Server Operational Cost Optimization for Cloud Computing Service Providers over a Time Horizon

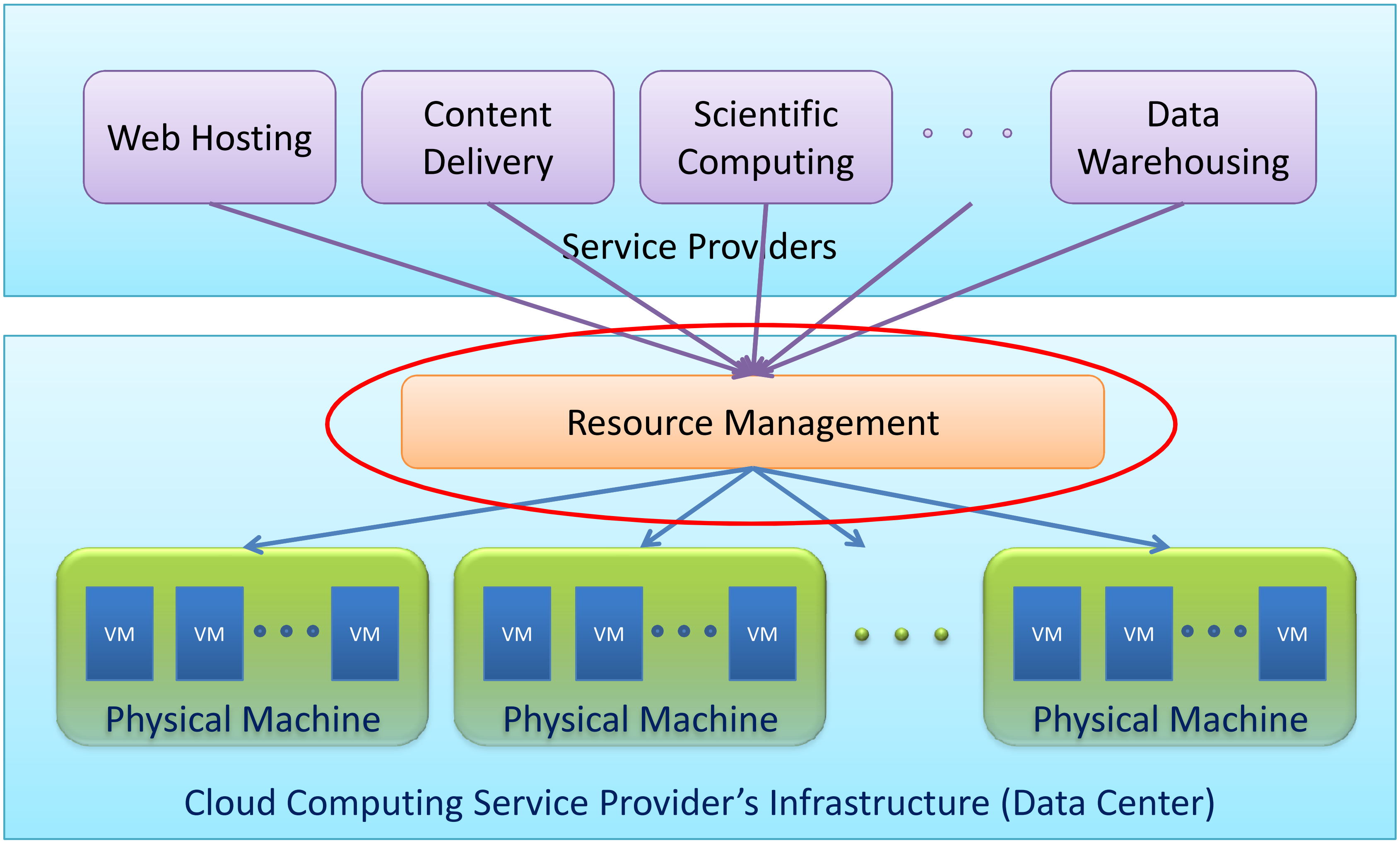
Haiyang(Ocean)Qian and Deep Medhi
Networking and Telecommunication Research Lab (NeTReL)
University of Missouri-Kansas City

