ELECTRICITY COST MINIMIZATION IN DATA CENTERS USING LIVE VM MIGRATION

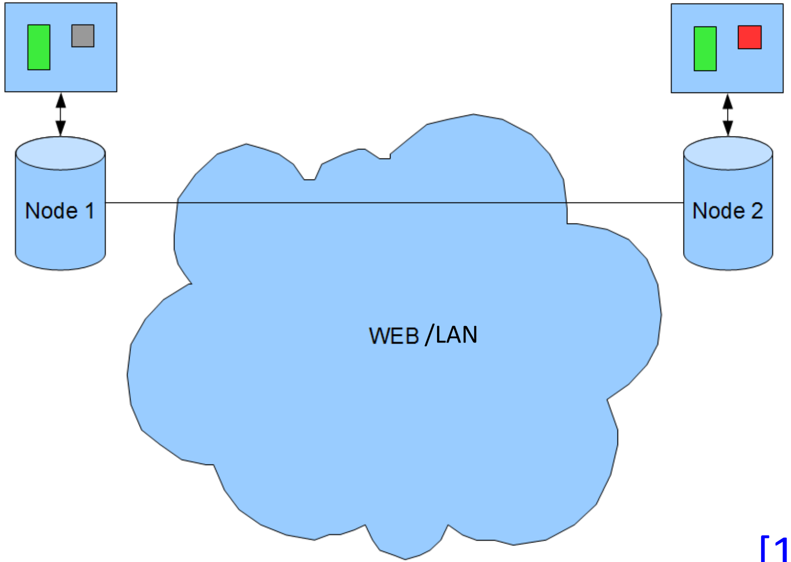
K M Sabidur Rahman
Friday Group Meeting, Netlab
UC Davis

UCDAVIS

Agenda

- VM Migration
- Electricity market
- Electricity consumption in Data centers
- Simulation topology
- Workload in Data Center
- Problem statement

VM Migration



[1] www.greenstarnetwork.com

Estimated annual electricity cost

Company	Servers	Electricity	Cost
eBay	16K	$\sim 0.6\times10^5$ MWh	~$3.7M
Akamai	40K	$\sim 1.7\times10^5$ MWh	~$10M
Rackspace	50K	$\sim 2\times10^5$ MWh	~$12M
Microsoft	>200K	$>6\times10^5$ MWh	>$36M
Google	>500K	$>6.3\times10^5$ MWh	>$38M
USA (2006)	10.9M	610×10^5 MWh	$4.5B
MIT campus		2.7×10^5MWh	$62M

[3] Qureshi, A., Weber, R., Balakrishnan, H., Guttag, J., Maggs, B., Cutting the electric bill for internet-scale systems. SIGCOMM39, 123–134 (2009)

