The Data Center as a Grid Load Stabilizer

Hao Chen*, Michael C. Caramanis** and Ayse K. Coskun*

*Department of Electrical and Computer Engineering

**Division of Systems Engineering

Boston University

{haoc, mcaraman, acoskun}@bu.edu

Power Grid & Market

- Power supply = demand ? (=> blackouts)
- Renewable energy sources: intermittent



- Lack of reliable, large-scale, economical energy storage solutions
- Independent System Operator (ISO):
  - New power market features:
    - Demand side regulation service (RS)
  - Credits provided to the participant who modulates its power consumption <u>dynamically</u> so as to track the ***RS signal***

Demand Side – Data Centers

- Electricity: 3% of the overall consumption in the US[1]
- Power capping /management techniques
  - Enable flexibility in power consumption
- Workload flexibility

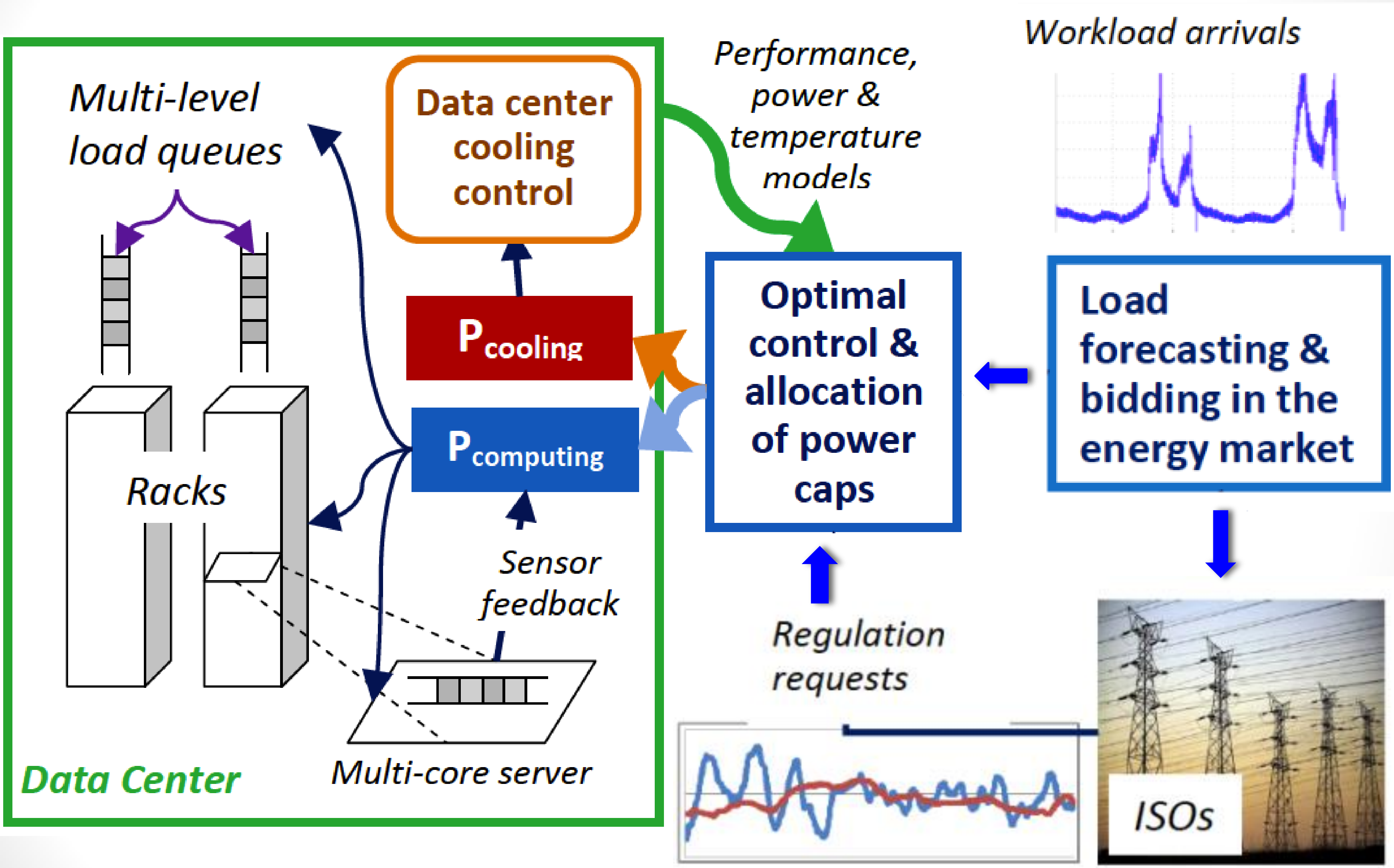
Data centers offer a unique opportunity for providing power regulation service (RS) reserves.

Benefits of Participation

- Help solve unstable renewable energy problem
- Provide additional reserves to accommodate other less flexible uses of electricity
- Achieve significant monetary savings

[1]: J. Koomey. Growth in Data Center Electricity Use 2005 to 2010. Oakland, CA: Analytics Press. August, 1, 2010.

Data Centers in Advanced Power Market



Contributions

- A dynamic control policy for solving ***server commitment*** problem, leveraging:
  - Server-level power capping techniques
  - Information on server power states and overheads
  - Job scheduling & allocation decisions
- RS provision bidding value estimation
- Data center level (compared to previous work on a single server)

- Our solution is able to accurately track the ISO signal, and:
  - We achieve 50%+ monetary savings
  - The proposed policy does not cause major QoS degradation
  - Policy is agnostic of the specific type of workloads running
  - Significant improvement in both monetary savings and QoS compared to prior results based on a single server *(Chen et al. ICCAD 2013)*