USENIX Hot-ICE 2011 workshop
March 29, 2011, Boston

Outline

- Motivation
- Problem Formulation
- Evaluation
- Conclusion and Future Work

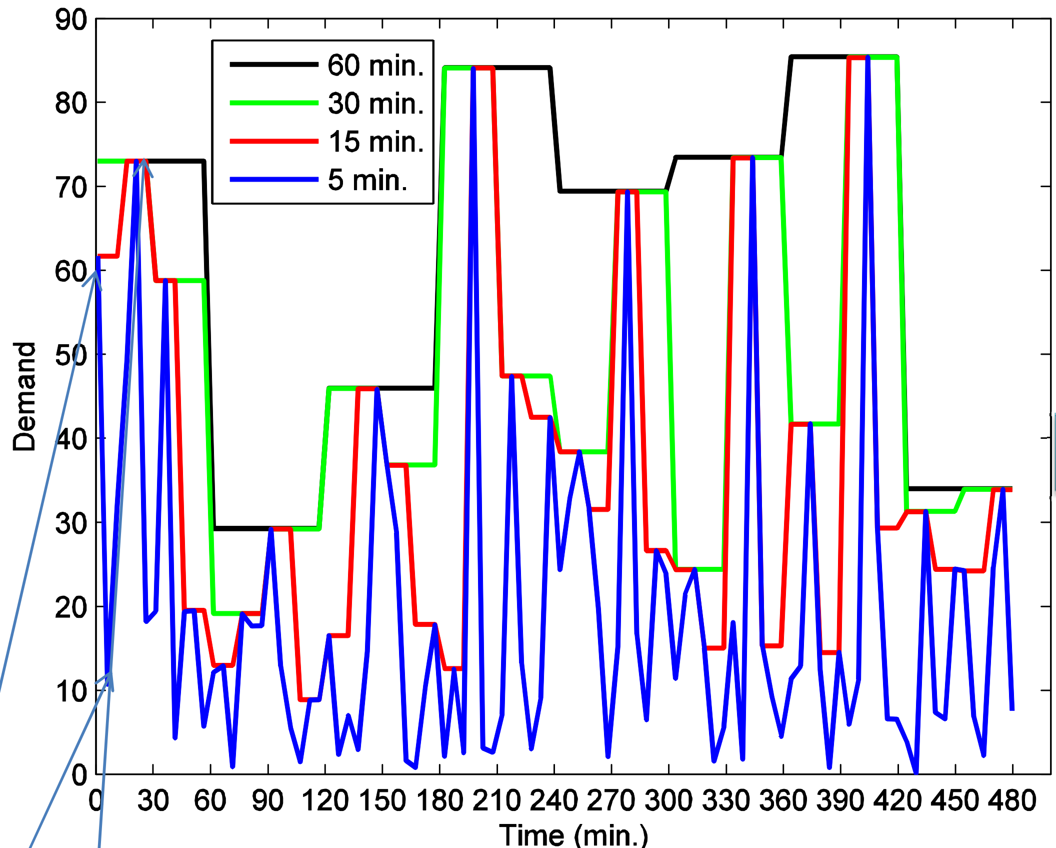
On-Demand Cloud Computing



UMKC

NETREL

Demand on CPU Resource



- Demand on CPU, Memory, I/O etc.