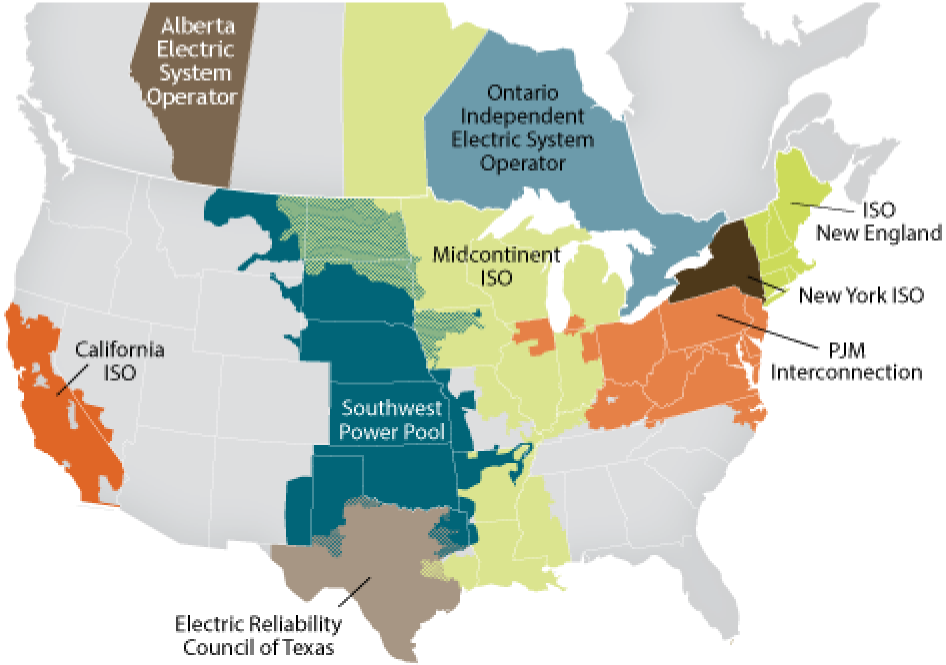
Microsoft

facebook

amazon web services

Akamai

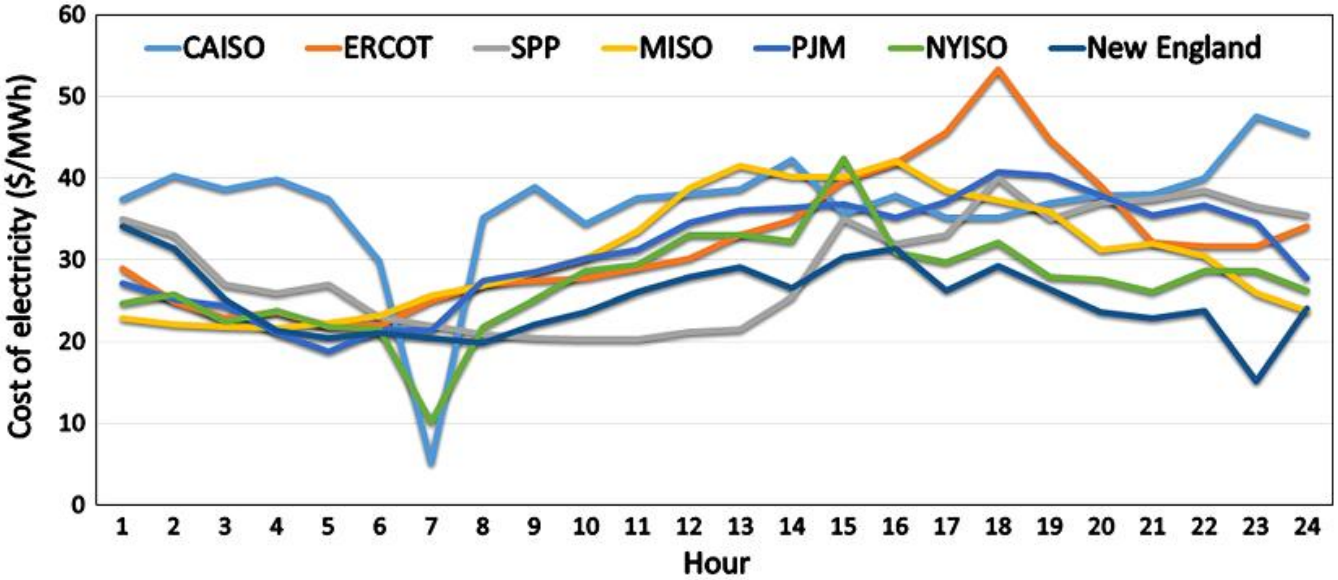
Electricity market



[4] http://www.isorto.org/about/default

UCDAVIS

Variable electricity cost



[2] A. Gupta, U. Mandal, P. Chowdhury, M. Tornatore and B. Mukherjee, “Cost-efficient live VM migration based on varying electricity cost in optical cloud networks”, Photonic Network Communications, Sep 2015

UCDAVIS

Market types

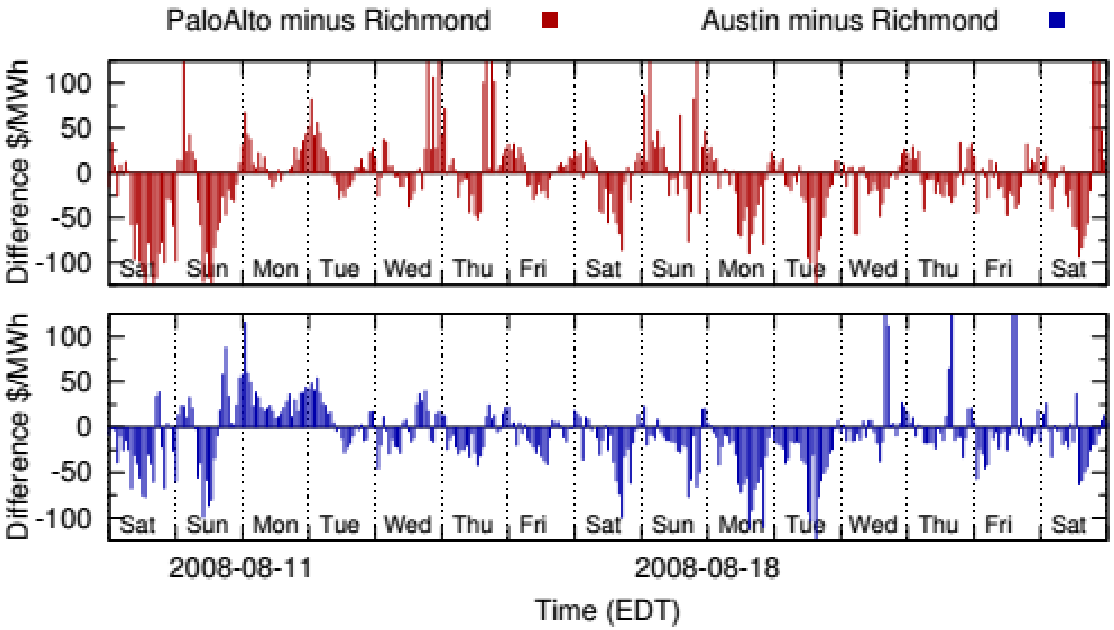
Each RTO operates multiple parallel wholesale markets.

Day-ahead markets (futures): provide hourly prices for delivery during the following day. The outcome is based on expected load

Real-time markets (spot) are balancing markets where prices are calculated every five minutes or so, based on actual conditions, rather than expectations. This market accounts for a small fraction of total energy transactions (less than 10% of total in NYISO)

[3] Qureshi, A., Weber, R., Balakrishnan, H., Guttag, J., Maggs, B., Cutting the electric bill for internet-scale systems. SIGCOMM39, 123–134 (2009)

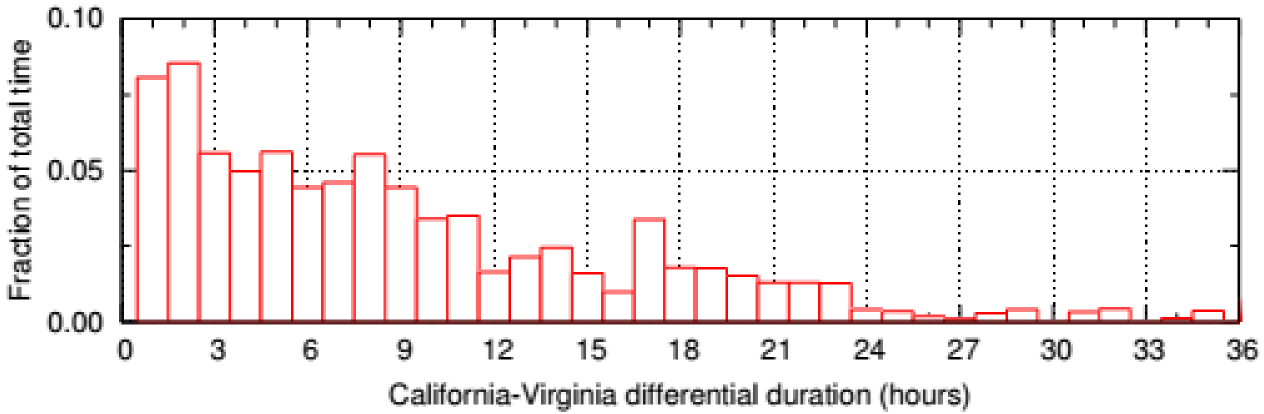
Hourly price differentials



[3] Qureshi, A., Weber, R., Balakrishnan, H., Guttag, J., Maggs, B., Cutting the electric bill for internet-scale systems. SIGCOMM39, 123–134 (2009)

UCDAVIS

Duration of price differential



[3] Qureshi, A., Weber, R., Balakrishnan, H., Guttag, J., Maggs, B., Cutting the electric bill for internet-scale systems. SIGCOMM39, 123–134 (2009)

UCDAVIS

Actual electricity bills

- Electricity bills varies, but, does the "Existing contracts" allow to be benefited from that?
- Real-time pricing program is getting popular this days. (PJM-MISO offers billing based on hourly consumptions) [5]
- Companies like Facebook/Akamai sometimes doesn't pay electricity bills directly. They lease "wholesale" data center space from third-party landlords

[5] "Midwest RTO Commonwealth Edison", www.comed.com
[6] "Facebook Datacenter", http://www.datacenterknowledge.com/the-facebook-data-center-faq/

Assumptions for simulation

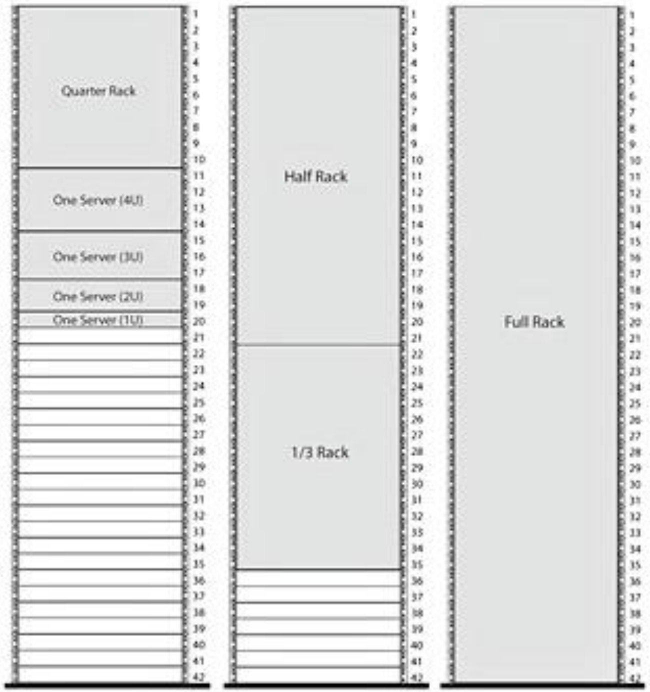
1. The company is enrolled in a Real-time Pricing Program for electricity billing
2. Real-time prices are known from RTO/ISOs and vary hourly
3. The electric bill paid by the service operator is proportional to consumption
4. Migration of work load using our method does not significantly alter prices and market behavior in RTO/ISOs

Rack and Server

42U(just over 6 feet) Racks are common [2][7]

U = 1.75 inch

Server size = multiple of U (1u/2U/3U/4U) [7]



[2] A. Gupta, U. Mandal, P. Chowdhury, M. Tornatore and B. Mukherjee, “Cost-efficient live VM migration based on varying electricity cost in optical cloud networks”, Photonic Network Communications, Sep 2015

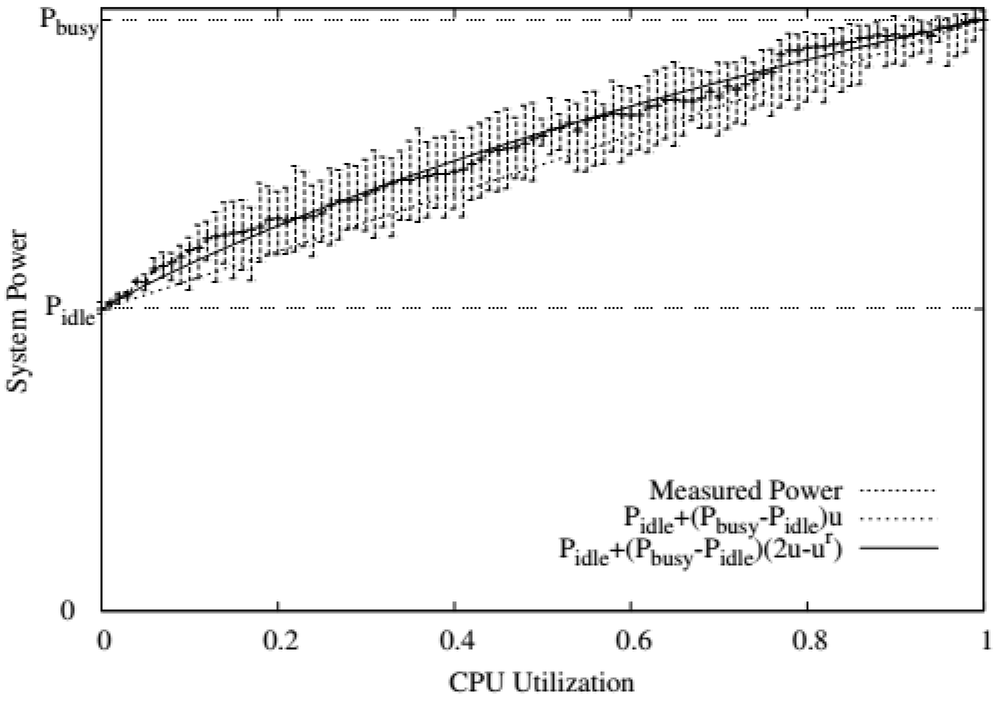
[7] https://en.wikipedia.org/wiki/Rack_unit

Sever power

Power usage characteristics of large collections of servers (up to 15 thousand)

r = constant = 1.4

u = CPU utilization



[8] X. Fan, W.-D. Weber, and L. A. Barroso, "Power Provisioning for a Warehouse-sized Computer," in ACM International Symposium on Computer Architecture, 2007

UCDAVIS

Power Model

Power consumed by a rack

$= n * (P_{idle} + (P_{busy} – P_{idle}) * (2u – u^r))$

where

n = number of servers in a rack

r = constant = 1.4

u = CPU utilization

Cooling and other extra power consumption

= n * (PUE - 1) * P_{busy}

where

PUE = a measure of data center energy efficiency: how much energy is used by the computing equipment (in contrast to cooling and other overhead)

$$\text{PUE} = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$$

[9] "PUE", https://en.wikipedia.org/wiki/Power_usage_effectiveness

Power Model

Total power consumed by a data center =

Number of racks * power consumed by a rack

VM Migration power consumption

Migrating a VM from a source DC to a destination DC consume power, in the backbone network

Migration power consumption =

*the total number of bits transferred in a VM migration **

*the power consumed by a core router in transmitting one bit **

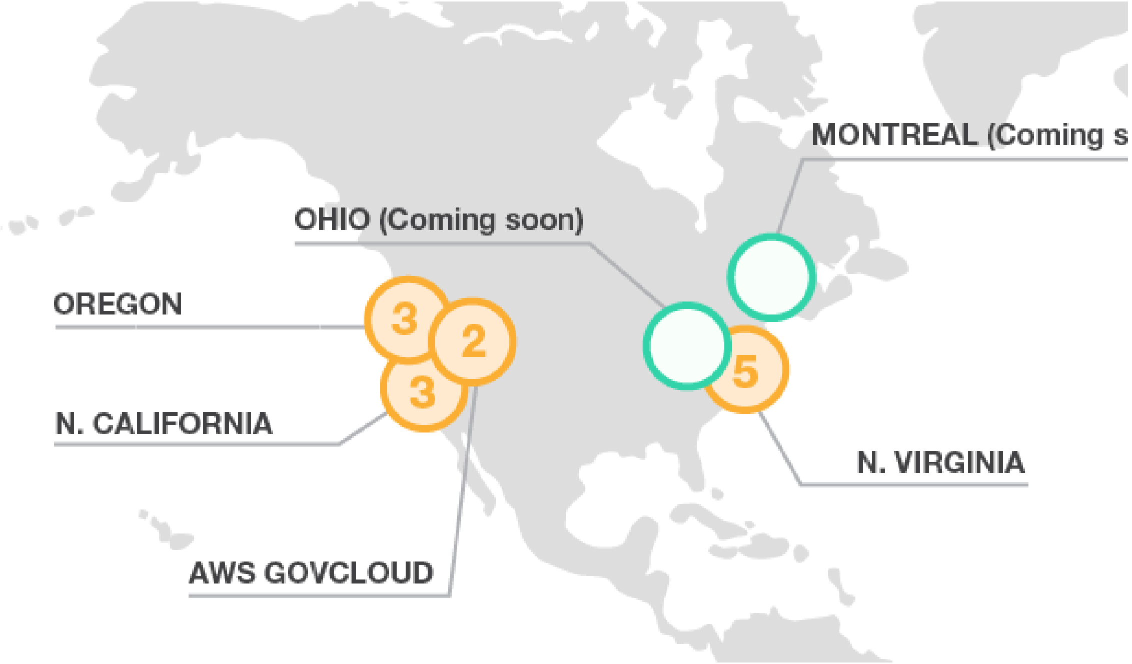
the cost of electricity at the core router

[2] A. Gupta, U. Mandal, P. Chowdhury, M. Tornatore and B. Mukherjee, "Cost-efficient live VM migration based on varying electricity cost in optical cloud networks", Photonic Network Communications, Sep 2015

Google datacenters

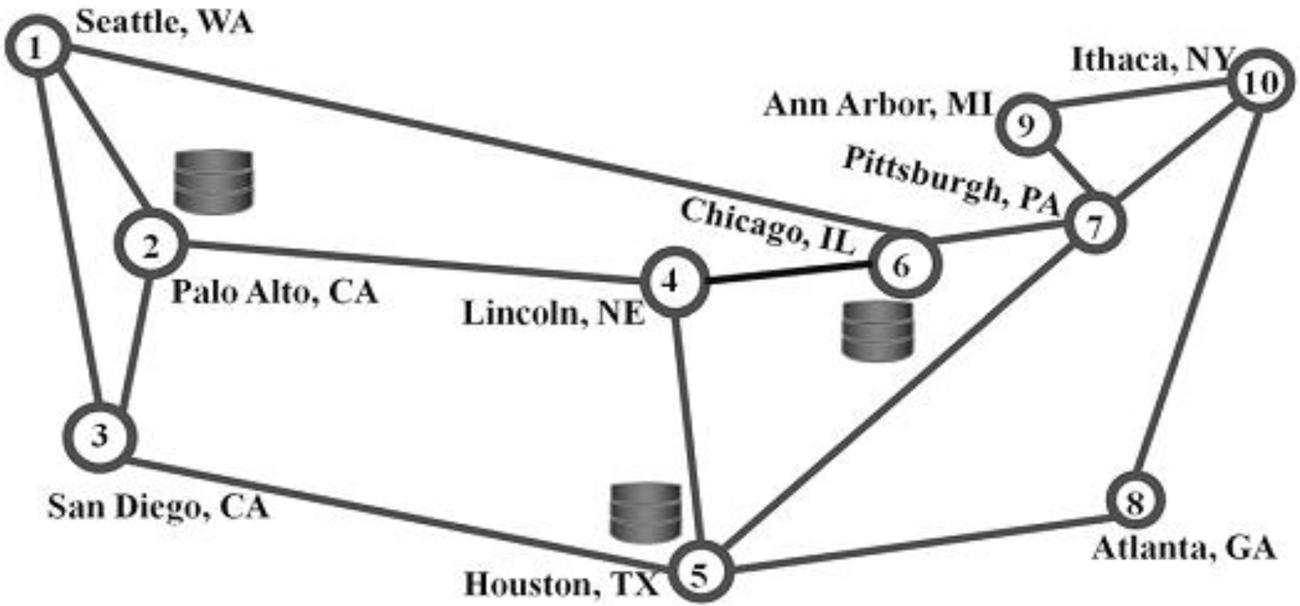
[10] http://royal.pingdom.com/2008/04/11/map-of-all-google-data-center-locations/

Amazon AWS



[11] https://aws.amazon.com/about-aws/global-infrastructure/

UCDAVIS



[2] A. Gupta, U. Mandal, P. Chowdhury, M. Tornatore and B. Mukherjee, “Cost-efficient live VM migration based on varying electricity cost in optical cloud networks”, Photonic Network Communications, Sep 2015

UCDAVIS

Characterization of workload

(a) the duration of task executions is bimodal in that tasks either have a short duration or a long duration

(b) most tasks have short durations

(c) Most resources are consumed by a few tasks with long duration that have large demands for CPU and memory

[12] A.K.Mishra, J.L. Hellerstein , W. Cirne and C.R Das, “Towards Characterizing Cloud Backend Workloads: Insights from Google Compute Clusters,” SIGMETRICS Perform. Eval. Rev., vol.37, pp. 34-41, March 2010

Task shape

Multi-dimensional representation of task resource usage or task shape. The dimensions are:

- time in seconds,
- CPU usage in cores, and
- memory usage in gigabytes

- Typically, core-hours or GB-hours are used as metrics

[12] A.K.Mishra, J.L. Hellerstein , W. Cirne and C.R Das, “Towards Characterizing Cloud Backend Workloads: Insights from Google Compute Clusters,” SIGMETRICS Perform. Eval. Rev., vol.37, pp. 34-41, March 2010

Characterization of workload

- Duration ranges from 300 to 86,400 seconds
- CPU usage ranges from 0 to 4 cores; and
- Memory usage varies from 0 to 8 gigabytes
- same clusters typically experiences same type of work loads
- Two types of long tasks: The first are computationally intensive, user-facing services such as work done by a **map reduce** master in processing web search results. The second kind of **long-running tasks relate to log-processing operations**, such as analysis of click throughs.

Google trace data

The clusterdata-2011-2 trace represents 29 day's worth of cell information from May 2011, on a cluster of about 12.5k machines

[13] https://github.com/google/cluster-data

Google data

The job events table contains the following fields:

1. timestamp
2. missing info
3. job ID
4. event type
5. user name
6. scheduling class
7. job name
8. logical job name

[13] https://github.com/google/cluster-data

Problem Statement

Given:

optical backbone network topology,

a set of DC nodes,

hourly prices of electricity at each node,

link capacities,

number of VMs a DC can host,

a multi-hour period and

dynamic requests with specific duration/priority

Problem Statement

Result:

Minimize the operating cost of the VMs over given period by deciding whether (including when and where) or not to migrate the VMs(also which VMs to migrate and in what sequence).

The services running on the VMs comes dynamically and have varying request duration. After we have decided that it's worth the migration, we'll calculate the cost saved by the migration.

Reference

[1] www.greenstarnetwork.com

[2] A. Gupta, U. Mandal, P. Chowdhury, M. Tornatore and B. Mukherjee, “Cost-efficient live VM migration based on varying electricity cost in optical cloud networks”, Photonic Network Communications, Sep 2015

[3] Qureshi, A., Weber, R., Balakrishnan, H., Guttag, J., Maggs, B., Cutting the electric bill for internet-scale systems. SIGCOMM39, 123–134 (2009)

[4] http://www.isorto.org/about/default

[5] “Midwest RTO Commonwealth Edison”, www.comed.com

[6] “Facebook Datacenter”, http://www.datacenterknowledge.com/the-facebook-data-center-faq/

[7] https://en.wikipedia.org/wiki/Rack_unit

[8] X. Fan, W.-D. Weber, and L. A. Barroso, “Power Provisioning for a Warehouse-sized Computer,” in ACM International Symposium on Computer Architecture, 2007

[9] “PUE”, https://en.wikipedia.org/wiki/Power_usage_effectiveness

[10] http://royal.pingdom.com/2008/04/11/map-of-all-google-data-center-locations/

[11] https://aws.amazon.com/about-aws/global-infrastructure/

[12] A.K.Mishra, J.L. Hellerstein , W. Cirne and C.R Das, “Towards Characterizing Cloud Backend Workloads: Insights from Google Compute Clusters,” SIGMETRICS Perform. Eval. Rev., vol.37, pp. 34-41, March 2010

[13] https://github.com/google/cluster-data


Thank You