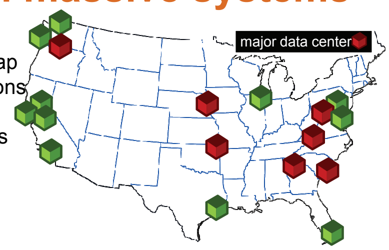
cutting the electric bill for internet-scale systems

Asfandyar Qureshi (MIT)
Rick Weber (Akamai)
Hari Balakrishnan (MIT)
John Guttag (MIT)
Bruce Maggs
(Duke/CMU/Akamai)

context: massive systems

Google:
- estimated map
- tens of locations in the US
- >0.5M servers



others
- thousands of servers / multiple locations
- Amazon, Yahoo!, Microsoft, Akamai
- Bank of America (≈50 locations), Reuters

electricity expenses

millions spent annually on electricity
- Google ~ 500k custom servers ~ $40 million/year
- Akamai ~ 40k off-the-rack servers ~ $10 million/year

electricity costs are growing
- systems are rapidly increasing in size
- outpacing energy efficiency gains

relative cost of electricity is rising
- 3-year server total cost of ownership by 2012:
  - electricity ≈ 2 × hardware
- bandwidth prices are falling

what is being done

reduce number of kWh
- energy efficient hardware
- virtualization and consolidation
- power off servers when possible
- cooling (air economizers instead of chillers, etc.)
- dc power distribution, etc.

reduce cost per kWh
- build data-centers where average price is low

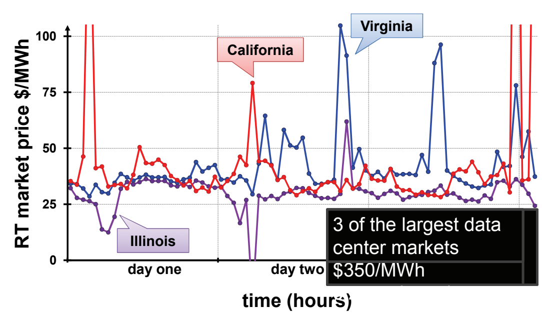
our proposal

exploit electricity market dynamics
- geographically uncorrelated price volatility
- monitor real-time market prices and adapt request routing

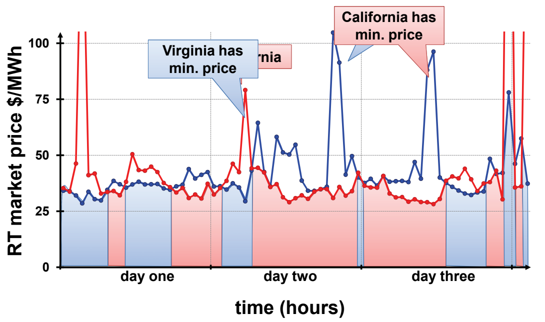
skew load across clusters based on prices
- leverage service replication and spare capacity

adapting to real-time prices is a new idea…
- complementary to energy efficiency work

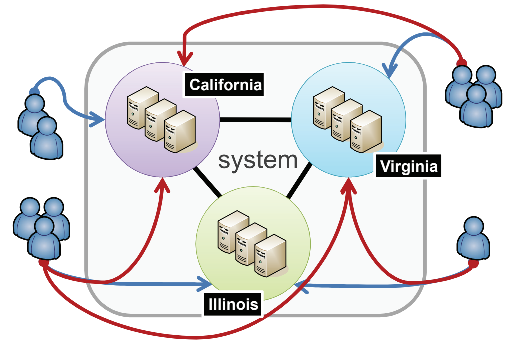
exploiting price volatility



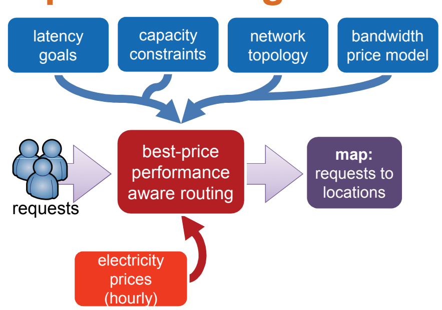
exploiting price volatility



system model (status quo)



request routing framework



will our proposal work?

will our proposal work?

does electricity usage depend on server load?

- how much can we reduce a location's electricity consumption by routing clients away from it?

will our proposal work?

does electricity usage depend on server load?

latency concerns

- how far away from a client is the cheap energy?

will our proposal work?

does electricity usage depend on server load?

latency concerns

bandwidth costs could rise

- cheaper electricity ~ more expensive bandwidth?

will our proposal work?

does electricity usage depend on server load?

latency concerns

bandwidth costs could rise

is there enough spare capacity?

how much can we save by exploiting price volatility?

