshold based [5]. In such methods, drive power states are changed after a given threshold time τ, which can be static or dynamic. This can be effective if the pattern of I/O requests is known but for the more common unknown circumstance, it can be shown that by using competitive analysis to set τ to the break-even time, this policy uses at most twice the power of the most optimal offline policy.

Another proposed method is to use variable speed disk drives [6]. Rather than having 2 discrete states (idle and standby), this approach lets drives spin at different speeds so that at least a few will be at full speed to service requests. It has been shown that this is a good way to manage the balance between performance and energy usage.

A more common approach is to use Dynamic Power Management (DPM) techniques, which seek to maximise energy savings by distributing I/O requests over time and across all the disks in the array [7]. It achieves energy savings by first controlling the frequency of I/O requests to each disk so that the mean idle time is large whilst maximising the idle-time variance across all the disks and through time. Techniques include power aware caching and buffering in memory [8], across disks [9], workload consolidation through data replication [10], popular data concentration [11] and diverting disk accesses by utilising data redundancy [12].

Previous work has been undertaken to analyse the performance of drives and storage clusters. Härder et al. [13] looked at the performance and power consumption of SSDs in a storage cluster and found significant differences in consumption. However, it is difficult to assess this as power consumption was not measured directly from the drives themselves and it was not stated whether the drives were performing maintenance actions at the time.

Attempts at energy reduction in storage systems are usually made at a higher system level. For example, improved methods of heat dissipation via different types of coolants or through different systems architecture. Our approach seeks to focus on the properties of the different drive types and the types of commands they are able to process more efficiently (with respect to energy) so that workload commands can be assigned optimally.

2.2. Optimisation methods

To the best of the authors knowledge, there have been no attempts to apply optimisation methods to assign commands to drives. The basic problem is to assign n commands to m drives where n is much larger than m and both are natural numbers. This could be tackled by means of a semi-assignment algorithm. Drwal [14] used this method to assign streams of requests to servers in order to minimise the worst case processing times although he offers no details as to how his algorithm improves the total time. Vasudevan [15] uses a semi-assignment algorithm to optimally assign applications to virtual machines in a data centre with respect to energy usage, CPU utilisation and application completion time which shows good results. For problems where there are more than one parameter to optimise, multi-objective goal attainment optimization has been suggested. This algorithm tries to satisfy a given set of well defined goals subject to multiple constraints. Jones [16] used the goal attainment algorithm to show how multiple goals regarding supply and demand could be satisfied. Oddoye [17] has applied goal attainment optimization in the health care sector and shown how it can optimally cut waiting times for patients whilst effectively assigning doctors, nurses and consultants to keep queues to a minimum. Ghoseiri [18] uses goal attainment to solve the vehicle routing problem and was successful in optimising multiple goals in a reasonable amount of computational time. Multi-objective goal attainment optimization will be used in this paper to solve this "command-to-drive" assignment problem.

Hybrid genetic algorithms [19] and evolutionary algorithms [20] have also been used to optimally schedule computing resources over a network where energy was one of the optimisation objectives. However, a weakness is that it did not verify the algorithm on real data.

3. Multi-objective goal attainment optimization

The main problem is to allocate n commands onto m drives for $n \gg m$ in order to achieve the goals of minimal energy usage, fast execution time and low storage cost. Let $\mathbf{x} = \{x_{ij}\}$ be the primary decision variable defined by: