# Cloud Computing Economies of Scale

## Mix 2010


**James Hamilton, 2010/3/15**

**VP & Distinguished Engineer, Amazon Web Services**

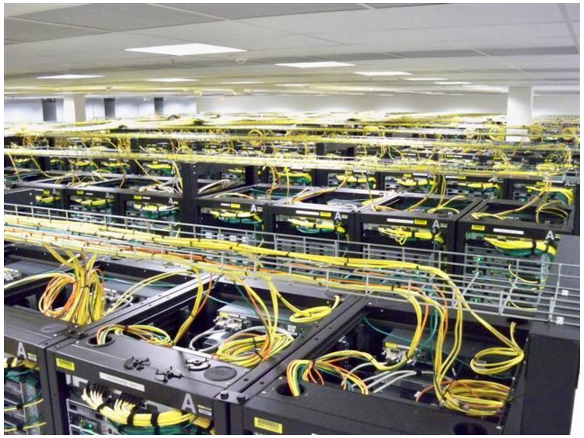
**email: James@amazon.com**

**web: mvdirona.com/jrh/work**

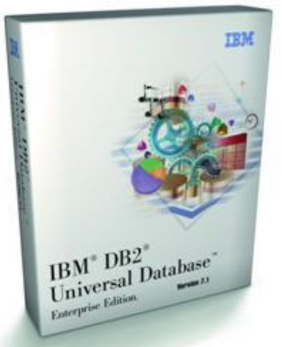
**blog: perspectives.mvdirona.com**

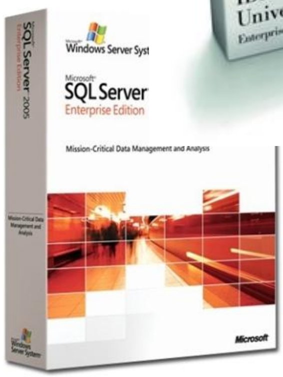
# Agenda

- Follow the money in infrastructure
  - Infrastructure cost breakdown
  - Where does the power go?
- Power Distribution Efficiency
- Mechanical System Efficiency
- Server Design & Utilization
- Cloud Computing Economics
  - Why utility computing makes sense economically
- Summary



# Background & Biases

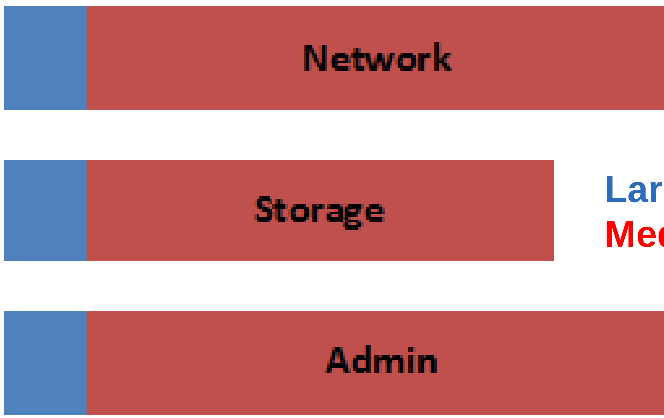



- 15 years database core engine dev.
  - Lead architect on IBM DB2
  - Architect on SQL Server
- Past 6 years in services
  - Led Exchange Hosted Services Team
  - Architect on the Windows Live Platform
  - Architect on Amazon Web Services
- Talk does not necessarily represent positions of current or past employers

# Economies of Scale

- 2006 comparison of very large service with mid-size: (~1000 servers):




Large Service [$13/Mb/s/mth]: $0.04/GB
Medium [$95/Mb/s/mth]: $0.30/GB (7.1x)

Large Service: $4.6/GB/year (2x in 2 Datacenters)
Medium: $26.00/GB/year* (5.7x)

Large Service: Over 1.000 servers/admin
Enterprise: ~140 servers/admin (7.1x)

- Large block h/w purchases significantly more economic
  - Large weekly purchases offer significant savings
  - H/W Manufacturers willing & able to do custom designs at scale
- Automation & custom s/w investments amortize well at scale
- Summary: scale economics strongly in play

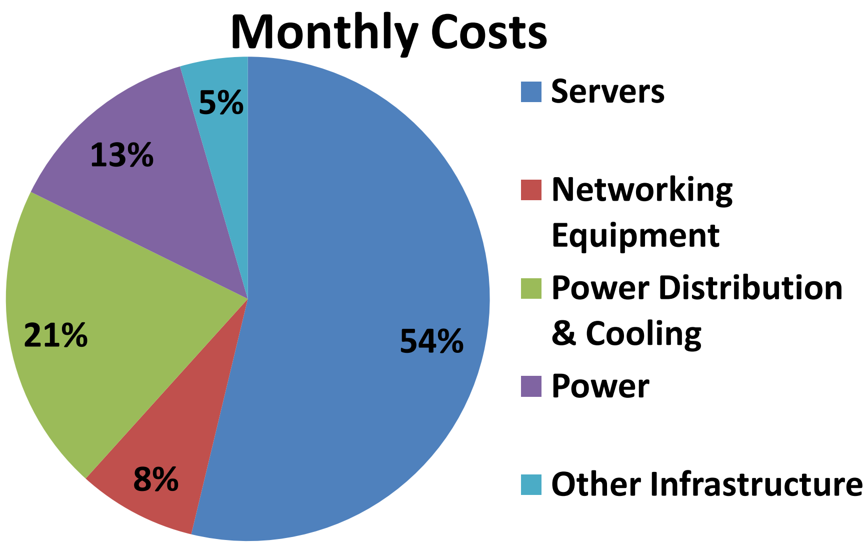
# PUE & DCiE

- Measure of data center infrastructure efficiency
- Power Usage Effectiveness
  - PUE = (Total Facility Power)/(IT Equipment Power)
- Data Center Infrastructure Efficiency
  - DCiE = (IT Equipment Power)/(Total Facility Power) * 100%


http://www.thegreengrid.org/en/Global/Content/white-papers/The-Green-Grid-Data-Center-Power-Efficiency-Metrics-PUE-and-DCiE

# Power & Related Costs [Will] Dominate

- **Assumptions:**
  - Facility: ~$88M for 8MW facility
  - Servers: Roughly 46k @ $1.45k each
  - Server power draw at 30% load: 80%
  - Commercial Power: ~$0.07/kWhr
  - PUE: 1.5




3yr server, 4yr net gear, & 10 yr infrastructure amortization

- **Observations:**
  - 34% costs functionally related to power (trending up while server costs down)
  - Networking high at 8% of costs & 19% of total server cost

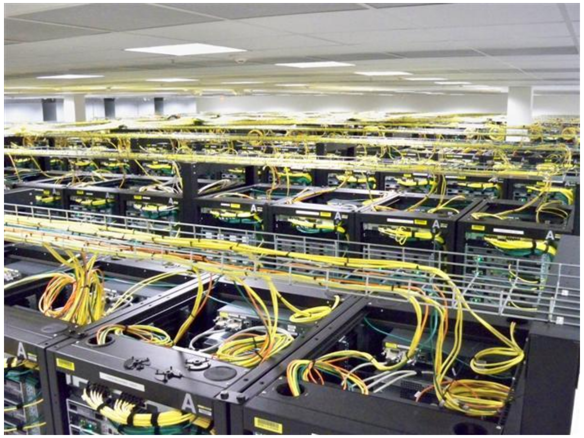
**Updated from**: http://perspectives.mvdirona.com/2008/11/28/CostOfPowerInLargeScaleDataCenters.aspx

# Where Does the Power Go?

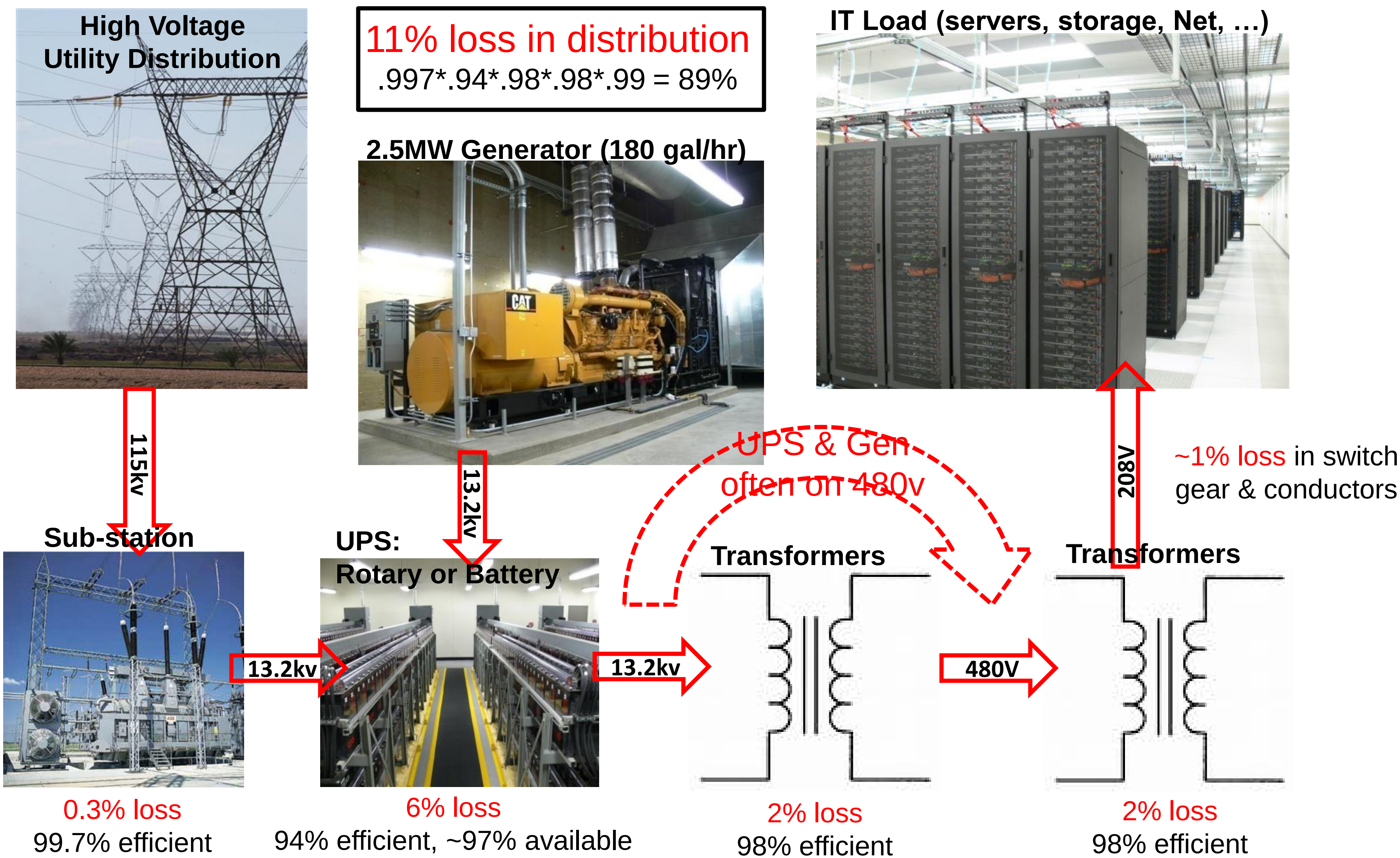
- **Assuming a good data center with PUE ~1.5**
  - Each watt to server loses ~0.5W to power distribution losses & cooling
  - IT load (servers & storage): 1/1.5 => 67%
  - Network gear <4% total power (5.8% of IT load)
- **Power losses are easier to track than cooling:**
  - Power transmission, conversion, & switching losses: 11%
    - Detailed power distribution losses on next slide
  - Cooling losses the remainder:100-(67+11) => 22%
- **Observations:**
  - **Utilization & server efficiency improvements very highly leveraged**
  - **Networking gear very power inefficient individually but not big problem in aggregate**
  - **Cooling costs unreasonably high**
  - **PUE improving rapidly**

# Agenda

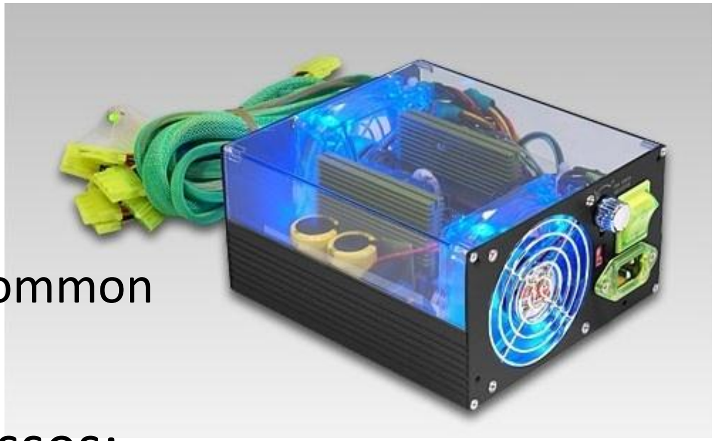
- Follow the money in infrastructure
  - Infrastructure cost breakdown
  - Where does the power go?
- Power Distribution Efficiency
- Mechanical System Efficiency
- Server Design & Utilization
- Cloud Computing Economics
  - Why utility computing makes sense economically
- Summary



# Power Distribution



# Power Distribution Efficiency Summary

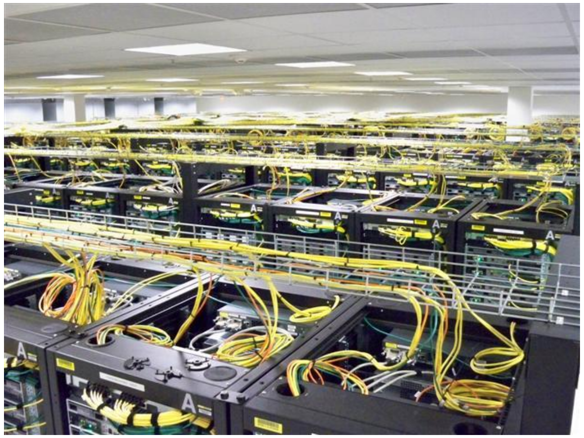


- Two additional conversions in server:
  1. Power Supply: often <80% at typical load
  2. On board step-down (VRM/VRD): <80% common
     - ~95% efficient both available & affordable
- Rules to minimize power distribution losses:
  1. Oversell power (more theoretic load than provisioned power)
  2. Avoid conversions (fewer transformer steps & efficient UPS)
  3. Increase efficiency of conversions
  4. High voltage as close to load as possible
  5. Size VRMs & VRDs to load & use efficient parts
  6. DC distribution a fairly small potential gain

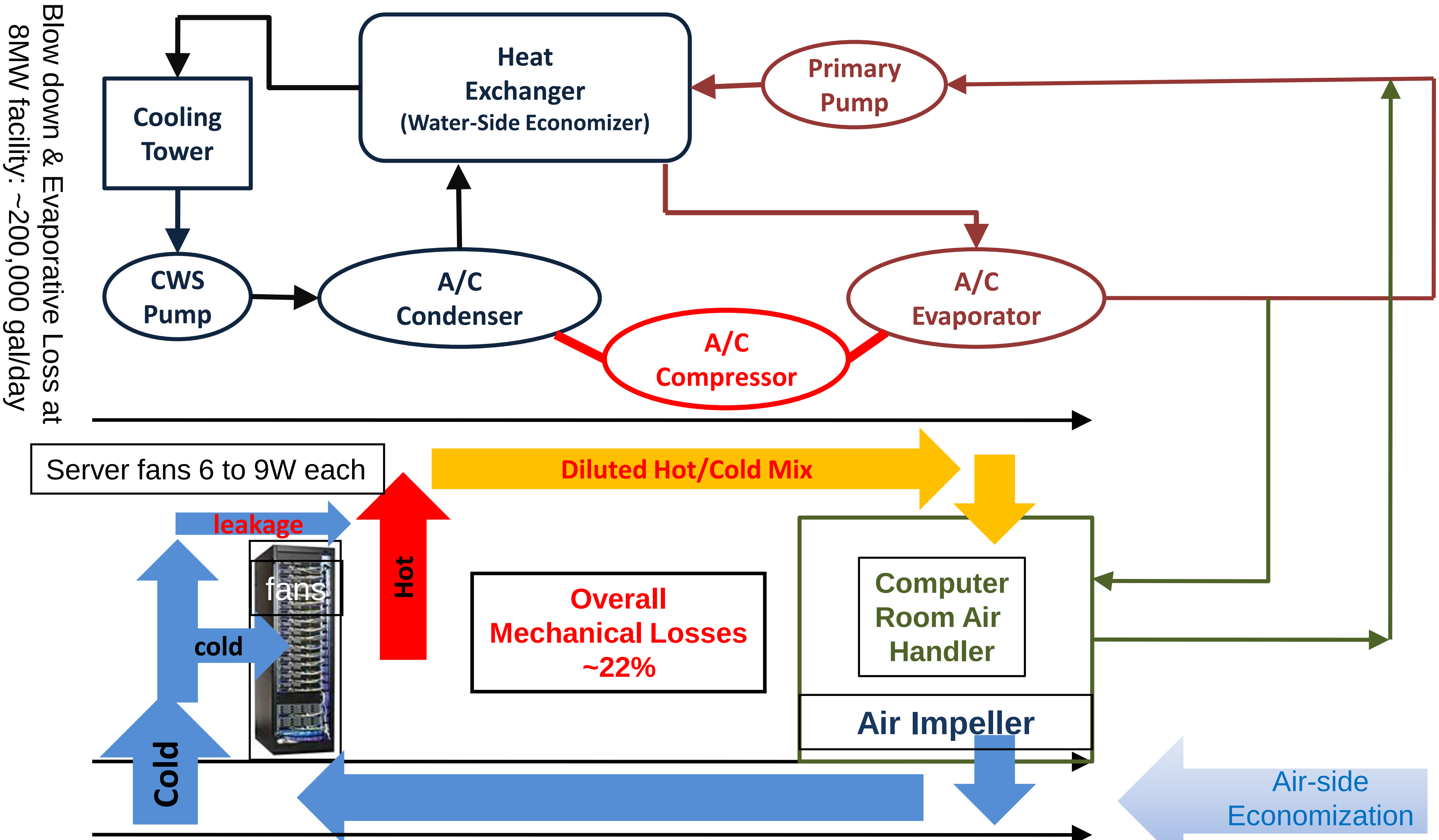
But power distribution improvements bounded to 11%