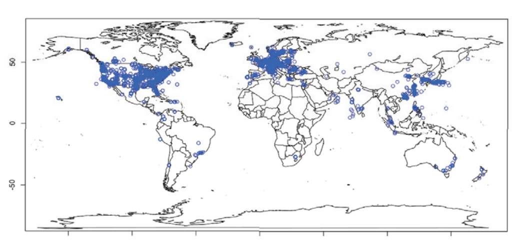
- **today: large companies more than $1M/year**
- **with better technology: more than $10M/year**
- **better than placing all servers in cheapest market**

Traffic Statistics

30,000+ domains
1.1 Tbps daily peak traffic
6,419 terabytes / day
274 billion hits / day
274 million unique client IP addresses / day
in 2009 expect to deliver more bits than in 1998-2008 combined

16

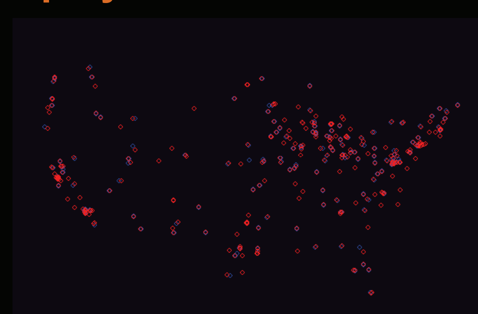
Network Deployment



40000+ Servers | 1450+ POPs | 950+ Networks | 67+ Countries

17

Deployment in U.S.



18

Assigning Clients to Servers

"High level": Map client to a cluster based on client's nameserver's IP address.
Algorithm: stable marriage with multi-dimensional hierarchical capacity constraints.

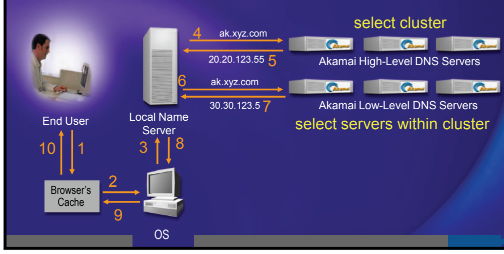
"Low level": Assign client to specific server or servers within cluster based on content requested. Algorithm: consistent hashing.

Embedded Image Delivery (e.g., Amazon)

Embedded URLs are Converted to ARLs

```
<html>
<head>
<title>Welcome to xyz.com!</title>
</head>
<body>                     ak
<img src=http://www.xyz.com/logos/logo.gif>
<img src=http://www.xyz.com/jpgs/navbar1.jpg>
<h1>Welcome to our Web site!</h1>
<a href="page2.html">Click here to enter</a>
</body>
</html>
```


Akamai DNS Resolution



Mapping

Maps IP address of client's name server and type of content being requested to an Akamai cluster.

Note: Doesn't depend on content provider (although indicated by ak.xyz.com).

Special cases: Akamai Accelerated Network Partners (AANPs)

General case: "Core Point" analysis

generality of results

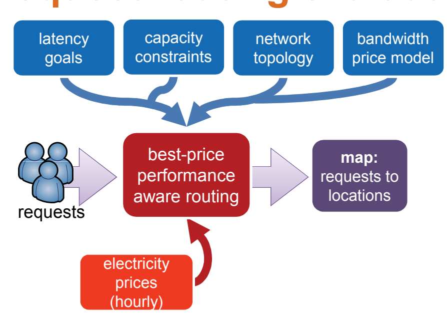
Akamai-specific inputs

- client workload
- geographic server distribution (25 cities / non-uniform)
- capacity & bandwidth constraints

results should apply to other systems

- realistic client workload
  - 2000 content providers
  - hundreds of billions of requests per day
- realistic server distribution
  - better than speculating…

request routing evaluation



request routing scheme

performance-aware price optimizer

- map client -> set of locations that meets latency goals
- rank locations based on electricity prices
- remove locations nearing capacity from set
- pick top-ranked location

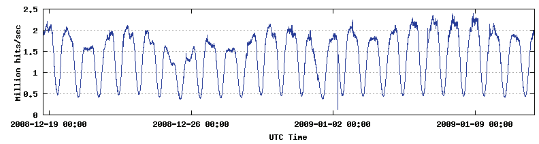
assumptions

- complete replication
- hourly route updates preserve stability
- uniform bandwidth prices (we will relax this later…)