D(t; t + Δ) = max{D(t); … ;D(t + Δ)}

Basic Review Point

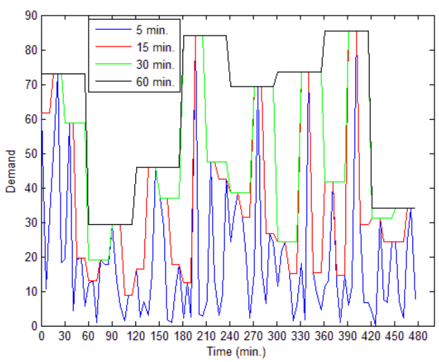
UMKC

NETREL

Server Operational Cost

Cost due to reconfiguration over a time horizon

Demand



Capacity

The # of servers and at which frequency at review points

Energy Consumption Cost

- Wear and Tear (turning on/off cost) most vulnerable part: hard disk

DVFS: Dynamic Voltage/Frequency Scaling

Proportional to the # of servers and the CPU frequency cubic

$V_e \sim f$ — V_e: Voltage, f: Frequency

$P \sim V_e^2 \times f \sim f^3$ — P: Power

$P = P_{fixed} + P_f \times f^3$ — P_{fixed}: Fixed component, P_f: Coefficient

$E = P \times t$ — E: Energy, t: Time

- Proportional to the # of servers
- Positively correlated to CPU frequency

UMKC

NETREL

Outline

- Motivation
- Problem Formulation
- Evaluation
- Conclusion and Future Work

Notations

System Variables

Options	Type	Set Notation	Element Notation	Range
Server	Z^+	$\boldsymbol{I}$	i	$[1,I]$
Frequency	Modular value	$\boldsymbol{J}$	J	$[1,J]$
Time	Z^+	$\boldsymbol{T}$	t	$[1,T]$

Cost Notations

C_{ij}	Power Consumption when server *i* is running at frequency option *j* (per time unit)
C_s^+	Cost of turning a server on at a review point
C_s^-	Cost of turning a server off at a review point

Capacity Notations

V_{ij}	Capacity of server *i* running at frequency option *j*.

Decision Variable:

$y_{ij}(t)$	if server *i* is turned on and operated at frequency *j* at time slot *t*

UMKC

NETREL

Minimize the Server Operational Cost over a Time Horizon

It is quadratic integer programming!

Minimize

$$\sum_{t\in T}\sum_{i\in I}\sum_{j\in J} C_{ij}\cdot y_{ij}(t)$$ +

$$\sum_{t\in T}\sum_{i\in I}(C_s^{+}\cdot\sum_{j\in J} y_{ij}(t)\cdot(\sum_{j\in J} y_{ij}(t)-\sum_{j\in J} y_{ij}(t-1)))$$ +

$$\sum_{t\in T}\sum_{i\in I}(C_s^{-}\cdot\sum_{j\in J} y_{ij}(t-1)\cdot(\sum_{j\in J} y_{ij}(t-1)-\sum_{j\in J} y_{ij}(t)))$$

server power consumption

Turning servers on cost

Dependency on immediate previous time slot

Turning servers off cost

Subject to

$$\sum_{j\in J} y_{ij}(t)\le 1,\ t\in T$$

One server can only be operated at one frequency at one time

$$\sum_{i\in I}\sum_{j\in J} V_{ij}y_{ij}(t)\ge D(t),\ t\in T$$

Demand requirement

Linearize the Objective Function

Introduce two binary variables to represent turning on/off

$$\sum_{j \in J} y_{ij}(t) - \sum_{j \in J} y_{ij}(t-1) - y+(t) + y^{-}(t) = 0$$

$y^{+}(t)$	$y^{-}(t)$
1	0
0	1
0	0
~~1~~	~~1~~

In case of "no change", two variables should be both 0

$$y_i^{+}(t) + y_i^{-}(t) \leq 1, \forall i \in I, \forall t \in T$$

Initialization (assume reshuffling at the beginning of planning)

$$y_i^{+}(1) = \sum_j y_{ij}(1) \qquad y_i^{-}(1) = 0$$

The objective function becomes

$$\sum_{t \in T} \sum_{i \in I} \sum_{j \in J} C_{ij} \cdot y_{ij}(t) + \sum_t \sum_{i \in I} (C_S^{+} \cdot y_i^{+}(t) + C_S^{-} \cdot y_i^{-}(t))$$

UMKC

NETREL

Re-formulate the Problem as Integer Linear Programming

Minimize

$$\sum_{t\in T}\sum_{i\in I}\sum_{j\in J} C_{ij}\cdot y_{ij}(t)+\sum_{t}\sum_{i\in I}(C_s^{+}\cdot y_i^{+}(t)+C_s^{-}\cdot y_i^{-}(t))$$

