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

Mohammad A. Islam, Xiaoqi Ren, Shaolei Ren, and Adam Wierman

This work was supported in part by the U.S. NSF under grants CNS-1551661, CNS-1565474, CNS-1518941, CPS-154471, ECCS-1610471, and AitF-1637598.
Multi-tenant data centers
Multi-tenant data centers

Managed by operator

Utility (Primary)

Generator

UPS

ATS

PDU

...
Multi-tenant data centers

Managed by operator

Utility (Primary)

Managed by individual tenants

Generator

ATS

UPS

PDU

PDU

PDU
Multi-tenant data centers are everywhere

2,000+ in U.S.
Multi-tenant data centers are everywhere

Google, Amazon, MS, Fb…: 7.8%

Multi-tenant: 37%

Enterprise: 53%

2,000+ in U.S.
Who are using multi-tenant data centers?
Who are using multi-tenant data centers?

Giant IT companies

25% of Apple’s servers are in multi-tenant data centers
Who are using multi-tenant data centers?

Giant IT companies: Google, Apple, Facebook, Microsoft, ...

25% of Apple’s servers are in multi-tenant data centers

Large IT companies: Box, Akamai, Salesforce, ...

Who are using multi-tenant data centers?

Giant IT companies

25% of Apple’s servers are in multi-tenant data centers

Large IT companies

Internet of things Hybrid-cloud
Data center costs breakdown

<table>
<thead>
<tr>
<th>Amortized Cost</th>
<th>Component</th>
<th>Sub-Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>~45%</td>
<td>Servers</td>
<td>CPU, memory, storage systems</td>
</tr>
<tr>
<td>~25%</td>
<td>Infrastructure</td>
<td>Power distribution and cooling</td>
</tr>
<tr>
<td>~15%</td>
<td>Power draw</td>
<td>Electrical utility costs</td>
</tr>
<tr>
<td>~15%</td>
<td>Network</td>
<td>Links, transit, equipment</td>
</tr>
</tbody>
</table>

## Data center costs breakdown

<table>
<thead>
<tr>
<th>Amortized Cost</th>
<th>Component</th>
<th>Sub-Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>~25%</td>
<td>Infrastructure</td>
<td>Power distribution and cooling</td>
</tr>
<tr>
<td>~15%</td>
<td>Power draw</td>
<td>Electrical utility costs</td>
</tr>
</tbody>
</table>

Data center costs breakdown

<table>
<thead>
<tr>
<th>Amortized Cost</th>
<th>Component</th>
<th>Sub-Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>~25%</td>
<td>Infrastructure</td>
<td>Power distribution and cooling</td>
</tr>
<tr>
<td>~15%</td>
<td>Power draw</td>
<td>Electrical utility costs</td>
</tr>
</tbody>
</table>

Data center costs breakdown

<table>
<thead>
<tr>
<th>Amortized Cost</th>
<th>Component</th>
<th>Sub-Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>~25%</td>
<td>Infrastructure</td>
<td>Power distribution and cooling</td>
</tr>
<tr>
<td>~15%</td>
<td>Power draw</td>
<td>Electrical utility costs</td>
</tr>
</tbody>
</table>

Data center costs breakdown

Capital Expenditure (CapEx)

Operational Expenditure (OpEx)

<table>
<thead>
<tr>
<th>Amortized Cost</th>
<th>Component</th>
<th>Sub-Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>~25%</td>
<td>Infrastructure</td>
<td>Power distribution and cooling</td>
</tr>
<tr>
<td>~15%</td>
<td>Power draw</td>
<td>Electrical utility costs</td>
</tr>
</tbody>
</table>

CapEx > 1.5×OpEx!

Cost of infrastructure

$10-25 per Watt

Utility (Primary)

Generator

ATS

UPS

PDU

PDU

PDU

...
Underutilization in data centers
Underutilization in data centers
Underutilization in data centers

![Graph showing power usage over time with spot capacity and power budget lines.]
Increase infrastructure utilization
Increase infrastructure utilization

Exploit the “spot capacity”
Some inspirations

• “Power routing” in ASPLOS’10 and “soft fuse” in EuroSys’09
Some inspirations

• “Power routing” in ASPLOS’10 and “soft fuse” in EuroSys’09
• “Spot instances” from Amazon
Some inspirations

• “Power routing” in ASPLOS’10 and “soft fuse” in EuroSys’09
• “Spot instances” from Amazon

• “Preemptible VM” from Google Cloud
Spot capacity in multi-tenant data centers
Spot capacity in multi-tenant data centers

No centralized control
Spot capacity in multi-tenant data centers

No centralized control $\rightarrow$ Power routing,…
Spot capacity in multi-tenant data centers

No centralized control $\rightarrow$ Power routing,…

A market for spot capacity
Spot capacity in multi-tenant data centers

No centralized control $\rightarrow$ Power routing,…

A market for spot capacity

Tenants buy spot capacity from the data center operator
Spot capacity in multi-tenant data centers

- Flexibility for cost conscious tenants

![Power vs Time Graph]
Spot capacity in multi-tenant data centers

- Flexibility for cost conscious tenants

Peak-based subscription

![Graph showing power over time with a dashed line indicating peak-based subscription.](image)
Spot capacity in multi-tenant data centers

- Flexibility for cost conscious tenants

![Graph showing power consumption over time with peak-based and conservative subscription lines.](image)
Spot capacity in multi-tenant data centers

- Flexibility for cost conscious tenants

![Graph](image.png)

- Peak-based subscription
- Conservative subscription
- Spot capacity
Spot capacity in multi-tenant data centers

• Tenants:
  • tenants with insufficient capacity reservations can temporarily process its workloads without power capping (or cap power less frequently/aggressively than it would otherwise).
Spot capacity in multi-tenant data centers

• Tenants:
  • tenants with insufficient capacity reservations can temporarily process its workloads without power capping (or cap power less frequently/aggressively than it would otherwise).

• Operator:
  • Revenue from guaranteed capacity: not affected
  • Extra revenue from spot capacity
Spot capacity in multi-tenant data centers

• Tenants:
  • tenants with insufficient capacity reservations can temporarily process its workloads without power capping (or cap power less frequently/aggressively than it would otherwise).

• Operator:
  • Revenue from guaranteed capacity: not affected
  • Extra revenue from spot capacity

Spot capacity market is a **win-win** for both tenants and operator
Challenges

• Spot capacity is **limited** and **intermittent**
Challenges

• Spot capacity is limited and intermittent

• Tenants’ spot capacity need is dynamic and invisible to the data center operator
Challenges

• Spot capacity is **limited** and **intermittent**

• Tenants’ spot capacity need is **dynamic** and **invisible** to the data center operator

• Infrastructure constraints require fine granularity in spot capacity allocation (e.g., rack level)
Goal: A **scalable** and **runtime** design for spot capacity allocation
Problem formulation

• Goal: operator profit maximization

\[
\max_{q(t)} \quad q(t) \cdot \sum_{r \in S(t)} D_r (q(t)) .
\]
Problem formulation

• Goal: operator profit maximization

\[
\text{maximize } q(t) \cdot \sum_{r \in S(t)} D_r(q(t)).
\]
Problem formulation

• Goal: operator profit maximization

\[
\text{maximize } q(t) \cdot \sum_{r \in S(t)} D_r (q(t))
\]
Problem formulation

- Goal: operator profit maximization

\[
\text{maximize} \quad q(t) \cdot \sum_{r \in S(t)} D_r (q(t))
\]

- Infrastructure constraints:
  \[
  \text{Rack:} \quad D_r (q(t)) \leq P_r^R, \quad \forall r \in S(t)
  \]
  \[
  \text{PDU:} \quad \sum_{r \in S(t) \cap \mathcal{R}_m} D_r (q(t)) \leq P_m(t), \quad \forall m \in \mathcal{M}
  \]
  \[
  \text{UPS:} \quad \sum_{r \in S(t)} D_r (q(t)) \leq P_o(t)
  \]
How to solve it?

\[
\text{maximize } q(t) \cdot \sum_{r \in S(t)} D_r(q(t)).
\]

Rack:

\[D_r(q(t)) \leq P_r^R, \ \forall r \in S(t)\]

PDU:

\[\sum_{r \in S(t) \cap \mathcal{R}_m} D_r(q(t)) \leq P_m(t), \ \forall m \in \mathcal{M}\]

UPS:

\[\sum_{r \in S(t)} D_r(q(t)) \leq P_o(t)\]
How to solve it?

maximize \( q(t) \cdot \sum_{r \in S(t)} D_r (q(t)) \).

Rack:
\[ D_r (q(t)) \leq P_r^R, \ \forall r \in S(t) \]

PDU:
\[ \sum_{r \in S(t) \cap R_m} D_r (q(t)) \leq P_m(t), \ \forall m \in \mathcal{M} \]

UPS:
\[ \sum_{r \in S(t)} D_r (q(t)) \leq P_o(t) \]
How to solve it?

maximize \( q(t) \cdot \sum_{r \in S(t)} D_r (q(t)) \).

Rack:
\[ D_r (q(t)) \leq P_r^R, \quad \forall r \in S(t) \]

PDU:
\[ \sum_{r \in S(t) \cap R_m} D_r (q(t)) \leq P_m(t), \quad \forall m \in M \]

UPS:
\[ \sum_{r \in S(t)} D_r (q(t)) \leq P_o(t) \]

- Soliciting the demand curve → privacy and overhead
How to solve it?

\[
\text{maximize } q(t) \cdot \sum_{r \in S(t)} D_r (q(t)).
\]

- **Rack:** \( D_r (q(t)) \leq P^R_r, \forall r \in S(t) \)
- **PDU:** \( \sum_{r \in S(t) \cap R_m} D_r (q(t)) \leq P_m(t), \forall m \in M \)
- **UPS:** \( \sum_{r \in S(t)} D_r (q(t)) \leq P_o(t) \)

• Soliciting the demand curve \(\rightarrow\) privacy and overhead
• Pre-set price \(\rightarrow\) low level demand prediction
How to solve it?

\[
\begin{align*}
\text{maximize} \quad & q(t) \cdot \sum_{r \in S(t)} D_r (q(t)) \\
\text{Rack:} \quad & D_r (q(t)) \leq P_r, \quad \forall r \in S(t) \\
\text{PDU:} \quad & \sum_{r \in S(t) \cap \mathcal{R}_m} D_r (q(t)) \leq P_m(t), \quad \forall m \in \mathcal{M} \\
\text{UPS:} \quad & \sum_{r \in S(t)} D_r (q(t)) \leq P_o(t)
\end{align*}
\]

• Soliciting the demand curve → privacy and overhead
• Pre-set price → low level demand prediction
• Market approach → an in-between solution
SpotDC: spot capacity management

Operator

Tenants
SpotDC: spot capacity management

Spot capacity predictions

Operator

Tenants
SpotDC: spot capacity management

Spot capacity predictions

Response (bids)

Operator

Tenants
SpotDC: spot capacity management

Spot capacity predictions

Response (bids)

Price and actual spot power allocation

Operator

Tenants
SpotDC: spot capacity management

Spot capacity predictions

Response (bids)

Price and actual spot power allocation

Gain spot power

Operator

Tenants
Timings in SpotDC

1. Demand Bidding
2. Spot Capacity Prediction
3. Market Clearing
4. Using Spot Capacity
Demand bidding

• A piece-wise-linear bid
Demand bidding

- A piece-wise-linear bid
- Tenants only submit four parameters
Demand bidding

- A piece-wise-linear bid
- Tenants only submit four parameters
- Captures tenants’ demand elasticity
Spot capacity prediction

• Available spot capacity prediction: max - predicted
  • UPS and PDU level predictions: Use previous time slot usage as references.
Spot capacity prediction

- Available spot capacity prediction: max - predicted
  - UPS and PDU level predictions: Use previous time slot usage as references.
Spot capacity prediction

- Available spot capacity prediction: max - predicted
  - UPS and PDU level predictions: Use previous time slot usage as references.

Less than $\pm 2.5\%$ change 99% of the time
Evaluation methodology

- 10 tenants with sprinting (delay sensitive) and opportunistic (delay tolerance) workloads
- Using Dynamic voltage and frequency scaling (DVFS) for power scaling.
Evaluation methodology

- 10 tenants with sprinting (delay sensitive) and opportunistic (delay tolerance) workloads
- Using Dynamic voltage and frequency scaling (DVFS) for power scaling.
Evaluation methodology

- 10 tenants with sprinting (delay sensitive) and opportunistic (delay tolerance) workloads
- Using Dynamic voltage and frequency scaling (DVFS) for power scaling.
Performance evaluation
Performance evaluation

Sprinting tenants drive up the price
Performance evaluation
Performance evaluation

Sprinting tenants avoid SLO violations
Performance evaluation

- Sprinting tenants avoid SLO violations
- Opportunistic tenants gain throughput boost
Tenants’ benefit from SpotDC

![Graph showing performance with and without SpotDC]
Tenants’ benefit from SpotDC

Performance boosts with SpotDC
Operator’s extra profit

![Graph showing extra profit vs available spot capacity for different bidding strategies: SpotDC, StepBid, and FullBid.]
Operator’s extra profit

SpotDC is close to optimal allocation with full information
SpotDC: Spot capacity management

A market-based approach for providing spot capacity to tenants and helping operator further increase data center utilization
SpotDC: Spot capacity management

A market-based approach for providing spot capacity to tenants and helping operator further increase data center utilization

Simple, Scalable & Efficient