DC INFRASTRUCTURES III: MEASURING AND BEST PRACTICES

David López
V.2.1
Updated Spring 2021

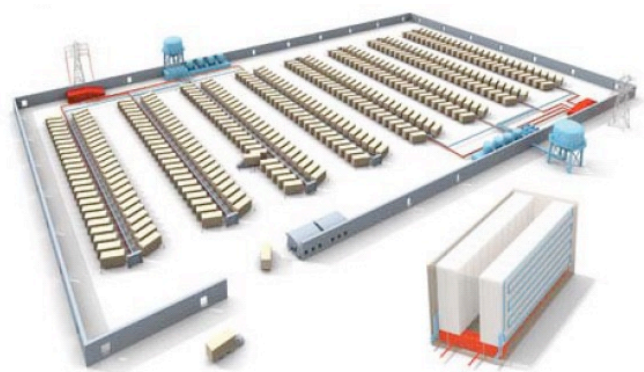
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Future buildings: time for containers?

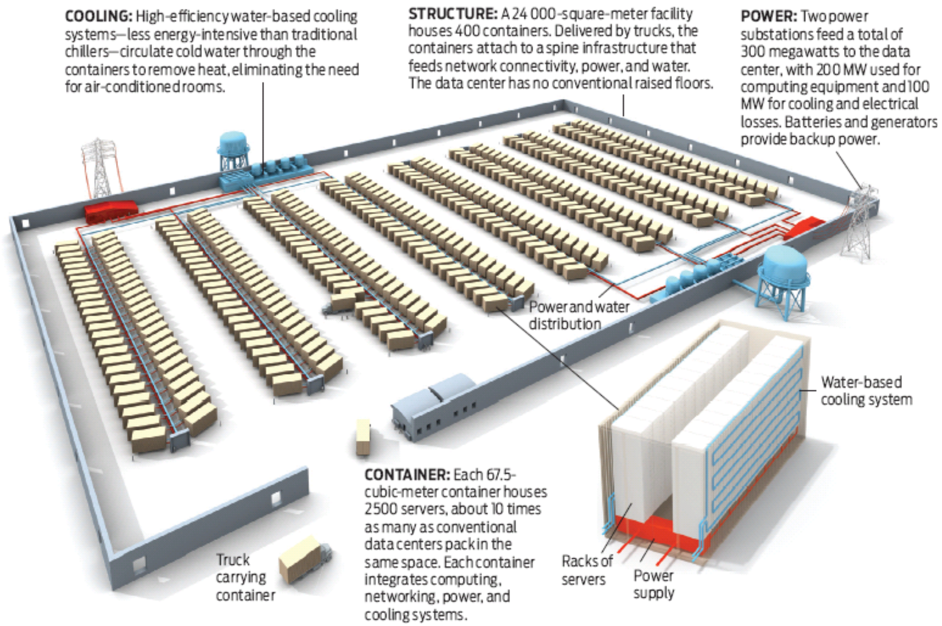
- **Provides excellent energy efficiency by offering more precise control of airflow within the container**
- **Examples: Microsoft and Google**

Estimated 14% consumption reduction

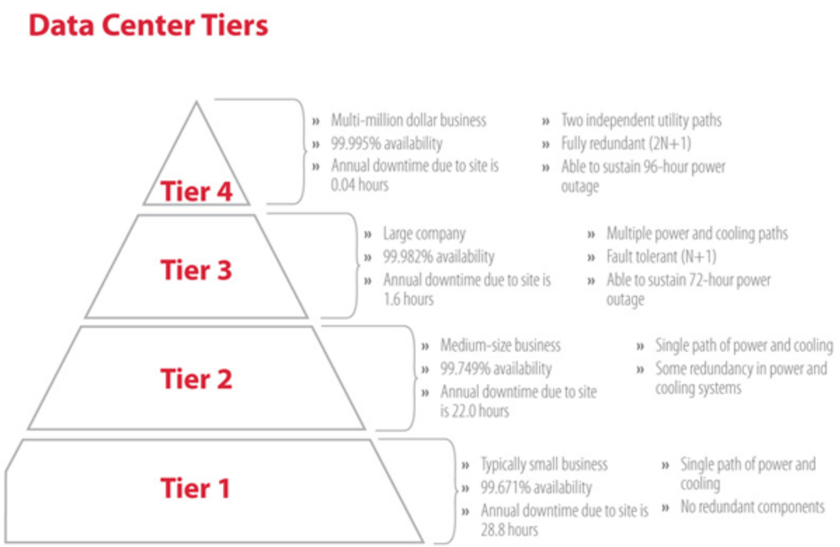
Next generation Datacenters?