Subject to

$$\sum_{j\in J} y_{ij}(t)\le 1,\forall i\in I,\forall t\in T$$

$$\sum_{i\in I}\sum_{j\in J} V_{ij}y_{ij}\ge D,\forall t\in T$$

$$\sum_{j\in J} y_{ij}(t)-\sum_{j\in J} y_{ij}(t-1)-y+(t)+y^{-}(t)=0,\forall i\in I,\forall t\in T$$

$$y_i^{+}(t)+y_i^{-}(t)\le 1,\forall i\in I,\forall t\in T$$

$$y_i^{+}(1)=\sum_{j\in J} y_{ij}(1),\forall i\in I$$

$$y_i^{-}(1)=0,\forall i\in I$$

Binary

$$y_{ij}(t),\forall I\in I,\forall j\in J,\forall t\in T \qquad y_i^{+}(t),y_i^{-}(t),\forall i\in I,\forall t\in T$$

Outline

- Motivation
- Problem Formulation
- Evaluation
- Conclusion and Future Work

UMKC

NETREL

Evaluation Setup

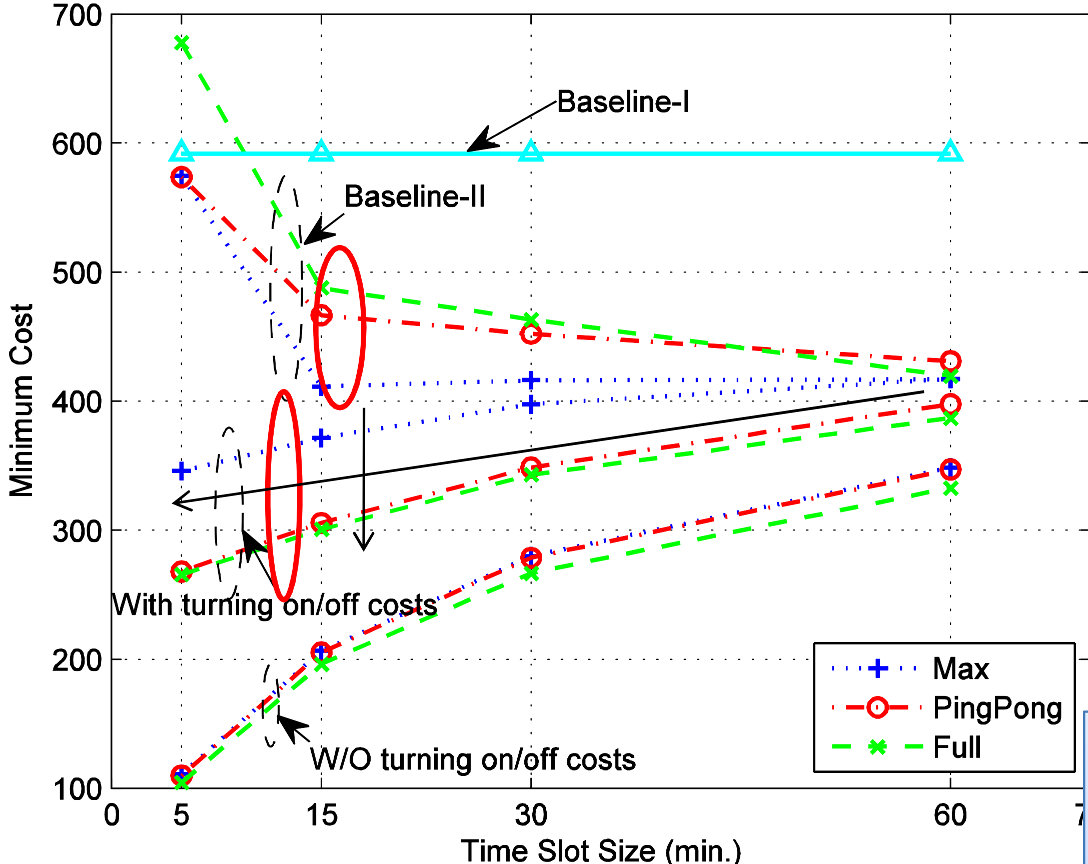
- A 100 homogeneous server cluster with DVFS capability*

#	*j*	1	2	3	4	5	6	7	8
Freq.	F_j	1.4	1.57	1.74	1.91	2.08	2.25	2.42	2.6
Cap.	V_j	.5385	.6038	.6692	.7346	.8	.8645	.9308	1
watts	P_j	60	63	66.8	71.3	76.8	83.2	90.7	100
cents	*Cj*	.42t	.441t	.467t	.4991t	.5376t	.5824t	.6349t	.7t

