A Spot Capacity Market to Increase Power Infrastructure Utilization in Multi-Tenant Data Centers

Mohammad A. Islam
University of California, Riverside

Xiaoqi Ren
California Institute of Technology

Shaolei Ren
University of California, Riverside

Adam Wierman
California Institute of Technology

1 INTRODUCTION

Scaling up power infrastructures to accommodate growing data center demand is one of the biggest challenges faced by data center operators today. It incurs a huge capital expense of US$10-25 to build each watt of IT critical power capacity. Additional constraints, such as local grid capacity and long time-to-market cycle, are also limiting the data center capacity expansion, as attested to by the recent data center supply shortfall in Silicon Valley.

Traditionally, data center power infrastructure is sized to support the servers' maximum aggregate power demand. Nonetheless, the power demands of servers rarely peak simultaneously, thus resulting in a low average utilization of the scarce power capacity.

More recently, data center operators have sought to overcome this inefficiency by aggressively using capacity oversubscription, i.e., by deploying more servers than what the power and/or cooling capacity allows and applying peak shaving techniques to handle the resulting emergencies [1]. However, data center power infrastructure is still largely under-utilized today, wasting more than 15% of the capacity on average. This is not due to the lack of data center capacity demand, but due to the fluctuation of the aggregate server power demand that does not always stay at high levels, whereas the infrastructure is provisioned to sustain a high demand. Consequently, there exists a varying amount of unused power capacity over time, which we refer to as *spot (power) capacity*. Fig. 1 illustrates the fluctuating power usage and resulting spot capacity in a large server cluster.

Spot capacity is common in data centers and, if properly utilized, can improve application performances at runtime. For example, recent studies have proposed to dynamically allocate spot capacity to servers/racks for performance boosting via power routing [2] and "soft fuse" [3].

Importantly, all the prior research on exploiting spot capacity for improving performance has focused on an owner-operated data center, where the operator fully controls the servers. In contrast, *our goal is to develop an approach for exploiting spot capacity in multi-tenant data centers.* Multi-tenant data centers (also commonly called colocation data centers) host multiple tenants in a shared facility, each managing their *own* physical servers while the operator is only responsible for non-IT infrastructure support (like power and cooling). There are nearly 2,000 multi-tenant data centers in the U.S., serving almost all industry sectors.

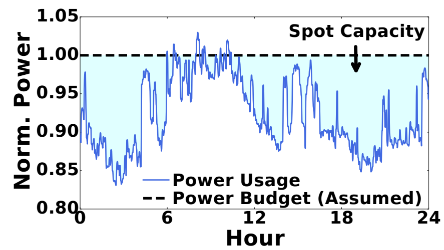


Figure 1: Power trace of a production server cluster in a top-brand IT company.

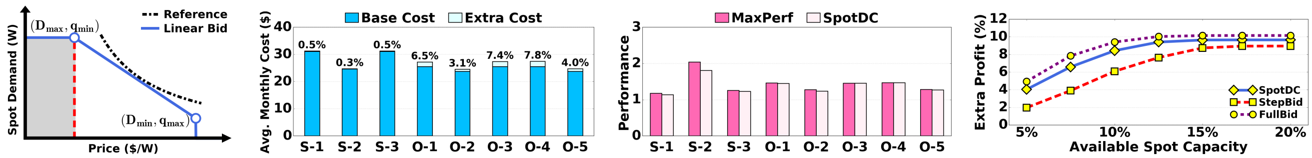
The potential value of spot capacity in multi-tenant data centers parallels that in owner-operated data centers. Concretely, in practice, more than 50% of the servers in data centers run delay-tolerant workloads (like batch data processing) [1] and hence are well poised to opportunistically utilize spot capacity for performance improvement. Additionally, the operator can make extra profit by offering spot capacity to tenants on demand. Despite these benefits, exploiting spot capacity is more challenging and requires a significantly different approach in multi-tenant data centers than in owner-operated data centers, because the operator has no control over tenants' servers, let alone the knowledge of which tenants need spot capacity and by how much.

In this abstract, we present a novel market approach, called Spot Data Center capacity management (SpotDC), which leverages demand function bidding and dynamically allocates spot capacity to tenants' racks to improve performance. Our work is motivated by other spot markets (e.g., cognitive radio and Amazon EC2). However, market design for spot power in multi-tenant data centers has a variety of multifaceted challenges. First, spot capacity is allocated to tenants on a rack level, but the operator does not know when/which racks need spot capacity and by how much. Second, tenants' rack-level power usage can vary flexibly to achieve different performances, but extracting the *elastic* spot capacity demand at scale at runtime can be very challenging, especially in a large data center with thousands of racks. Finally, practical constraints (e.g., multi-level power capacity and heat density) require the operator to set market prices in a new way.

Our design of SpotDC addresses each of these challenges. First, it has a low overhead: only soliciting four bidding parameters for each rack that needs spot capacity. Second, it quickly computes spot capacity allocation under practical constraints, without compromising reliability. Finally, as demonstrated in realistic settings, SpotDC benefits both the tenants and the operator: tenants improve

SIGMETRICS '17, June 5–9, 2017, Urbana-Champaign, IL, USA

ACM ISBN ACM ISBN 978-1-4503-5032-7/17/06.
https://doi.org/http://dx.doi.org/10.1145/3078505.3078542



(a) Demand function (b) Tenant's cost (c) Normalized performance (d) Operator's profit

Figure 2: Piece-wise linear demand function and evaluation result.

performance by 1.2–1.8x (on average) at a marginal cost increase, while the operator can increase its profit by 9.7%.

2 DESIGN OF SPOTDC

The core of SpotDC is to leverage a new demand function bidding approach to extract tenants' rack-level spot capacity demand elasticity at runtime.

Demand function captures how the demand varies as a function of the price. To allocate spot capacity, the operator needs to know tenants' rack-wise demand functions for spot capacity at runtime. One might think that the operator can predict tenants' rack-level spot capacity demand functions, but this is very difficult because the demand can be highly dynamic and there can be hundreds or even thousands of racks. Another straightforward approach is to solicit tenant's complete rack-level demand curve under all possible prices (labeled as "Reference" in Fig. 2a), but this has a high communication overhead and is rarely used in real markets.

Demand function bidding. In practice, *parameterized* demand function bidding is commonly applied in different markets, such as Amazon EC2. Here, we propose a new parameterized demand function which, as illustrated by "Linear Bid" in Fig. 2a, *approximates* the actual demand curve using three line segments: first, a horizontal segment: tenant specifies its maximum spot capacity demand for a rack as well as the market price it is willing to pay; second, a linearly decreasing segment: the demand decreases linearly as the market price increases; and third, a vertical segment: the last segment indicates tenant's maximum acceptable price and the corresponding minimum demand. As shown in Fig. 2a, our piece-wise linear demand function for rack r is uniquely determined by four parameters: $\mathbf{b}_r = \{(D_{\max,r}, q_{\min,r}), (D_{\min,r}, q_{\max,r})\}$, where $D_{\max,r}$ and $D_{\min,r}$ are the maximum and minimum spot capacity demand, and $q_{\min,r}$ and $q_{\max,r}$ are the corresponding prices, respectively.

Our proposed demand function is simple yet can reasonably extract tenants' demand elasticity. In fact, it is a demand function widely studied for theoretical analysis. It also represents a *midpoint* between soliciting the complete demand curve and StepBid used by Amazon, which is simpler but cannot extract spot capacity demand elasticity (i.e., a tenant's spot capacity demand can only be either 100% or 0% satisfied). Our experiment results show that, using our demand function, the operator's profit is much higher than that using StepBid and also fairly close to the optimal profit when the complete demand curve is solicited.

Spot capacity allocation. The following three steps describe the spot capacity allocation process.

Step 1: *Demand function bidding.* Participating tenants decide, at their own discretion, bidding parameters for each rack that needs spot capacity and submit the bids to the operator.

Step 2: *Market clearing.* Upon collecting the bids, the operator sets the market price $q(t)$ to maximize profit subject to multi-level power and heat density constraints. This can be done very quickly through a simple search over the feasible price range.

Step 3: *Actual spot capacity allocation.* Given the market price $q(t)$ plugged into the demand function, each tenant knows its per-rack spot capacity and can use additional power up to the allocated spot capacity.

3 EVALUATION

We build a scaled-down testbed with 10 tenants grouped into two clusters. We consider delay-sensitive web search and web service workloads, as well as delay-tolerant Hadoop and graph analysis workloads. We model each tenant's power and performance to devise their demand bidding strategies for different workloads in order to improve performances at a low cost. We use real-world traces and conduct a one-month long simulation to evaluate SpotDC.

Figs. 2b and 2c show the additional cost incurred by each tenant participating in SpotDC and the corresponding performance improvement, respectively. Here, the base cost is tenant's power-related cost without SpotDC. The performance is normalized to that without SpotDC, while MaxPerf assumes a complete control of tenants' servers to allocate spot capacity for total performance maximization as if in an owner-operated data center. We see that for a marginal cost increase, the participating tenants are able to significantly improve their performance. Fig. 2d shows that the operator can make extra profit by exploiting spot capacity using SpotDC. We also see that SpotDC outperforms StepBid in terms of the operator's profit and is very close to FullBid that solicits a complete demand curve from each tenant.

To conclude, SpotDC *is a novel market approach for exploiting spot capacity to increase power infrastructure utilization in multi-tenant data centers and turn spot capacity into a "win-win" resource: the operator makes additional profit, while the tenants improve performance at a low cost.*

ACKNOWLEDGMENTS

This work was supported in part by the U.S. NSF under grants CNS-1551661, CNS-1565474, ECCS-1610471, AitF-1637598 and CNS-1518941, and the Resnick Sustainability Institute at Caltech.

REFERENCES

[1] M. A. Islam, X. Ren, S. Ren, A. Wierman, and X. Wang, "A market approach for handling power emergencies in multi-tenant data center," in *HPCA*, 2016.
[2] S. Pelley, D. Meisner, P. Zandevakili, T. F. Wenisch, and J. Underwood, "Power routing: Dynamic power provisioning in the data center," in *ASPLOS*, 2010.
[3] S. Govindan, J. Choi, B. Urgaonkar, and A. Sivasubramaniam, "Statistical profiling-based techniques for effective power provisioning in data centers," in *EuroSys*, 2009.