Source: Tech Titans Building Boom , Randy H. Katz. IEEE Spectrum, February 2009
http://spectrum.ieee.org/green-tech/buildings/tech-titans-building-boom



Tiers



Costs of a Datacenter

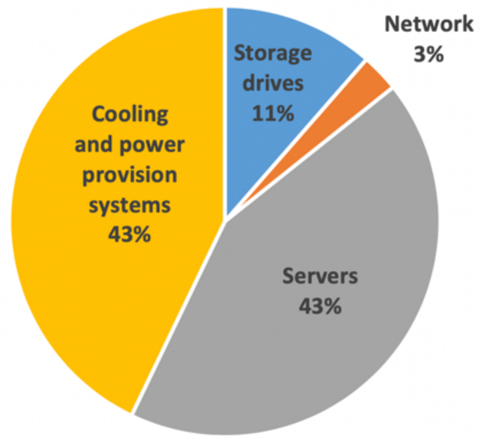
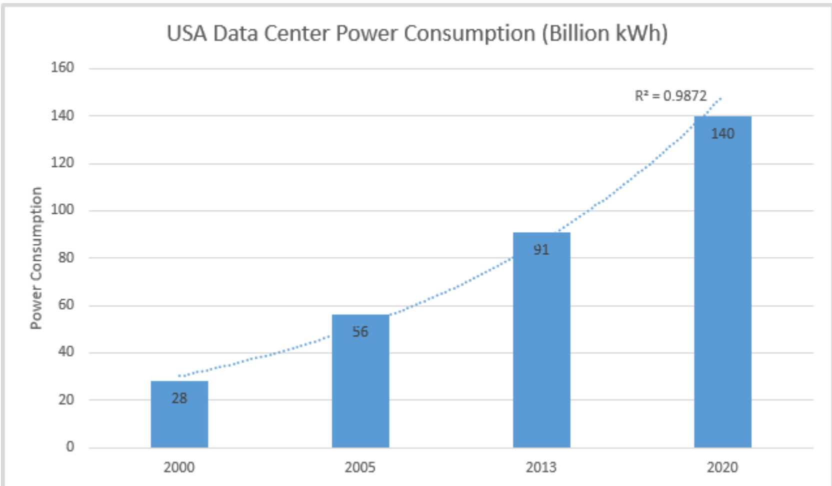
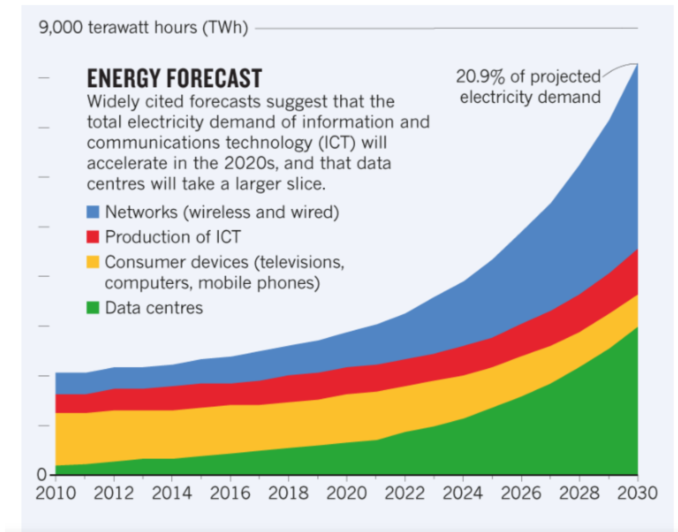


Figure 1. Fraction of U.S. data center electricity use in 2014, by end use. Source: Shehabi 2016.

How much does a DC consume?



Energy in DC forecast



PUE: Data Center efficiency metrics

Data center supporting infrastructure has a major impact on the energy use

A common measure of how efficiently a DC uses its power is called power usage effectiveness ratio (**PUE**).

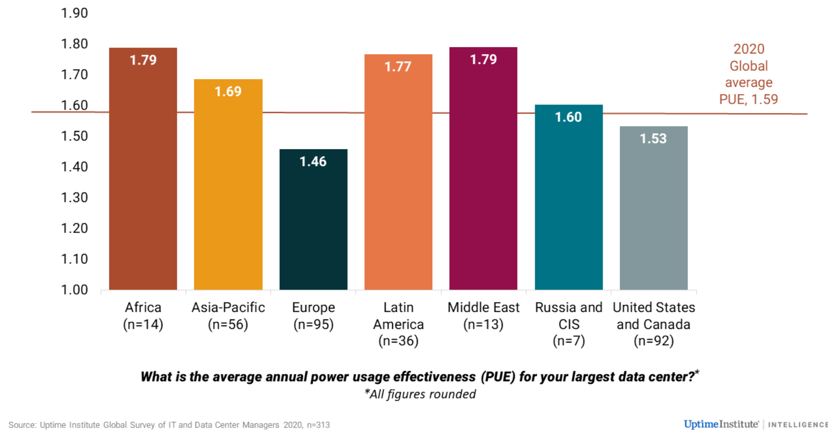
PUE DC UPC= 1.7 (*)

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

(*) Source: Javier Hidalgo

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

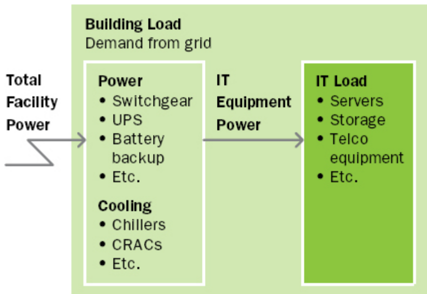
Are we efficient?



UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

PUE & DCE

PUE: Power Usage Effectiveness
DCE: Data Center Efficiency



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

$$DCE = \frac{1}{PUE} = \frac{\text{IT Equipment Power}}{\text{Total Facility Power}}$$

Source: The Green Grid, 2008, "The Green Grid Metrics: Data Center Infrastructure Efficiency (DCiE) Detailed Analysis"

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

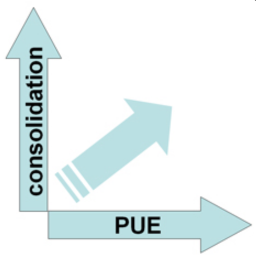
Is PUE a good measure?

Usually the goal is to reduce the data center PUE

- Nevertheless any project that improves an IT load alone will yield a worse PUE!

Example:

- 100 Mw coming into a facility and 50 Mw are taken up by the IT load → PUE = 2
- A consolidation strategy reduce the IT part to 40 Mw
- The PUE is now 2.25 (90/40), which is worse than the PUE of 2 we had before our virtualized/consolidated strategy



PUE is not enough!

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

So?

No actual metric has been defined yet

Ideas:

Efficiency = Computation / Total energy

$$\text{Efficiency} = \underbrace{\frac{1}{\text{PUE}}}_{(a)} \times \underbrace{\frac{1}{\text{SPUE}}}_{(b)} \times \underbrace{\frac{\text{Computation}}{\text{Total energy to Electronic Components}}}_{(c)}$$

We know the (a) term of equation (DCE)

SPUE

Term (b): PUE does not account for inefficiences within the server (server’s power supply, voltage regulator modules -VRMs-, cooling fans)

- SPUE= Server PUE = total server input power / power consumer by components directly involved in computation (motherboards, disks, CPUs, DRAM, I/O cards, and so)
- Current SPUE values (Barroso & Hölzle)= 1.6 – 1.8

Some companies use **tPUE** (total PUE)

$$\text{tPUE} = \underbrace{\frac{1}{\text{PUE}}}_{(a)} \times \underbrace{\frac{1}{\text{SPUE}}}_{(b)}$$

CPU utilization in Google servers

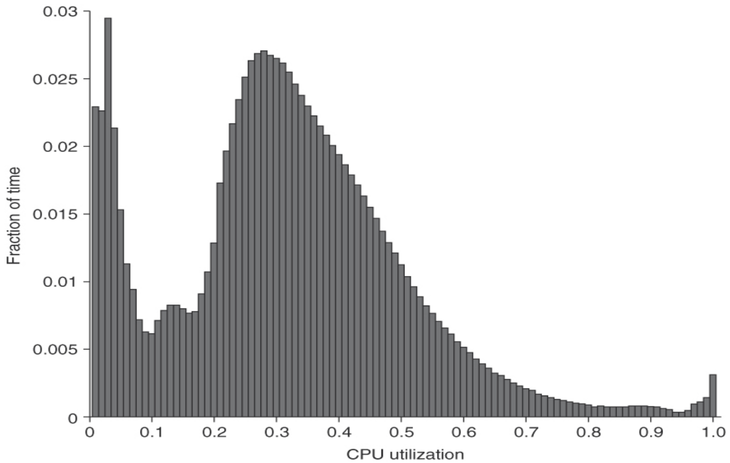


Figure 6.3 Average CPU utilization of more than 5000 servers during a 6-month period at Google. Servers are rarely completely idle or fully utilized, in-stead operating most of the time at between 10% and 50% of their maximum utilization. (From Figure 1 in Barroso and Hölzle [2007].)

Borrowed from Hennessy & Patterson: Computer Architecture, A Quantitative Approach

TCO (Total Cost of Ownership)

TCO analysis was popularized by the Gartner Group in 1987

- TCO tries to quantify the financial impact of deploying an information technology product over its life cycle
- An associate idea is the Return of Investment (**ROI**)

Technology deployment can include the following as part of TCO:

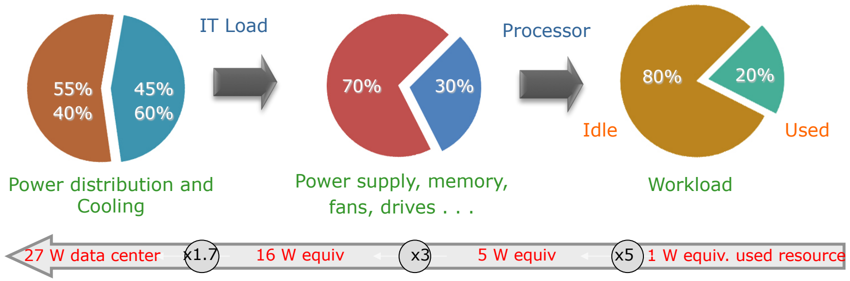
- Computer hardware and programs
  - Network hardware and software
  - Server hardware and software
  - Workstation hardware and software
  - Installation and integration of hardware and software
  - Purchasing research
  - Warranties and licenses
  - License tracking – compliance
  - Migration expenses
  - Risks: susceptibility to vulnerabilities, availability of upgrades, patches and future licensing policies, etc.

(continue)

TCO (Total Cost of Ownership)

- Operation expenses
  - Infrastructure (floor space)
  - Electricity (for related equipment, cooling, backup power)
  - Testing costs
  - Downtime, outage and failure expenses
  - Diminished performance (i.e. users having to wait, diminished money-making ability)
  - Security (including breaches, loss of reputation, recovery and prevention)
  - Backup and recovery process
  - Technology training
  - Audit (internal and external)
  - Insurance
  - Information technology personnel
  - Corporate management time
- Long term expenses
  - Replacement
  - Future upgrade or scalability expenses
  - Decommissioning

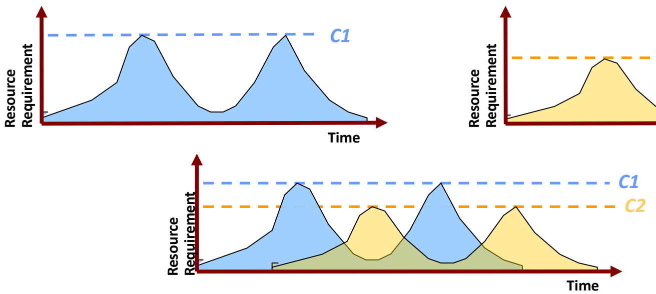
How do datacenters spent energy?