- The demand is forecasted and profiled every 5 minutes based on the traces of the demand on CPU
  - Assume the distribution is exponential with the mean of 20 (20% utilization)
- How optimal solution is effected by (and how good it is?)
  - Granularity: 5 min, 15 min, 30 min, 60 min
  - DVFS capability: Full, PingPong, Max
  - Relations between power consumption and turning on/off cost

* The CPU frequency is adopted from Chen. *et. al*. SIGMETRICS 2005 paper [6]

Minimum Cost in a 100 Server Cluster



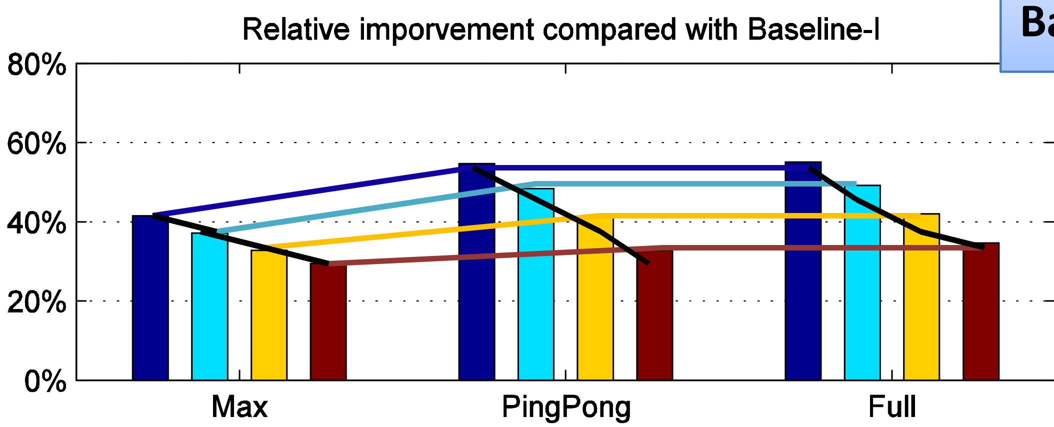
- Outperforms Baseline cases
  - Σ local optimum (BL-II) ≠ global optimum (our solution)
- Finer time granularity, better optimum
  - Partial gain cancelled out because of the existence of turn on/off cost
- More frequency options improves optimum. But, the improvement from PingPong to Full is marginal.

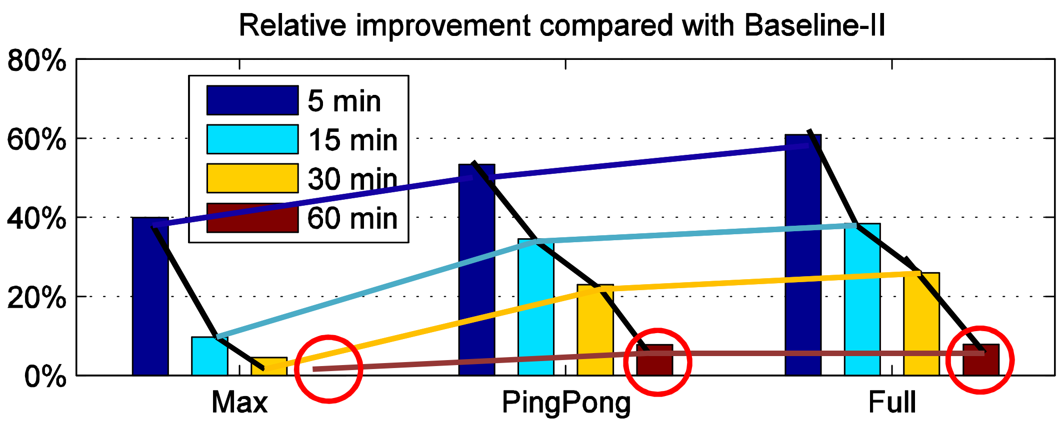
Max: operated at maximum frequency only
PingPong: operated at maximum and minimum freq.
Full: operated at full spectrum (discrete)

Baseline-I: all servers are always on and operated at maximum frequency (static allocation)
Baseline-II: the optimization is executed for each time slot independently (tuning on/off cost is ignored) (independent optimization)

UMKC

Relative Improvement (R)





Baseline-I*:* static allocation
Baseline-II*:* independent optim.