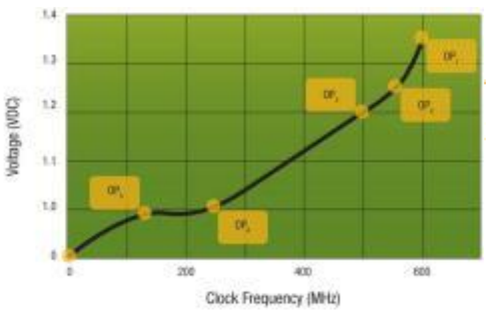
Outline

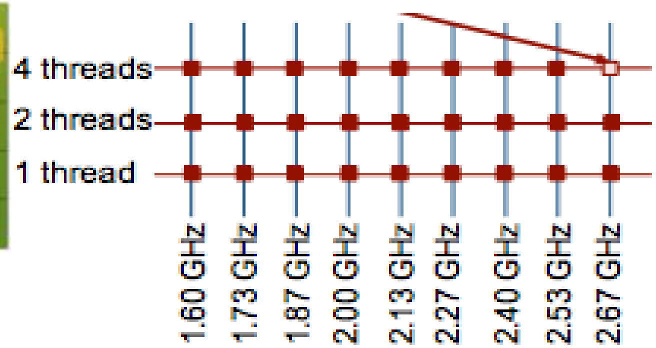
- Background
  - Data Center Power Management
  - Power Market and Data Center Participation
  - Regulation Service (RS)
- Data Center Model
- Dynamic Power Control Policy
- Regulation Reserves Bidding
- Results

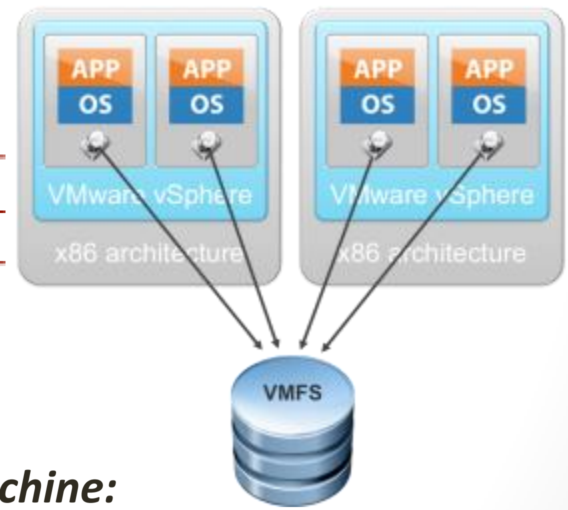
Data Center Power Management

Data Center Server Farms:

- Power and resource budgeting [Zhan DAC13][Gandhi SIGMETRICS09];
- Server Commitment: sleep and idle [Meisner Sigplan Not09][Isci ISCA13][Gandhi IGCC12].







Single Server Level:

- DVFS [Li HPCA06]
- Power Capping: DVFS + multi-thread allocation/migration [Cochran et al. Micro11][Rangan et al. ISCA09][Reda et al. Micro12]

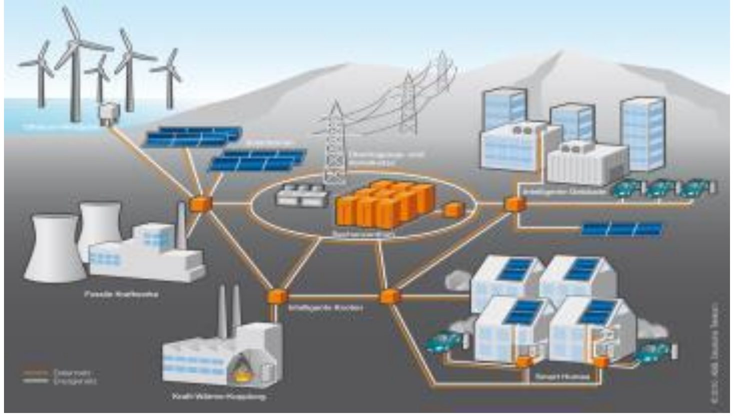
Virtual Machine:

- Power allocation [Nathuji et al. HPDC08]
- Resource consolidation policy [Hwang et al. ISLPED12]

Power Market and Data Center Participation

Power Market:

- Dynamic pricing policy for RS bidding [Caramanis CDC12]
- Smart building RS provision [Paschalidis CDC-ECC11]



Data Center Participation:

- Analytical profit model of data center participation [Ghamkhari SmartGridComm12]
- Analysis of different advanced power market for data centers to participate [Aikema IGCC12]
- Workload allocation among geographically distributed data centers [Wang ICDCS13][Wang SIGMETRICS13]

⇦ Smart Grid

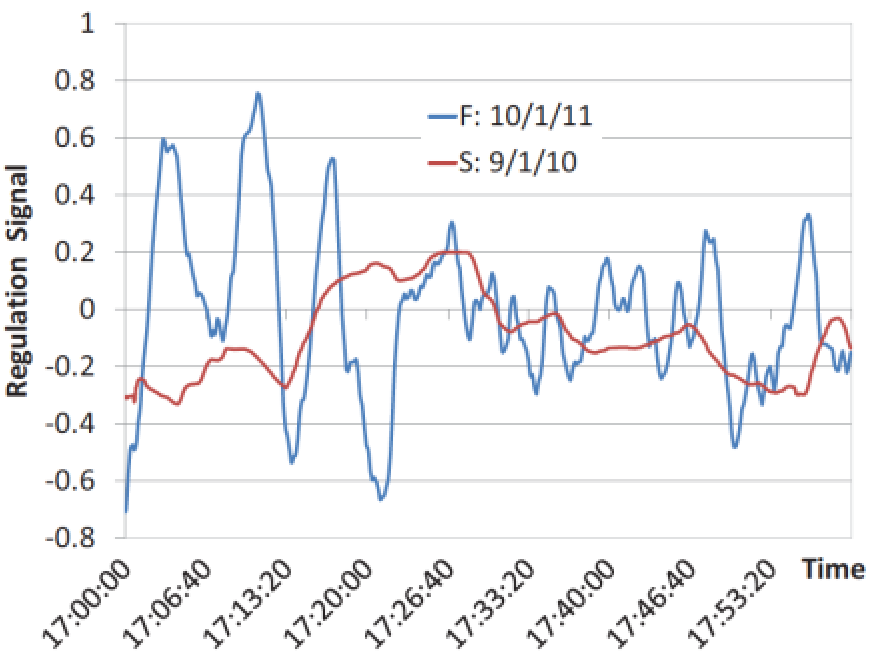
This work is the first to design policies for the data center for:

- *Power budgeting and management*
- *Server commitment*

to enable the data center to participate in the advanced power market programs.