Idea: IBM and Dynamic Infrastructure, Doug Neilson, IBM Systems Group, 2009

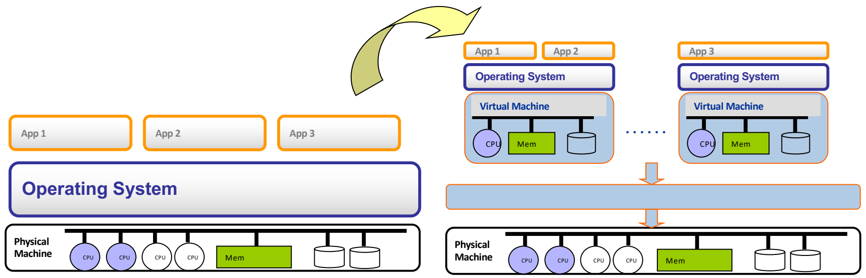
Consolidation

- Servers are underused: around 10 to 20 percent
- Consolidate into a single machine
  - Resource multiplexing if different peak times



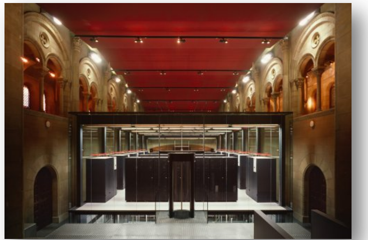
Virtualization

Virtualization divides a physical server into isolated virtual environments, enabling organizations to run multiple applications or OS on a single server



Best Practice at Marenostrum

Can we Minimize losses and thermal/cooling overheads in Marenostrum?



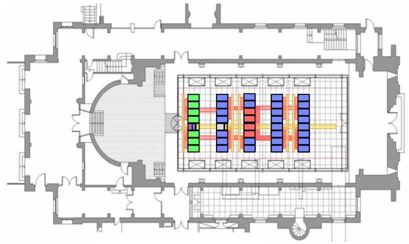
Source: BSC

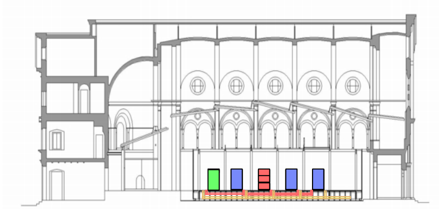
Source:
"El reto operacional de dirigir el tercer supercomputador más grande de Europa"
Sergi Girona,
Director Operaciones
BSC-CNS, 2009.

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Some Marenostrum I facts

- Peak Perf: 94.21 Teraflops
- 10,240 Power PC 970MP at 2.3 GHz (2560 JS21 blades)
- 20 TB of main memory
- 280 + 90 TB of disk storage
- Interconnection networks:
  - Myrinet and Gigabit Ethernet
- Linux: SuSe Distribution
- Space 160 square meters
- About 26 tons of steel (19 glass).
- Power aprox. 1.071 kW

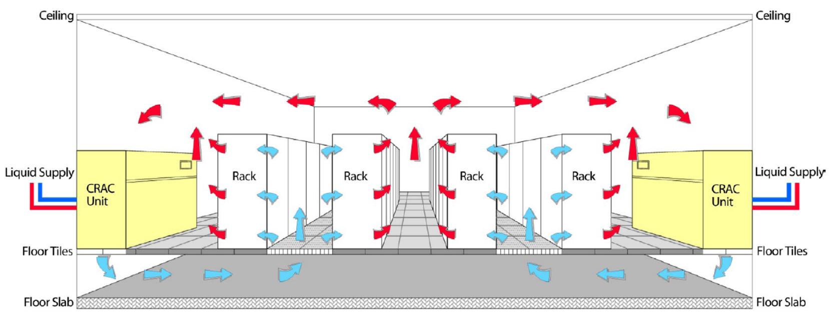




UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Improving air flow management