C_b: Cost of baseline
C_{op}: Optimal cost
$R=(C_b - C_{op})/C_{op}$

- Finer granularity, more improvement
- Improvement over Baseline-II diminishes as time granularity gets coarser
- Improvement from PingPong to Full is marginal

Max: operated at maximum frequency only
PingPong: operated at maximum and minimum freq.
Full: operated at full spectrum (discrete)

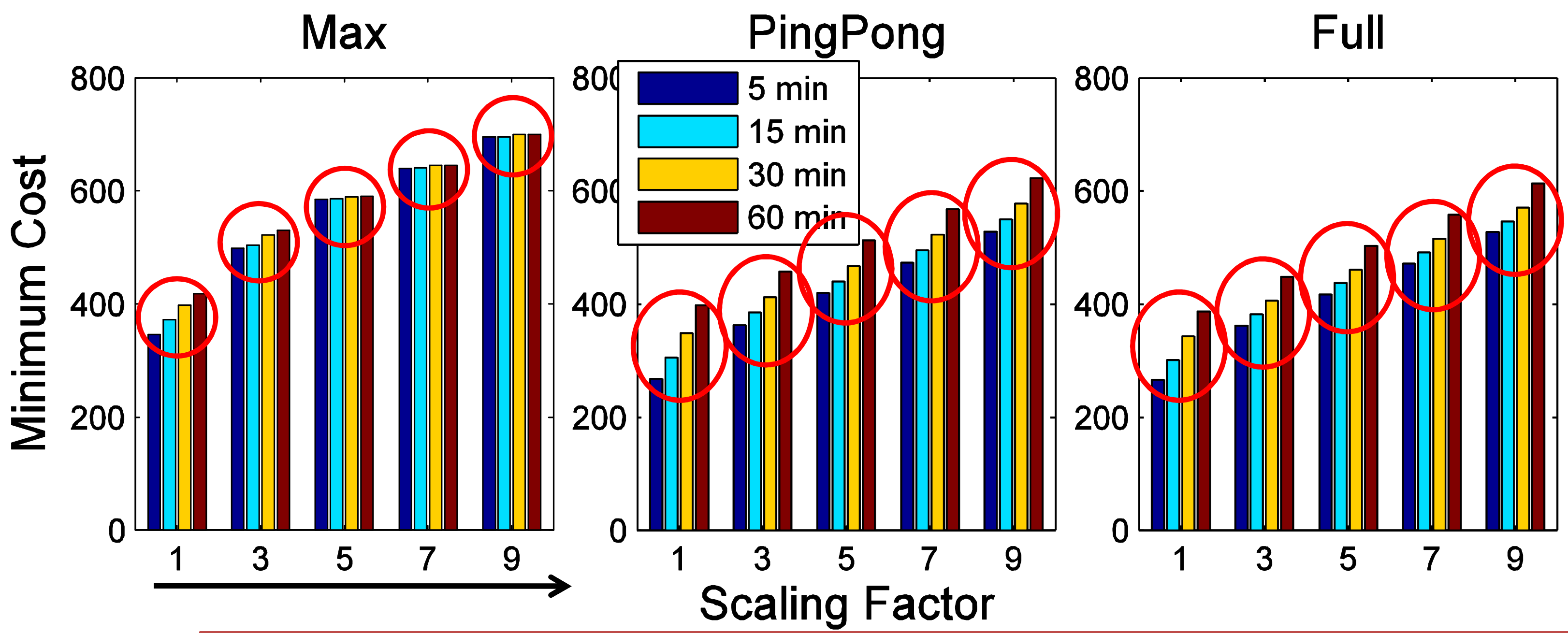
UMKC

NETREL

Scaling Factor Vesus Minimum Cost

Scaling Factor: the ratio between turning on/off cost and power consumption cost

Max: operated at maximum frequency only
PingPong: operated at maximum and minimum frequent
Full: operated at full spectrum (discrete)



- The gain obtained Finer time granularity goes down as SF increase
- Turning on/off cost dominant, less significant impact of time granularity
- Power consumption dominant, more significant impact

Outline

- Motivation
- Problem Formulation
- Evaluation
- Conclusion and Future Work

UMKC

NETREL

Conclusion

- The demand is dynamic over time horizon due to the nature of provisioning service
- Multi-time period mathematical model to optimize server operational cost
- Leverage turning servers on/off and DVFS in synchronous manner
- Significantly reduce the server operational cost compared with static allocation and local optimization
- Finer time slot granularity results in better optimum, but the improvement depends on relationships of cost components
- Optimization aspects for DVFS chip design and operating system software management

UMKC

NETREL

Future Work

- Heuristics for large scale cloud clusters
- Management overhead (such as migration) for reconfiguration cost besides turn on/off cost
- Communication cost when allocating resources
- Leverage turning on/off and DVFS asynchronously
- Uncertainty in demand
- *We need demand trace/profile/workload in real cloud/cluster computing environment*
  - *The demand for resources from individual customers*
  - *Customer information*

References

[1] Barroso, L. A., AND HOLZLE, U. *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines.* Morgan and Claypool Publishers, 2009.

[2] BERTINI, L., LEITE, J. C. B., AND MOSS´E , D. Power optimization for dynamic configuration in heterogeneous web server clusters. *J. Syst. Softw. 83, 4 (2010), 585–598.*

[3] BIANCHINI, R., AND RAJAMONY, R. Power and energy management for server systems. *IEEE Computer 37 (2004), 2004.*

[4] BICHLER, M., SETZER, T., AND SPEITKAMP, B. Capacity planning for virtualized servers. In *Workshop on Information Technologies and Systems (WITS) (Milwaukee, Wisconsin, 2006).*

[5] BOHRER, P., ELNOZAHY, E. N., KELLER, T., KISTLER, M., LEFURGY, C., MCDOWELL, C., AND RAJAMONY, R. *The case for power management in web servers. Kluwer Academic Publishers,* Norwell, MA, USA, 2002, pp. 261–289.

[6] CHEN, Y., DAS, A., QIN, W., SIVASUBRAMANIAM, A., WANG, Q., AND GAUTAM, N. Managing server energy and operational costs in hosting centers. *SIGMETRICS Perform. Eval. Rev. 33, 1 (2005), 303–314.*

[7] FILANI, D., HE, J., GAO, S., RAJAPPA, M., KUMAR, A., SHAH, R., AND NAAPPAN, R. Dynamic data center power management: Trends, issues and solutions. *Intel Technology Journal* (2008).

[8] GREENBERG, A., HAMILTON, J., MALTZ, D. A., AND PATEL, P. The cost of a cloud: research problems in data center networks. *SIGCOMM Comput. Commun. Rev. 39, 1 (2009), 68–73.*

[9] JOHNSON, L. A., AND MONTGOMERY, D. C. *Operations Research in Production Planning, Scheduling, and Inventory Control.* John Wiley & Sons, 1974.

[10] MENG, X., PAPAS, V., AND ZHANG, L. Improving the scalability of data center networks with traffic-aware virtual machine placement. In *INFOCOM (2010).*

[11] PETRUCCI, V., LOQUES, O., AND MOSS´E, D. Dynamic optimization of power and performance for virtualized server clusters, Technical Report, 2009.

[12] PINHERIO, E., BIANCHINI, R., CARRERA, E. V., AND HEATH, T. Dynamic cluster reconfiguration for power and performance. In *Compilers and Operating Systems for Low Power (2003),* L. Benini, M. Kandemir, and J. Rammanujam, Eds., Kluwer.

[13] PI´O RO, M., AND MEDHI, D. *Routing, Flow, and Capacity Design in Communication and Computer Networks. Morgan Kaufmann* Publishers, 2004.

[14] VISHWANATH, K. V., AND NAGAPPAN, N. Characterizing cloud computing hardware reliability. In *Proc. of 1st ACM Symposium on Cloud Computing (June 2010).*

Thank you!

Questions?