Future Generation Computer Systems 29 (2013) 196–207

Contents lists available at SciVerse ScienceDirect

Future Generation Computer Systems

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

An adaptive energy-conserving strategy for parallel disk systems

Mais Nijim [a], Xiao Qin [c,*], Meikang Qiu [b], Kenli Li [d]

[a] Department of Electrical Engineering and Computer Science, Texas A&M University-Kingsville, Kingsville, TX 78363, United States
[b] Department of Electrical and Computer Engineering, University of Kentucky, Lexington, KY 40506-0046, United States
[c] Department of Computer Science and Software Engineering, Auburn University, Auburn, AL 36849, United States
[d] School of Computer and Communication, Hunan University, Changsha, 410082, China

ARTICLE INFO

Article history:
Received 20 April 2010
Received in revised form
30 January 2012
Accepted 14 May 2012
Available online 11 June 2012

Keywords:
Energy saving
Parallel disks
Adaptive strategy

ABSTRACT

Although various parallel disk systems have been developed to achieve high I/O performance and energy efficiency, most existing parallel disk systems lack an adaptive way to conserve energy in dynamically changing workload conditions. To solve this problem, we develop an adaptive energy-saving scheme or DCAPS in parallel disk systems. We show that adaptability in energy conservation can be achieved through the integration of a dynamic disk scheduling scheme and power management in parallel disk systems. DCAPS consists of a data partitioning mechanism, a response time estimator, and an adaptive energy-conserving mechanism. The Data partitioning mechanism allows DCAPS to adjust the parallelism degrees of write requests based on dynamic workload conditions. Apart from supporting the data partitioning mechanism, the response time estimator makes it possible for the adaptive energy-conserving mechanism to dynamically adjust voltage supply levels while guaranteeing desired response times. We conducted extensive experiments to quantitatively evaluate the performance of the proposed energy-conserving strategy. Experimental results consistently show that DCAPS significantly reduces energy consumption of parallel disk systems in a dynamic environment over the same disk systems without using DCAPS.

Published by Elsevier B.V.

1. Introduction

Since the performance gap between processors and disks has widened over the last two decades [1], the performance of data-intensive applications has been significantly affected by disk systems [2,3]. Because parallel disks are highly scalable, parallel disk systems can alleviate I/O bottleneck problems in many data-intensive systems like video surveillance [4], remote-sensing database systems [5,6], digital libraries [7,8], simulation tools [9], and long running simulations [10].

Although parallel disks play an important role in achieving high-performance for data-intensive applications, a substantial amount of energy consumed in data-intensive systems is contributed by parallel disk systems. A recent industry report reveals that storage devices account for almost 27% of the total energy consumed by a data center [11]. For example, the power consumption of today's data center ranges from 75 to 200 W/ft^2. Since this trend will undoubtedly continue in the near future [12], the energy-consumption problem in data centers will become even more serious. Therefore, we are motivated to extensively investigate energy-conservation software techniques for parallel disk systems.

Modern parallel disk systems have increasingly become energy efficient (see, for example, [13]); however, there is a lack of approaches for adaptively conserving energy in parallel disks. Adaptive energy-saving techniques are important in parallel disk systems because of two reasons. First, many data centers have dynamically changing workload characteristics. For example, I/O loads of web servers are known to dynamically change with time [14]. Second, real-world data-intensive applications tend to have performance and resource requirements. For example, disk requests issued by data-intensive applications need to be completed within specified response times [15]. Adaptively conserving energy in parallel disk systems becomes particularly critical for data-intensive applications running in dynamically changing computing environments.

In this study, we are inspired by the needs of data-intensive applications to develop a way of flexibly and adaptively reducing energy consumption caused by the data-intensive applications. We show that adaptability in energy conservation can be achieved through the integration of a dynamic disk scheduling scheme and power management in parallel disk systems. We focus

* Corresponding author.
E-mail addresses: mais.nijim@tamuk.edu (M. Nijim), xqin@cs.nmt.edu, xqin@auburn.edu (X. Qin), mqiu@engr.uky.edu (M. Qiu), lkl510@263.net (K. Li).
URL: http://www.eng.auburn.edu/~xqin (X. Qin).

0167-739X/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.future.2012.05.003

on disk scheduling in this research, because disk scheduling algorithms (e.g., shortest seek time first (SSTF) and SCAN) can significantly improve disk performance [16–20]. For example, the SSTF algorithm aims to minimize seek time of disk requests [21]; the SCAN algorithm solves the unfairness problem and reduces seek time [21]; and Reist and Daniel developed a parameterized generalization of SCAN and SSTF to seamlessly integrate the two algorithms [22]. Most existing disk scheduling algorithms are inadequate for improving energy efficiency of parallel disk systems. To address this problem, we develop an adaptive energy-conservation technique or DCAPS that incorporate power management with disk scheduler in parallel disks systems. More importantly, our DCAPS scheme can offer significant energy savings while guaranteeing desired response times of disk requests.

Our DCAPS scheme can manage two types of disk I/O parallelisms, namely, inter-request parallelism and intra-request parallelism. The inter-request parallelism allows multiple independent disk requests to be served simultaneously by multiple disks, whereas the intra-request parallelism enables a single disk request to be responded by an array of disks in parallel. The parallelism degree of a data request is the number of disks in which the requested data is residing [23]. DCAPS adjusts the parallelism degrees of write requests based on dynamic workload conditions (see Section 4.1). After determining parallelism degrees, DCAPS optimizes disk supply voltage levels to reduce power consumption while guaranteeing requests' desired response times by utilizing the disk scheduling mechanism (see Sections 4.2 and 4.3).

Experimental results show that DCAPS significantly reduces energy consumption of parallel disk systems in a dynamic environment over the same disk systems without using DCAPS. In addition, DCAPS improves energy efficiency of parallel disks without reducing satisfied ratio of requests having desired response times.

The rest of the paper is organized as follows. We summarize related work in the next section. Section 3 describes the system architecture for energy-efficient parallel disk systems. In Section 4, we propose the adaptive energy-conservation scheme called DCAPS. Section 5 evaluates the performance of the proposed energy-saving technique by comparing an existing approach. Section 6 concludes the paper with summary and future directions.

2. Related work

Disk I/O has become a performance bottleneck for data-intensive applications due to the widening gap between processor speeds and disk access speeds [1,24,25]. To help alleviate the problem of disk I/O bottleneck, a large body of work has been done on parallel disk systems. For example, Kallahalla and Varman designed an on-line buffer management and scheduling algorithm to improve performance of parallel disks [26]. Scheuermann et al. addressed the problem of making use of striping and load balancing to tune performance of parallel disk systems [23]. Rajasekaran and Jin developed a practical model for parallel disk systems [27]. Kotz and Ellis proposed investigated several write back policies used in a parallel file system implementation [28]. Our research is different from the previous studies in that we focused on energy savings for parallel disk systems. Additionally, our strategy is orthogonal to the existing techniques in the sense that our scheme can be readily integrated into existing parallel disk systems to substantially improve energy efficiency and performance of the systems.

Abundant research has been done to improve energy efficiency of mobile devices (e.g., smart phones) to increase battery life of the devices. Unlike mobile devices where battery life is critical, disk systems have to be energy efficient in order to lower electricity bills in data centers. In the past decade, much attention has been paid to the development of energy-efficient disk systems. For example, the conventional wisdom of saving energy in disks is to place idle disks into the low-power (e.g., standby) mode. Because significant energy in disks is consumed by spindle motors, dynamic power management schemes are proved to be very effective [29]. Therefore, the dynamic power management techniques have been widely applied to reduce energy dissipation in disk drives of both PCs and high-performance computers [30].

Most of the previous research regarding conserving energy focuses on single disk system in laptops and mobile devices to extend the battery life. Recently, a handful of techniques proposed to save energy in disk systems include dynamic power management [30–32], I/O workload-skew schemes [33,34], power-aware cache management [35–37], power-aware perfecting schemes [38–40], software-directed power management techniques [41], redundancy techniques [41], multi-speed disks [42–44], and data placement technique [45]. However, the research on energy-efficient parallel disk systems is still in its infancy. It is imperative to develop new energy conservation techniques that can provide significant energy savings for parallel disk systems while maintaining high performance.

The dynamic voltage scaling technique or DVS is a widely adopted approach to conserving energy in processors. The DVS technique can dynamically reduce the voltage supplies of processors to conserve energy consumption in processors (see, for example, [46,47]). Thus, processor voltage supplies are scaled down to the most appropriate levels, thereby quadratically reducing power whenever possible. Compared with traditional computer systems with fixed voltage supplies, DVS-enabled systems can achieve high energy efficiency. Our approach differs from the conventional DVS methods, because ours is the first technique of its kind designed exclusively for energy-efficient parallel disk systems aiming to guarantee desired response times requests issued by of data-intensive applications. Our adaptive energy-conserving strategy seamlessly incorporates DVS-enabled dynamic power management into the disk scheduler to achieve high energy efficiency in parallel disk systems while satisfying desired response times.

3. System architecture and model

In this section, we first present a framework within which we develop an adaptive energy-conservation technique for parallel disk systems. Then, we describe the power consumption model used to estimate the power consumption of large-scale parallel disk systems.

3.1. System architecture

Fig. 1 outlines the framework of energy-efficient parallel disk systems. The framework is general enough to accommodate a wide range of storage systems, including both network attached storage devices (NAS) and storage area networks (SAN). The framework in Fig. 1 embraces a parallel disk system, networks, and the DCAPS scheme. In this study, we focus on the development of DCAPS that consists of a data partitioning mechanism (see Section 4.1), a response time estimator (see Section 4.2), and an adaptive energy-conserving mechanism (see Section 4.3).

The data partitioning mechanism (see Section 4.1) is geared to divide a large amount of data into fixed-size of data units stored on a number of disks. We consider file striping – a generic method for various data types. To optimize the parallelism degree (a.k.a., stripe unit size) for each write request, the data partitioning mechanism relies on the response time estimator to predict requests' response times.

The response time estimator (see Section 4.2) not only supports the data partitioning mechanism, but also is indispensable for the adaptive energy-conserving mechanism. The response time esti-