Data Center Power Management: Thoughts on the Future

Dan Reed
reed@microsoft.com
Multicore and Scalable Computing Strategist
Managing Director, Cloud Computing Futures

The Real Science of Power

- For background, see
  - *Science*, Volume 306, November 14, 2004
  - Margolus-Levitin theorem (*Physica* D 120 188, 1998)
- Suppose we fill space time with clocks
  - All exchange signals and measure arrival times
- How many events N are in a volume of radius R and time T?
  - i.e., How much computation is possible in that volume?
- We're constrained to avoid gravitational collapse

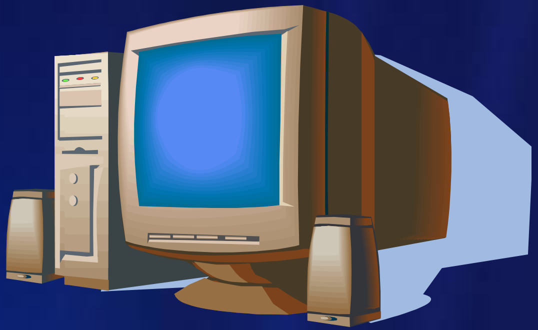
$$N \equiv \frac{1}{\pi}\frac{T}{t_p}\frac{R}{l_p} = \frac{TR}{\pi(t_p l_p)} = \frac{TR}{\pi c t_p^2} = \frac{TRc^5}{\pi \hbar G}$$

- This is a BIG number!

Microsoft

Sapir–Whorf: Context and Research

- Sapir–Whorf Hypothesis (SWH)
  - language influences the habitual thought of its speakers
- Computing analog
  - available systems shape research agendas
- Consider some past examples
  - VAX 11/780 and UNIX
  - workstations and Ethernet
  - PCs and web
  - inexpensive clusters



- Today's examples
  - multicore, clouds and services …

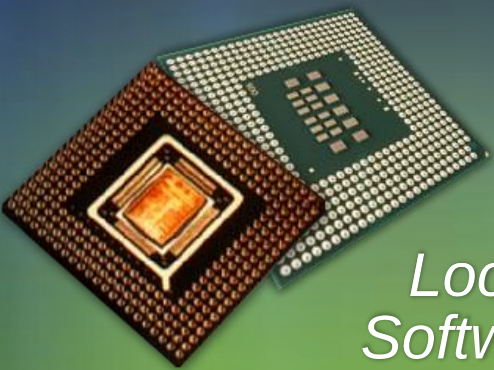
Microsoft

Today's Truisms (2009)

- Bulk computing is almost free
  - … but software and power are not
- Inexpensive sensors are ubiquitous
  - … but scientific data fusion remains difficult
- Moving lots of data is {still} hard
  - … because we're missing trans-terabit/second networks
- People are really expensive!
  - … and robust software remains extremely labor intensive
- Scientific challenges are complex
  - … and social engineering is not our forte
- Our political/technical approaches must change
  - … or we risk solving irrelevant problems