Regulation Service (RS)

Bidding: ($\overline{P}$, R) → Price Settling: Get contract → ISO: RS signal → Data Center Regulation



Typical PJM 150sec ramp rate (F) and 300sec ramp rate (S) regulation signal trajectories

$$P_{cap}(t) = \overline{P} + z(t)R$$

Error: $$\epsilon(t) = \frac{|P_{real}(t) - P_{cap}(t)|}{R}$$

ε(t) needs to be small:

ε(t) > threshold => lose license

Costs: $\Pi^E \overline{P} - \Pi^R R$ **Credit Earned**

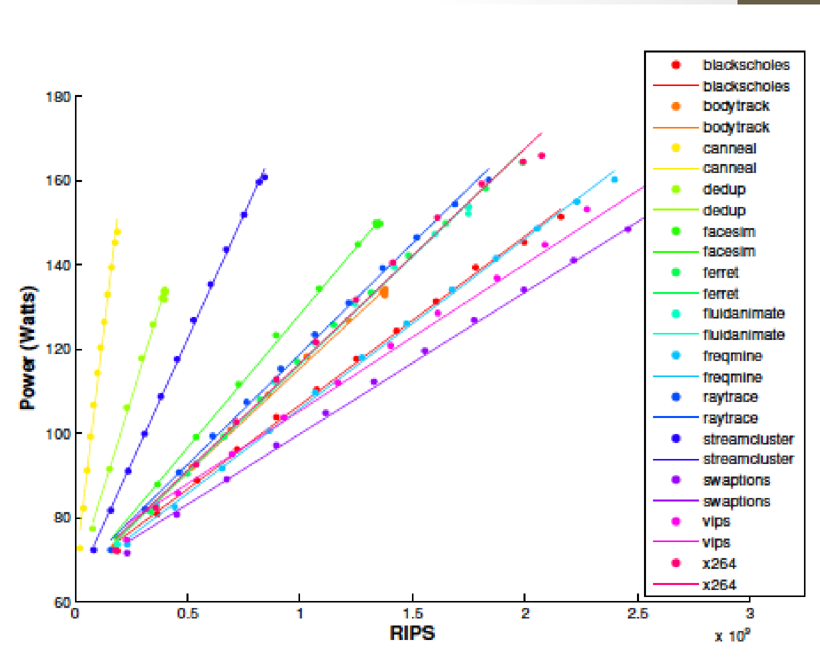
- Π^E and Π^R : market clearing prices
- Credits are reduced based on statistics of ε(t)

Outline

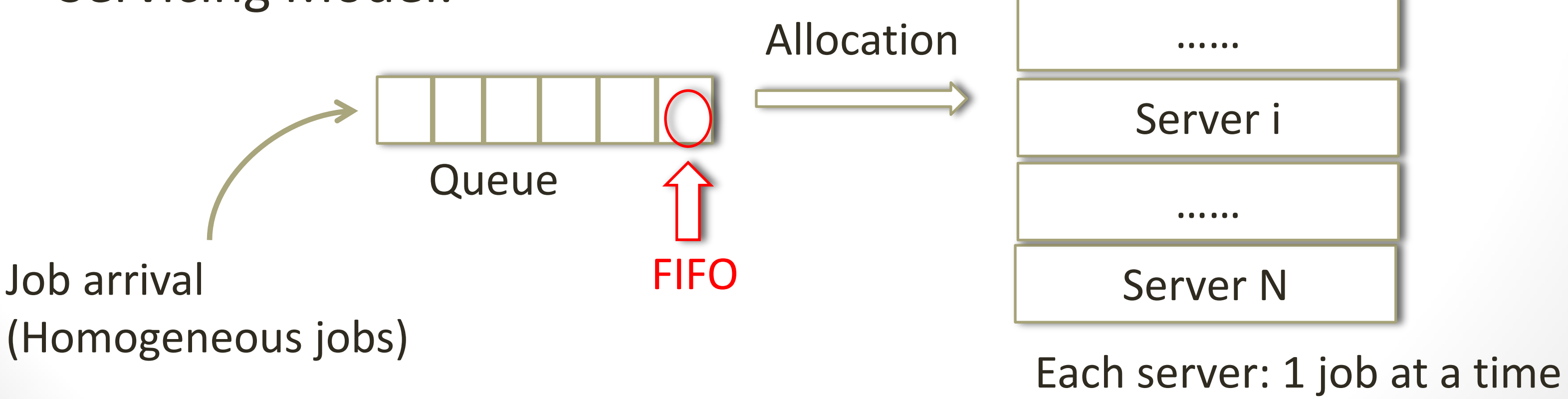
- Background
- Data Center Model
- Dynamic Power Control Policy
- Regulation Reserves Bidding
- Results

Data Center Model

- Server States:
  - Active: $P_{server} = P_{dyn} + P_{static}$
    - P_{dyn} can be modulated by DVFS or CPU resource limits
    - $P_{dyn} = k * RIPS$
  - Idle: $P_{server} = P_{static}$
  - Sleep: $P_{server} = P_{sleep}$
    - Constant low power, but resuming from sleep has time delay (t_{res}) and energy cost (E_{loss})



- Servicing Model:



Outline

- Background
- Data Center Model
- Dynamic Power Control Policy
  - Goals and Optimization Problem
  - Designed Rules and Policies
- Regulation Reserves Bidding
- Results

Dynamic Power Control Policy

- Goals:
  - Reduce the tracking error $\varepsilon(t) = \dfrac{\left|P_{real}(t) - P_{cap}(t)\right|}{R}$
  - Improve the energy efficiency, including:
    - reduce the energy waste during the server state transition period
    - reduce the static energy waste
  - Reduce the workload QoS performance degradation
- Optimization:

$$\min_{u(t) \in U(x(t))} J(x(t), u(t)) = \alpha_1 \left|P_{real}(t) - P_{cap}(t)\right| + \alpha_2 N_{tran}(t) - \alpha_3 N_{sleep}(t) - \alpha_4 N_{peak}(t)$$

Tracking Error — Transition Energy Waste — Static Energy Waste

  - x(t): data center states at t (including server states and workload states);
  - u(t): available control set at t;
  - N_{tran}(t): # of servers that are suspending to or resuming from the sleep state at t;
  - N_{sleep}(t): # of servers in sleep at t;
  - N_{peak}(t): # of servers running at their peak capacities at t.

Dynamic Power Control Policy

Additional Designed Rules:

- For a server that is running a job:

  => keep active at a power rate at least P_{min} until job finished, to guarantee QoS;
- When no jobs are waiting in the queue:

  => no idle server is activated.

- **Server State Transition Rules** [Gandhi IGCC12]:
  - A server that has been in idle > t_{tout} (timeout threshold):

    => goes to sleep;
  - When a new job arrives:

    => select the server with the smallest current $t_{idle}(t)$ to activate;
  - When we need to force servers to sleep:

    => select the servers with the largest current $t_{idle}(t)$ to put to sleep.

$t_{idle}(t)$: the time that a server has been in the idle state at time t.

Dynamic Power Control Policy

Control Policy:

- **Case 1**: $P_{real}(t-1) < P_{cap}(t)$
  1. Active servers with $P_{server} < P_{peak}$: $P_{server} \rightarrow P_{peak}$;
  2. Existing waiting jobs and idle servers: activate idle servers → P_{peak};
  3. Sleeping servers: resume using *server state transition rules*.

  Do the above three steps in order until $P_{real}(t) = P_{cap}(t)$.

- **Case 2**: $P_{real}(t-1) > P_{cap}(t)$
  1. Active servers with $P_{server} < P_{peak}$: P_{server} -> P_{min};
  2. Active servers with $P_{server} = P_{peak}$: P_{server} -> P_{min};
  3. Idle servers: suspend using *server state transition rules*.

  Do the above three steps in order until $P_{real}(t) = P_{cap}(t)$.

Outline

- Background
- Data Center Model
- Dynamic Power Control Policy
- Regulation Reserves Bidding
  - Estimate $(\overline{P}, R)$
- Results

Regulation Reserves Bidding

Average Power Consumption:

Avg. # of Servers in diff. states

Transition power waste

$$\overline{P} = \frac{\int_0^{1h} (\overline{P} + Rz(t))dt}{1h} = \overline{N}_{active} * P_{active} + \overline{N}_{idle} * P_{idle} + \overline{N}_{sleep} * P_{sleep} + \frac{E_{loss,1h}}{1h} \quad (1)$$

Power of Servers in diff. states

# of state transitions in 1h

$$E_{loss,1h} = E_{loss} * N_{res} \approx (t_{res} * N_{tran}) \times (p_b * N_{dc}) \quad (2)$$

Energy waste of each transition

$$N_{dc} = \overline{N}_{active} + \overline{N}_{idle} + \overline{N}_{sleep} \quad (3)$$

Total dynamic energy for processing jobs

$$\overline{N}_{active} = \frac{\int_0^{1h} N_{active}(t)dt}{1h} \approx \frac{E_{dyn}}{P_{dyn,\max} * 1h} \approx \frac{/ * kI}{P_{dyn,\max} * 1h} \quad (4)$$

$$\overline{N}_{idle} = \overline{N}_{sleep}$$

Slack

Min, Max power of servers

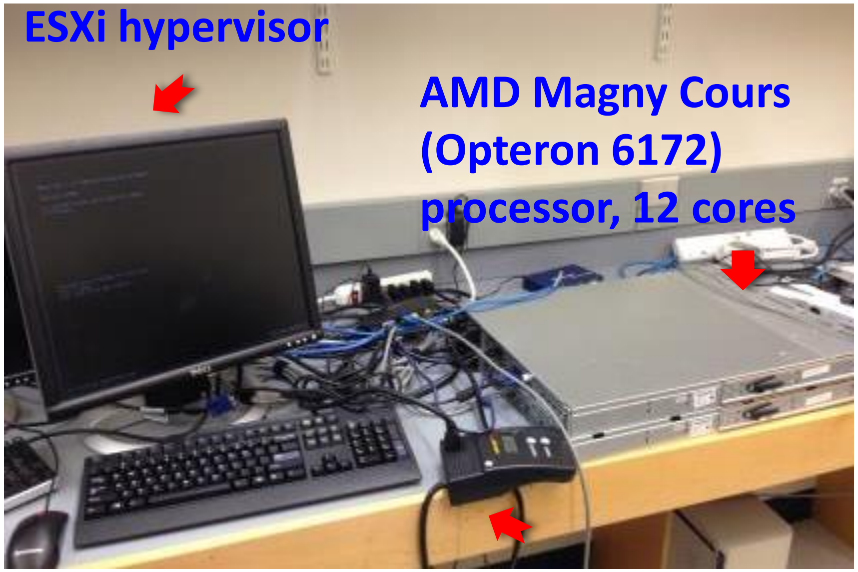
Regulation Reserve: $R \le \min\{N_{dc}P_{peak} - \overline{P},\ \ \overline{P} - N_{dc}P_{sleep}\}$

