COST EFFICIENT RESOURCE MANAGEMENT FRAMEWORK FOR HYBRID JOB SCHEDULING UNDER GEO DISTRIBUTED DATA CENTERS

Ms. K. Narmatha¹ and Mr. V.P. Muthukumar²

¹M.Phil Full time Research Scholar, PG & Research Department of Computer Science, Vivekanandha College of Arts and Sciences for Women (Autonomous) Tiruchengode, Tamilnadu, Namakkal-637 205
²Head of the Department of Computer Applications Vivekanandha College of Arts and Sciences for Women (Autonomous) Tiruchengode, Tamilnadu, Namakkal-637 205

Abstract—The cloud data centers are used to provide data sharing with the users. Data center energy efficiency is achieved with demand based resource provisioning model and server power management mechanism. Data center jobs are categorized into two types based on their requirements. They are delay sensitive jobs (SENs) and delay tolerant jobs (TOLs). Immediate resource requirements are provided under the delay sensitive jobs model. Backup and maintenance operations are carried out under the delay tolerant jobs model. Joint SEN and TOL resource provisioning scheme is employed to manage delay sensitive and delay tolerant jobs. Joint server provisioning, SEN load dispatching, TOL load shifting and SEN/TOL capacity allocation scheme manages the data center jobs. The OrgQ scheme handles the resource allocation for the TOL jobs. The Geo-Distributed Data Center (GDC) management system is build to handle job allocation with spatial and temporal parameters. Priority features are integrated to manage the delay sensitive jobs and delay tolerant jobs. Energy cost management process is carried out with traffic, time and location dynamics. The system manages the virtualization tasks against the data center allocation process.

Keywords— Cloud Data Centers, Resource Provisioning Schemes, Delay Sensitive Jobs, Delay Tolerant Jobs and Energy Cost Management

I. INTRODUCTION

Internet-scale cloud services that deploy large distributed systems of servers around the world are revolutionizing all aspects of human activity. The rapid growth of such services has lead to a significant increase in server deployments in data centers around the world. Energy consumption of data centers account for roughly 1.5% of the global energy consumption and is increasing at an alarming rate of about 15% on an annual basis. The surging global energy demand relative to its supply has caused the price of electricity to rise, even while other operating expenses of a data center such as network bandwidth have decreased precipitously. Consequently, the energy costs now represent a large fraction of the operating expenses of a data center today and decreasing the energy expenses has become a central concern for data center operators.

The emergence of energy as a central consideration for enterprises that operate large server farms is drastically altering the traditional boundary between a data center and a power utility. Traditionally, a data center hosts servers but buys electricity from an utility company through the power grid. However, the criticality of the energy supply is leading data centers to broaden their role to also generate much of the required power on-site, decreasing their dependence on a third-party utility. While data centers have always had generators as a short-term backup for when the grid fails, on-site generators for sustained power supply is a newer trend. For instance, Apple recently announced that it will build a massive data
center for its iCloud services with 60% of its energy coming from its onsite generators that use “clean energy” sources such as fuel cells with biogas and solar panels [5]. As another example, eBay recently announced that it will add a 6 MW facility to its existing data center in Utah that will be largely powered by on-site fuel cell generators [7]. The trend for hybrid data centers that generate electricity on-site with reduced reliance on the grid is driven by the confluence of several factors. This trend is also mirrored in the broader power industry where the centralized model for power generation with few large power plants is giving way to a more distributed generation model where many smaller onsite generators produce power that is consumed locally over a “micro-grid”.

A key factor favoring on-site generation is the potential for cheaper power than the grid, especially during peak hours. On-site generation also reduces transmission losses that in turn reduce the effective cost, because the power is generated close to where it is consumed. In addition, another factor favoring on-site generation is a requirement for many enterprises to use cleaner renewable energy sources, such as Apple’s mandate to use 100% clean energy in its data centers. Such a mandate is more easily achievable with the enterprise generating all or most of its power on-site, especially since recent advances the fuel cell technology of Bloom Energy make on-site generation economical and feasible.

Finally, the risk of service outages caused by the failure of the grid, as happened recently when thunderstorms brought down the grid causing a denial-of-service for Amazon’s AWS service for several hours [8], has provided greater impetus for on-site power generation that can sustain the data center for extended periods without the grid.