Microsoft

Next-Generation Applications



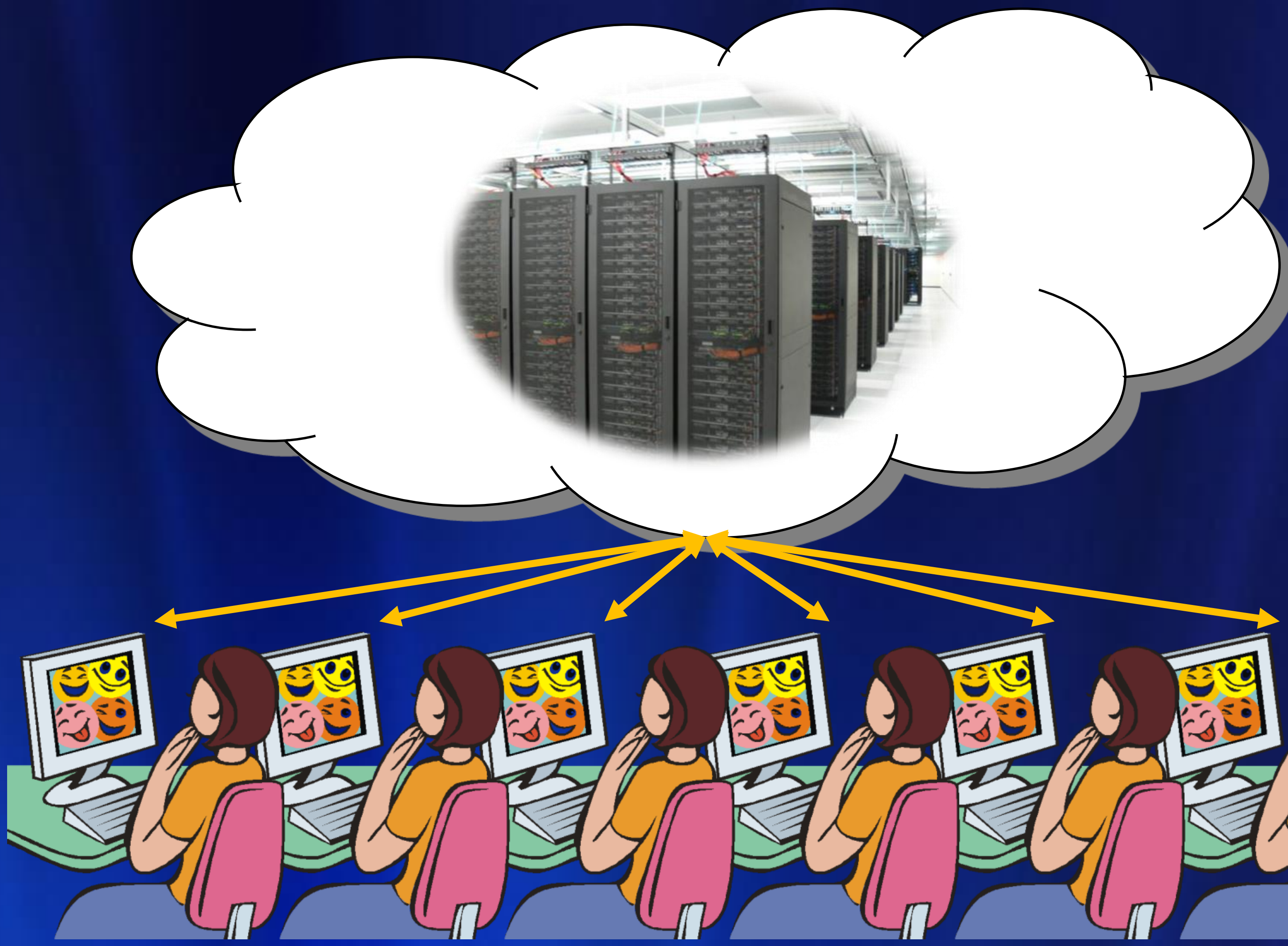
Concurrency Spectrum

Local Software

Software
+
Services

Global Services

New Software Architecture



Microsoft

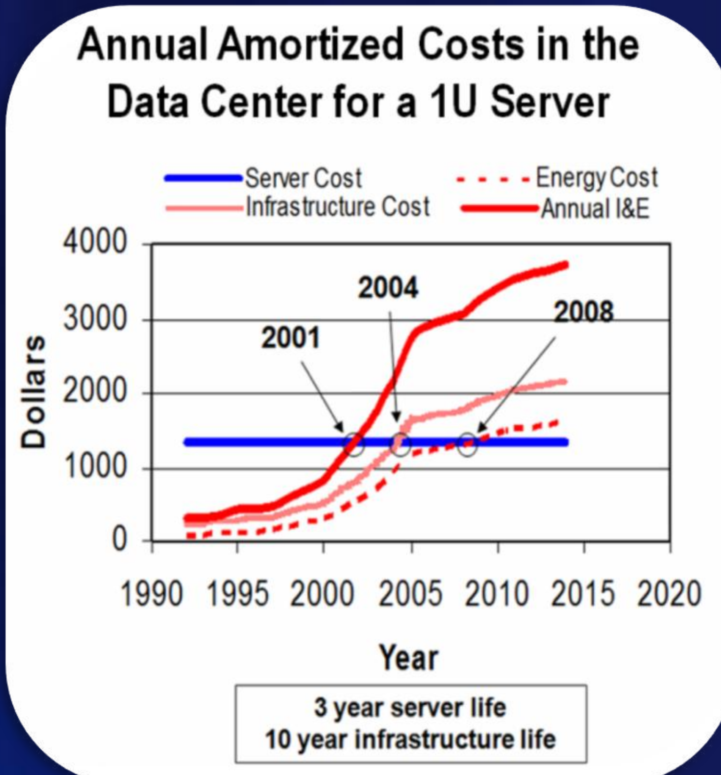
Generic Cloud Data Centers

- Massive commodity servers
- Energy intensive infrastructure
- Cooling inefficiencies
- Environmental issues
- Expensive UPS support
- Enterprise TCP/IP networks
- Long deployment times
    - Construction and integration
- Diverse services and SLAs

- *Explosive growth*
    - *demand and expectations*

Data Center "PacMan"

- Land - 2%
- Core and shell costs – 9%
- Architectural – 7%
- Mechanical/Electrical – 82%
  - 16% increase/year since 2004



Belady, C., "In the Data Center, Power and Cooling Costs More than IT Equipment it Supports", *Electronics Cooling Magazine* (February 2007)

Source: Christian Belady

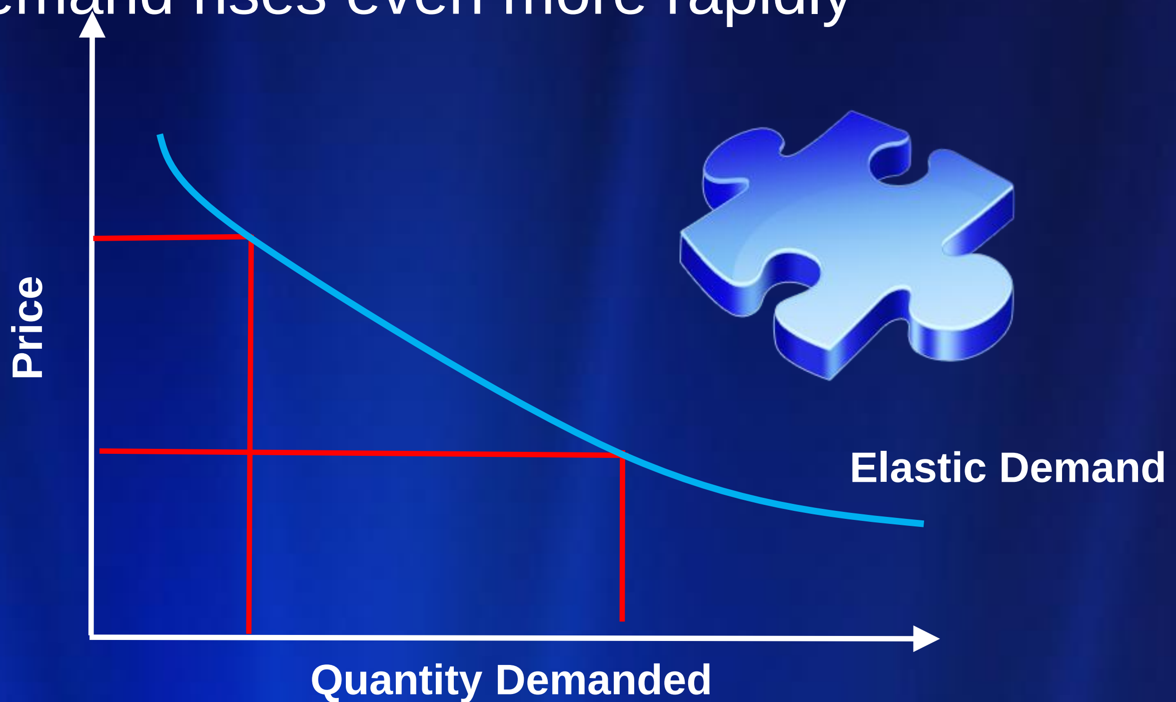
Consider These Services Challenges

- Environmental responsibility
  - Managing under a large power envelope
  - Adaptive systems management
- Provisioning 100,000 servers
  - Hardware: at most one week after delivery
  - Software: at most a few hours
- Resilience during a blackout/disaster
  - Data center failure
  - Service rollover for 20M customers
- Programming the entire data center
  - Power, environmentals, provisioning
  - Component tracking, resilience, …

Microsoft

Remember Jevon's Paradox

- Improved technology
  - Doubles work produced for given cost
  - Sounds a bit like Moore's Law doesn't it?
- Jevon's paradox
  - Demand rises even more rapidly



Microsoft

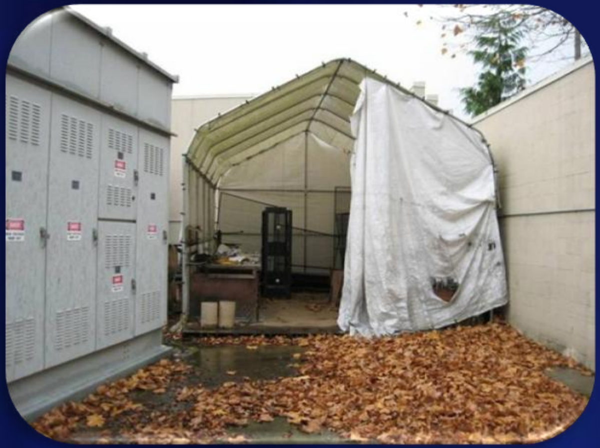
It's Not Just Watts …

- Watts alone are irrelevant
  - Turn off the equipment and declare victory
- The real metric is the following …

$$\frac{\mathit{Operations}}{\mathit{Watt} \times \mathit{Dollars}}$$

- Many convolved ideas
  - Application execution efficiency
  - Microarchitecture and system design
  - Power supply efficiency
  - Packaging and cooling overhead
  - Market costs for power and hardware
  - Cost of people and money

The Purity of Minimalism



- Configuration
  - 5 HP DL585 systems
  - 11/2007-6/2008
  - Zero failures
- Incidents
  - Water dripped from the tent onto the rack
  - A windstorm blew a fence section onto the rack
  - A leaf was sucked onto the server fascia

