Energy-efficient and Thermal-aware Data Placement, Replication, and Scheduling in Data Centers
Amol Deshpande, Samir Khuller, University of Maryland, College Park, MD 20742

In this white paper, we lay out a research agenda that aims to reconsider a broad variety of optimization problems that arise in normal data center operation with a focus on energy efficiency and temperature management. Data centers are increasingly becoming the key infrastructure pieces to handle rapidly growing data management needs in domains ranging from scientific data analysis to large-scale web-based applications. Our work focuses on two related, yet distinct issues that arise in data centers. First, given the sheer amount of computing machinery, energy efficient computing is rapidly becoming a key concern – computing equipment in the U.S. alone is estimated to consume more than 20 million gigajoules of energy per year, the equivalent of four million tons of CO₂ emissions into the atmosphere. Low-power and energy proportional hardware is clearly an important step toward reducing these energy demands. However, that alone may not provide a complete solution and we must also develop supporting software level techniques such as energy efficient algorithms and data structures that complement the energy-friendly hardware. Second, there is an increasing trend toward tightly packing computing machinery in data centers to reduce the total cost of ownership.¹ A similar trend is also seen at micro-scale with an increasingly larger number of processor cores or other units per chip. However, higher density results in higher operating temperatures, and often advanced cooling techniques are needed to cool the equipment.² Higher temperatures not only result in increasing failures and possibility of errors, but also higher energy consumption for the same amount of work. Software optimizations can play an important role in regulating the temperatures in data centers, allowing the data center to be run at higher overall temperature, and thus resulting in better energy efficiency.

Next we outline a set of research challenges and our initial work in developing energy-efficient and thermal-aware optimization algorithms.

Modeling: To be able to formulate and solve the optimization problems, a key challenge is modeling both the data center behavior (especially how the thermal profile changes in response to workload), and the expected workload. With the former, there are several spatial effects that must be taken into consideration, and the most accurate way to computing the thermal profile is using numerical methods like computational fluid dynamics, which may be intractable in large data centers. One option is to use approximate models that attempt to capture the relationship between the workloads and the temperatures; however, there is lack of work in developing such models so far. On the other hand, the query workload is somewhat easier to model and we can build predictive models for it using history traces.

Task Assignment and Scheduling: Most data centers replicate data for fault tolerance (e.g., Hadoop file system (HDFS) maintains at least 3 copies of each data item by default), and incoming jobs can be often be assigned to one of several servers containing the data required to run the computation. Traditional scheduling mechanisms use the current server loads to make this assignment. However, new optimization algorithms are needed for two reasons. First, the temperature of a server depends on both the load on the server itself and the load on other close-by servers (thermal coupling). Thus we should do task assignment in a way that spreads out the load across the

²http://www.computerworld.com/s/article/9144466/Data_center_density_hits_the_wall
data center and avoids hotspots.\(^3\) Secondly, inherent redundancy through replication enables a significant opportunity to schedule the tasks so that some of the machines can be deactivated to conserve energy. In our prior work,\(^4\) we have designed several algorithms for a metric that combines the total response time and the number of inactive machines.

**Data Replication, Placement, and Migration:** A related issue is that of data replication and placement. Given an expected workload, we can aim to find a data placement that can help avoid thermal hotspots and/or result in better energy efficiency. For hotspot reduction, one can argue that placing popular data items as far as possible (physically) is the judicious choice. Also popular data items should be replicated more enabling us to balance out the load. However, for energy efficiency, it is beneficial to cluster data items together so that fewer machines or disks are involved in answering a query – that would also reduce the cost of maintaining the consistency of replicas. In our preliminary work,\(^5\) we have developed several algorithms for energy-efficient replication for read-only workloads. However, understanding the trade-offs here is a rich area of further research.

**Controlling Disk and Processor Speeds:** We need to design algorithms for adjusting and tuning the parameters of the computing machinery, perhaps the key ones being the disk and processor speeds. Modern disks and processors are equipped with multiple speed states. In general a higher rotating disk speed results in higher data rate at the cost of higher power dissipation. Based on the data replication across disks and task assignment, we can allocate the disk speeds such that all requests get processed by their deadlines and thermal hotspots are avoided. However, tuning the speed of thousands of disks is a major challenge. Similar challenges exists in controlling the processor speeds (often called speed scaling). A related issue is that of designing algorithms to appropriately use a mix of different hardware, with differing energy characteristics (e.g., in a system that contains both SSDs and hard disks, choosing which data to keep in SSDs can significantly impact the total energy consumption).

**Participant Backgrounds**

Dr. Khuller has developed approximation algorithms for many problems that are closely related to the research agenda laid out above. As part of an NSF-funded project (CCF-0113192), he developed algorithms for computing almost optimal data layouts, as well as considered the problems of quickly re-organizing the layout to adapt to changing demand, where it was assumed that the total workload was unchanged, but just the distribution for demand had changed. This work has led to a variety of conference and journal publications. Dr. Deshpande has worked on many problems in database systems including distributed query optimization, and adaptive query processing. He has also extensively worked in the area of energy-efficient computing in wireless sensor networks, designing algorithms for query processing and data collection with a focus on minimizing the total energy consumed by the sensor network. His recent research has focused on algorithms for data replication and layout. As part of their joint NSF funded grant (CCF-0937865), Dr. Khuller and Dr. Deshpande are working on optimization algorithms for thermal management in data centers.

---

\(^3\) A. Deshpande, S. Khuller, A. Srivastava; A case for spatially aware optimization in data centers; *Manuscript.*


\(^5\) A. K. Kayyoor, A. Deshpande and S. Khuller; Energy efficient data placement and replication in data centers; Submitted.