Raised floor with hot – cold aisles
Estimated PUE around 1,45



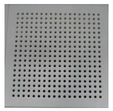
Font: Luiz Andre Barroso, Urs Hoelzle, "The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines", 2009.
(Image courtesy of DLB associates , ref [23] of the book)

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Under-floor pressure

Problem measured: to much under-floor pressure

- Solution: Move some floor tiles



Benefits observed:

- Improvement of AC equipment performance
- Improved the rack-bottom temperature

Source: BSC

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Substituting floor tiles

- Composite
  - 20% opening

- Metallic tile
  - 40% opening

Source: BSC

Benefits observed:

- Less working pressure for the Cooling components
- All bladecenters temperature reduced by 2º C
- Cold barrier that prevents the reflux of hot air

Temperature map

Problem measured:

- not all the Racks have the same temperature

Idea:

- Force the air flow

28.50	28.00	25.50	28.00	25.50	25.50	25.50	31.00
27.00	24.00	24.00	25.00	23.00	24.00	23.00	28.50
26.50	25.50	23.50	25.00	24.50	23.00	24.00	29.50
28.00	24.50	23.50	26.00	24.00	24.50	23.50	28.50
22.50	26.00	25.00	27.00	24.50	25.00	25.00	27.50
27.50	27.00	27.00	27.00	25.50	28.50	25.00	27.50
27.50	26.50	27.00	27.50	26.50	26.00	28.00	27.50
27.50	25.00	23.00	25.00	24.00	26.00	24.50	30.00
29.00	24.00	24.50	23.50	23.00	23.50	24.50	29.00
28.50	24.00	23.50	25.50	24.50	25.50	24.50	28.00
27.00	23.50	25.00	25.50	21.50	25.00	25.50	26.00
27.00	24.50	24.50	22.50	22.50	25.50	24.00	27.50
26.00	25.50	MYRI	MYRI	MYRI	MYRI	24.00	30.00
28.50	24.00	MYRI	MYRI	MYRI	MYRI	25.00	27.50
26.00	24.00	MYRI	MYRI	MYRI	MYRI	23.50	30.50
27.00	24.00	19	19	19	19	25.50	31.00
28.50	21.50	19	19	19	19	22.00	29.00
25.50	25.50	19	19	19	19	21.50	27.00
28.50	25.50	26.00	26.00	NET	26.50	26.00	27.50
29.00	25.50	24.50	23.50	NET	24.50	23.00	25.50
26.50	24.50	24.50	24.00	NET	24.50	23.00	26.50
27.50	25.50	25.50	25.00	NET	22.50	23.00	23.00
25.50	26.00	26.50	27.50	NET	24.50	25.50	21.00
27.00	26.00	27.00	28.00	NET	26.50	27.00	24.00

Forcing air flow

Methacrylate screens installed in front of each rack

- to guide the cold air flow directly to the computer, instead of each rack having to take cooling air from the general environment

Source: BSC

Benefits observed:

- All BladeCenters Rack has an equal temperature +/- 1ºC
- BladeCenter fan speed reduced

Results

Including other improvements:

- reduction of **10%** of power consumption (and CO2)

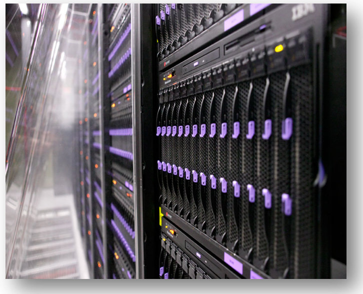


Image courtesy of UPC

Marenostrum **power consumption**: approx. 1.2 Mw

→ aprox 1.100.000 €/year

PUE ~ 1,3

DATA CENTER INFRASTRUCTURES

David López

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH