Cloud for the Enterprise 2010

James Hamilton, 2010/11/11
VP & Distinguished Engineer, Amazon Web Services
email: James@amazon.com
web: mvdirona.com/jrh/work
blog: perspectives.mvdirona.com
Agenda

• Quickening pace of DC infrastructure innovation
• Where does the money go?
• Power distribution infrastructure
• Mechanical systems
• Sea change in networking
• Server innovations
• Cloud Computing Economics

Talk does not necessarily represent positions of current or past employers

2010/11/11  http://perspectives.mvdirona.com
Pace of Innovation

• Datacenter pace of innovation increasing
  – Driven by cloud service providers and very high-scale internet applications like search
  – Cost of datacenter & H/W infrastructure dominates
  – Not just a cost center

• High focus on infrastructure innovation
  – Driving down cost
  – Increasing aggregate reliability
  – Reducing resource consumption footprint
Where Does the Money Go?

**Assumptions:**
- Facility: ~$88M for 8MW critical power
- Servers: 46,000 @ $1.45k each
- Commercial Power: ~$0.07/kWhr
- Power Usage Effectiveness: 1.45

**Observations:**
- 31% costs functionally related to power (trending up while server costs down)
- Networking high at 8% of costs & 19% of total server cost (many pay more)

Power Distribution

11% lost in distribution
\[0.997 \times 0.94 \times 0.98 \times 0.98 \times 0.99 = 89\%\]

High Voltage Utility Distribution

Generators

Sub-station

UPS: Rotary or Battery

13.2kv

13.2kv

13.2kv

13.2kv

2% loss
98% efficient

2% loss
98% efficient

2% loss
98% efficient

6% loss
94% efficient, ~97% available

~1% loss in switch gear & conductors

IT Load (servers, storage, Net, ...)

Note: Two more levels of power conversion in the server
Mechanical Systems

- Cooling Tower
- CWS Pump
- A/C Condenser
- Heat Exchanger (Water-Side Economizer)
- Primary Pump
- A/C Evaporator
- A/C Compressor
- Blow down & Evaporative Loss at 8MW facility: ~200,000 gal/day

Server fans 6 to 9W each

Diluted Hot/Cold Mix

Overall Mechanical Losses ~22%

Computer Room Air Handler

Air Impeller

2010/11/11
http://perspectives.mvdirona.com
Hot Aisle/Cold Aisle Containment
Most data center run in this range

ASHRAE 2008 Recommended Class 1

64-81°F
ASHRAE Allowable

Most data center run in this range

ASHRAE Allowable Class 1

ASHRAE 2008 Recommended Class 1

59F-90F
Air-Side Economization & Evaporative Cooling

• Limiting factors to high temp operation
  – Higher fan power trade-off
  – More semiconductor leakage current
  – Possible negative failure rate impact

• Avoid direct expansion cooling entirely
  – Air side economization
  – Higher data center temperatures
  – Evaporative cooling

• Requires Filtration
  • Particulate & chemical pollution
Sea Change in Networking

- Current networks over-subscribed
  - Forces workload placement restrictions
  - Goal: all points in datacenter equidistant

- Mainframe model goes commodity
  - Competition at each layer rather than vertical integration
Server Innovation

• Shared Infrastructure Racks
  – Shared redundant PSUs & fans
  – e.g. Dell Fortuna & Rackable CloudRack
• Next Level: Multi-server on board
  – Intel Atom: SeaMicro
  – ARM: SmoothStone
• Very Low-Cost, Low-Power Servers
  – ARM, Atom, client & embedded CPUs
  – Cold storage (reduce CPU $ to GB)
  – Highly partitionable workloads: Web services, memcached
• Low utilization is still the elephant in room

2010/11/11
http://perspectives.mvdirona.com
Infrastructure at Scale

• Datacenter design efficiency
  – Average datacenter efficiency low with PUE over 2.0 (Source: EPA)
    • Many with PUE well over 3.0
  – High scale cloud services in the 1.2 to 1.5 range
  – Lowers computing cost & better for environment

• Multiple datacenters
  – At scale multiple datacenters can be used
    • Close to customer
    • Cross datacenter data redundancy
    • Address international markets efficiently

• Avoid massive upfront data cost & years to fully utilize
  – Scale supports pervasive automation investment
Utilization & Economics

• Server utilization problem
  – 30% utilization VERY good & 10% to 20% common
    • Expensive & not good for environment
  – Solution: pool number of heterogeneous services
    • Single reserve capacity pool far more efficient
    • Non-correlated peaks & law of large numbers

• Pay as you go & pay as you grow model
  – Don’t block the business
  – Don’t over buy
  – Transfers capital expense to variable expense
  – Apply capital for business investments rather than infrastructure

• Charge back models drive good application owner behavior
  – Cost encourages prioritization of work by application developers
  – High scale needed to make a market for low priority work
Amazon Cycle of Innovation

• 15+ years of operational excellence
  – Managing secure, highly available, multi-datacenter infrastructure

• Experienced at low margin cycle of innovation:
  – Innovate
  – Listen to customers
  – Drive down costs & improve processes
  – Pass on value to customers

• AWS price reductions expected to continue
AWS Approach

• Broad set of services:
  • Infrastructure Services
    • SimpleDB
    • Simple Storage Service
    • CloudFront
    • Simple Queue Service
    • Elastic MapReduce
    • Relational Database Service
    • Elastic Block Store
    • Premium Support
    • Virtual Private Cloud
  • Payments & Billing
    • Flexible Payment Services
    • DevPay
  • On Demand Workforce
    • Mechanical Turk
  • Alexa Web Services
    • Web Information Service
    • Top Sites
  • Merchant Services
    • Fulfillment Web Service

• “Open the hood” approach
  • Simple, layerable building block services
  • Component services are substitutable
H/W Cost & Efficiency Optimization

- Service optimized hardware
  - Custom cloud-scale design teams:
    - Contract manufacturers, Dell DCS, SGI (Rackable), ZT Systems, HP, …

- Purchasing power at volume

- Supply chain optimization
  - Shorter chain drives much higher server utilization
    - Predicting next week easier than 4 to 6 months out
    - Less overbuy & less capacity risk

- Networking transit costs strongly rewards volume

- Cloud services unblocks new business & growth
  - Remove dependence on precise capacity plan
AWS Pace of Innovation

- Amazon EC2 with Windows Server 2008, Spot Instances, Boot from Amazon EBS
- Amazon CloudFront Streaming
- Amazon VPC enters Unlimited Beta
- AWS Region in Northern California
- International Support for AWS Import/Export

2009
- EBS Shared Snapshots
- SimpleDB in EU Region
- Monitoring, Auto Scaling & Elastic Load Balancing in EU

2010
- Amazon EC2 Reserved Instances with Windows, Extra Large High Memory Instances
- Amazon S3 Versioning Feature
- Consolidated Billing for AWS
- Lower pricing for Outbound Data Transfer

2010 Jan
- Amazon CloudFront Private Content
- SAS70 Type II Audit
- AWS SDK for .NET

2010 Feb
- Amazon VPC: Europe launch
- Amazon RDS: Northern California Region, Multi-AZ Deployments, AWS Management Console support
- Amazon CloudFront: Access logs for streaming
- Amazon S3 Reduced Redundancy Storage

2010 Mar
- AWS SDK for Java
- Windows server ‘Bring Your Own License’ pilot program
- Amazon CloudFront: Singapore edge location, private content for streaming

2010 Apr
- Amazon CloudFront: HTTPS support, lower request pricing, NYC edge location
- AWS Import/Export exits beta; web service support
- AWS Management Console for Amazon S3
- Amazon CloudWatch monitoring for Amazon EBS volumes