# Agenda

- Follow the money in infrastructure
  - Infrastructure cost breakdown
  - Where does the power go?
- Power Distribution Efficiency
- Mechanical System Efficiency
- Server Design & Utilization
- Cloud Computing Economics
  - Why utility computing makes sense economically
- Summary

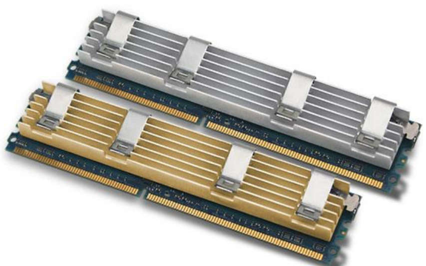


# Conventional Mechanical Design



# Air Cooling

- Allowable component temps higher than historical hottest place on earth
  - Al Aziziyah, Libya: 136F/58C (1922)
- So, it's just a mechanical engineering problem
  - More air & better mechanical designs
  - Tradeoff: power to move air vs cooling savings & semi-conductor leakage current
  - Partial recirculation when external air too cold
- Currently available equipment temp limits:
  - 40C: CloudRack C2 & most net gear
  - 35C: Most of the server industry



**Memory: 3W - 20W**
**Temp Spec: 85C-105C**

**Hard Drives: 7W- 25W**
**Temp Spec: 50C-60C**

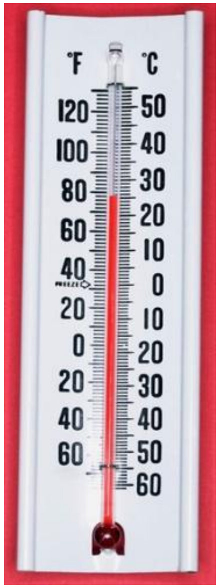
**I/O: 5W - 25W**
**Temp Spec: 50C-60C**


**Processors/Chipset: 40W - 200W**
**Temp Spec: 60C-70C**

**Thanks to Ty Schmitt, Dell Principle Thermal/Mechanical Arch. & Giovanni Coglitore, Rackable Systems CTO**

# Mechanical Efficiency Summary

- Prioritized mechanical System optimizations:
  1. Raise data center temperatures
  2. Tight airflow control, short paths & large impellers
  3. Cooling towers rather than A/C
  4. Air-side economization & evap cooling
     - outside air rather than A/C & towers

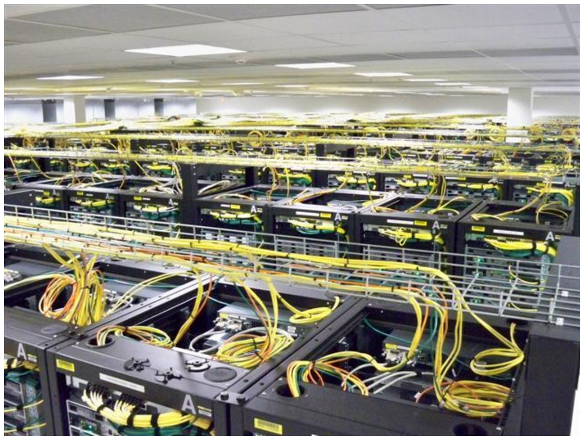


# Server Design & Utilization

- 75% of total power is delivered to the IT equipment
  - All but 4% delivered to servers & storage
- Clearly server & storage efficiency important
- But, server utilization is the elephant in the room
  - 10% to 20% common
  - 30% unusually good
- Conclusion:
  - most of the resources in the datacenter are **unused** more than they are doing productive work

# Agenda

- Follow the money in infrastructure
  - Infrastructure cost breakdown
  - Where does the power go?
- Power Distribution Efficiency
- Mechanical System Efficiency
- Server Design & Utilization
- Cloud Computing Economics
  - Why utility computing makes sense economically
- Summary



# 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

# H/W Cost & Efficiency Optimization

- Service optimized hardware
  - Custom cloud-scale design teams:
    - Dell DCS, SGI (aka Rackable), ZT Systems, Verari, 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 rewards volume
- Cloud services unblocks new business & growth
  - Remove dependence on precise capacity plan

# Investments at Scale

- Deep automation only affordable when amortized over large user base
  - Lack of automation drives both cost & human error fragility

- S/W investments at scale
  - Massive distributed systems investments such as Amazon Simple Storage Service & Elastic Block Store hard to justify without scale


- Special Skills with deep focus
  - Distributed systems engineers, power engineering, mechanical engineering, server h/w design, networking, supply chain, 24x7 operations staff, premium support,…

# 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 Web Services
# Pace of Innovation

**2009**

**Jan**
- » Lower CloudFront pricing tiers
- » AWS Management Console

**Feb**
- » New SimpleDB Features
- » FPS General Availability

**Mar**
- » Amazon EC2 with Windows
- » Amazon EC2 in EU Region
- » AWS Toolkit for Eclipse
- » Amazon EC2 Reserved Instances

**Apr**
- » Reserved Instances in EU Region
- » Elastic MapReduce
- » SQS in EU Region

**May**
- » AWS Import/Export
- » New CloudFront Feature
- » Monitoring, Auto Scaling & Elastic Load Balancing

**Jun**
- » AWS Security Center

**Jul**
- » Elastic MapReduce in EU

**Aug**
- » AWS Multi-Factor Authentication
- » Virtual Private Cloud
- » Lower Reserved Instance Pricing

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

**Oct**
- » Relational Database Service
- » High-Memory Instances
- » Lower EC2 Pricing

**Nov**
- » CloudFront private content
- » SAS70 Type II Audit
- » AWS SDK for .NET

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

**2010**

**Jan**

**Feb**
- » EC2 Reserved Instances with Windows, Extra Large High Memory Instances
- » S3 Versioning
- » AWS Consolidated Billing
- » Lower pricing for Outbound Data

# Summary

- Measure efficiency using work done/dollar & work done/joule
  - Server costs dominate all other DC infrastructure & admin at scale
  - 2/3 of total data center power is delivered to servers
  - Utilization poor: Servers are idle more than not
  - Conclusion: nearly ½ the provisioned power not doing useful work
- Considerable room for DC cooling improvements
- Cloud services drive:
  - Higher resource utilization
  - Innovation in power distribution & mechanical systems
  - Lower cost, higher reliability, & lower environmental impact

# More Information

- **This Slide Deck:**
  - I will post all but last slide to http://mvdirona.com/jrh/work this week
- **Power and Total Power Usage Effectiveness (tPUE)**
  - http://perspectives.mvdirona.com/2009/06/15/PUEAndTotalPowerUsageEfficiencyTPUE.aspx
- **Berkeley Above the Clouds**
  - http://perspectives.mvdirona.com/2009/02/13/BerkeleyAboveTheClouds.aspx
- **Degraded Operations Mode**
  - http://perspectives.mvdirona.com/2008/08/31/DegradedOperationsMode.aspx
- **Cost of Power**
  - http://perspectives.mvdirona.com/2008/11/28/CostOfPowerInLargeScaleDataCenters.aspx
  - http://perspectives.mvdirona.com/2008/12/06/AnnualFullyBurdenedCostOfPower.aspx
- **Power Optimization:**
  - http://labs.google.com/papers/power_provisioning.pdf
- **Cooperative, Expendable, Microslice Servers**
  - http://perspectives.mvdirona.com/2009/01/15/TheCaseForLowCostLowPowerServers.aspx
- **Power Proportionality**
  - http://www.barroso.org/publications/ieee_computer07.pdf
- **Resource Consumption Shaping:**
  - http://perspectives.mvdirona.com/2008/12/17/ResourceConsumptionShaping.aspx
- **Email**
  - James@amazon.com