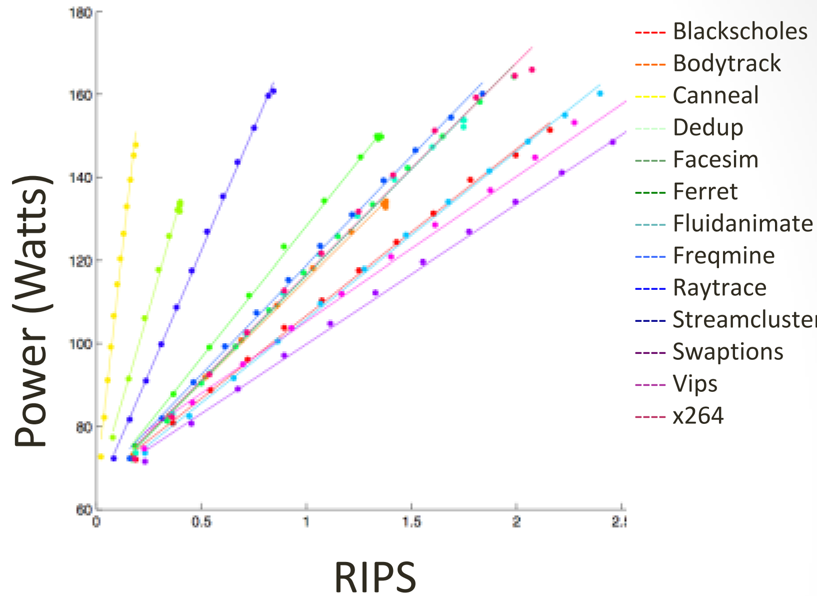
# of servers in the data center

Outline

- Background
- Data Center Model
- Dynamic Power Control Policy
- Regulation Reserves Bidding
- Results
  - Methodology
  - Single Server vs. Data Center
  - Fast Sleep vs. Deep Sleep
  - Impact of Cluster Utilization
  - Impact of Different Workloads

Methodology





Linear Regression: $P_{dyn,j}$

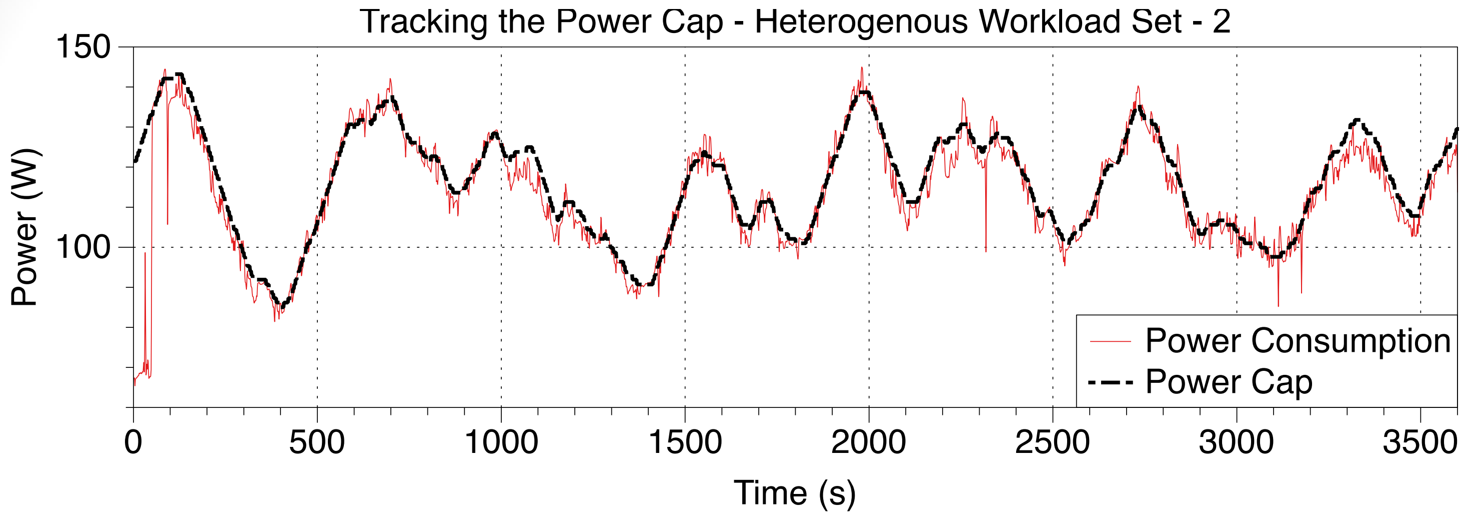
$$P_{server,j} = \boxed{C_j * RIPS_j} + P_{static}$$

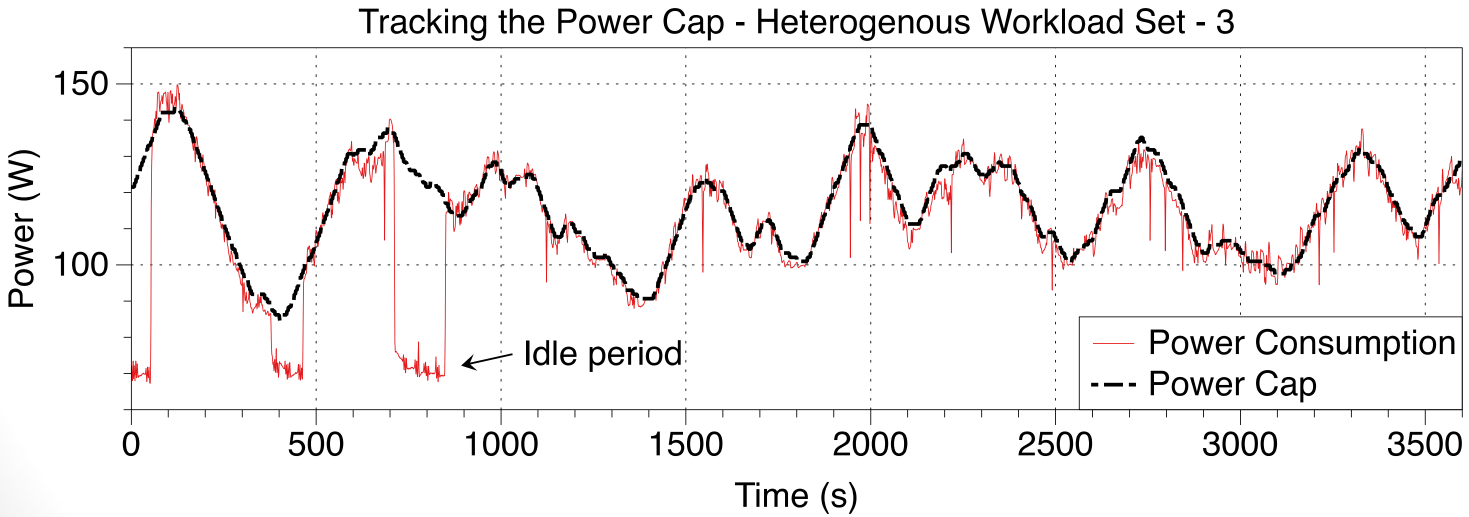
1-hour long HPC type workload *(run 10 times)*

- Applications from PARSEC 2.1 multi-threaded benchmark suite
- Job arrivals follow a Poisson process
- Generated by Monte Carlo method

Data Center: 100 Servers

Power Tracking – Single Server (ICCAD'13)

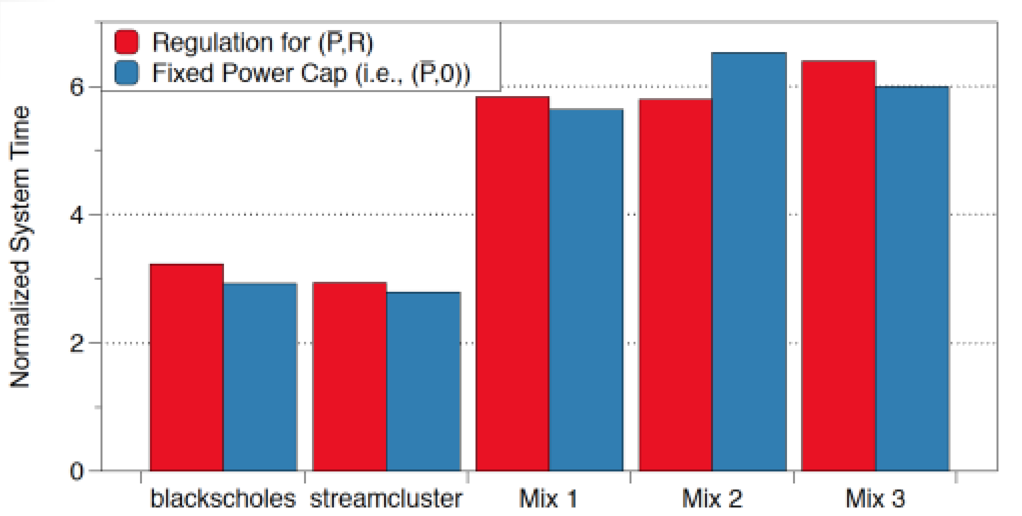




Error 7- 8%

Synthetic workload can fill in the idle periods.

QoS & Monetary Savings (ICCAD'13)

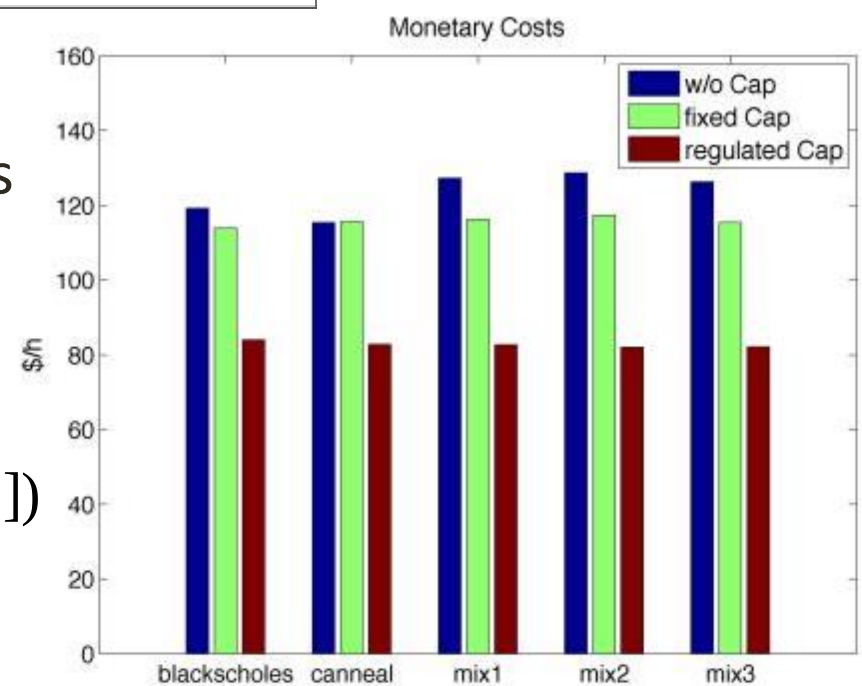


Similar Performance…

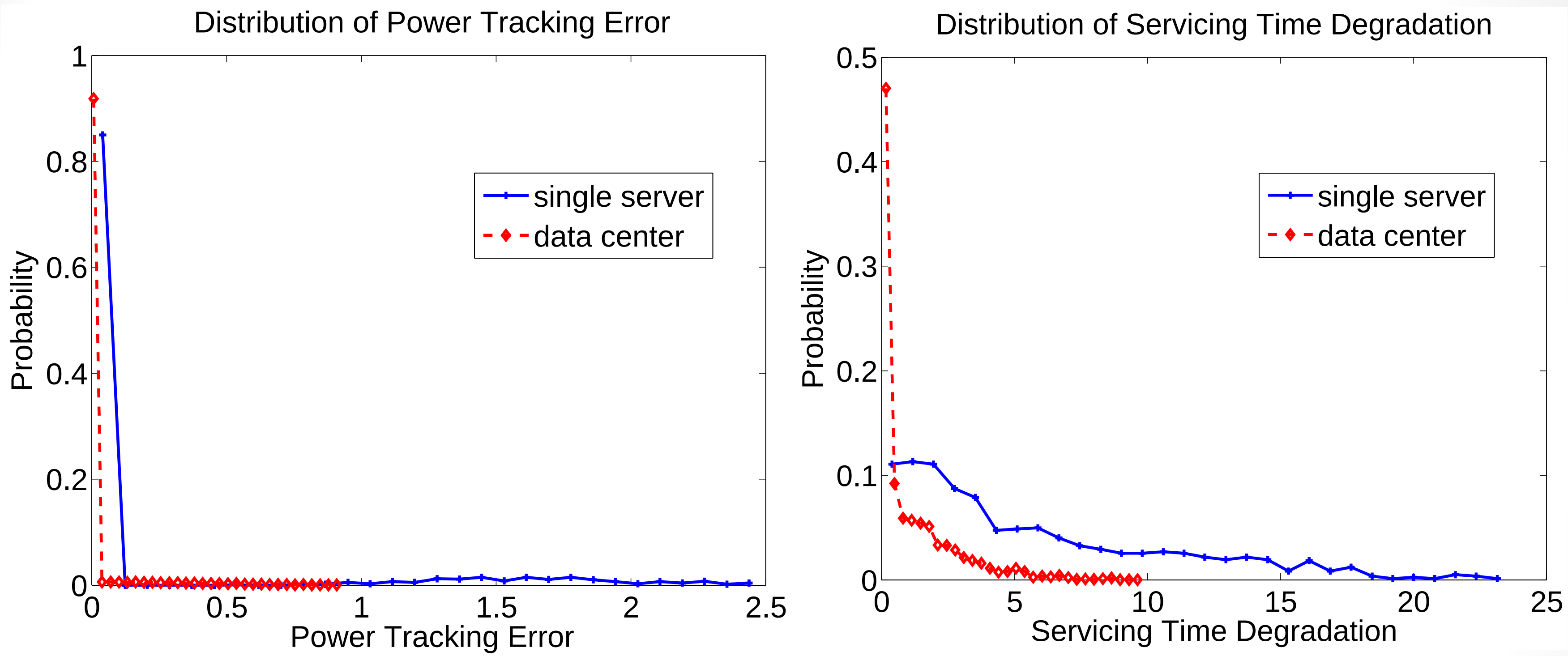
29% Monetary Savings!!!

- 10,000 identical servers
- w/o Cap: $P^{E}\,\text{å}\,P(t)$
- Fixed Cap: $P^{E}\overline{P}$
- Regulation:

$$P^{E}\overline{P} - (P^{R}R - P^{R}c[s_{e}^{2} + (\overline{e})^{2}])$$



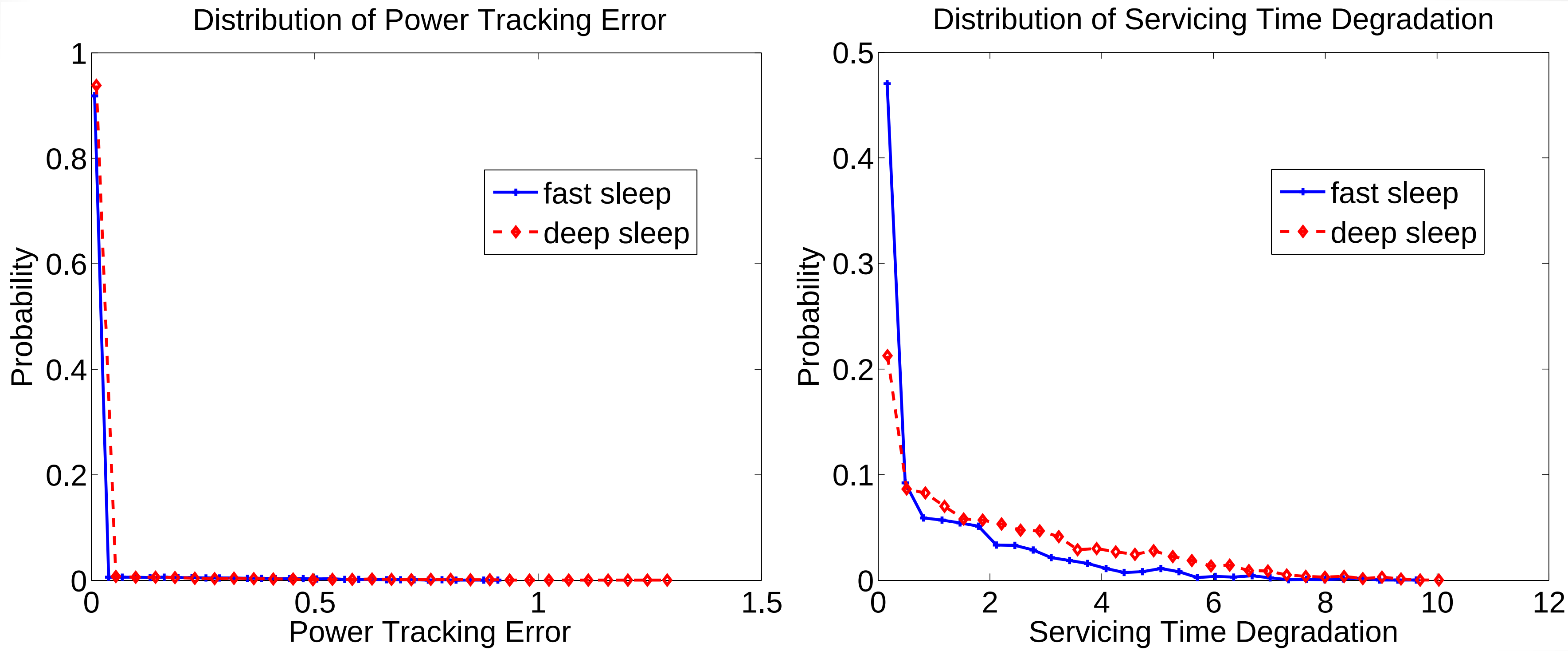
Results - Single Server vs. Data Center



Regulation Reserves (R) /Avg. Power Consumption ($\overline{P}$):