II. RELATED WORK

Cassandra is similar to Galileo in its network layout and storage capabilities, as both systems are designed around the DHT paradigm. Cassandra’s primary use case is the high throughput management of tabular, multidimensional information. The system allows users to create their own partitioning schemes, but the partitioning algorithm used directly affects information retrieval as well; for instance, using the random data partitioner backed by a hash algorithm does not allow range queries or future reconfiguration of the partitioning algorithm. Cassandra scales out linearly as more hardware is added and supports MapReduce computations.

Apache Hive is a data warehouse that runs on the Hadoop and HDFS platform. As an analysis platform, it is capable of a wide range of functionality, including summarizing datasets and performing queries [2]. Unlike Galileo and a number of other storage frameworks, the system is intended for batch use rather than online transaction processing (OLTP). In Hive, users can perform analysis using the HiveQL query language, which transforms SQL-like statements into MapReduce jobs that are executed across a number of machines in a Hadoop cluster. The Metastore, a system catalog, provides an avenue for storing pre-computed information about the data stored in the system. Hive emphasizes scalability and flexibility in its processing rather than focusing on low latency.

A considerable amount of research has been conducted on supporting query types beyond the standard get and put operations of DHTs. For instance, Gao and Steenkiste maps a logical, sorted tree containing data points to physical nodes, enabling range queries. Chawathe et. al outlines a layered architecture for DHTs wherein advanced query support is provided by a separate layer that ultimately decomposes the queries into get and put operations, decoupling the query processing engine from the underlying storage framework.

Popivanov and Miller explores the issues surrounding managing and performing similarity searches over large quantities of time series information. To effectively summarize large datasets, this approach employs several wavelets that can outperform the commonly-used Haar wavelet and accurately estimates values for a wide range of data types. Methods that rely on wavelets or synopses are generally very problem-or dataset-specific and can limit arbitrary queries.
BlinkDB [4] is an approximate query processing (AQP) system that extends the functionality found in the Hive query engine to support responsive approximate queries. The system performs sampling at two scales: a broad, random sample of the dataset, along with focused sampling over frequently accessed items. This hybrid approach requires much less information to be read from disk to compute query responses. Like Galileo, BlinkDB queries can specify time and error bounds.

FastRAQ [3] considers both the storage and retrieval aspects associated with range-aggregate queries. To manage the error bounds of these approximate queries, partitioning is based on stratified sampling: a threshold is used to control the maximum relative error for each segment of the dataset. Like Galileo, data is assigned hierarchically to groups and then physical nodes. Queries are resolved using adaptive summary statistics that are built dynamically based on the distributions of the data.

The Approximate QUery Answering System (AQUA) intends to provide estimated responses to queries by avoiding direct access to the data itself. The system collects synopsis data in a number of ways: observing new information as it arrives, periodically inspecting the underlying data warehouse, or directly contacting the data warehouse during a query [4]. AQUA returns its query results alongside an accuracy measure and can support continuous reporting, wherein more results can be streamed to a client as accuracy increases. Unlike Galileo, AQUA does not support time bounds or target error rates in queries. It is also designed for batch processing rather than online transaction processing and cannot respond to real-time changes in the dataset.

III. JOINT RESOURCE PROVISIONING FOR INTERNET DATACENTERS

Cloud computing based Internet applications have been increasingly popular in recent years. Meanwhile, cloud service providers such as Google and Microsoft have to budget many millions of dollars for their Internet datacenters (IDCs) annually, in particular, for energy costs. Thus, how to provide desirable cloud services at a low cost is an important issue to be addressed.

Researchers have proposed various schemes to reduce IDC energy consumption. Among them, the so-called “capacity right-sizing” is a promising direction, e.g., in [11]. The key idea is to provision servers dynamically based on the load of requests. Extra servers are proposed to be shut down or scheduled in a sleeping mode to save energy. In this paradigm, to determine a proper number of active servers, it is important to know the volume of load. For example, sophisticated statistical models are used to predict the load of a Microsoft datacenter that provides Live messenger services.

Just obtaining the load size information, however, is still far from a fine-grained load-awareness. In datacenters, there exist various jobs that have different traffic patterns and service requirements. The existing capacity right-sizing schemes mentioned above often focus on request-response interactive applications, which require a small service latency. In datacenters, besides those delay-sensitive jobs (SENs), there are also a large amount of delay-tolerant batch jobs, e.g., scientific computing jobs. Giving a higher priority to SENs, the “extra” servers can be utilized to process those delay tolerant jobs (TOLs) rather than shut down, which is often referred to as trough/valley filling. Some existing work has considered resource provisioning for TOLs jobs only, e.g., in [12] [1]. There are also some literatures considering both SENs and TOLs, e.g., [9], joint resource provisioning for SENs and TOLs has not been studied in-depth yet. For example, in [10], the authors consider capacity for interactive workloads as a given variable and optimize capacity for batch jobs only. In our paper, we fully consider energy costs and service requirements by SENs and TOLs, respectively, as well as their interactions. We design joint SEN and TOL provisioning schemes, where capacity for SENs and capacity for TOLs are both control variables. This is our first key contribution.

In addition to the prioritized service requirements of datacenter jobs, there are many other challenges in datacenter resource provisioning. On one hand, capacity demand of SENs and TOLs is time varying. Short-term SEN traffic dynamics cannot be avoided since SENs need to be served
promptly. Turning on/off servers incurs a large time latency, i.e., up to several minutes. Thus, one cannot tune the number of active servers based on the instantaneous capacity demand of SENs. More importantly, given a higher priority to serving SENs and the relatively static total server resource, available capacity for TOLs is random and usually difficult to predict or learn in statistics. Thus, joint capacity allocation for SENs and TOLs is challenging. Joint resource provisioning in datacenters with both SEN and TOL traffic dynamics is our second key contribution.

In this paper, we consider a set of geo-distributed IDCs. For distributed IDCs, load shifting brings both opportunities and constraints. First, due to service agility, different classes of SENs or TOLs may require different sets of IDCs. Moreover, IDCs may be heterogeneous in service rates and energy consumption for each class of SENs or TOLs. Thus, a wise load shifting scheme can improve service efficiency and reduce energy consumption. Second, electricity prices exhibit diversity in both location and time. As studied in [13], price aware load shifting can reduce energy costs significantly. In this paper, we leverage both location and temporal price diversity to reduce IDC energy costs. Different from server provisioning, load shifting can be performed in a small time scale, e.g., on the order of hundreds of milliseconds [14]. How to jointly and efficiently use server provisioning, load shifting and SEN/TOL capacity allocation, which have different time granularities, to provision SENs and TOLs for distributed IDCs is a challenging problem, which is our another key contribution.

We study joint resource provisioning for SENs and TOLs. Our goal is to guarantee QoS of SENs, i.e., by constraining SEN overloading probability and achieve a good delay performance for TOLs at a low cost. To achieve this goal, we design joint server provisioning, SEN load dispatching, TOL load shifting and SEN/TOL capacity allocation. The joint schemes are configured and optimized by a decision maker based on an integrated convex optimization model. The decision-maker determines the number of active servers, SEN load dispatching ratios, TOL load shifting amount and TOL capacity sharing ratios in a large time scale, e.g., on the order of tens of minutes. Then the joint schemes are executed at different time granularities. Server provisioning is performed with a large time interval, i.e., the same as that of the decision-maker computing system parameters. In a smaller time scale, e.g., hundreds of milliseconds, instantaneous SEN load dispatching is performed based on the current dispatching ratios computed by the decision-maker. When SENs arrive an IDC, capacity allocation is performed instantaneously to serve the SENs. TOL load shifting is also performed in a small time scale following the optimal configurations. Then, capacity allocation is performed instantaneously to provision TOLs based on both the remaining instantaneous capacity for TOLs and TOL capacity sharing ratios at each IDC.

We explicitly differentiate SENs and TOLs in IDCs. We consider joint SEN and TOL resource provisioning, with traffic dynamics of both SENs and TOLs considered. Both the large-time-scale, i.e., hourly and small-time-scale traffic dynamics, i.e., hundreds of milliseconds, of SENs are considered to capture the real-world traffic models.

We design joint server provisioning, SEN load dispatching, TOL load shifting and SEN/TOL capacity allocation schemes for geo-distributed IDCs with different time granularities. Our schemes minimize the total energy costs, assure the QoS for SENs and guarantee TOL queue stability. Note that we focus on TOL provisioning in this paper. Specifically, we propose a queue-based trough-filling scheme, named OrgQ. We also consider a back-pressure routing based TOL provisioning scheme, named SubQ and find its disadvantages in the scenario of geo-distributed IDCs with electricity price diversity. Moreover, to show the advantages of OrgQ, we also design benchmark schemes which do not leverage any TOL queue information, in both a stationary ergodic setting and a non stationary ergodic setting.

We perform extensive simulations to compare the performance of OrgQ to other schemes based on simulated traffic trace and real traffic trace. Our results show that OrgQ outperforms both the
benchmarks and SubQ, since it can achieve a better tradeoff between costs and queue delay. We also show various properties of our proposed schemes which help people better understand datacenter resource provisioning.

**IV. ISSUES ON RESOURCE PROVISIONING SCHEME**

Data center energy efficiency is achieved with demand based resource provisioning model and server power management mechanism. Data center jobs are categorized into two types based on their requirements. They are delay sensitive jobs (SENs) and delay tolerant jobs (TOLs). Immediate resource requirements are provided under the delay sensitive jobs model. Backup and maintenance operations are carried out under the delay tolerant jobs model. Joint SEN and TOL resource provisioning scheme is employed to manage delay sensitive and delay tolerant jobs. Joint server provisioning, SEN load dispatching, TOL load shifting and SEN/TOL capacity allocation scheme manages the data center jobs. The OrgQ scheme handles the resource allocation for the TOL jobs. The following issues are identified from the current resource provisioning schemes.

- The server manages the single application based queue only
- Virtualization effects are not analyzed
- Complex power cost management mechanism
- Priority factors are not focused in the queue management process

**V. COST EFFICIENT RESOURCE MANAGEMENT FRAMEWORK FOR HYBRID JOBS**

The Geo-Distributed Data Center (GDC) management system is build to handle job allocation with spatial and temporal parameters. Priority features are integrated to manage the delay sensitive jobs and delay tolerant jobs. Energy cost management process is carried out with traffic, time and location dynamics. The system manages the virtualization tasks against the data center allocation process. Delay sensitive and delay tolerant jobs are managed under the joint resource provisioning scheme. Priority and spatio temporal factors are used in the energy cost estimation process. Fault tolerant and durability features are provided in the resource provisioning process. The system is divided into five major modules. They are Cloud Data Centers, Geo-Distributed Data Centers, Job Request Management, Joint resource provisioning scheme and Multi Constrained Resource Provisioning Scheme. Cloud data centers are deployed with storage and data resources. Geo-distributed data centers are build with spatial parameters. The job request management process collects the job requests from the clients. Delay sensitive and tolerant jobs are managed under the joint resource provisioning scheme. Spatial, temporal, energy cost and priority factors are used in the multi constrained resource management scheme.

Cloud data centers are constructed with a set of servers. Storage space and data resources are provided under the cloud data centers. Power management of the cloud data center is achieved with active and idle mode based operations. Cloud data center resources are provided with reference to the request and its priority levels. Cloud data centers are grouped to build Geo Distributed Data Centers (GDC) to provide resources for the users. Spatial parameters are considered in the Geo distributed data center construction process. Location and time factors are considered in the request redirection process. All the resources are provided with reference to the energy cost criteria. The job requests are collected from the clients for cloud resource allocation process. Delay sensitive jobs (SEN) and delay tolerant jobs (TOL) are received by the cloud data centers. The resources are allocated with reference to the job request category information. Jobs are also classified with spatial, temporal and priority values. The joint resource provisioning process is carried out with Joint SEN and TOL resource provisioning technique. The server provisioning is carried out with SEN/TOL load handler and SEN/TOL capacity allocation models. The OrgQ scheme is used to assign resources for the delay tolerant jobs. Energy cost
values are calculated with similar levels for all data servers. The cloud resources are allocated with job
category, internal queue status, priority, location, time and energy cost constraints. The multi
constrained resource provisioning scheme is build to handle the multiple application based server
allocation process. Traffic, priority and spatio temporal factors are used in the energy cost estimation
process. Fault tolerant support is provided to transfer jobs during resource failures.

VI. EXPERIMENTAL ANALYSIS

The resource scheduling schemes are build to allocate resources for the jobs in Geo distributed
data centers. The Joint Resource Provisioning Scheme (JRPS) is build to allocate resources for the
sensitive and tolerant jobs. The Multi Constrained Resource Provisioning Scheme (MCRPS) is used to
allocate resources with multi constraint model. The system is tested with three performance measures.
They are Makespan time, Resource Utilization Rate and Power Saving rate parameters.

![Makespan Time analysis between JRPS and MCRPS](image)

Figure No: 6.1. Makespan time analysis between Joint Resource Provisioning Scheme (JRPS) and
Multi Constrained Resource Provisioning Scheme (MCRPS)

The time interval between the job submission process and job completion process is estimated as
makespan time for the computational jobs. The makespan time between Joint Resource Provisioning
Scheme (JRPS) technique and Multi Constrained Resource Provisioning Scheme (MCRPS) technique is
shown in figure 6.1. The Multi Constrained Resource Provisioning Scheme (MCRPS) technique reduces
the makespan time 20% than the Joint Resource Provisioning Scheme (JRPS) technique. The resource
utilization rate is estimated with available resource and shared resource levels for the computational
jobs. The Resource Utilization Rate between Joint Resource Provisioning Scheme (JRPS) technique and
Multi Constrained Resource Provisioning Scheme (MCRPS) technique is shown in figure 6.2. The Multi
Constrained Resource Provisioning Scheme (MCRPS) technique increases the Resource Utilization Rate
10% than the Joint Resource Provisioning Scheme (JRPS) technique. The power saving rate analysis is
carried out to estimate the energy consumption for the shared resources under the cloud environment.
The power saving rate analysis rate between Joint Resource Provisioning Scheme (JRPS) technique and
Multi Constrained Resource Provisioning Scheme (MCRPS) technique is shown in figure 6.3. The Multi
Constrained Resource Provisioning Scheme (MCRPS) technique increase the Power Saving rate 45%
than the Joint Resource Provisioning Scheme (JRPS) technique.
Figure No: 6.2. Resource Utilization Rate Analysis between Joint Resource Provisioning Scheme (JRPS) and Multi Constrained Resource Provisioning Scheme (MCRPS)

Figure No: 6.3. Power Saving Analysis between Joint Resource Provisioning Scheme (JRPS) and Multi Constrained Resource Provisioning Scheme (MCRPS)

VII. CONCLUSION

Geo-distributed data centers (GDC) are built to provide resources for the data access services. Delay sensitive jobs (SENs) and delay tolerant jobs (TOLs) are managed by the joint resource provisioning scheme. The joint resource provisioning scheme is enhanced to manage multiple applications under a single server environment. Energy cost management, priority and virtualization factors are integrated in the resource provisioning process. The resource provisioning scheme is built for the Geo-Distributed Data Center (GDC) environment. Delay sensitive jobs and delay tolerant jobs are managed with priority factors. The energy cost management is upgraded with spatial and temporal features. The data center allocation time is reduced with high utilization rates.

REFERENCES

[1] M. Polverini, A. Cianfrani, S. Ren and A. V. Vasilakos, “Thermal-Aware Scheduling of Batch Jobs in Geographically Distributed Data Centers,” IEEE Transactions on Cloud Computing, VOL. 2, NO. 1, Jan-Mar, 2014.

[2] Matthew Malensek, Sangmi Pallickara and Shrideep Pallickara, “Fast, Ad Hoc Query Evaluations over Multidimensional Geospatial Datasets”, IEEE Transactions on Cloud Computing, January-March 2017.
[3] X. Yun, G. Wu, G. Zhang, K. Li and S. Wang, “Fastraq: A fast approach to range-aggregate queries in big data environments,” Cloud Computing, IEEE Transactions on, 2014.

[4] S. Agarwal, B. Mozafari, A. Panda, H. Milner, S. Madden and I. Stoica, “BlinkDB: queries with bounded errors and bounded response times on very large data,” in Proceedings of the 8th ACM European Conference on Computer Systems. ACM, 2013, pp. 29–42.

[5] J. Lowesohn. Apple’s main data center to go fully renewable this year. 2012.

[6] Z. Liu, et.al, “Greening geographical load balancing”, in ACM SIGMETRICS, 2011.

[7] K. Fehrenbacher. ebay to build huge bloom energy fuel cell farm at data center. 2012.

[8] K. Fehrenbacher. Is it time for more off-grid options for data centers?. 2012.

[9] H. Xu and B. Li, “Temperature Aware Workload Management in Geo-distributed Datacenters,” in ICAC 2013.

[10] Z. Liu, et.al, “Renewable and Cooling Aware Workload Management for Sustainable Data Centers,” in ACM SIGMETRICS, 2012.

[11] D. Xu, X. Liu and B. Fan, “Efficient Server Provisioning and Offloading Policies for Internet Datacenters with Dynamic Load Demand,” in IEEE Transactions on Computers, 2014.

[12] D. Xu and X. Liu, “Geographic Trough Filling for Internet Datacenters,” in IEEE INFOCOM, Mini-Conference, 2012.

[13] N. Buchbinder, N. Jain and I. Menache, “Online Job-Migration for Reducing the Electricity Bill in the Cloud,” in IEEE INFOCOM, 2011.

[14] R. Stanojevic and R. Shorten, “Distributed dynamic speed scaling,” in IEEE INFOCOM, 2010.