Empirical Workload and Energy Management for the Cloud

G. Kesidis and B. Urgaonkar
CSE and EE Depts, The Pennsylvania State University, {gik2,buu1}.psu.edu

I. Overview of Our Research

We have been pursuing research in two broad areas in cloud computing, partly funded by past and ongoing NSF grants: (i) power management in data centers [8], [9], [10], [7], and (ii) performance modeling and optimization of cloud workloads [6], [11].

Power management for data centers: Given the large and growing contribution of electric power to the costs of building and operating data centers, power management for data centers is now widely recognized to be an important research area. We have been working on several threads in the area of power management for data centers.

As one example, we have been studying online control techniques to lower the electric utility costs for both data centers and their tenants with an emphasis on the complexities raised by (i) real-world electricity pricing schemes, (ii) difficulty in predicting workloads effectively, and (iii) scalability limitations inherent in state-of-the-art work. Whereas a significant amount of research exists on data center power cost optimization, most existing formulations have tended to oversimplify the tariff schemes offered by electricity utilities. A salient example is a pricing scheme that charges the consumer not only for the energy consumption (which might itself be based on time-varying prices) but additionally for the peak power consumption over the billing cycle (say a month), and the latter may be further complicated by tiering[1] Such pricing schemes result in significantly more complicated power cost optimization problems (and systems design and implementation issues). Another limitation in the state-of-the-art is that most of the proposed works are generally "point solutions" that consider a very small subset of the overall gamut of power control knobs (e.g., Dynamic Voltage-Frequency Scaling for CPUs, scheduling and load balancing, shutting IT equipment off, migration, modulating the quality of solution for some applications such as MPEG video streaming, among others). Perhaps a key reason behind this is that incorporating the many knobs that exist in today's data centers results in control problems that suffer from the well-known "curse of dimensionality." We are working towards control techniques and system design that overcome all of these shortcomings. Our control techniques make use of a variety of "knobs" ranging from virtual machine pause/delay, server shutdown/restart, to workload migration and consolidation.

Performance modeling and resource allocation: We have also been studying problems related to performance modeling and control of cloud workloads. One specific example concerns the analytical modeling of MapReduce-like parallel-processing applications. Advanced and customized techniques of quality-of-service (QoS) management are being proposed by *industry* for cloud applications, particularly for data centers operating Software Defined Networking (SDN) [2]. Here, we plan to tap into a vast prior literature on modeling, performance evaluation and control of parallel-processing systems. In a preliminary study [6], we extend the compact (hence practically compelling) "network calculus" framework of QoS management to MapReduce, including a numerical study based on (limited disclosed) Facebook datasets [4], [1].

II. Why NSF's Experimental Clouds for Our Research?

A common concern cross-cutting both of our research thrusts described above is to *let real-world data inform our modeling and design*. We have only been partially successful in our efforts

[1] Such a tariff may be directly sold by the electric utility (ISO) or via electricity brokers offering hedges to simplify spot-pricing.

to achieve this, and are hoping to benefit from access to NSF's upcoming cloud platforms in this regard.

Briefly, to validate the efficacy of our solutions, to date we have used the following "real-world" datasets: Facebook [4], [1] in [10], [6]; Google [5] in [10], [7]; and more detailed workload data of a private, commercial data center through our IBM Zurich collaborators in [11]. For workloads with demonstrated short-term predictability, we have devised novel control techniques based on Markov Decision Theory and Model Predictive Control and evaluated our techniques with publicly available power consumption traces for data centers [8], [9]. However, many data center workloads exhibit very complex temporal behavior that brings into question the efficacy of traditional prediction techniques and power/resource management informed by such predictions. For example, we find strong evidence for such complexity in our ongoing collaboration with researchers at IBM Research Zurich, wherein we are analyzing traces from IBM's private commercial clouds that host a wide variety of tenants (with resource allocations ranging from just a few machines to several hundreds) [11]. When workload predictors are of limited use (as we find for several tenants in IBM's private clouds), it may be desirable to consider "myopic" control techniques that do not rely upon predictions based on long-term history.

Existing publicly available datasets, *e.g.*, [5], [4], [1] have only limited information, so it is difficult to accurately characterize workload based on them for performance study. The CMU experimental cloud [3] is rather small scale. The work with IBM Zurich work is on-going and *may* eventually lead to experimental prototypes tested on large-scale commercial data centers; however, the associated datasets are confidential.

III. Proposed Use of NSF's Experimental Clouds

NSF's developing experimental cloud testbed offers the possibility for real-world *dynamic* experimentation with *prototypes* of proposed mechanisms for performance and cost effective and secure workload and energy management in the cloud. We also want to mount our prototypes on a centralized SDN platform of network configuration and control (*e.g.*, OpenDaylight (ODL), OP-NFV).

References

[1] Y. Chen, A. Ganapathi, R. Griffith, and R. Katz. The Case for Evaluating MapReduce Performance Using Workload Suites. In *Proc. IEEE MASCOTS*, 2011.

[2] Cisco's Network Architecture Geeks (NAG) Workshops, Nov. 2013 and Oct. 2014.

[3] CMU's OpenCloud Hadoop cluster trace. http://ftp.pdl.cmu.edu/pub/datasets/hla/dataset.html.

[4] Facebook workload repository. https://github.com/SWIMProjectUCB/SWIM/wiki/Workloads-repository, 2009-2010.

[5] Google Cloud Platform Blog, Sneak peek: Google Cloud Dataflow, a Cloud-native data processing service. http://googlecloudplatform.blogspot.com/2014/06/sneak-peek-google-cloud-dataflow-a-cloud-native-data-processing-service.html, June 26, 2014.

[6] G. Kesidis, B. Urgaonkar, Y. Shan, S. Kamarava, J. Liebeherr. Network calculus for parallel processing. *in preparation*; earlier version at Arxiv Technical Report 1409.0820, Sept. 2 2014.

[7] N. Nasiriani, C. Wang, B. Urgaonkar, G. Kesidis. Fair peak-workload management by a virtual electric utility in the cloud. *in preparation*.

[8] C. Wang, B. Urgaonkar, Q. Wang, G. Kesidis and A. Sivasubramaniam. Data Center Cost Optimization Via Workload Modulation Under Real-World Electricity Pricing. In *Proc. ACM/IEEE Int'l Conf. on Utility and Cloud Computing (UCC)*, Desden, Germany, Dec. 2013.

[9] C. Wang, B. Urgaonkar, G. Kesidis, U.V. Shanbhag, Q. Wang. A Case for Virtualizing the Electric Utility in Cloud Data Centers. In *Proc. USENIX HotCloud*, Philadelphia, June 2014.

[10] C. Wang, B. Urgaonkar, G. Kesidis, and Q. Wang. A Hierarchical Demand Response Framework for Data Center Power Cost Optimization Under Real-World Electricity Pricing. In *Proc. IEEE MASCOTS*, Paris, Sept. 2014.

[11] C. Wang, R. Birke, L.Y. Chen, B. Urgaonkar, G. Kesidis. Myopic control of workload in the cloud. *in preparation*.