Akamai workload

measured traffic on Akamai's CDN

- large subset of Akamai's servers (~20K) in 25 cities
- collected over 24 days (Dec 2008 – Jan 2009)
- 5-min samples
  - number of hits and bytes transferred
  - track how Akamai routed clients to clusters
  - group clients by origin state
- also derived a synthetic workload



electricity prices

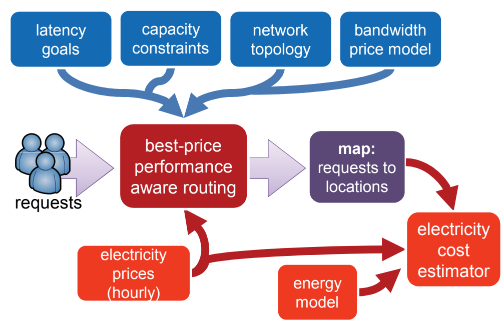
extensive survey of US electricity markets

- regional wholesale markets (both futures and spot)
- nature and causes of price volatility (see paper…)

data collection

- 39 months worth of historical hourly prices
  - January 2006 through March 2009
- 6 different regional wholesale markets
- 30 locations

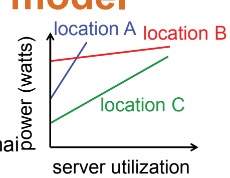
request routing evaluation



location energy model

linear model (roughly)

- server utilization -> watts
- scaling: number of servers
- based on a Google study
- power measurements at Akamai



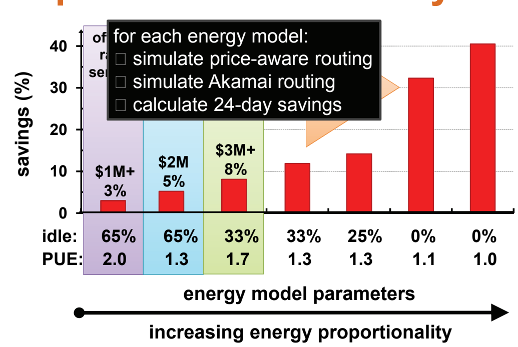
important parameters

(a) $\frac{\text{idle server power}}{\text{peak server power}}$ (b) PUE = $\frac{\text{power entire data center}}{\text{power used by IT equip.}}$

critical: how proportional is power to load?

- server power management? are idle servers turned off?
- the 'energy elasticity' of the system

importance of elasticity



bandwidth costs

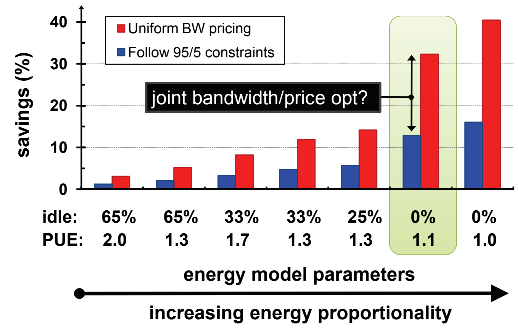
are we increasing bandwidth costs?
- problematic: bandwidth prices are proprietary

uniform bandwidth price model
- fixed cost per bit regardless of time and place

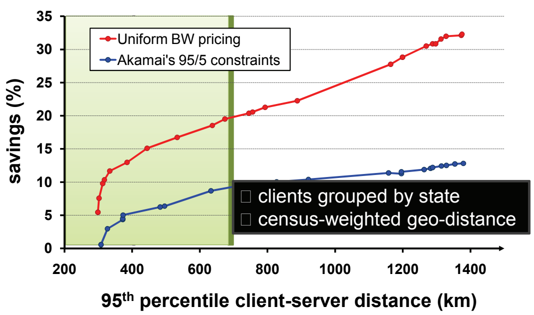
95/5 bandwidth pricing model
- prices set per network port
- network traffic is divided into 5-minute windows
- 95th percentile of traffic is used for billing

approach: 95th percentiles from Akamai data
- constrain routing so that 95th percentiles are unchanged
- Akamai's routing factors in bandwidth prices…

bandwidth constraints



latency constraints



limitations

Akamai doesn't use geographic distance as a primary metric in assigning clients to servers

Akamai's power consumption is typically not metered

practical implications

who can use this approach?
- servers in multiple locations
- some energy proportionality

complications
- electric billing based on peak power
- we need prices w/ time-varying uncorrelated volatility
  - e.g., wholesale market prices in the US

current energy sector trends are favorable

conclusion

significant value in price volatility
- large systems today: save more $1M/year
- increased energy elasticity: more than $10M/year

required mechanism already mostly in place
- minimal incremental changes required
- integrate real-time market information

extensions
- other cost functions (carbon, NO_x)
- other inputs (weather)
- active market participation (demand response, etc.)

market diversity