Microsoft

www.thegreengrid.org

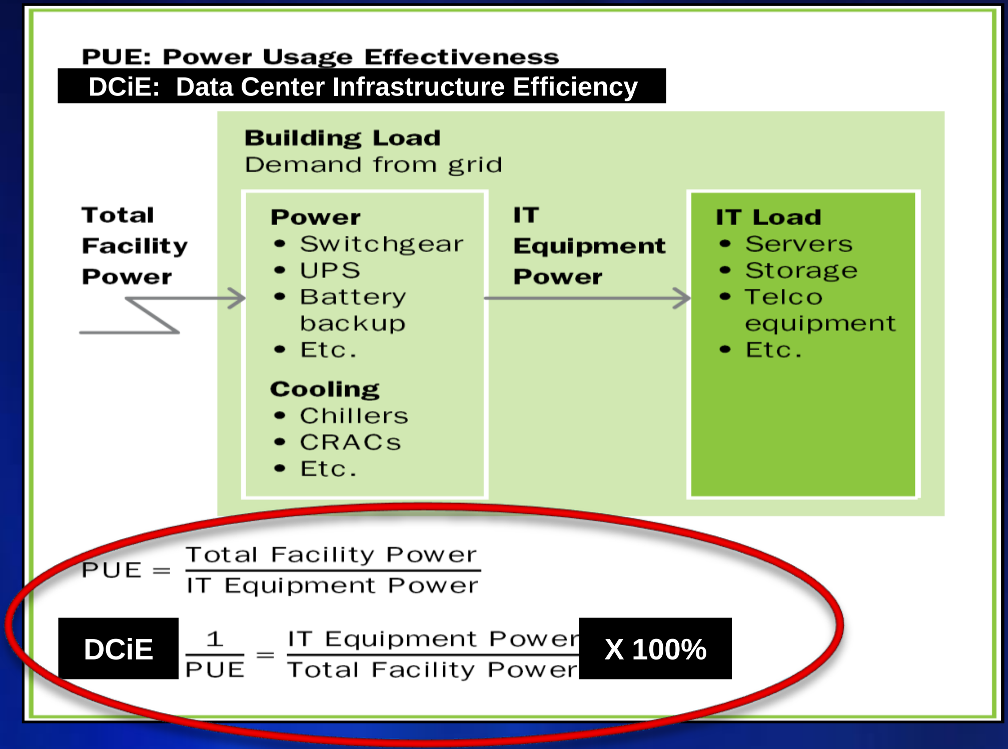
www.energystar.gov

www.microsoft.com/environment/

Microsoft

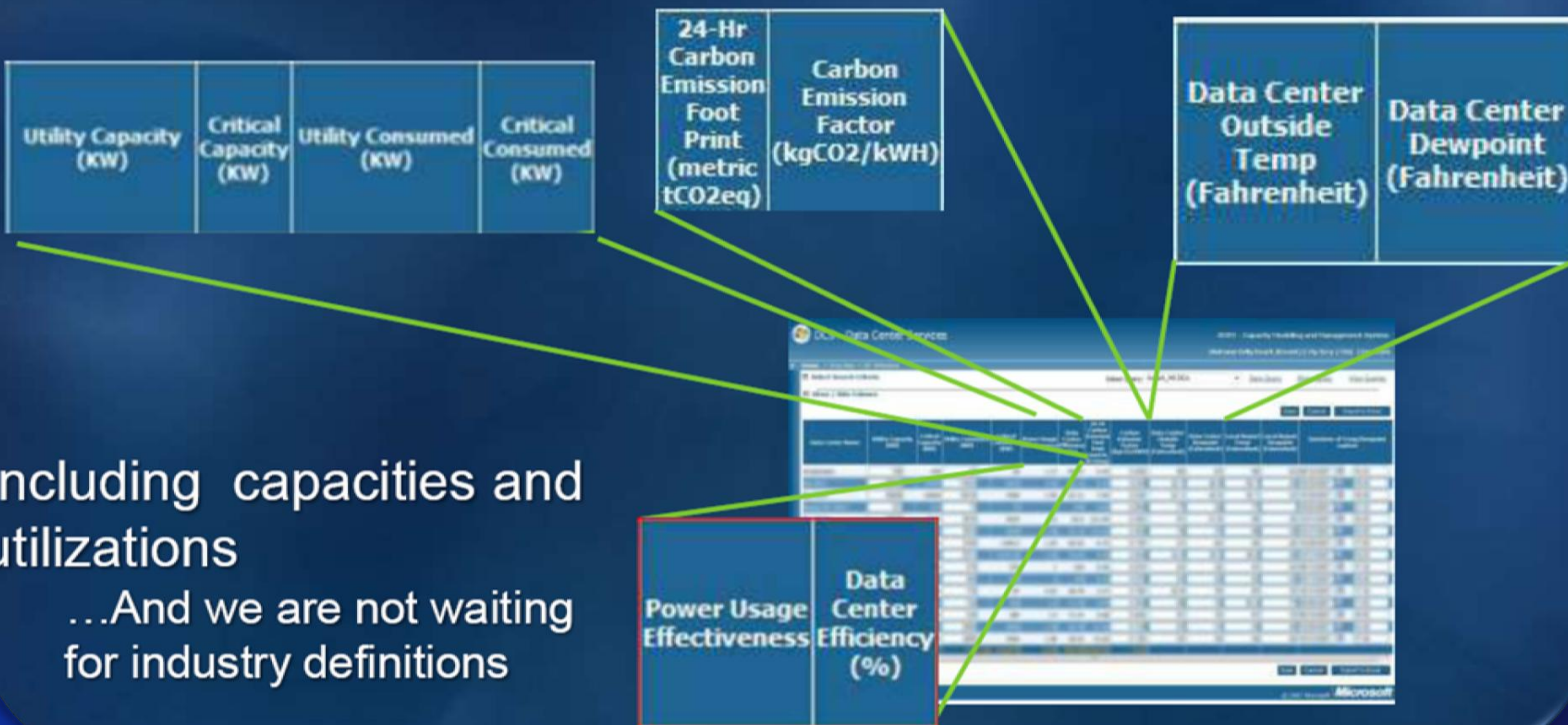
Industry Initiatives – The Green Grid

- Microsoft is founder/board member and leads two work groups
  - Technology & Strategy
  - Metrics Measurement



Microsoft

Microsoft Tracks Many Things



Including capacities and utilizations

…And we are not waiting for industry definitions

Microsoft

Microsoft Best Practices

- Engineer the data center for cost and energy efficiency
- Optimize the design to assess multiple factors
- Optimize provisioning for maximum efficiency and productivity
- Monitor and control data center performance in real time
- Make data center operational excellence part of organizational culture
- Measure power usage effectiveness (PUE)
- Use temperature control and airflow distribution
- Eliminate the mixing of hot and cold air
- Use effective air-side or water-side economizers
- Share and learn from industry partners

The Container Rationale

Legacy Data Center

Customer

Chilled Water → Conditioned Air

ASHRAE Guidelines

Manufacturer

→ Server or Rack

Commoditized Data Center

Customer

Chilled Water

Interface Guidelines

Chilled Water

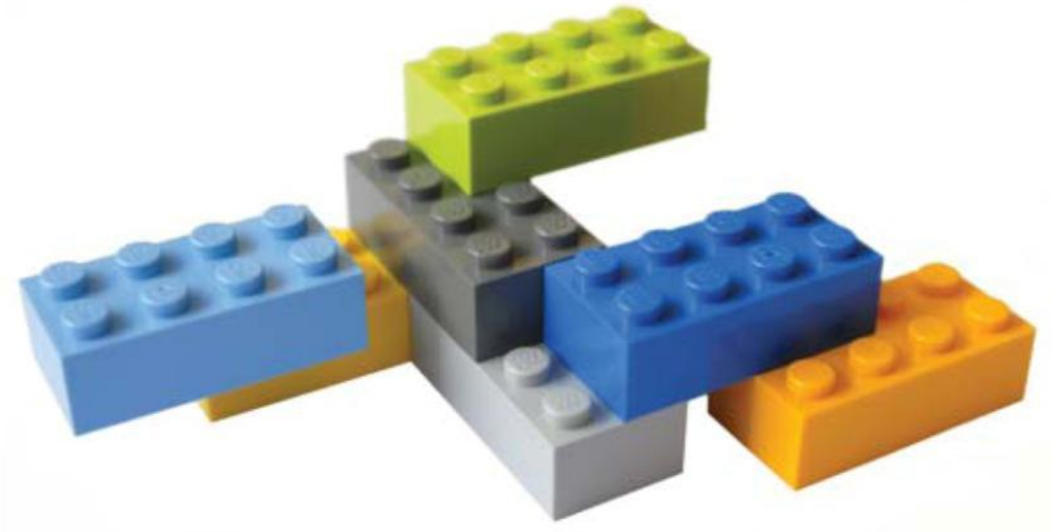
Manufacturer

40' Dry Van Container
40' x 8' x 8'6"

Raise Operating Temperatures

- Data center temperature zones
  - As we've seen, they're not chiseled in stone
- Operate at >50C inlet temperature instead
  - Perhaps accept more component failures
  - But, this is not necessarily true
- Aggressively embrace airside economization
  - Ambient air, with minimal conditioning
- Recognize the power of alternative cooling
  - Liquid, phase change, …

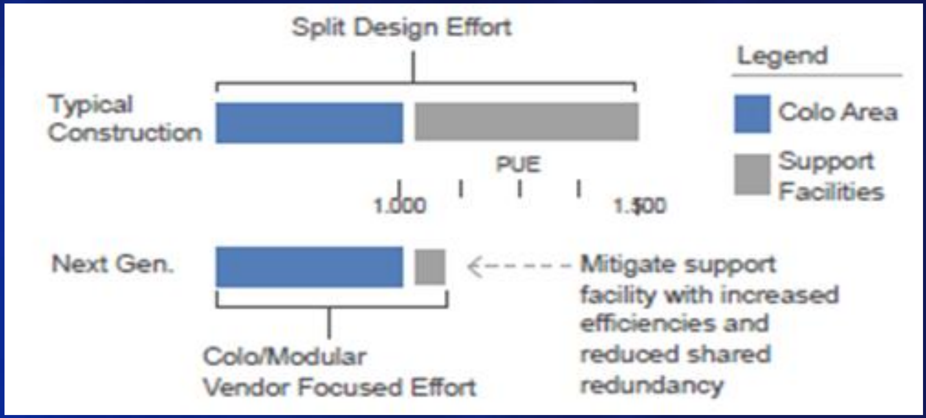
Microsoft Gen4 Data Center Vision



The design aspiration of the Generation 4.0 Data Center is to provide a flexible, modular, and scalable solution consisting of inter-changeable components utilized in a plug-and-play, just-in-time configuration.

Microsoft Gen4 Data Centers

- Scalable
- Plug-and-play spine infrastructure
- Factory pre-assembled
  - Pre-assembled containers (PACs)
  - Pre-manufactured buildings (PMBs)
- Rapid deployment
- De-mountable
- Reduced construction
- Sustainable measures
- Map applications to class



Microsoft

Let's Look Longer Term …

Microsoft

Cloud Computing Futures Vision

Philosophy: *The data center is a computer that must be designed and programmed as an integrated system*

Processors

Low power
Virtualization
Optical Interconnect
Multicore
Heterogeneity

Storage

On-chip
FLASH
PCM

Networks

Optics
Distributed routing
Non-TCP/IP

Packaging

Chip stacking
Modularity
Liquid cooling
Over-provisioned

Software

Introspection
Tier-splitting
Adaptation
Resilience

Move Redundancy to Geo-Level

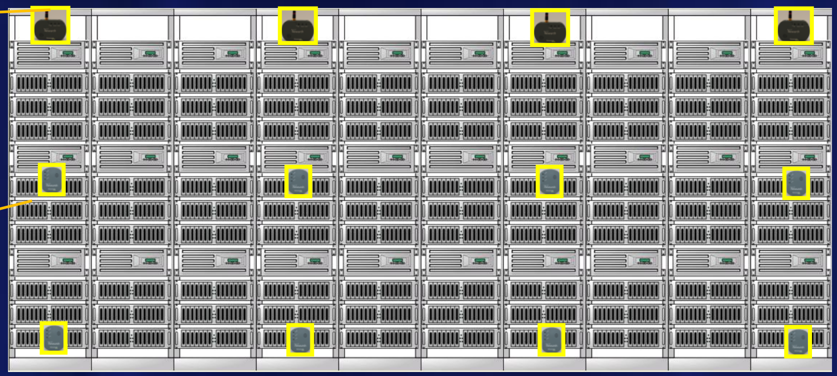
- Power redundancy is a major cost
  - Batteries to supply up to 15 minutes at some facilities
- Instead use more, smaller, cheaper data centers
  - This is the Gen4 concept
- Focus on multi-terabit WAN interconnect
  - Weak consistency and workload rollover

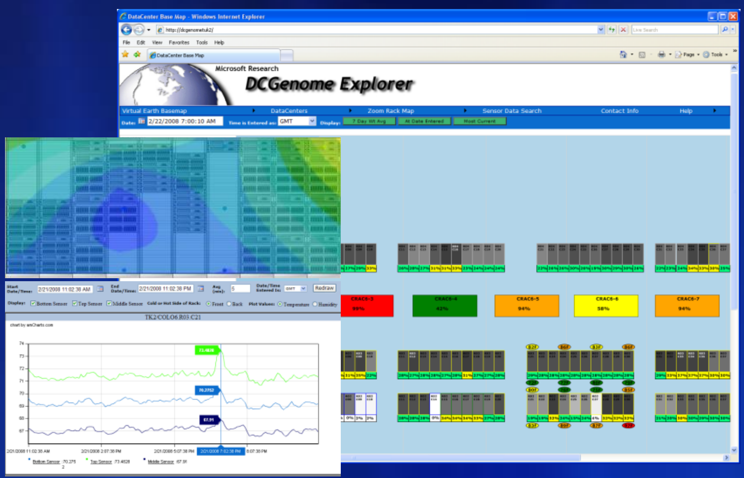
Microsoft

Data Center Genome

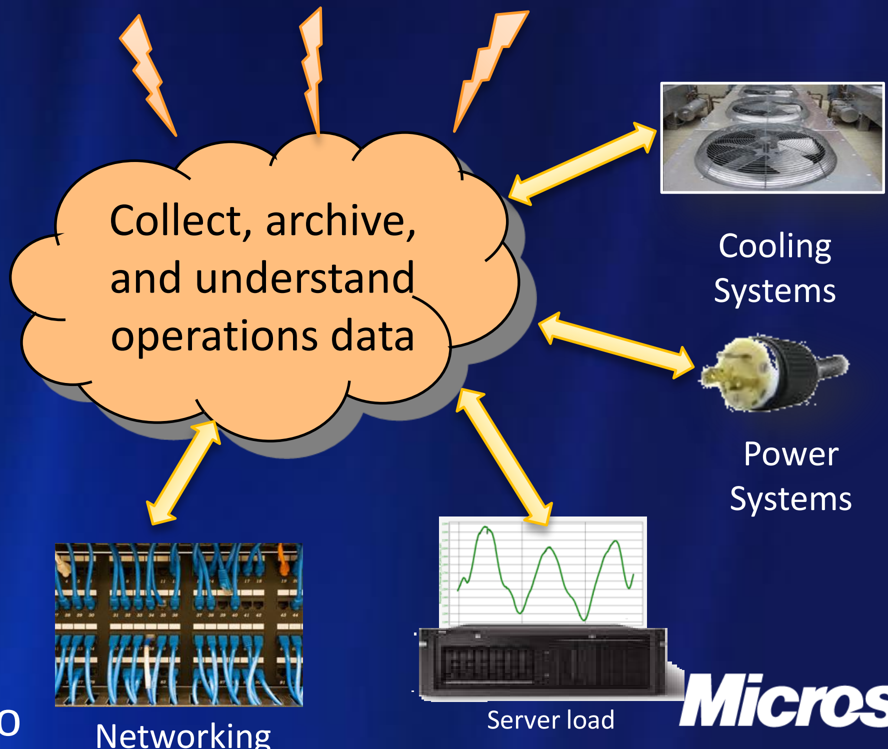
Saving energy & improving operation efficiency by networked sensing, data mining, and control.

MSR *Genomotes*





Operation monitoring,
Capacity planning,
Device provisioning,
Resource control



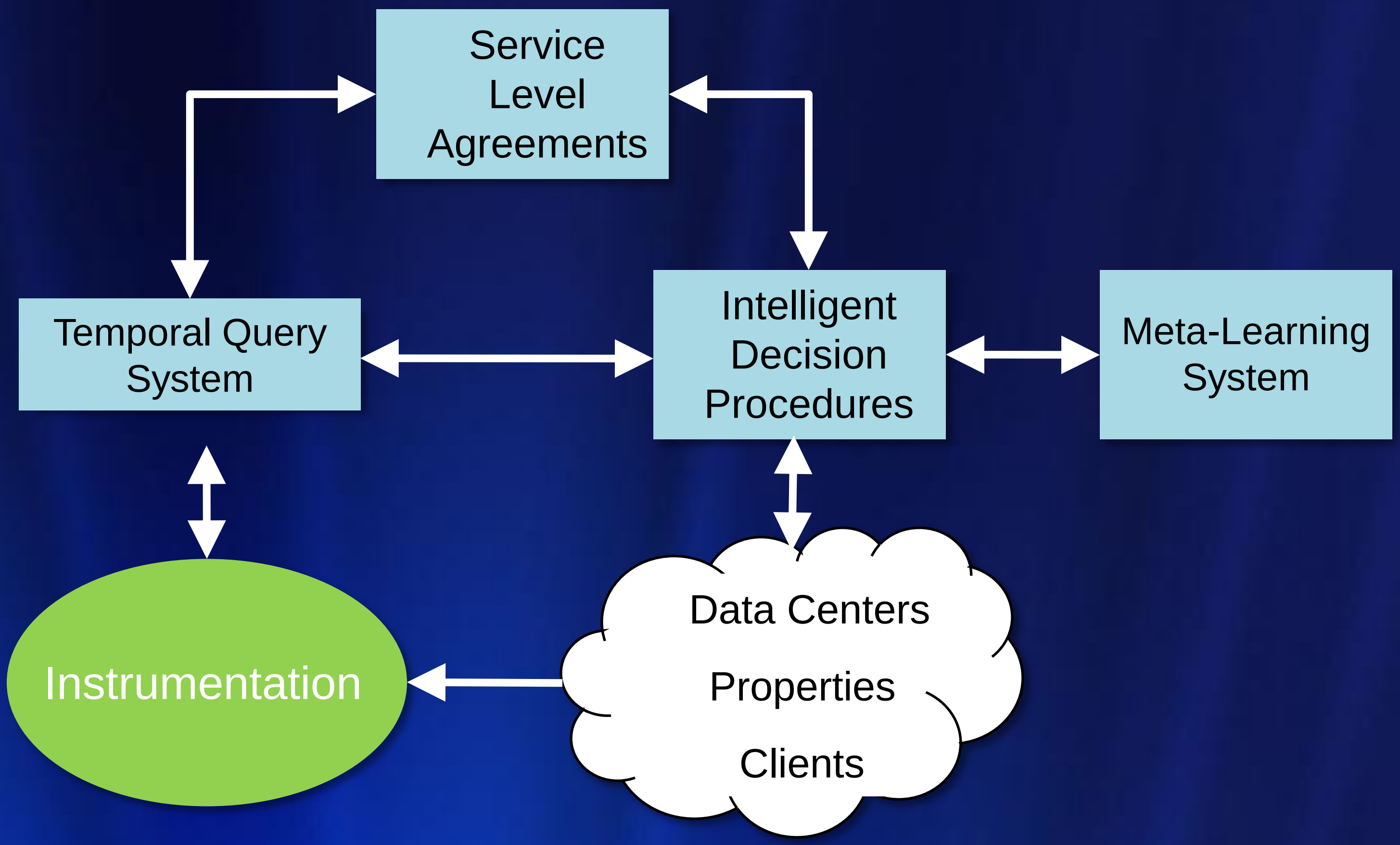
Source: Feng Zhao

Think Low Power

- Two 50 node prototypes
  - Intel Atom dual core
    - 1.6 GHz, 512 KB L2, 533 MHz FSB
    - 2 GB DRAM
  - Either FLASH SSD or disk
  - Genomote environmental monitors
- Microsoft software stack
  - Configuration baseline
- Marlowe control software
  - Intelligent power management

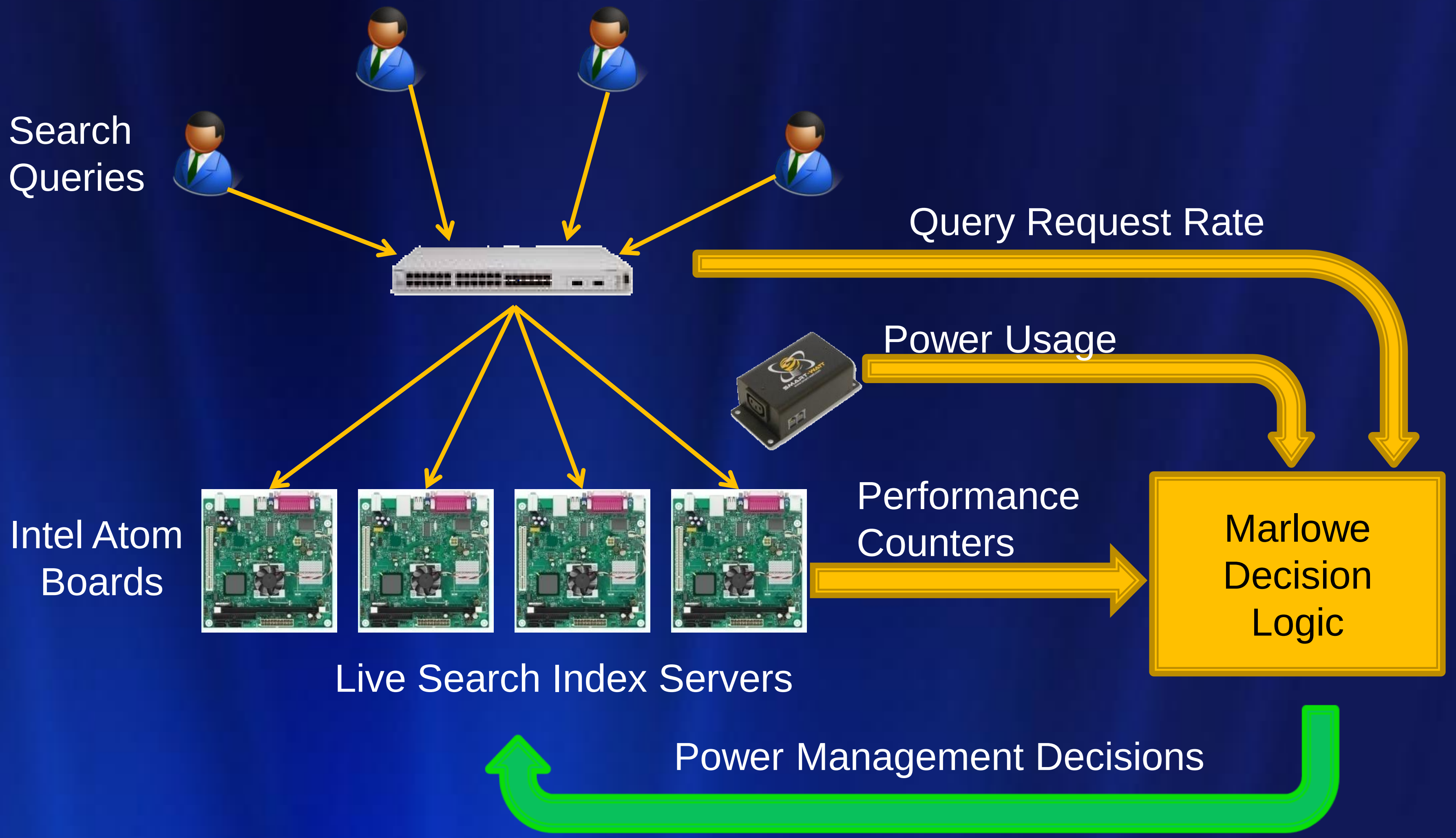
Microsoft

Marlowe Adaption Layer



Microsoft

Recent Demonstration



Some Research/Design Thoughts

- Draw the right bounding box
  - It defines the problem you solve
- Avoid high density packaging
  - It complicates cooling and design
- Use ambient air cooling wherever possible
  - Water is precious
- Run hot (really hot), if you can
  - Don’t believe the vendor specifications ☺
- Embrace component failure
  - Hardware is cheap and readily recyclable
- Machines and people do not mix well
  - Consider sealing hardware at the factory

Some Research/Design Thoughts

- One size does not fit all
  - Workloads vary widely
- Understand your workload
  - Use only the hardware you need
- Power down whenever possible
  - Off is the best solution
- Metrics reward and punish
  - Choose carefully what you measure
- Engage multidisciplinary solutions
  - Mechanical, electrical, economic, social …
- Culture shapes behavior
  - Implicit versus explicit costs

Microsoft®

Your potential. Our passion.™

Microsoft®