- Single Server: 29.7%
- Data Center: 56.8%

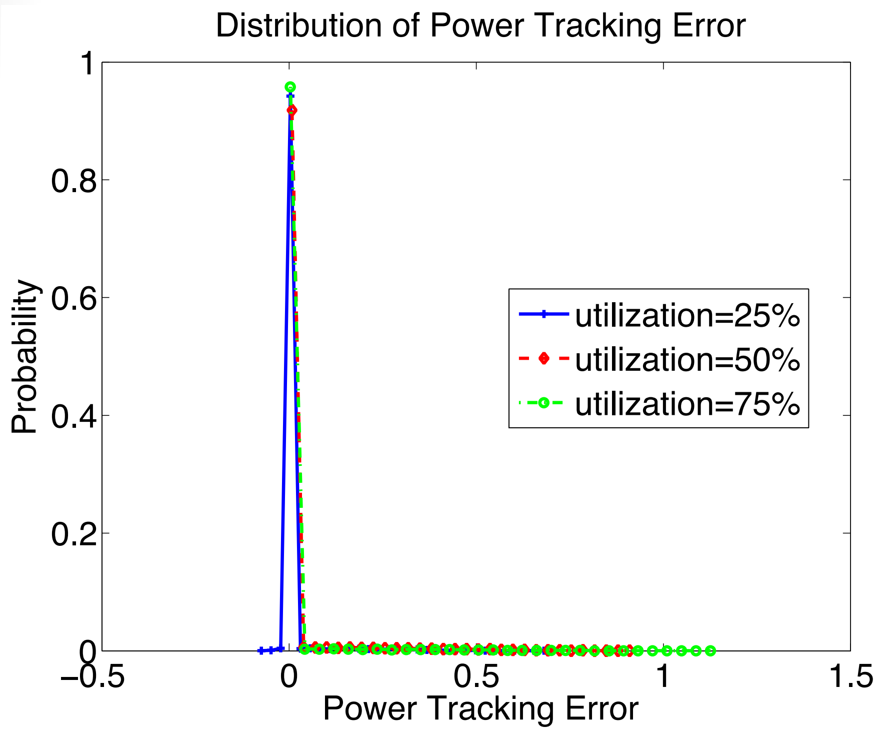
Results - Fast Sleep vs. Deep Sleep

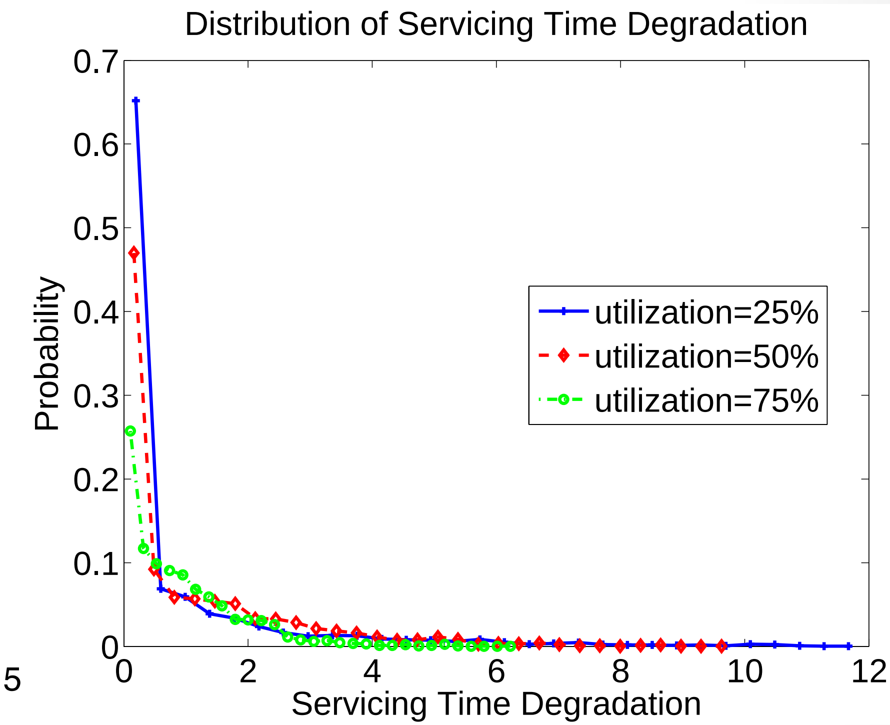


R/ $\overline{P}$:

- Fast Sleep (**t_{res}=10s**, P_{sleep}=10%*P_{peak}, P_{tran}=P_{peak}): 56.8%
- Deep Sleep (**t_{res}=200s**, P_{sleep}=5%*P_{peak}, P_{tran}=P_{peak}): 36.9%

Results - Impact of Cluster Utilization





R/ $\overline{P}$:

- 25% Utilization: 78.0%
- 50% Utilization: 56.8%
- 75% Utilization: 21.8%

Results - Impact of Different Workloads

CLUSTER LEVEL POWER REGULATION ON DIFFERENT WORKLOADS

	Blackscholes	Canneal	Streamcluster	Facesim
$\bar{P}$	$9.75 * 10^3$	$9.71 * 10^3$	$9.84 * 10^3$	$9.84 * 10^3$
R	$5.54 * 10^3$	$4.98 * 10^3$	$5.46 * 10^3$	$5.11 * 10^3$
$\bar{D}$	1.13	1.13	0.21	0.22
σ_D	1.54	0.69	0.26	0.27
$\bar{\epsilon}$	0.03	0.03	0.03	0.03
σ_ϵ	0.10	0.09	0.09	0.09
$R/\bar{P}$	56.8%	51.3%	55.5%	52.0%

QoS Degradation ($\bar{D}$, σ_D)

Tracking Error ($\bar{\epsilon}$, σ_ϵ)

[a] D and σ_D are mean and standard deviation of performance degradation; $\bar{\epsilon}$ and σ_ϵ are mean and standard deviation of tracking error.

Conclusion & Future Work

- A dynamic control policy for the data center RS provision
- An estimation method to calculate the RS provision bidding value
- Data centers are promising candidates for RS provisioning:
  - Accurately track the RS signal;
  - Achieve 50%+ monetary savings;
  - With no major QoS degradation;
  - Regardless of types of workloads.
- Significant improvement of data center vs. prior single server results, taking sleep states, utilization, etc. into account
- Future work: