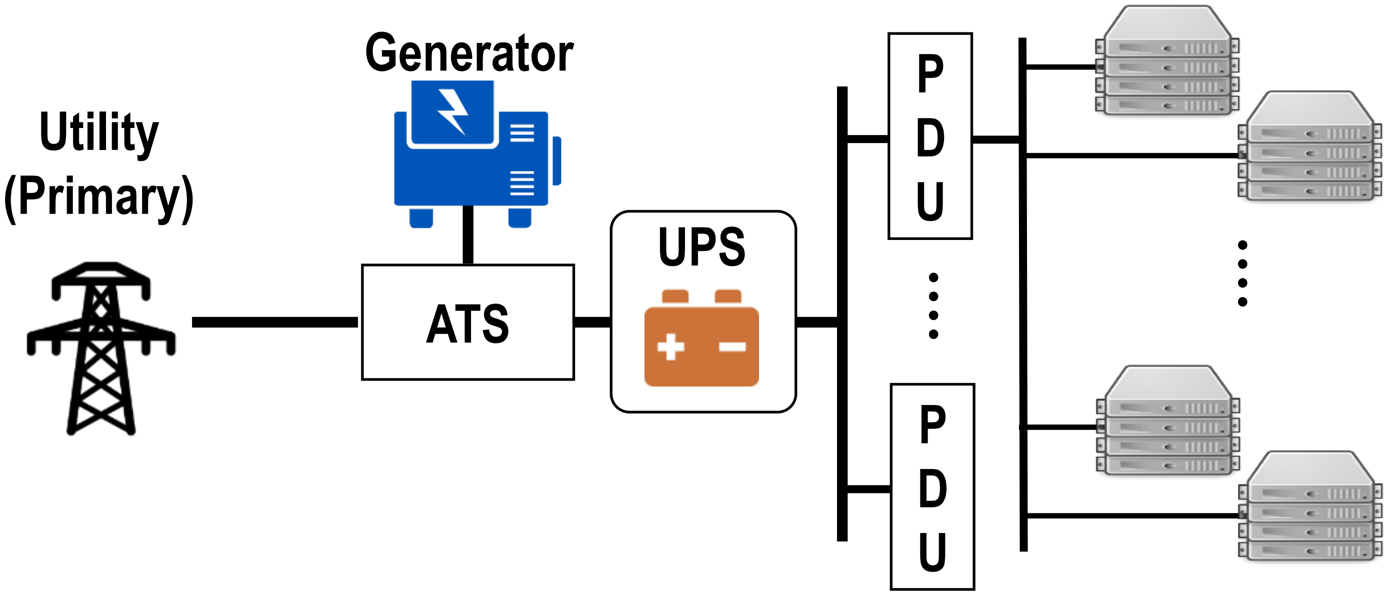
A Spot Capacity Market to Increase Power Infrastructure Utilization in Multi-Tenant Data Centers

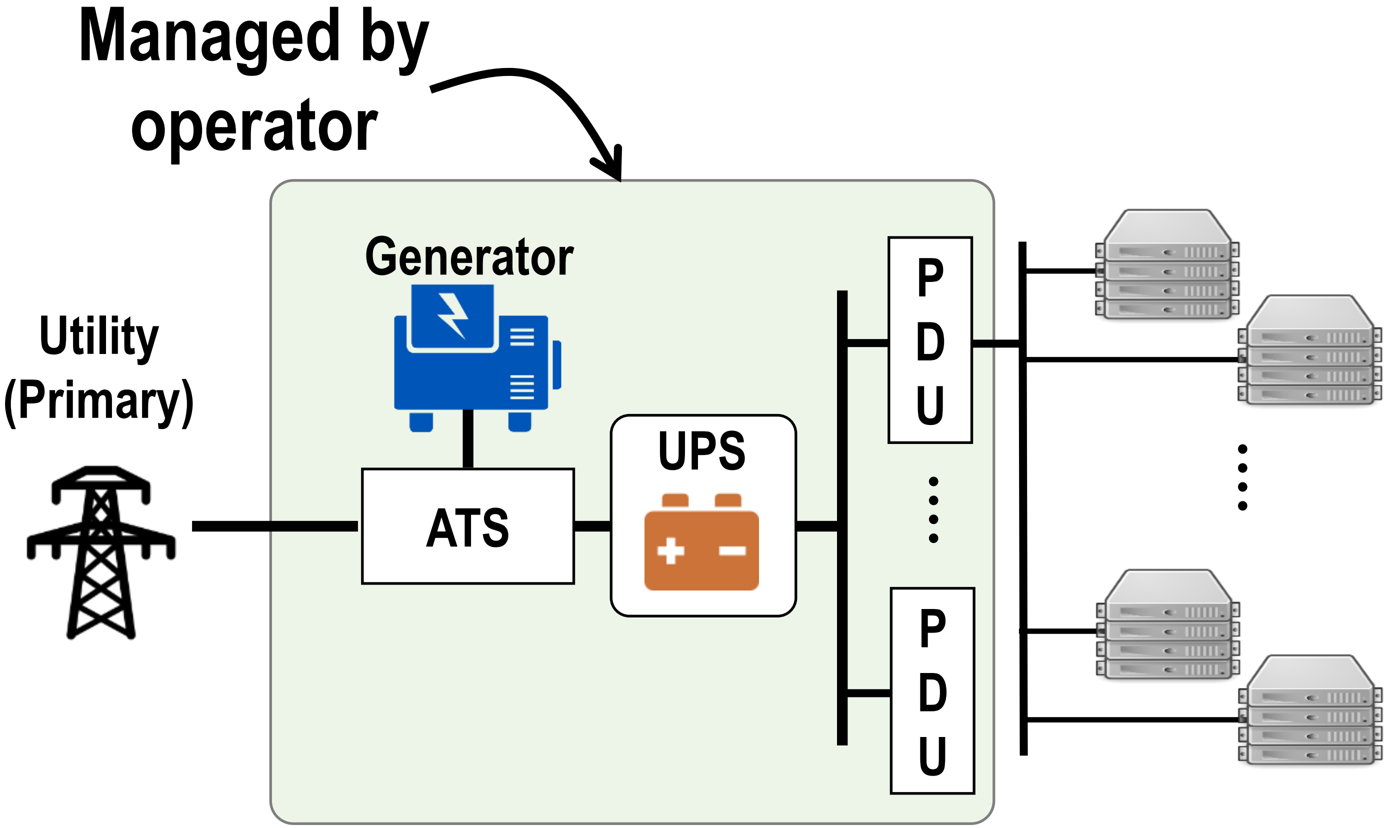
Mohammad A. Islam, Xiaoqi Ren, Shaolei Ren, and Adam Wierman

This work was supported in part by the U.S. NSF under grants CNS-1551661, CNS-1565474, CNS-1518941, CPS154471, ECCS-1610471, and AitF-1637598.

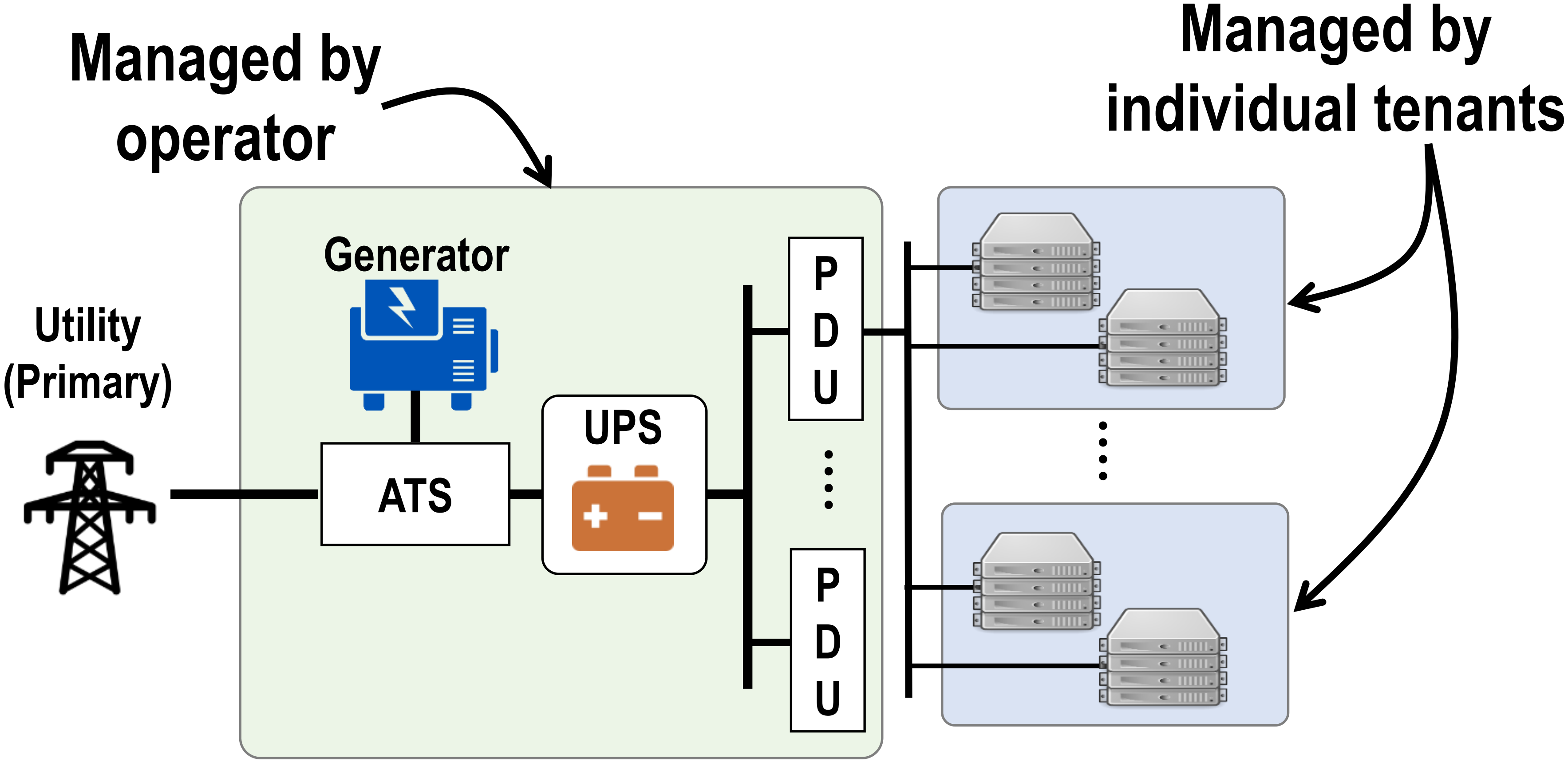
Multi-tenant data centers



Multi-tenant data centers

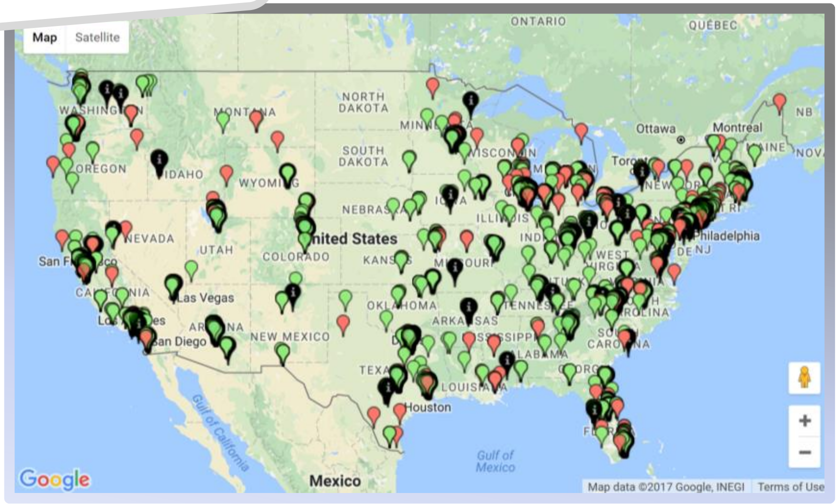


Multi-tenant data centers

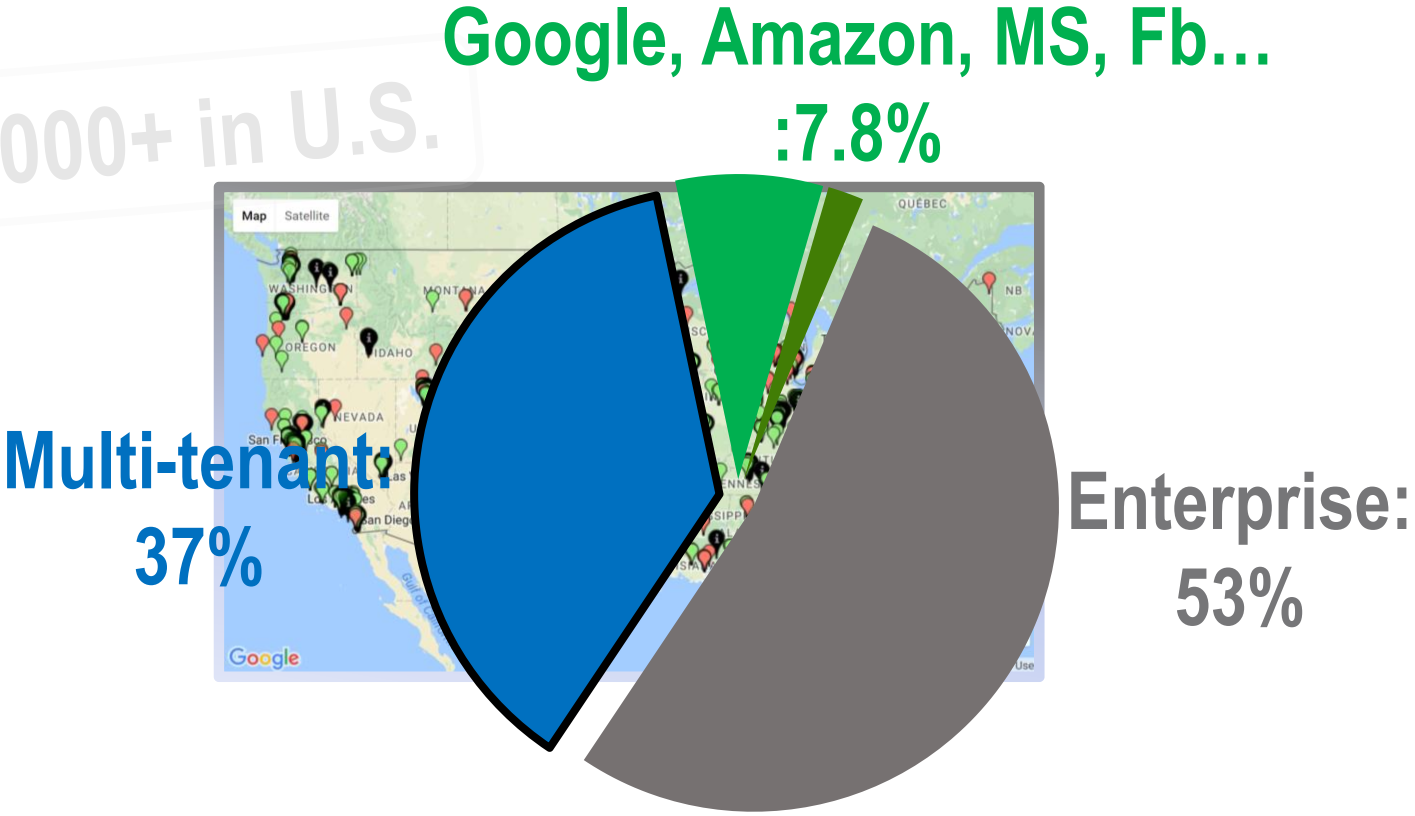


Multi-tenant data centers are everywhere

2,000+ in U.S.



Multi-tenant data centers are everywhere



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 ate in multi-tenant data centers

Who are using multi-tenant data centers?

Giant IT companies

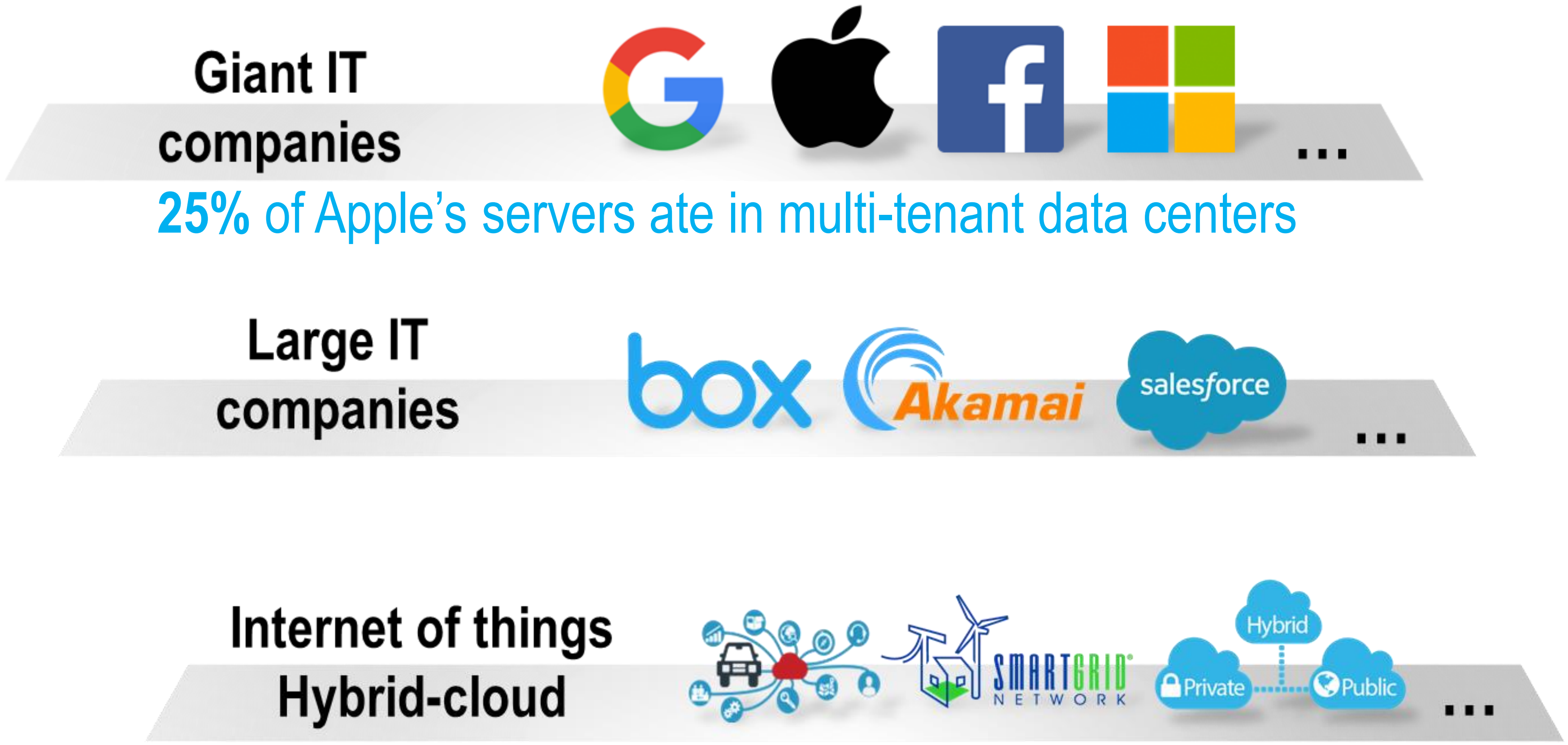
...

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

Large IT companies

...

Who are using multi-tenant data centers?



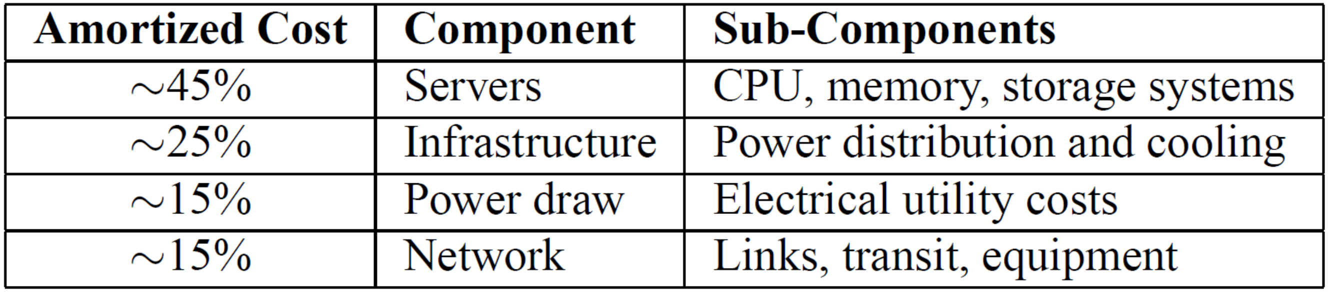
Giant IT companies

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

Large IT companies

Internet of things
Hybrid-cloud

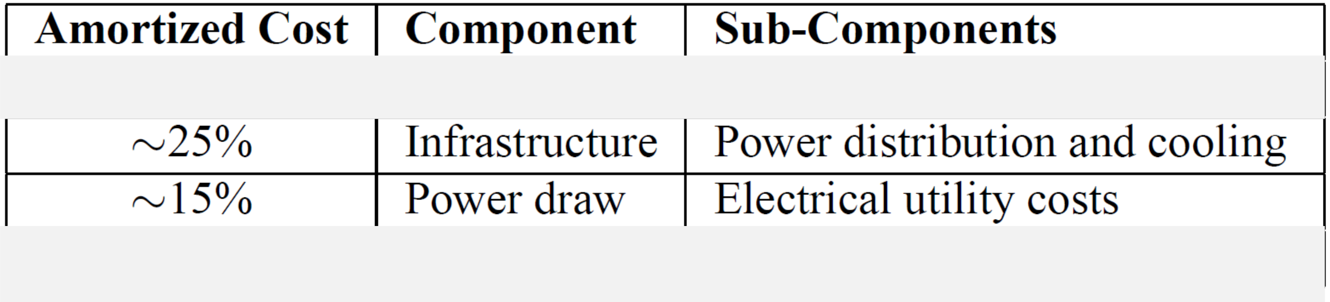
Data center costs breakdown



Amortized Cost	Component	Sub-Components
∼45%	Servers	CPU, memory, storage systems
∼25%	Infrastructure	Power distribution and cooling
∼15%	Power draw	Electrical utility costs
∼15%	Network	Links, transit, equipment

Source: A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. 2008. The cost of a cloud: research problems in data center networks. SIGCOMM Comput. Commun. Rev.

Data center costs breakdown

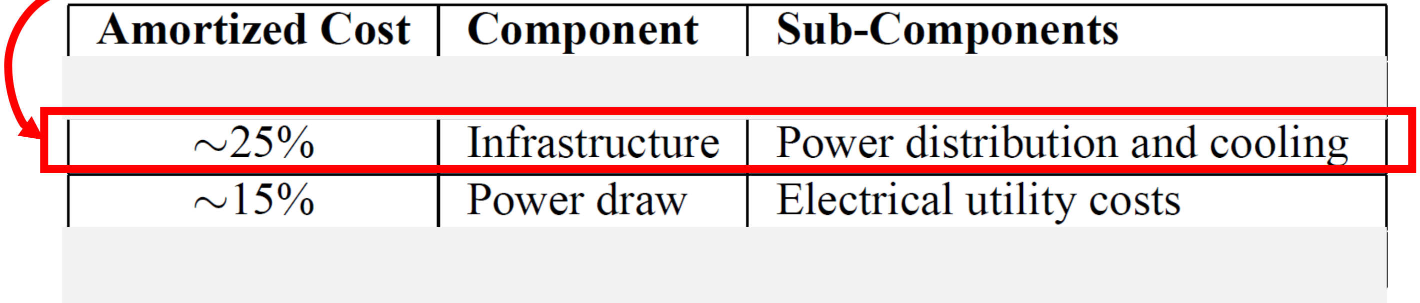


Amortized Cost	Component	Sub-Components
∼25%	Infrastructure	Power distribution and cooling
∼15%	Power draw	Electrical utility costs

Source: A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. 2008. The cost of a cloud: research problems in data center networks. SIGCOMM Comput. Commun. Rev.

Data center costs breakdown

Capital Expenditure
(CapEx)



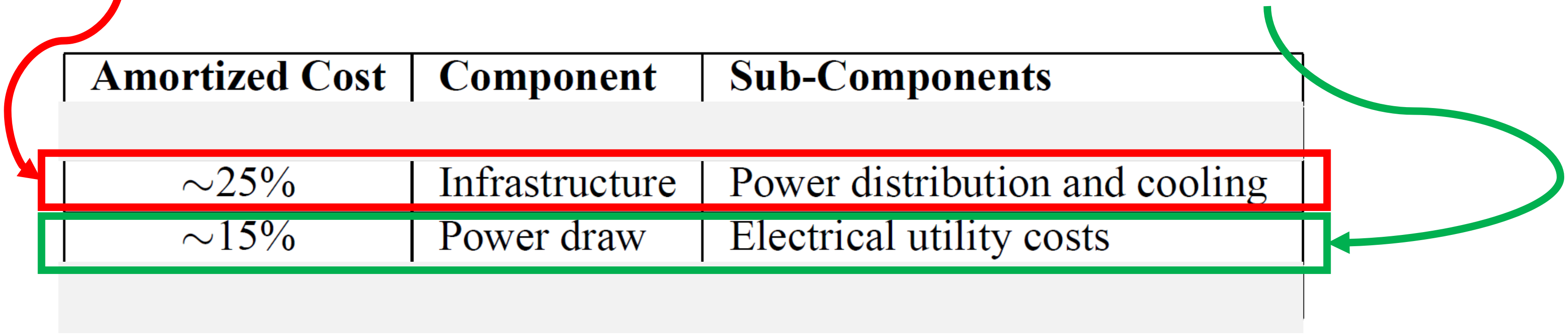
Amortized Cost	Component	Sub-Components
∼25%	Infrastructure	Power distribution and cooling
∼15%	Power draw	Electrical utility costs

Source: A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. 2008. The cost of a cloud: research problems in data center networks. SIGCOMM Comput. Commun. Rev.

Data center costs breakdown

Capital Expenditure
(CapEx)

Operational Expenditure
(OpEx)



Amortized Cost	Component	Sub-Components
∼25%	Infrastructure	Power distribution and cooling
∼15%	Power draw	Electrical utility costs

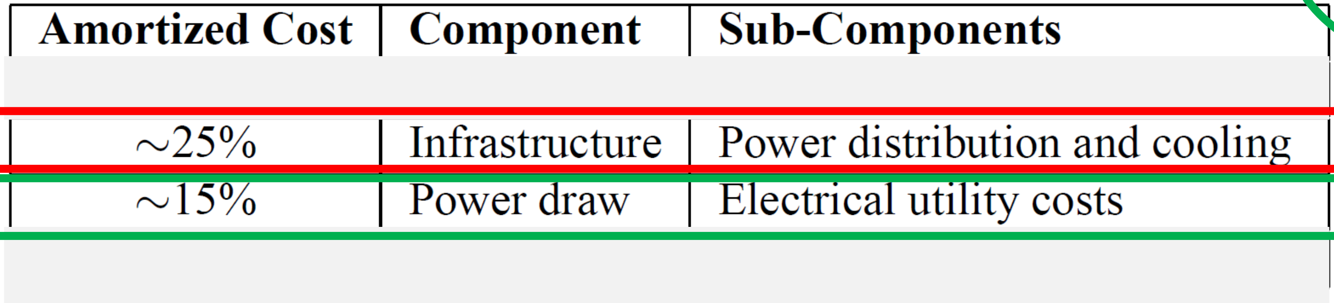
Source: A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. 2008. The cost of a cloud: research problems in data center networks. SIGCOMM Comput. Commun. Rev.

Data center costs breakdown

Capital Expenditure (CapEx)

CapEx > 1.5×OpEx !

Operational Expenditure (OpEx)

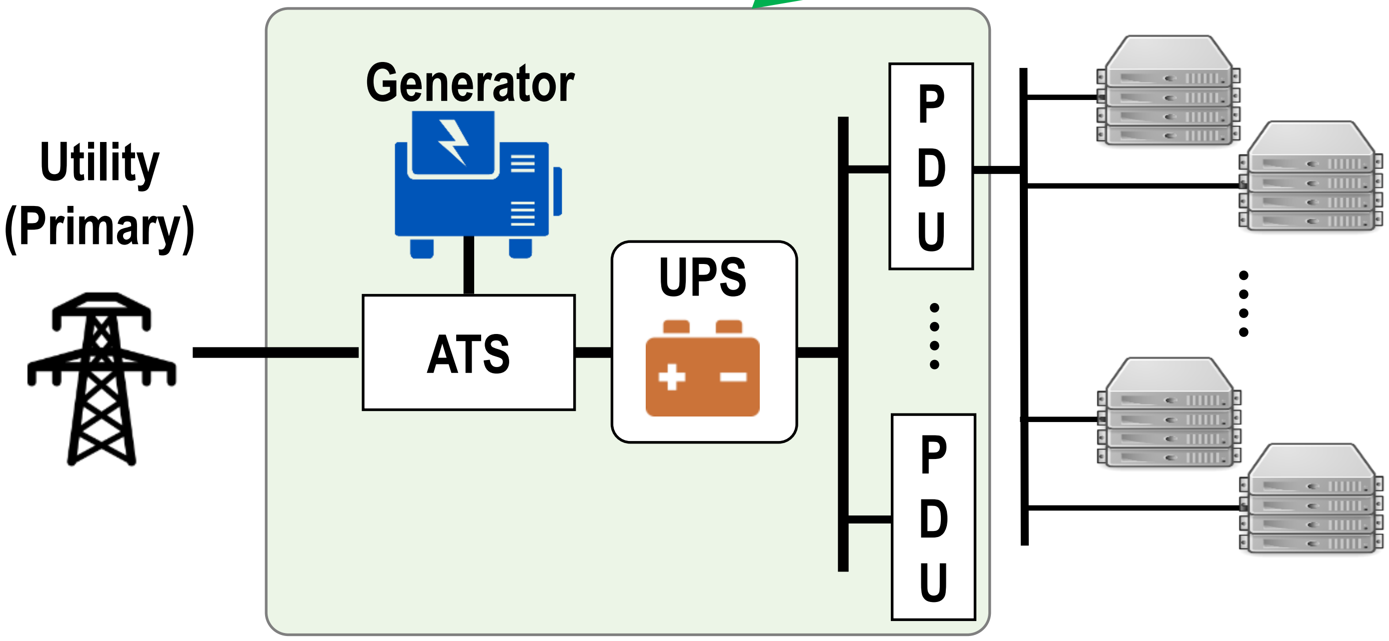


Amortized Cost	Component	Sub-Components
∼25%	Infrastructure	Power distribution and cooling
∼15%	Power draw	Electrical utility costs

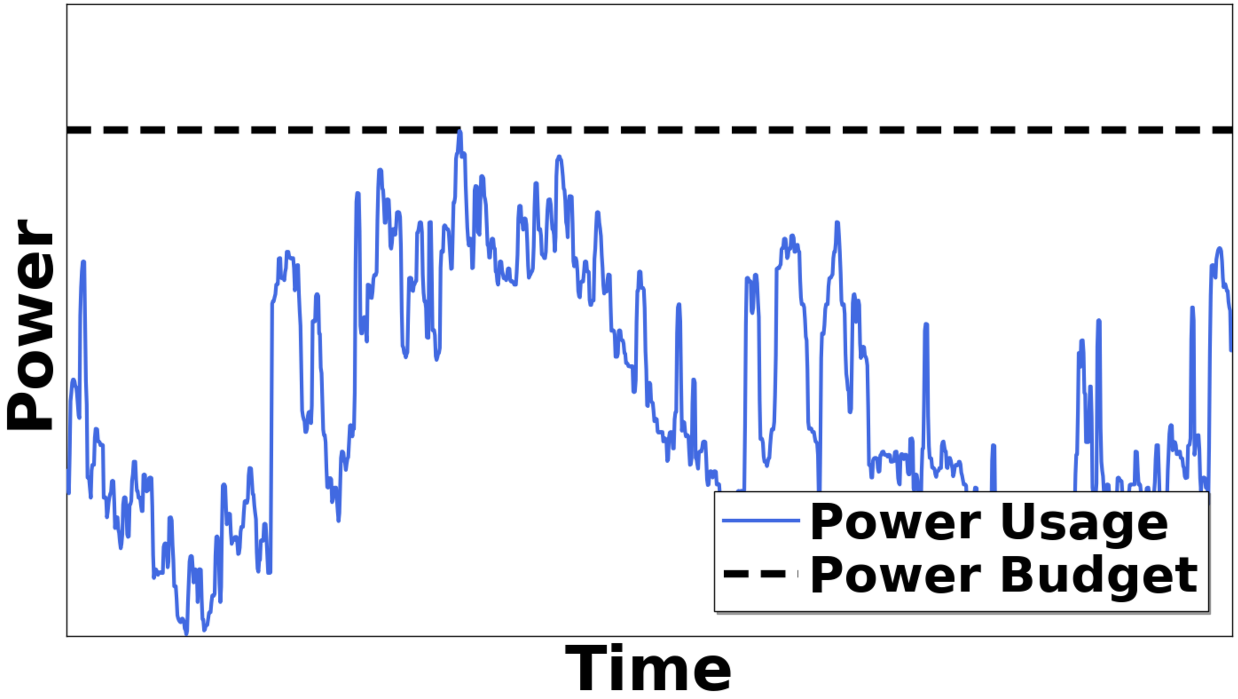
Source: A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel. 2008. The cost of a cloud: research problems in data center networks. SIGCOMM Comput. Commun. Rev.

Cost of infrastructure

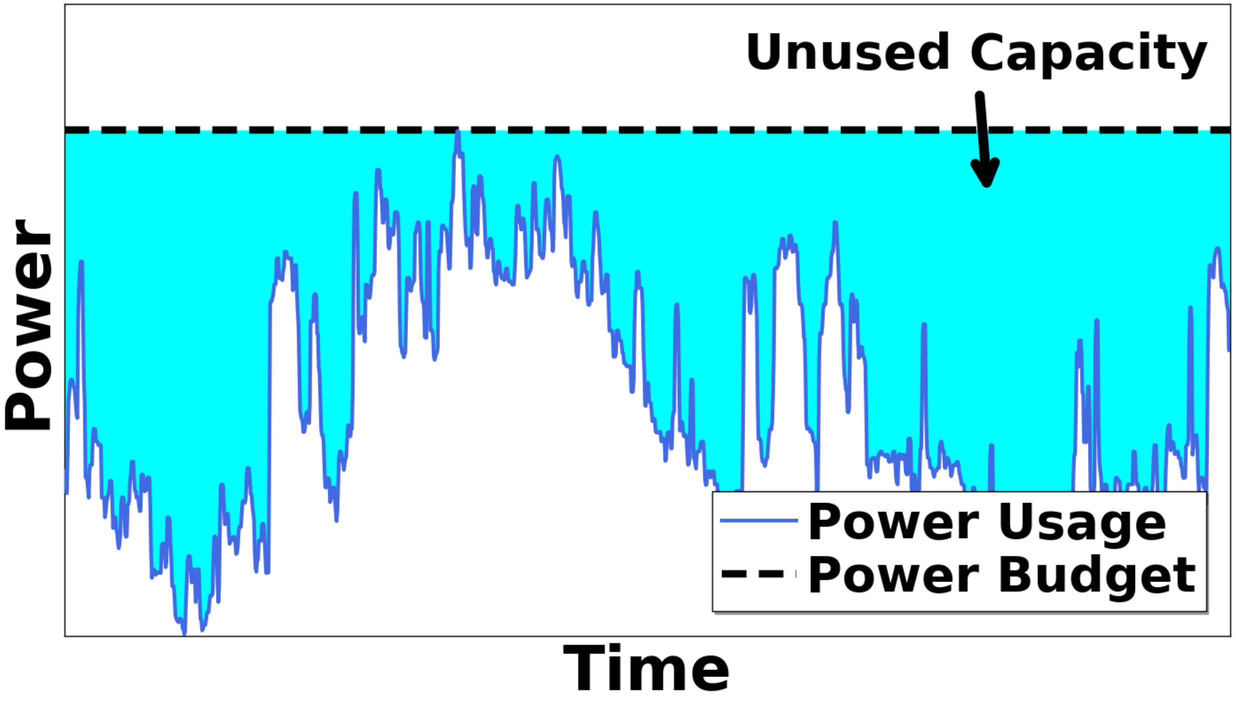
$10-25 per Watt



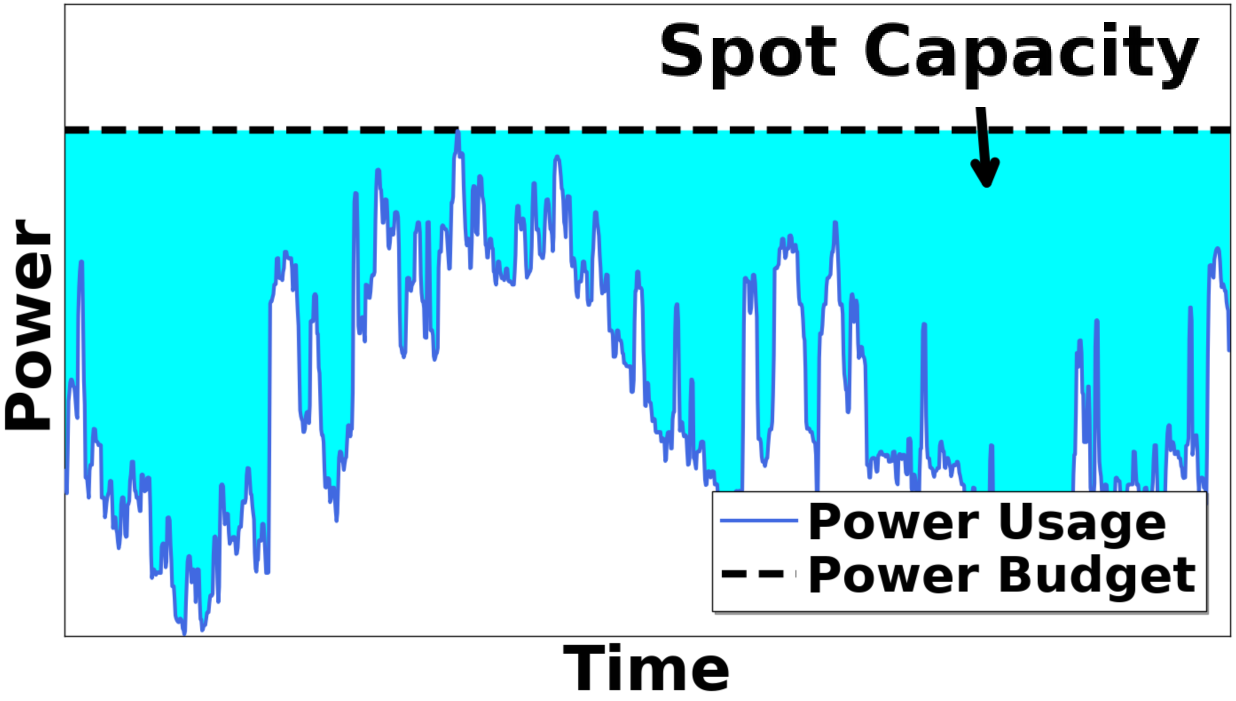
Underutilization in data centers



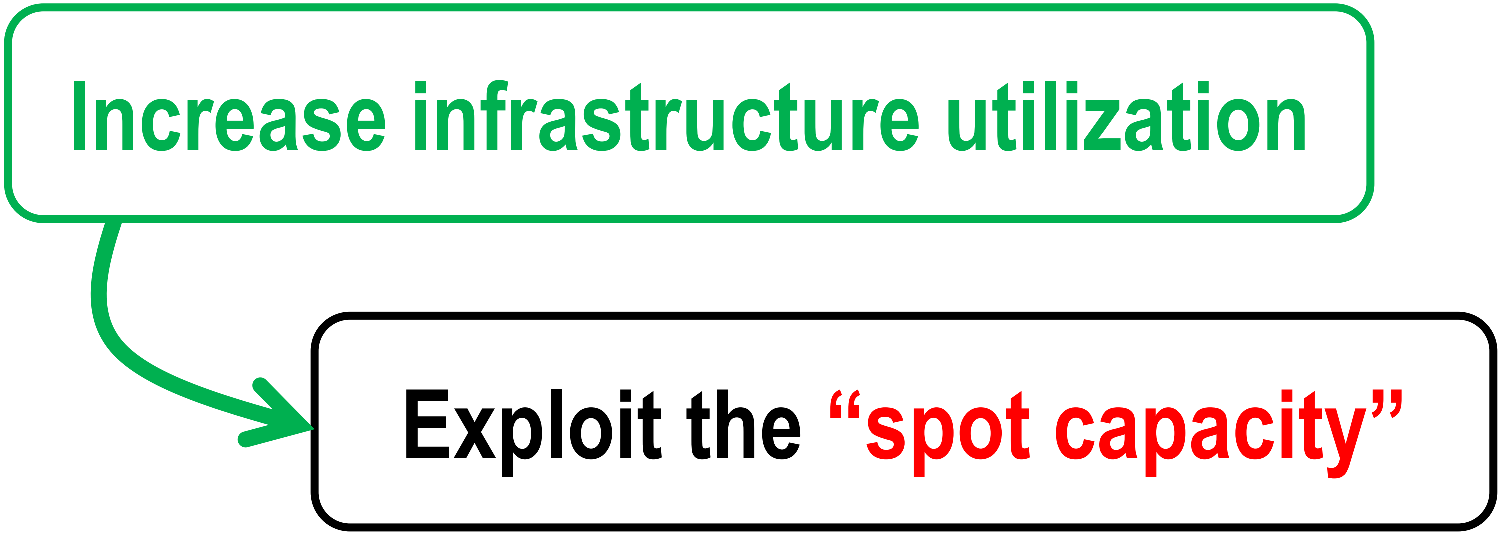
Underutilization in data centers



Underutilization in data centers



Increase infrastructure utilization



Increase infrastructure utilization

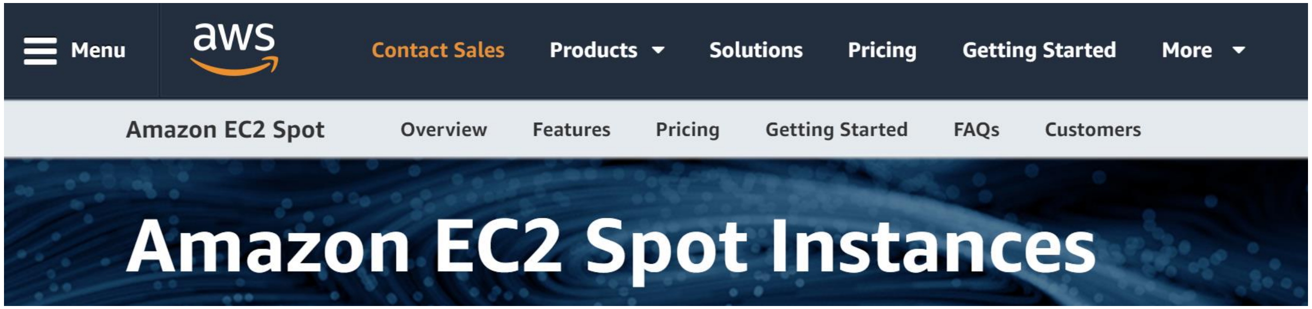
Exploit the “spot capacity”

Some inspirations

- “Power routing” in ASPLOS’10 and “soft fuse” in EuroSys’09

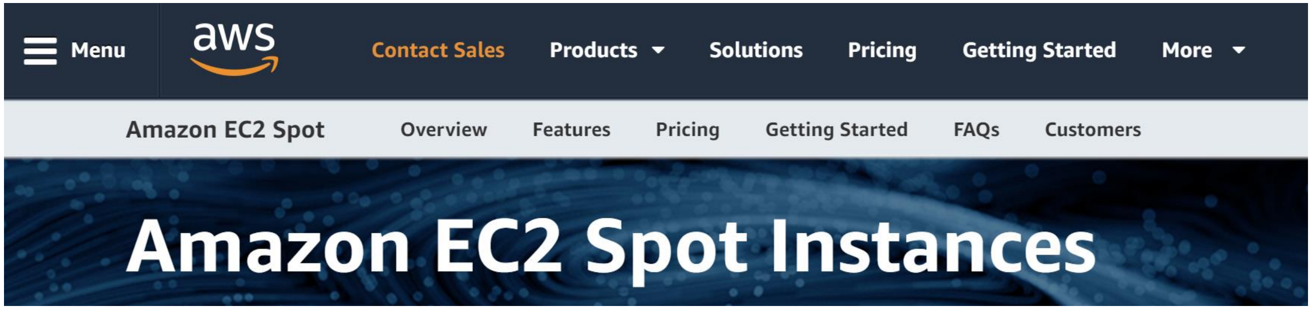
Some inspirations

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



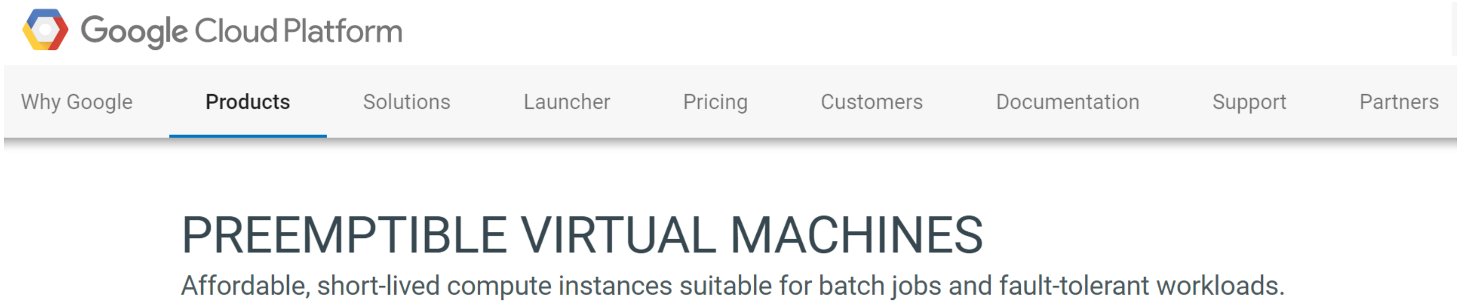
Some inspirations

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



Amazon EC2 Spot Instances

- “Preemptible VM” from Google Cloud



PREEMPTIBLE VIRTUAL MACHINES

Affordable, short-lived compute instances suitable for batch jobs and fault-tolerant workloads.

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→ Power routing,…

Spot capacity in multi-tenant data centers

No centralized control→ 🚫 Power routing,…

A market for spot capacity

Spot capacity in multi-tenant data centers

No centralized control→ 🚫Power routing,…

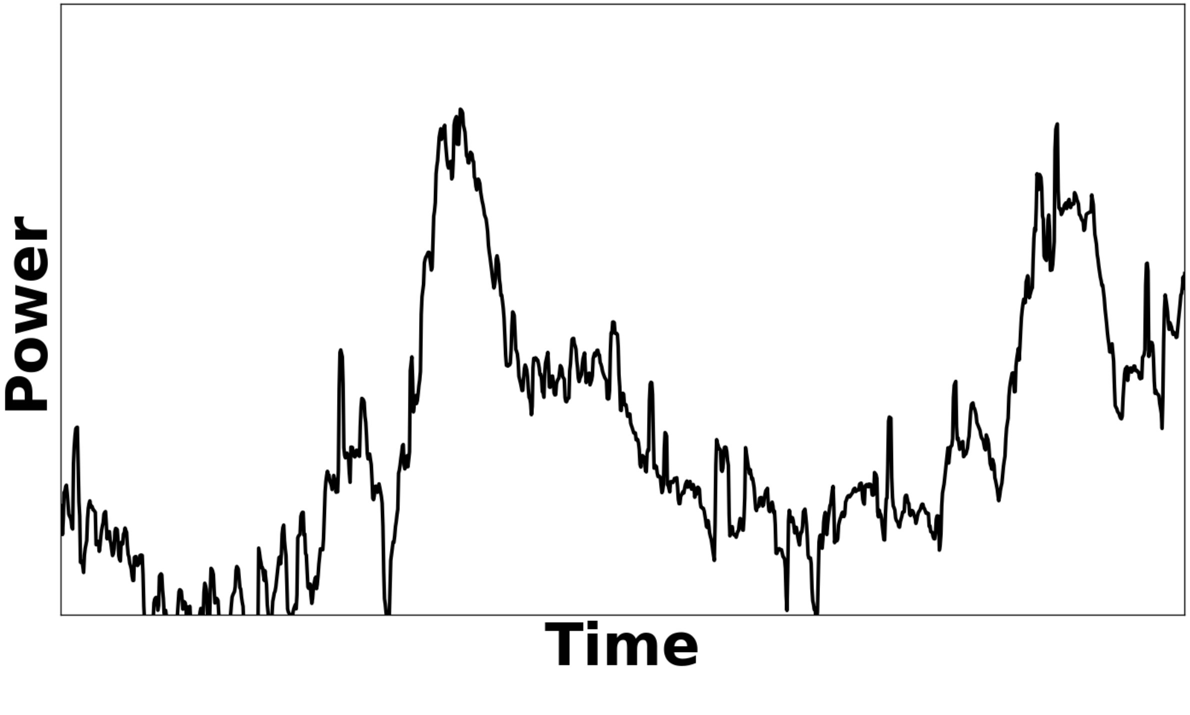
⬇

A market for spot capacity

Tenants buy spot capacity from
the data center operator

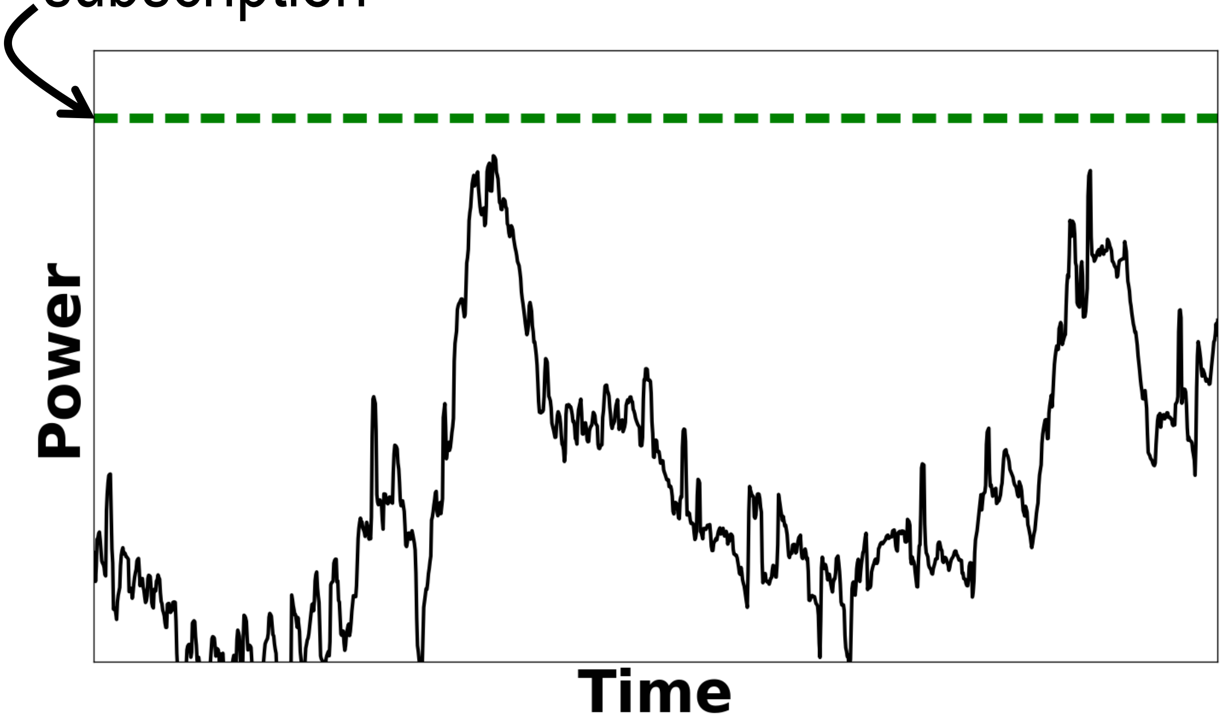
Spot capacity in multi-tenant data centers

- Flexibility for cost conscious tenants



Spot capacity in multi-tenant data centers

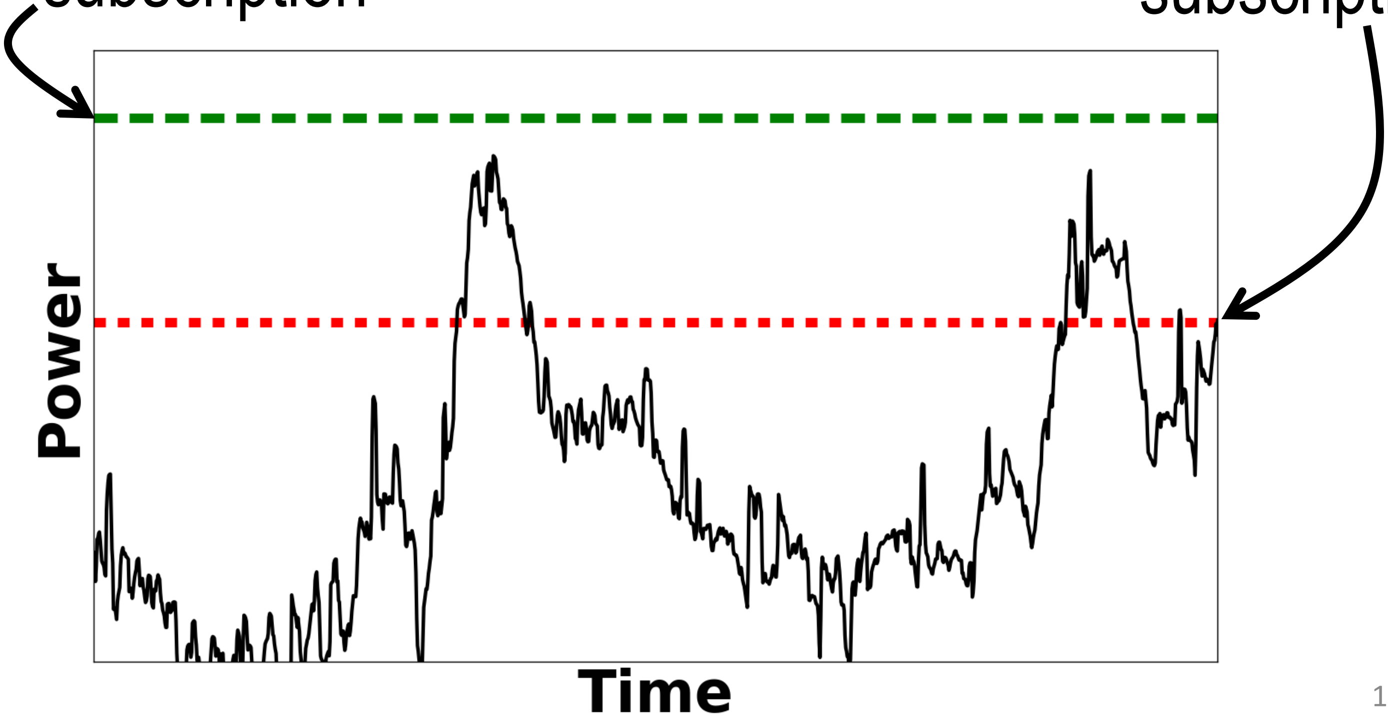
- Flexibility for cost conscious tenants



Peak-based subscription

Spot capacity in multi-tenant data centers

- Flexibility for cost conscious tenants



Peak-based subscription

Conservative subscription

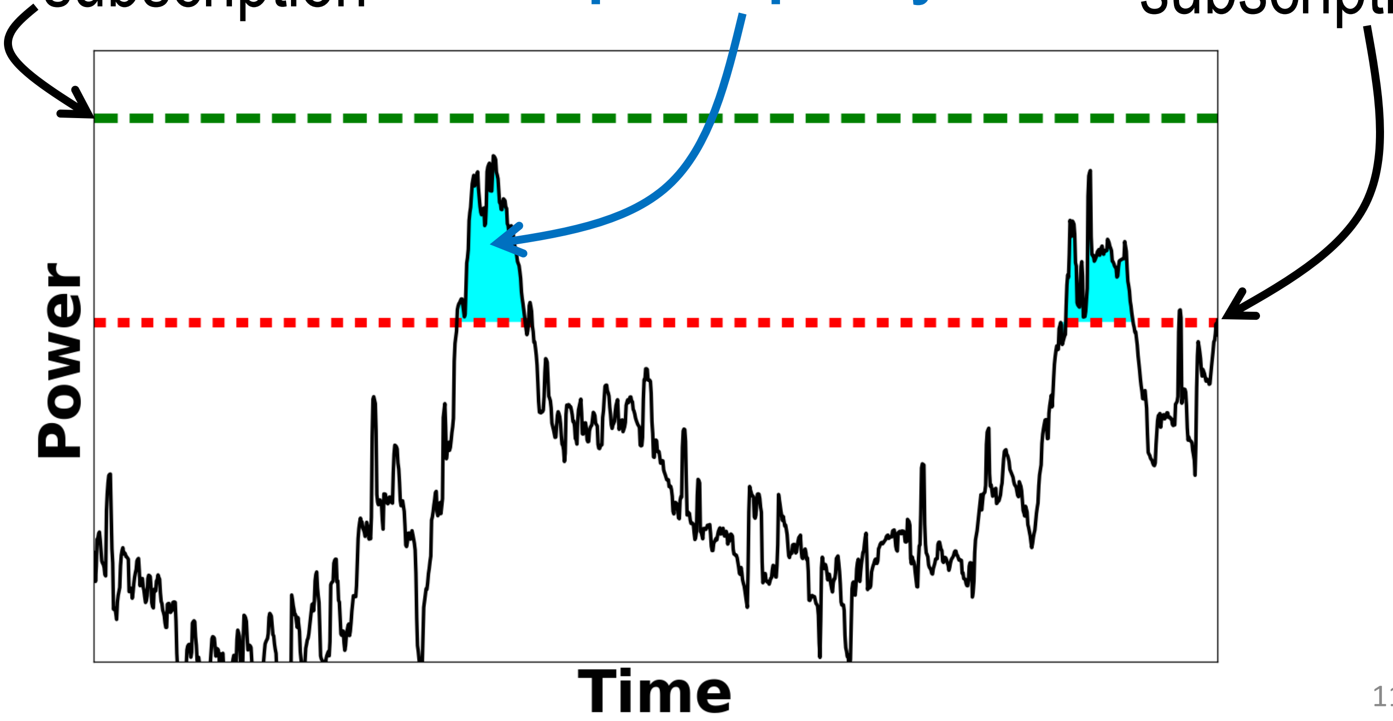
Spot capacity in multi-tenant data centers

- Flexibility for cost conscious tenants

Peak-based subscription

Spot capacity

Conservative subscription



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

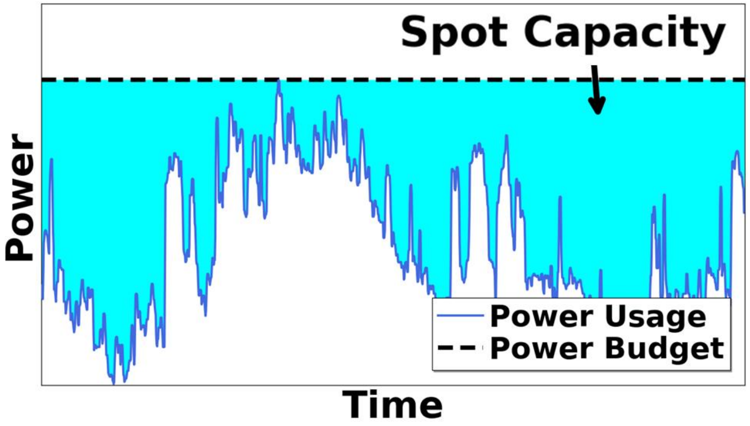
Spot capacity in multi-tenant data centers

- 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

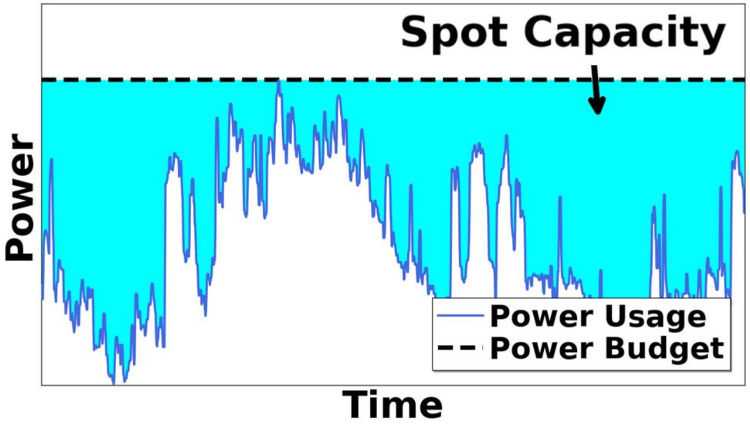
Challenges

- Spot capacity is **limited** and **intermittent**



Challenges

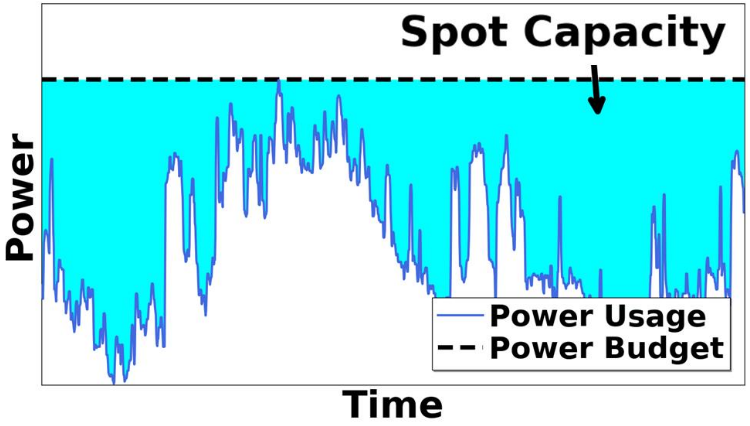
- Spot capacity is **limited** and **intermittent**



- 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

$$\underset{q(t)}{\text{maximize}} \quad q(t) \cdot \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right).$$

Problem formulation

- Goal: operator profit maximization

Rack level demand

$$\underset{q(t)}{\text{maximize}} \quad q(t) \cdot \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right).$$

Problem formulation

- Goal: operator profit maximization

Rack level demand

Price of spot capacity

$$\underset{q(t)}{\text{maximize}} \quad q(t) \cdot \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right).$$

Problem formulation

- Goal: operator profit maximization

Rack level demand

Price of spot capacity

$$\underset{q(t)}{\text{maximize}} \quad q(t) \cdot \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right).$$

Infrastructure constraints

$$\begin{aligned}
&\underline{\text{Rack}}: && D_r\left(q(t)\right) \leq P_r^R, \ \forall r \in \mathcal{S}(t) \\
&\underline{\text{PDU}}: && \sum_{r \in \mathcal{S}(t) \cap \mathcal{R}_m} D_r\left(q(t)\right) \leq P_m(t), \ \forall m \in \mathcal{M} \\
&\underline{\text{UPS}}: && \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right) \leq P_o(t)
\end{aligned}$$

How to solve it?

$$\underset{q(t)}{\text{maximize}} \quad q(t) \cdot \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right).$$

$$\underline{\text{Rack}}: \qquad D_r\left(q(t)\right) \leq P_r^R,\ \forall r \in \mathcal{S}(t)$$

$$\underline{\text{PDU}}: \qquad \sum_{r \in \mathcal{S}(t) \cap \mathcal{R}_m} D_r\left(q(t)\right) \leq P_m(t),\ \forall m \in \mathcal{M}$$

$$\underline{\text{UPS}}: \qquad \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right) \leq P_o(t)$$

How to solve it?

Unknown

$$\underset{q(t)}{\text{maximize}} \quad q(t)\cdot\sum_{r\in\mathcal{S}(t)} \boxed{D_r\left(q(t)\right).}$$

$$\underline{\text{Rack}}: \qquad D_r\left(q(t)\right) \leq P_r^R, \; \forall r \in \mathcal{S}(t)$$

$$\underline{\text{PDU}}: \qquad \sum_{r\in\mathcal{S}(t)\cap\mathcal{R}_m} D_r\left(q(t)\right) \leq P_m(t), \; \forall m \in \mathcal{M}$$

$$\underline{\text{UPS}}: \qquad \sum_{r\in\mathcal{S}(t)} D_r\left(q(t)\right) \leq P_o(t)$$

How to solve it?

Unknown

$$\underset{q(t)}{\text{maximize}} \quad q(t) \cdot \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right).$$

$$\begin{aligned}
&\underline{\text{Rack}}: && D_r\left(q(t)\right) \leq P_r^R, \ \forall r \in \mathcal{S}(t) \\
&\underline{\text{PDU}}: && \sum_{r \in \mathcal{S}(t) \cap \mathcal{R}_m} D_r\left(q(t)\right) \leq P_m(t), \ \forall m \in \mathcal{M} \\
&\underline{\text{UPS}}: && \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right) \leq P_o(t)
\end{aligned}$$

- Soliciting the demand curve → privacy and overhead

How to solve it?

Unknown

$$\underset{q(t)}{\text{maximize}} \quad q(t) \cdot \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right).$$

$$\begin{aligned}
&\underline{\text{Rack}}: && D_r\left(q(t)\right) \leq P_r^R, \ \forall r \in \mathcal{S}(t) \\
&\underline{\text{PDU}}: && \sum_{r \in \mathcal{S}(t) \cap \mathcal{R}_m} D_r\left(q(t)\right) \leq P_m(t), \ \forall m \in \mathcal{M} \\
&\underline{\text{UPS}}: && \sum_{r \in \mathcal{S}(t)} D_r\left(q(t)\right) \leq P_o(t)
\end{aligned}$$

- Soliciting the demand curve → privacy and overhead
- Pre-set price → low level demand prediction

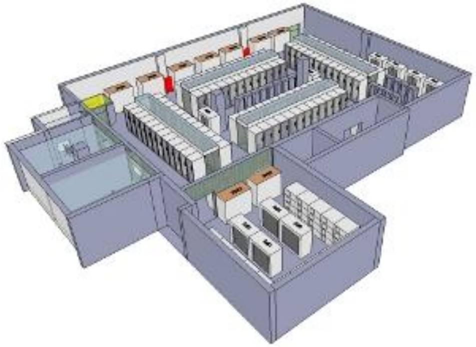
How to solve it?

$$\underset{q(t)}{\text{maximize}} \quad q(t)\cdot \sum_{r\in\mathcal{S}(t)} \overset{\text{Unknown}}{\boxed{D_r\left(q(t)\right).}}$$

$$\begin{aligned}
&\underline{\text{Rack}}: && D_r\left(q(t)\right) \le P_r^R,\ \forall r \in \mathcal{S}(t)\\
&\underline{\text{PDU}}: && \sum_{r\in\mathcal{S}(t)\cap\mathcal{R}_m} D_r\left(q(t)\right) \le P_m(t),\ \forall m \in \mathcal{M}\\
&\underline{\text{UPS}}: && \sum_{r\in\mathcal{S}(t)} D_r\left(q(t)\right) \le P_o(t)
\end{aligned}$$

- Soliciting the demand curve → privacy and overhead
- Pre-set price → low level demand prediction
- Market approach → an in-between solution

SpotDC: spot capacity management

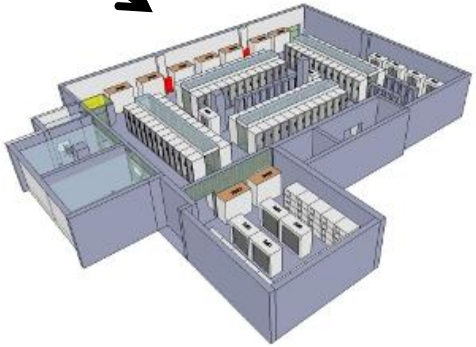


Operator

Tenants

SpotDC: spot capacity management

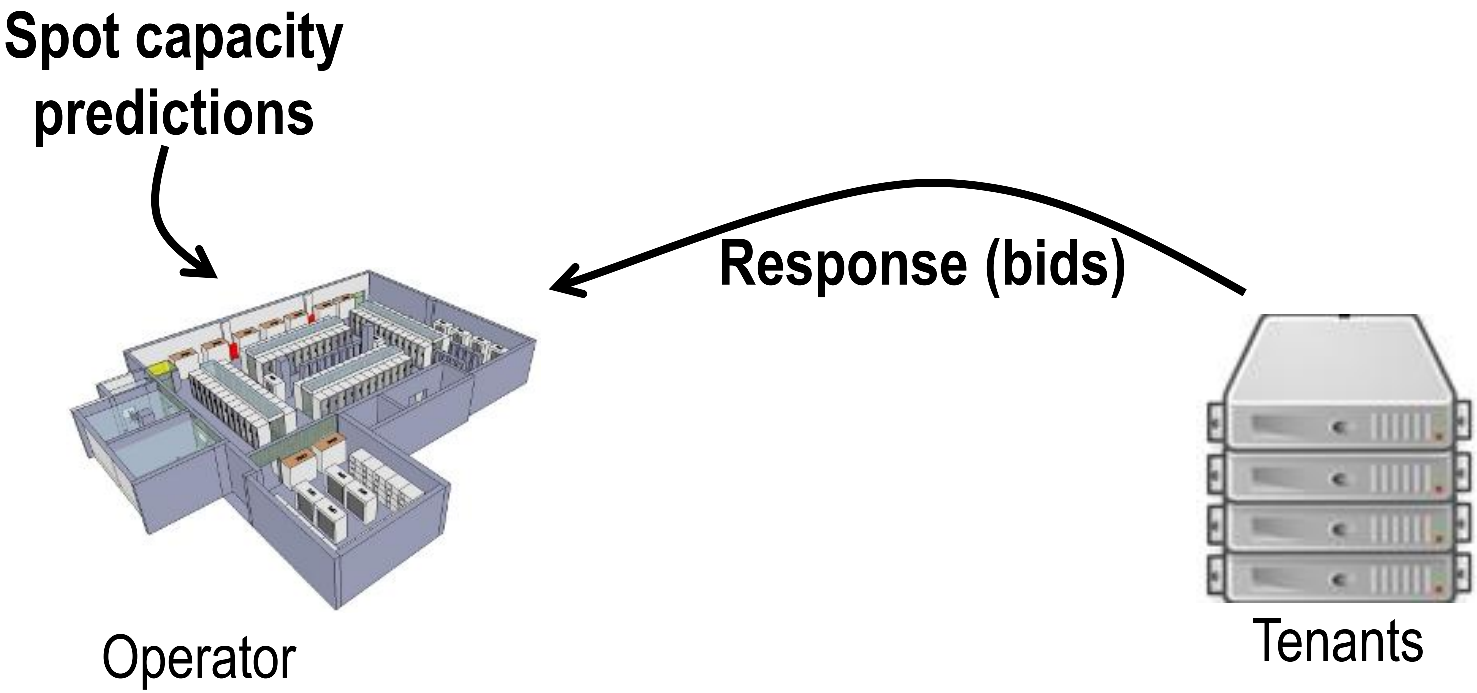
Spot capacity predictions



Operator

Tenants

SpotDC: spot capacity management



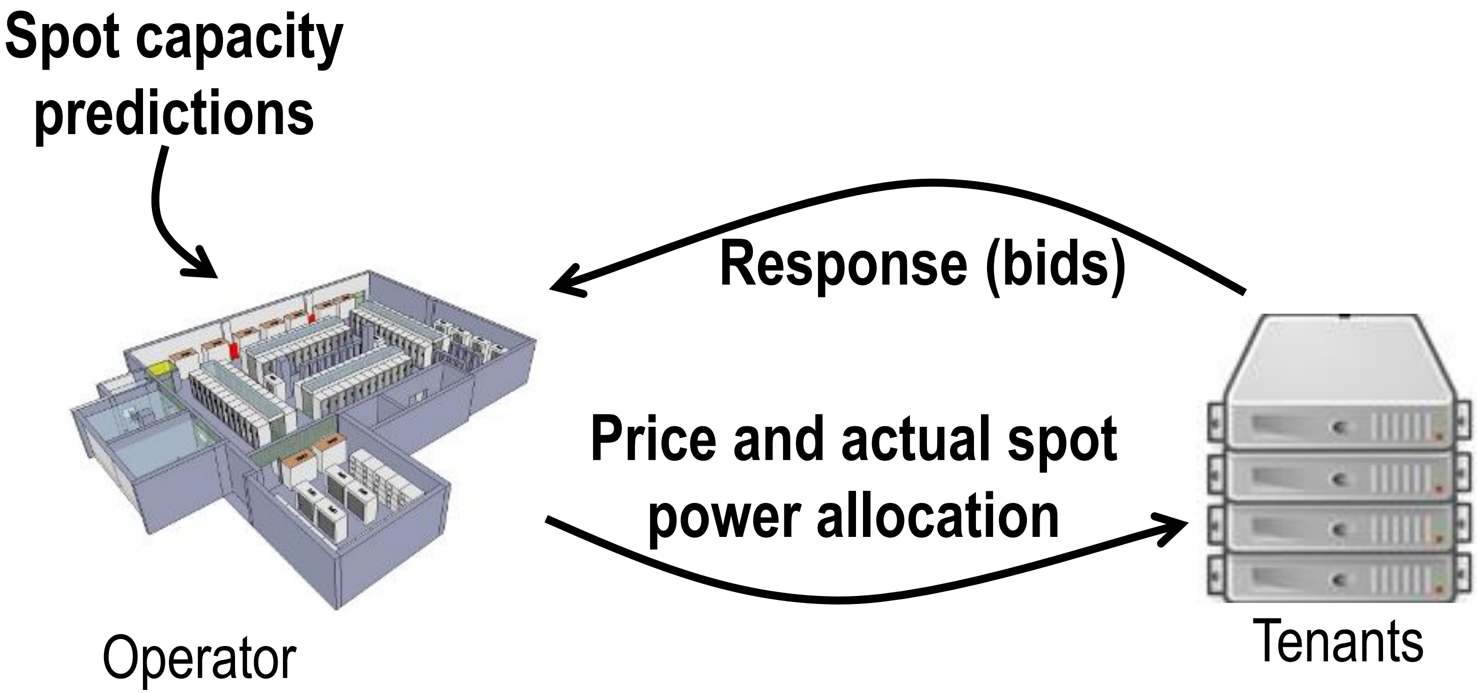
Spot capacity predictions

Response (bids)

Operator

Tenants

SpotDC: spot capacity management



Spot capacity predictions

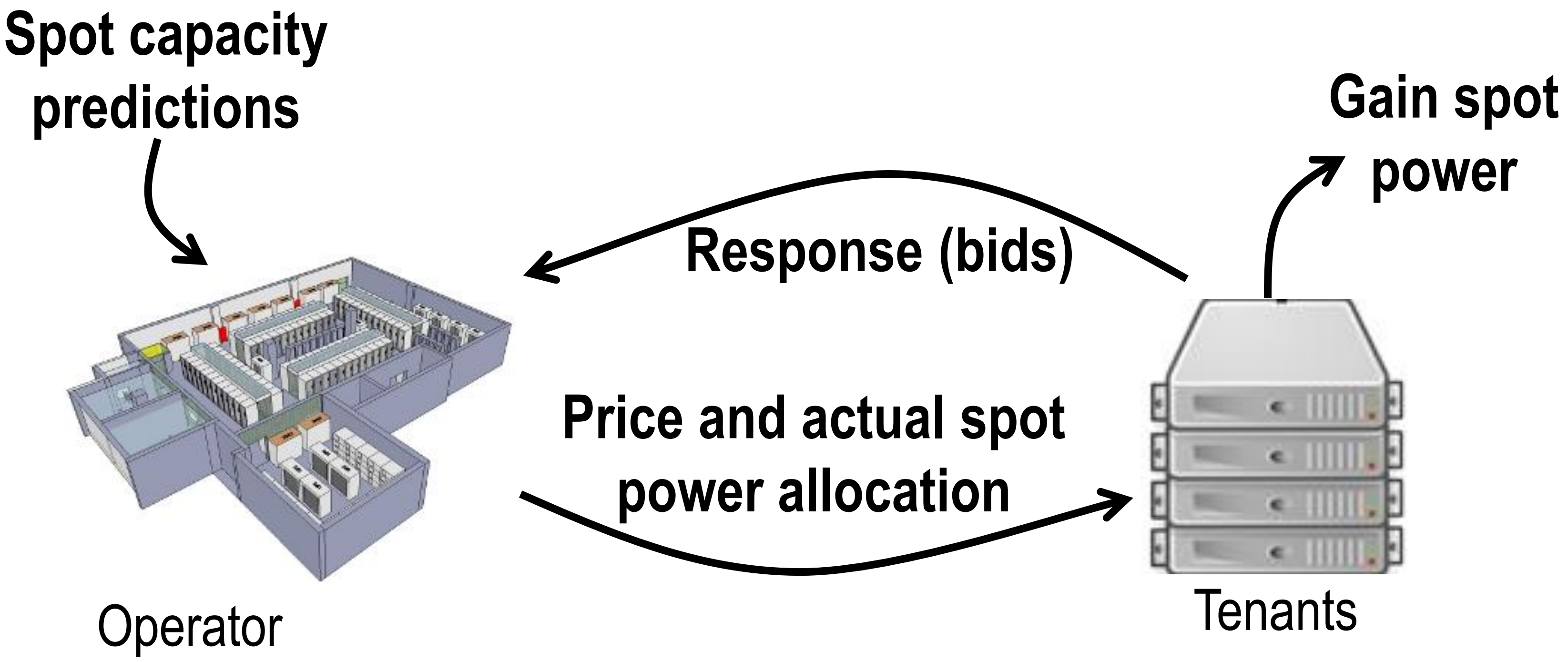
Response (bids)

Price and actual spot power allocation

Operator

Tenants

SpotDC: spot capacity management



Spot capacity predictions

Response (bids)

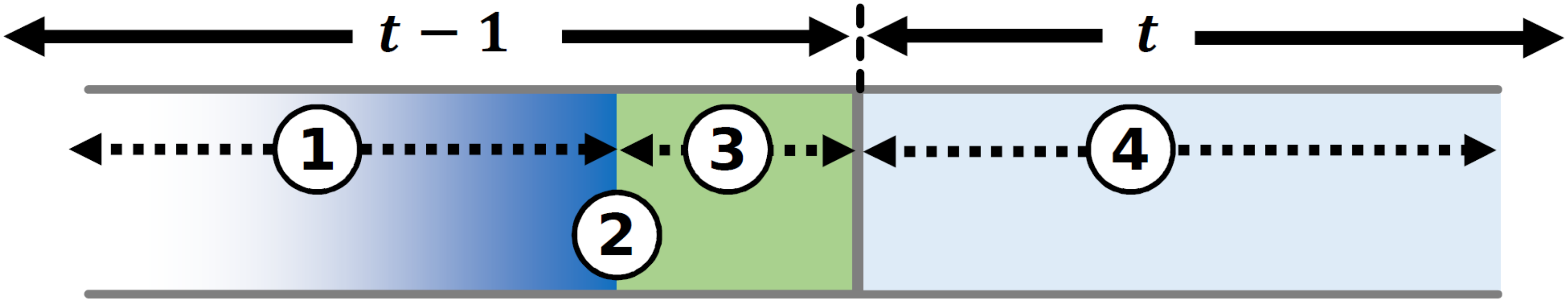
Gain spot power

Price and actual spot power allocation

Operator

Tenants

Timings in SpotDC



① Demand Bidding

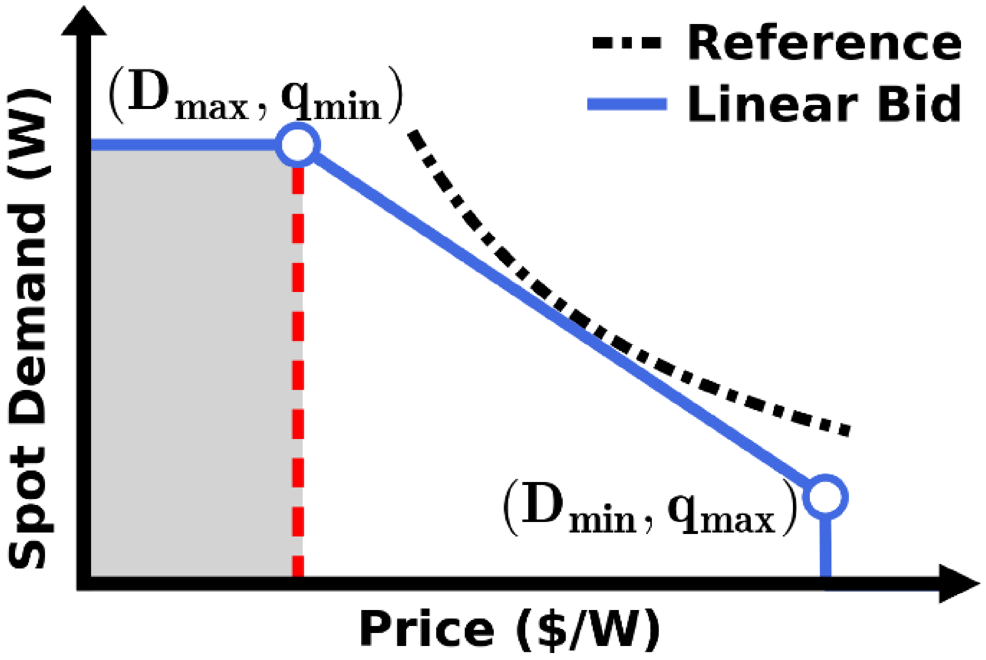
② Spot Capacity Prediction

③ Market Clearing

④ Using Spot Capacity

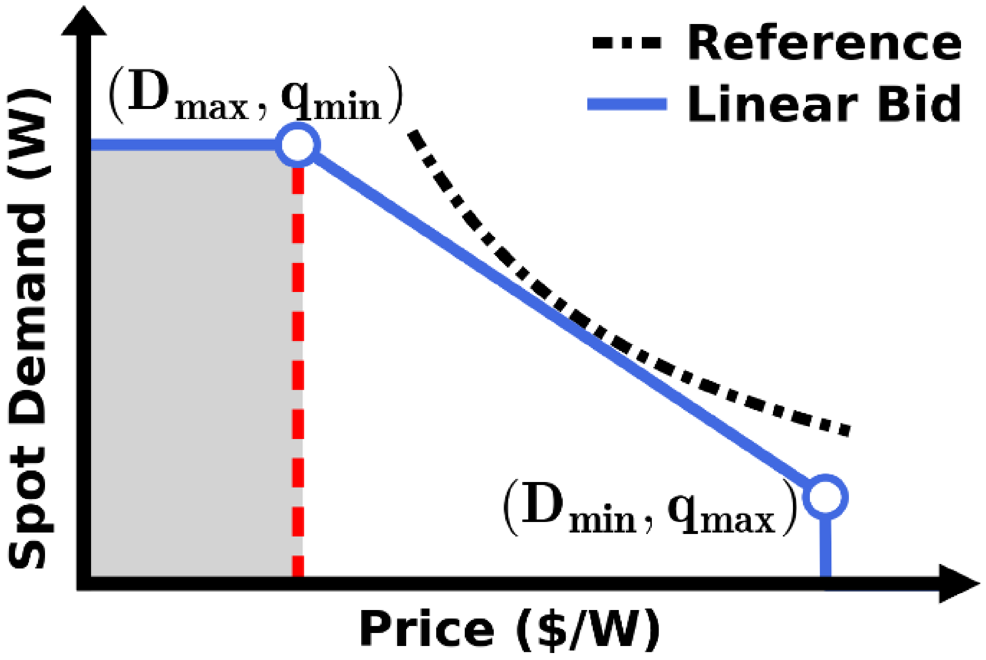
Demand bidding

- A piece-wise-linear bid



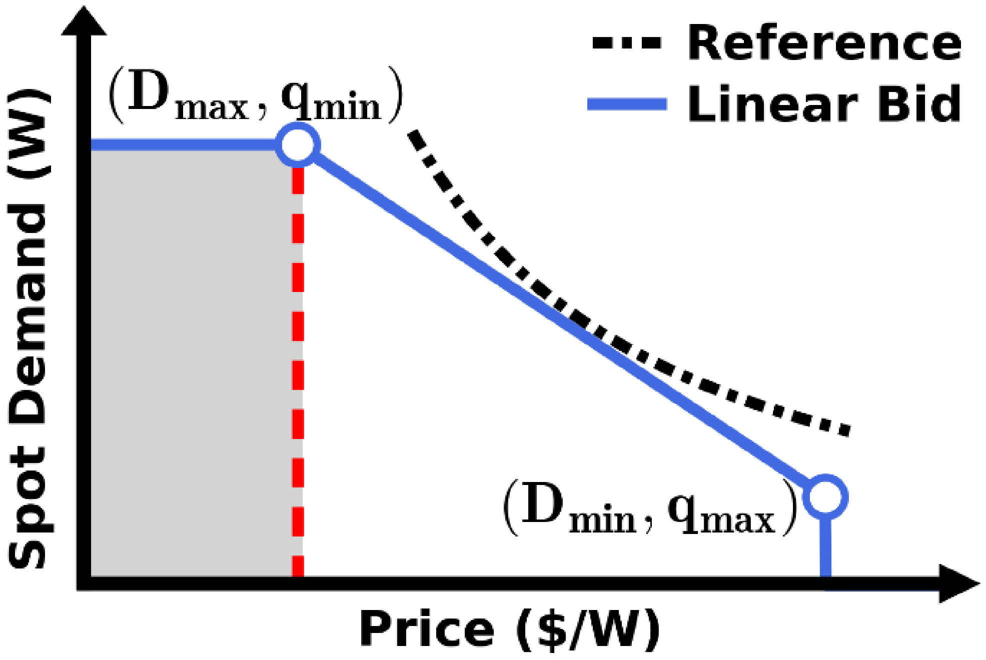
Demand bidding

- 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**

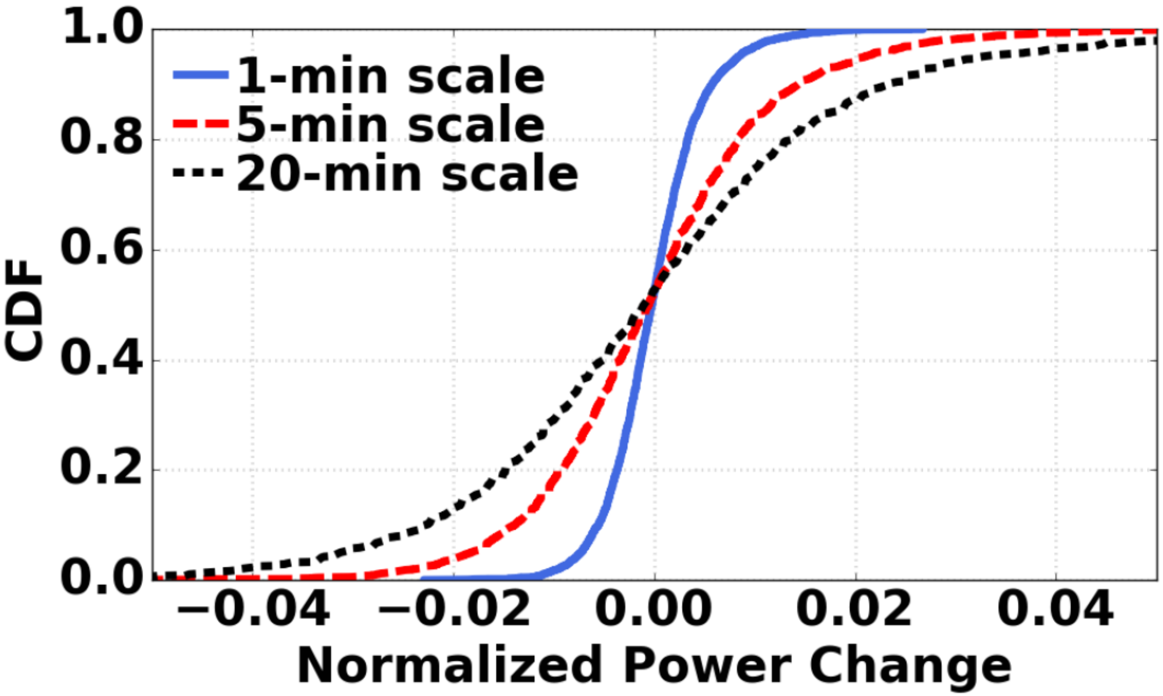


Spot capacity prediction

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

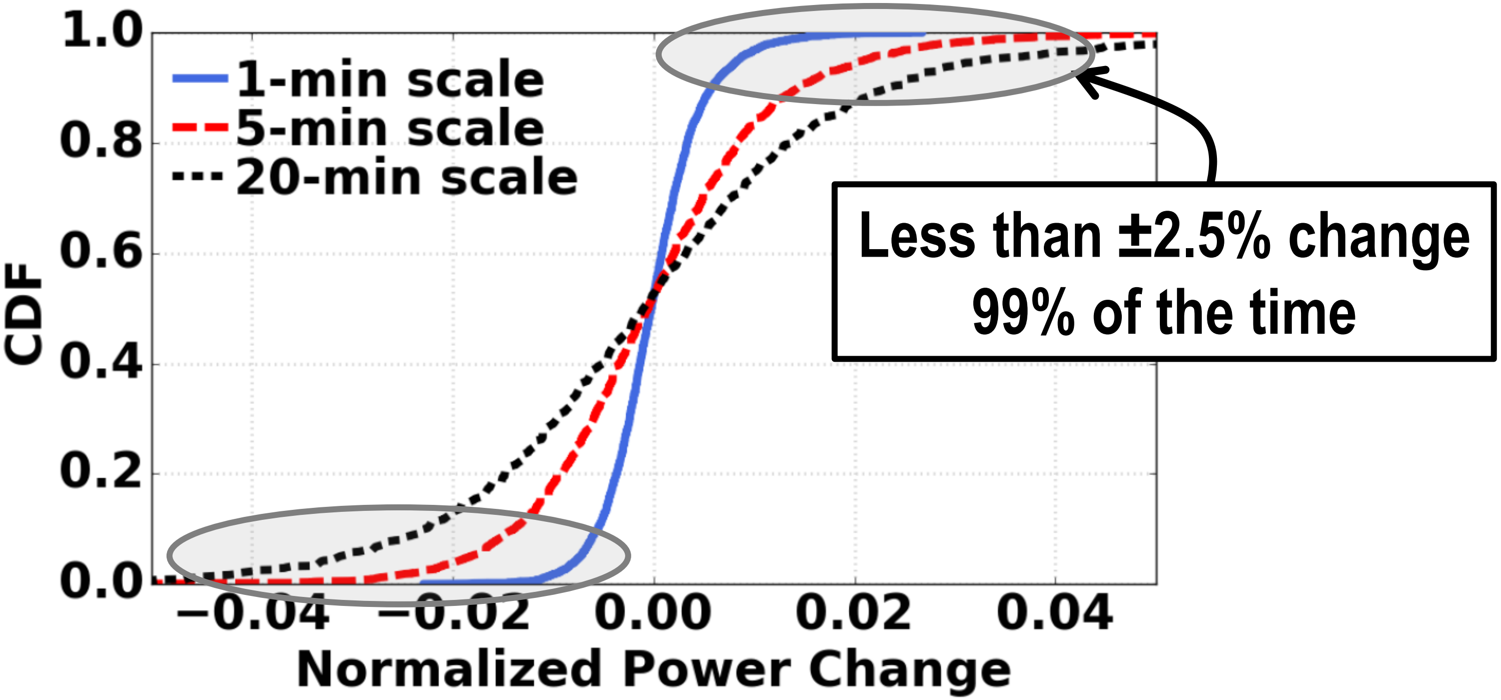
Spot capacity prediction

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



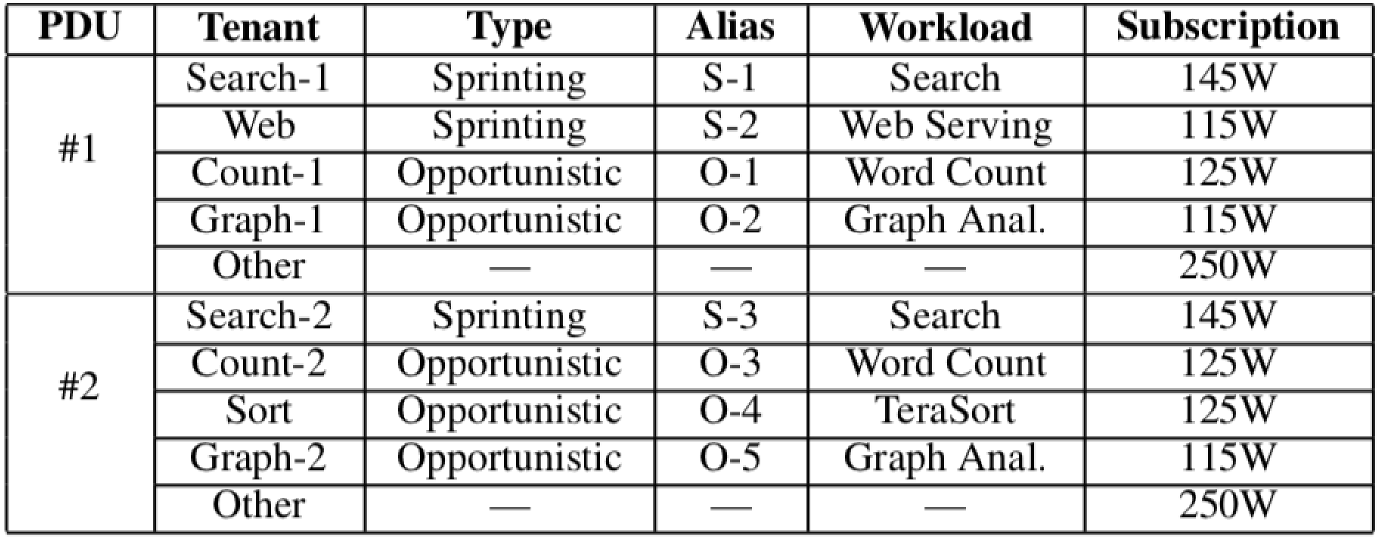
Spot capacity prediction

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



Less than ±2.5% change 99% of the time

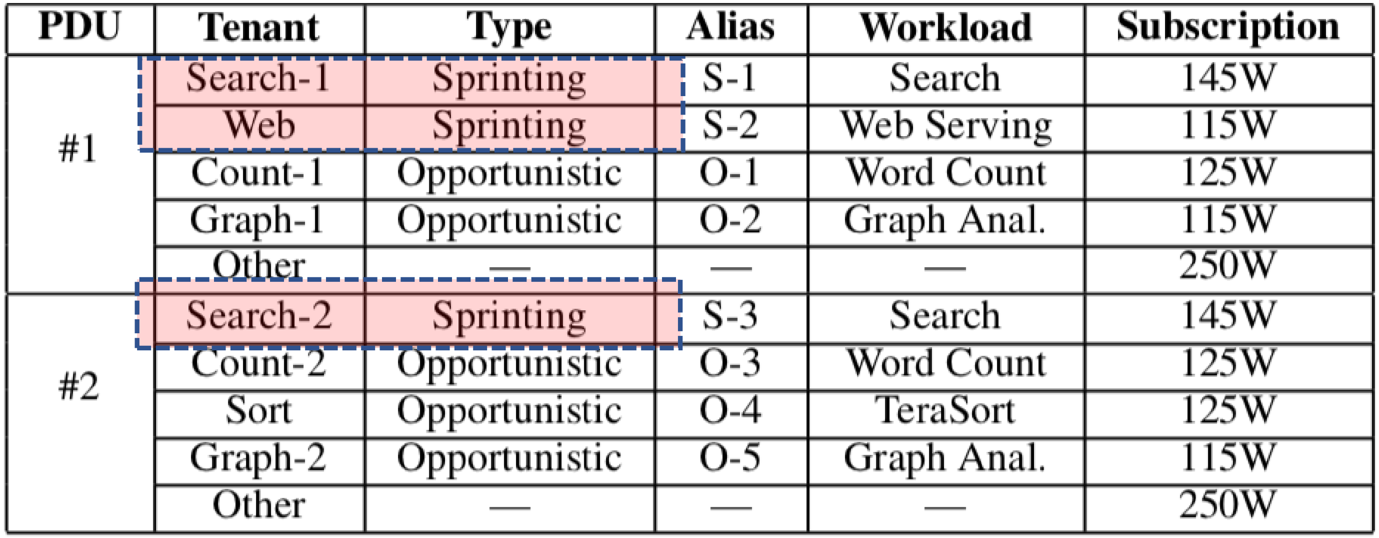
Evaluation methodology



PDU	Tenant	Type	Alias	Workload	Subscription
#1	Search-1	Sprinting	S-1	Search	145W
	Web	Sprinting	S-2	Web Serving	115W
	Count-1	Opportunistic	O-1	Word Count	125W
	Graph-1	Opportunistic	O-2	Graph Anal.	115W
	Other	—	—	—	250W
#2	Search-2	Sprinting	S-3	Search	145W
	Count-2	Opportunistic	O-3	Word Count	125W
	Sort	Opportunistic	O-4	TeraSort	125W
	Graph-2	Opportunistic	O-5	Graph Anal.	115W
	Other	—	—	—	250W

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

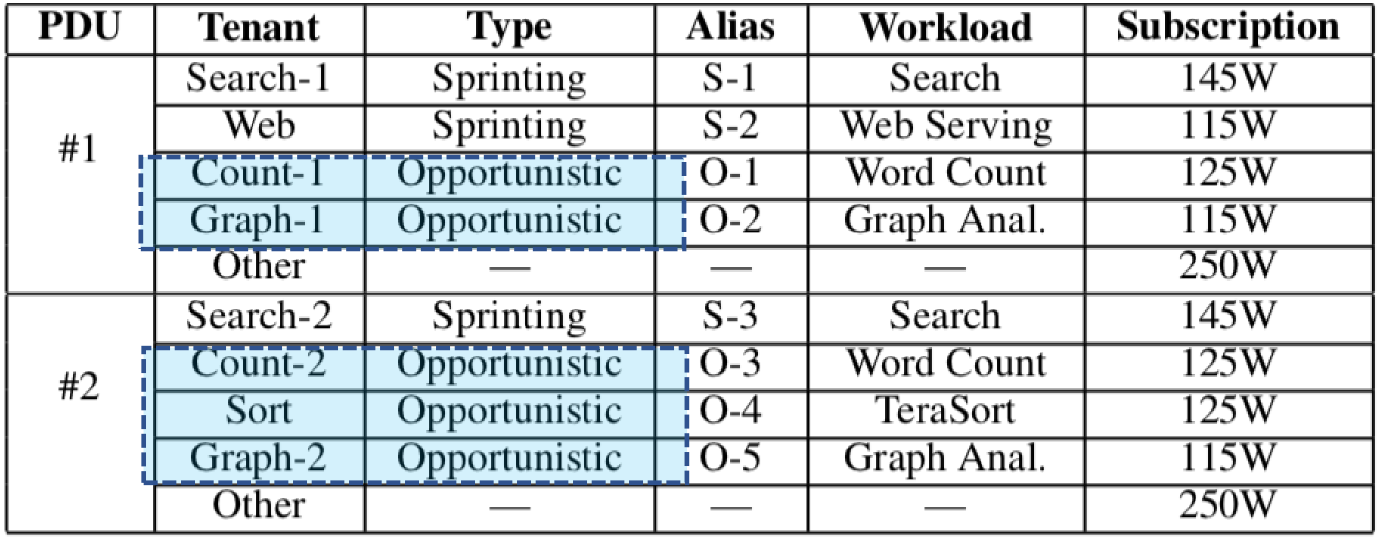
Evaluation methodology



PDU	Tenant	Type	Alias	Workload	Subscription
#1	Search-1	Sprinting	S-1	Search	145W
	Web	Sprinting	S-2	Web Serving	115W
	Count-1	Opportunistic	O-1	Word Count	125W
	Graph-1	Opportunistic	O-2	Graph Anal.	115W
	Other	—	—	—	250W
#2	Search-2	Sprinting	S-3	Search	145W
	Count-2	Opportunistic	O-3	Word Count	125W
	Sort	Opportunistic	O-4	TeraSort	125W
	Graph-2	Opportunistic	O-5	Graph Anal.	115W
	Other	—	—	—	250W

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

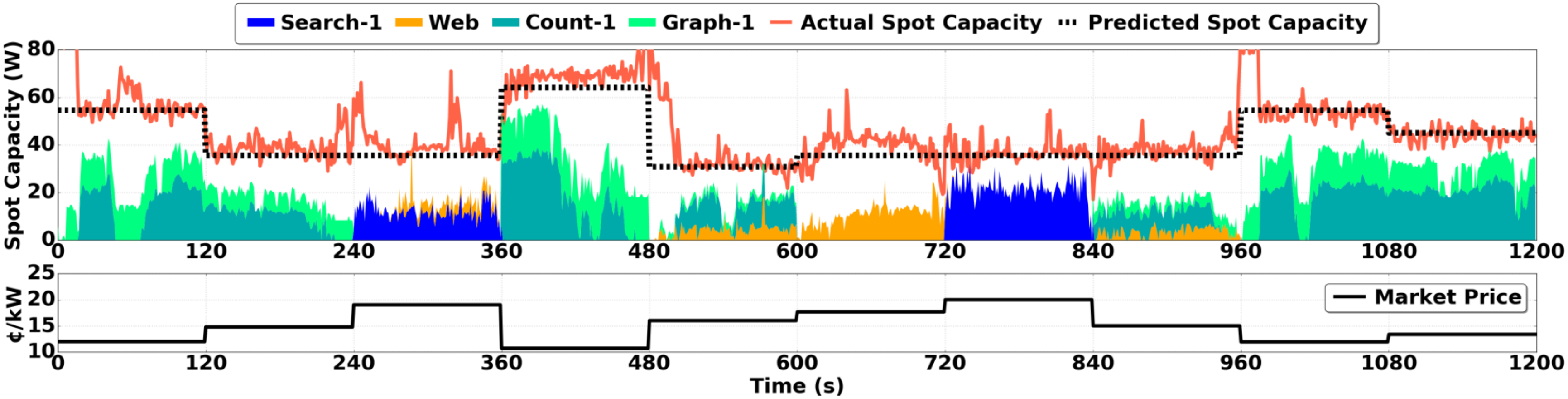
Evaluation methodology



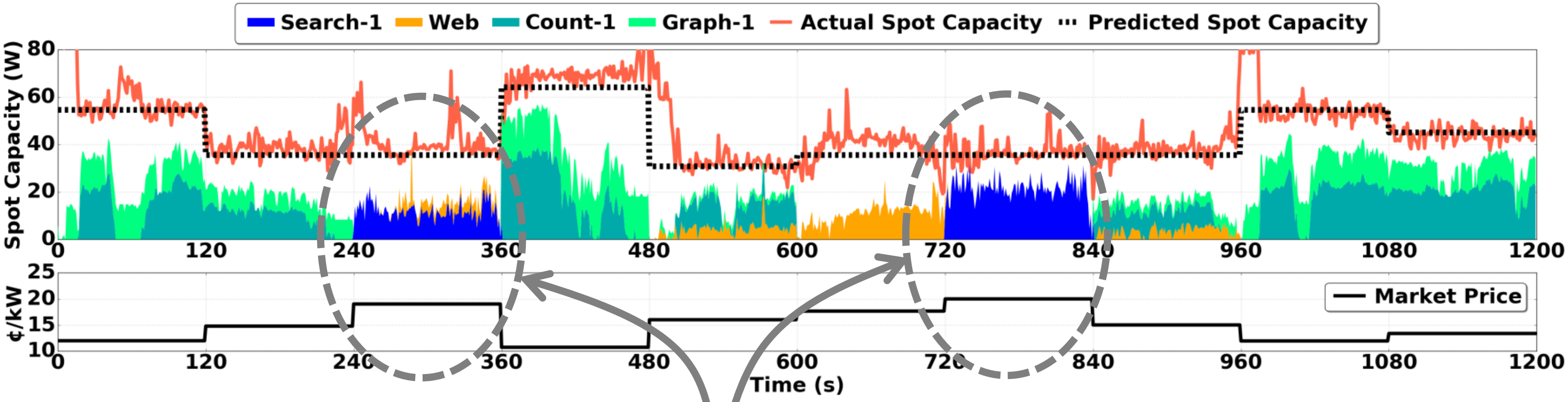
PDU	Tenant	Type	Alias	Workload	Subscription
#1	Search-1	Sprinting	S-1	Search	145W
	Web	Sprinting	S-2	Web Serving	115W
	Count-1	Opportunistic	O-1	Word Count	125W
	Graph-1	Opportunistic	O-2	Graph Anal.	115W
	Other	—	—	—	250W
#2	Search-2	Sprinting	S-3	Search	145W
	Count-2	Opportunistic	O-3	Word Count	125W
	Sort	Opportunistic	O-4	TeraSort	125W
	Graph-2	Opportunistic	O-5	Graph Anal.	115W
	Other	—	—	—	250W

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

Performance evaluation

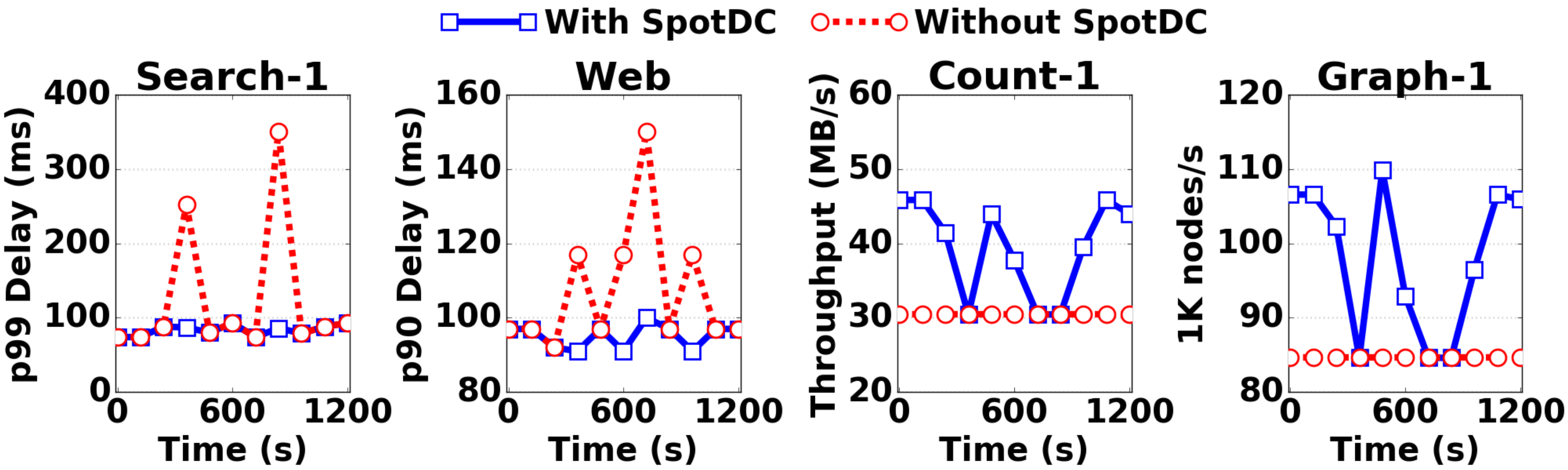


Performance evaluation



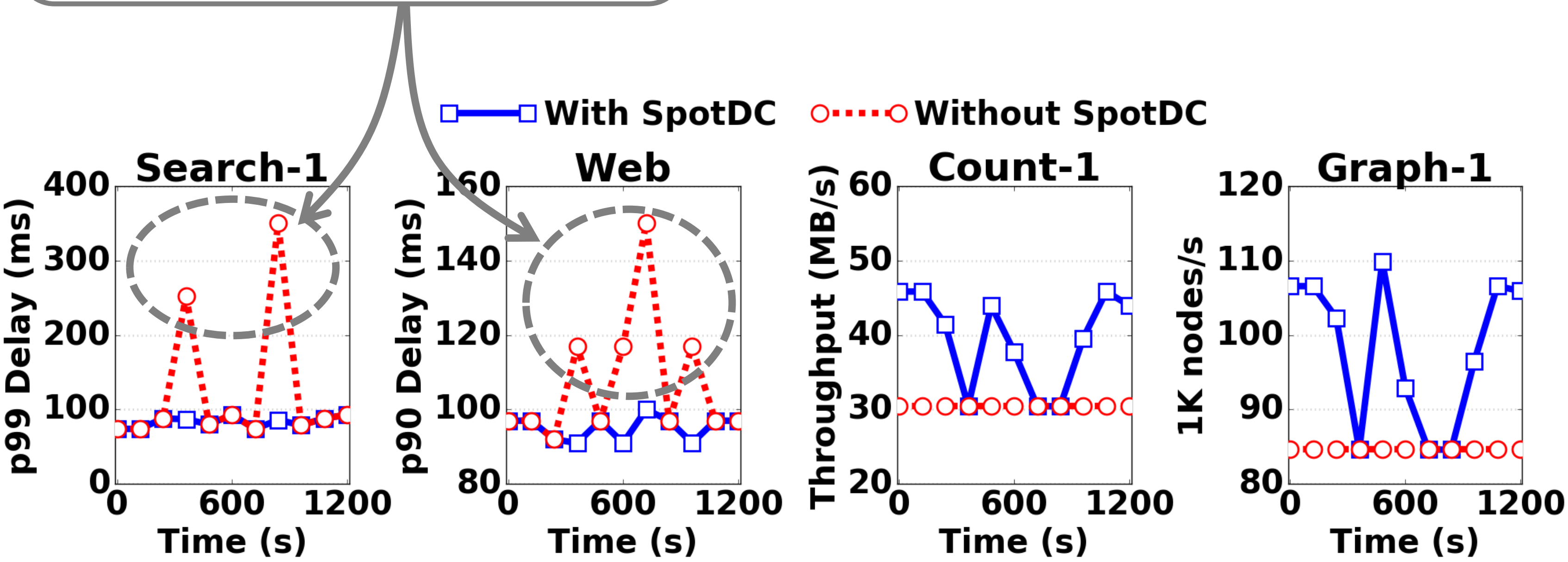
Sprinting tenants drive up the price

Performance evaluation



Performance evaluation

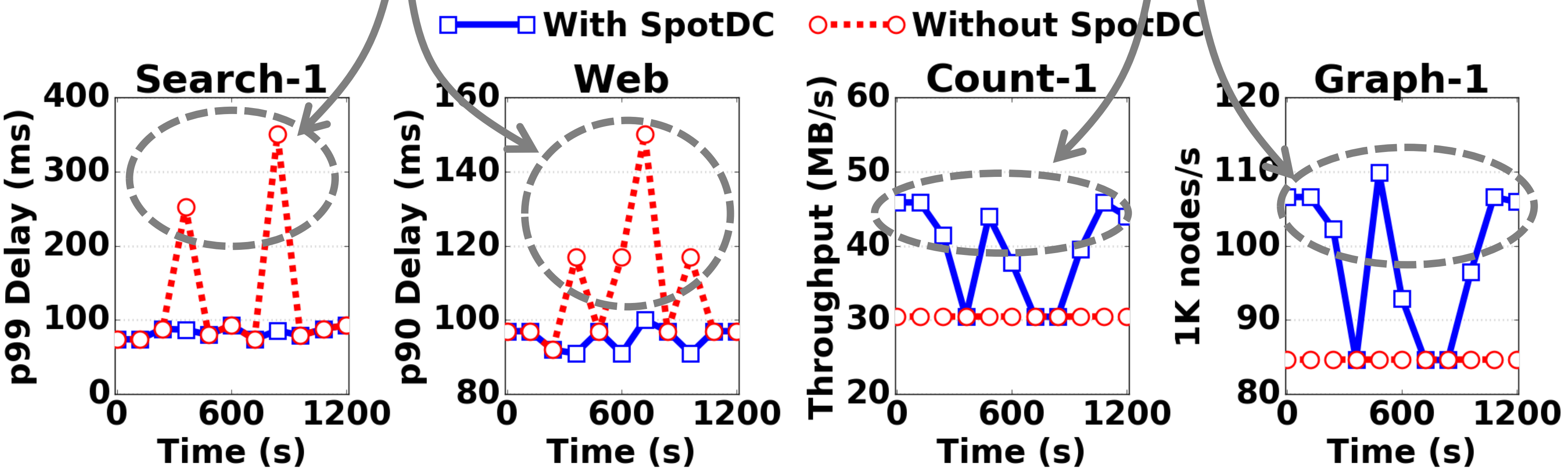
Sprinting tenants avoid SLO violations



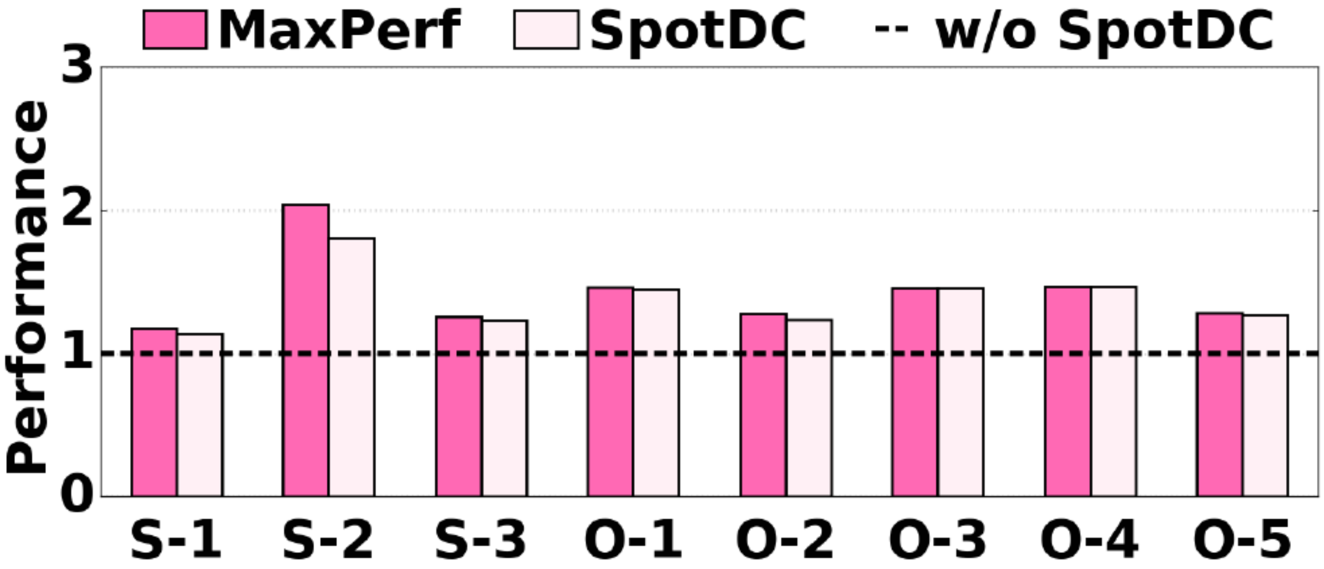
Performance evaluation

Sprinting tenants avoid SLO violations

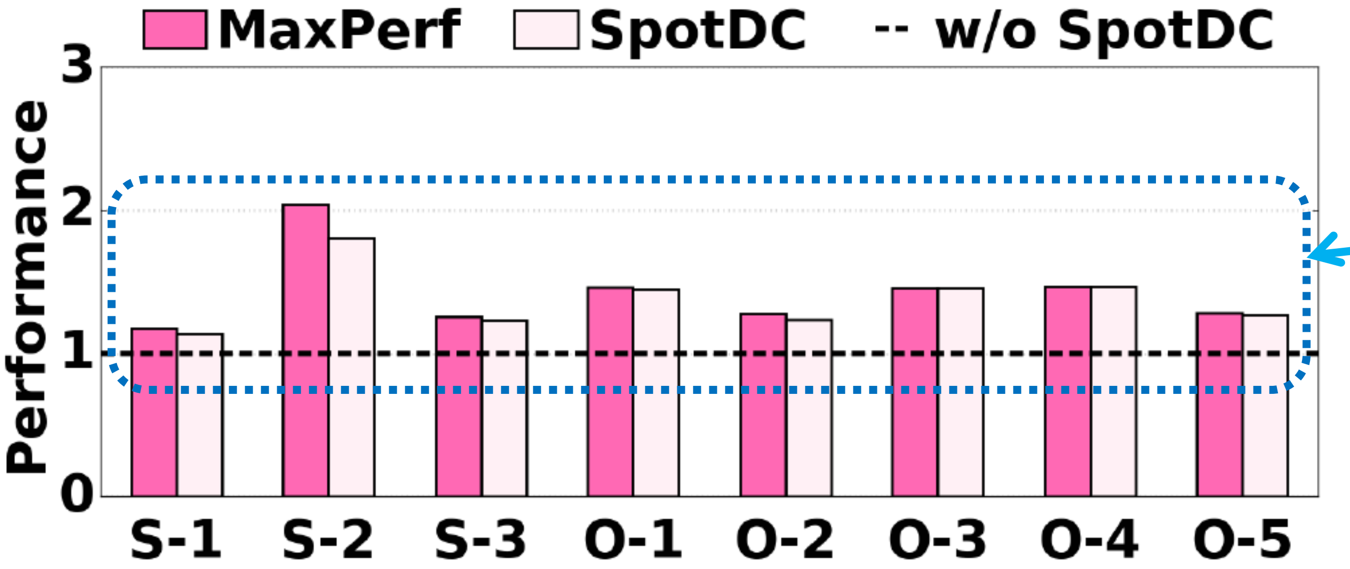
Opportunistic tenants gain throughput boost



Tenants’ benefit from SpotDC

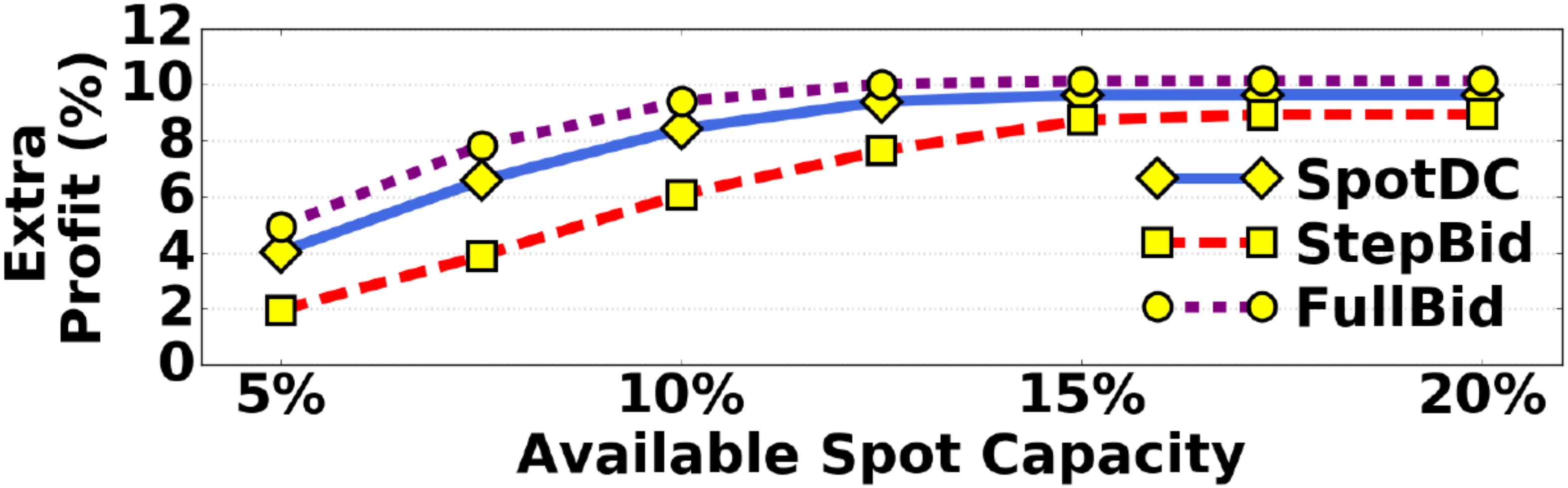


Tenants’ benefit from SpotDC



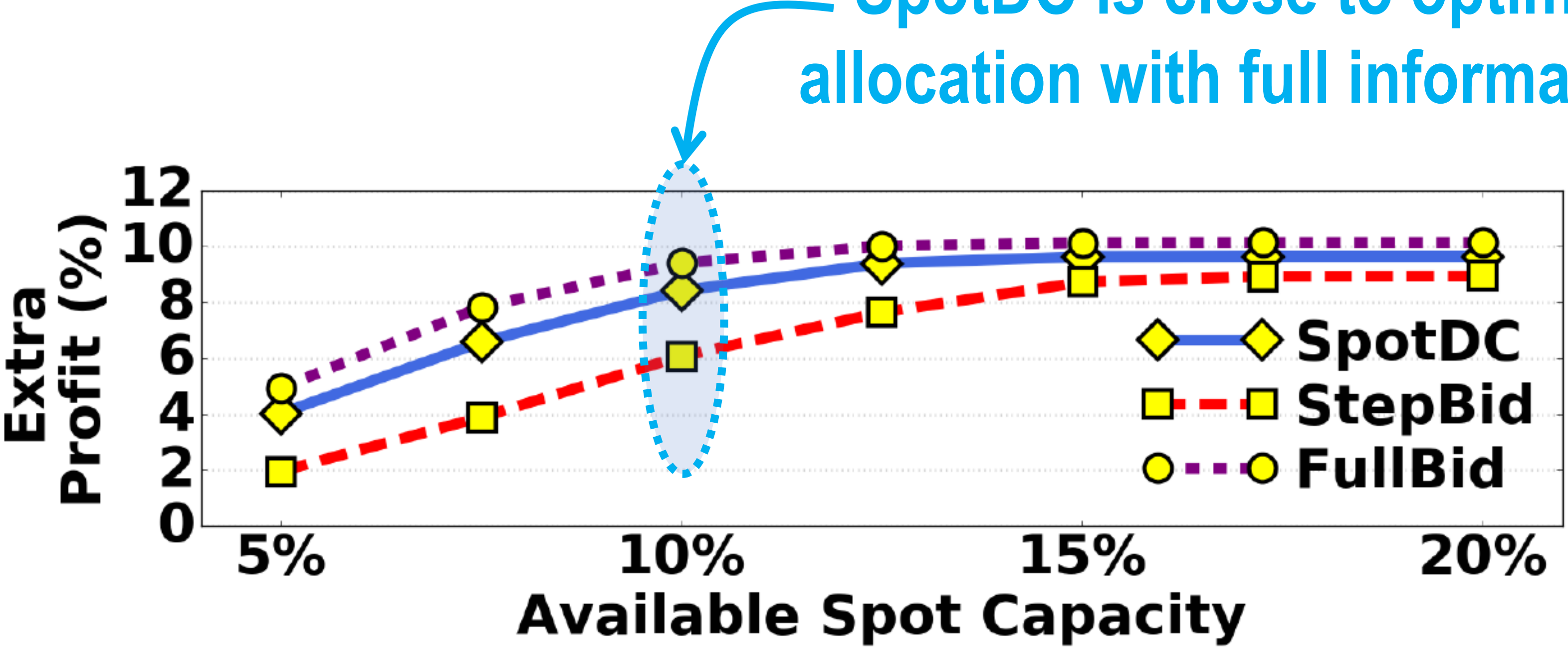
Performance boosts with SpotDC

Operator’s extra profit



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