2010 May
- Amazon SQS Free Tier
- Amazon S3 Bucket Policies
- Amazon VPC IP Address Assignment
- Amazon EC2 Cluster Compute Instances
- Amazon S3 Enhanced Support for RRS

2010 Jun
- Combined AWS Data Transfer Pricing
- Amazon SNS
- Amazon Elastic MapReduce: custom cluster configuration option
- Amazon RDS: EU Region launch
- AWS Asia Pacific (Singapore) Region

2010 Jul
- Lower Amazon EC2 Pricing
- AWS IAM Preview Beta
- AWS Console Support for Amazon VPC
- Amazon EC2 Micro Instances
- Amazon Linux AMI
- Amazon EC2 Tagging, Filtering, Import Key Pair, Idempotency
- Oracle certifies enterprise software on Amazon EC2
- AWS SDK on PHP

2010 Aug
- Amazon CloudFront Default Root Objects
- Amazon RDS Reserved Instances
- Amazon CloudFront Invalidation

2010 Sep
- Amazon RDS Read Replicas; lowers prices
- Amazon EC2 running SUSE Linux
- AWS Console supports Amazon SNS
- Amazon ELB support for HTTPS

2010 Oct
- Amazon VPC: Europe launch
- Amazon RDS: Northern California Region, Multi-AZ Deployments, AWS Management Console support
- Amazon CloudFront: Access logs for streaming
- Amazon S3 Reduced Redundancy Storage

2010 Nov
- Amazon CloudFront Streaming
- Amazon VPC enters Unlimited Beta
- AWS Region in Northern California
- International Support for AWS Import/Export

2010 Dec
- Amazon EC2 with Windows Server 2008, Spot Instances, Boot from Amazon EBS
- Amazon CloudFront Streaming
- Amazon VPC enters Unlimited Beta
- AWS Region in Northern California
- International Support for AWS Import/Export

» Amazon CloudFront Default Root Objects
» Amazon RDS Reserved Instances
» Amazon CloudFront Invalidation

» Amazon CloudFront: HTTPS support, lower request pricing, NYC edge location
» AWS Import/Export exits beta; web service support
» AWS Management Console for Amazon S3
» Amazon CloudWatch monitoring for Amazon EBS volumes

» Amazon SQS Free Tier
» Amazon S3 Bucket Policies
» Amazon VPC IP Address Assignment
» Amazon EC2 Cluster Compute Instances
» Amazon S3 Enhanced Support for RRS

» Combined AWS Data Transfer Pricing
» Amazon SNS
» Amazon Elastic MapReduce: custom cluster configuration option
» Amazon RDS: EU Region launch
» AWS Asia Pacific (Singapore) Region

» Lower Amazon EC2 Pricing
» AWS IAM Preview Beta
» AWS Console Support for Amazon VPC
» Amazon EC2 Micro Instances
» Amazon Linux AMI
» Amazon EC2 Tagging, Filtering, Import Key Pair, Idempotency
» Oracle certifies enterprise software on Amazon EC2
» AWS SDK on PHP

» Amazon CloudFront Streaming
» Amazon VPC enters Unlimited Beta
» AWS Region in Northern California
» International Support for AWS Import/Export
More Information

• These Slides:
  • I’ll post the slides to http://mvdirona.com/jrh/work later this week
• Power and Total Power Usage Effectiveness
  • http://perspectives.mvdirona.com/2009/06/15/PUEAndTotalPowerUsageEfficiencyTPUE.aspx
• Berkeley Above the Clouds Paper
• Degraded Operations Mode
• Cost of Power
  • http://perspectives.mvdirona.com/2008/12/06/AnnualFullyBurdenedCostOfPower.aspx
• Power Optimization
  • http://labs.google.com/papers/power_provisioning.pdf
• Cooperative, Expendable, Microslice Servers
• Power Proportionality
  • http://www.barroso.org/publications/ieee_computer07.pdf
• Resource Consumption Shaping:
• Email & Blog
  • James@amazon.com & http://perspectives.mvdirona.com