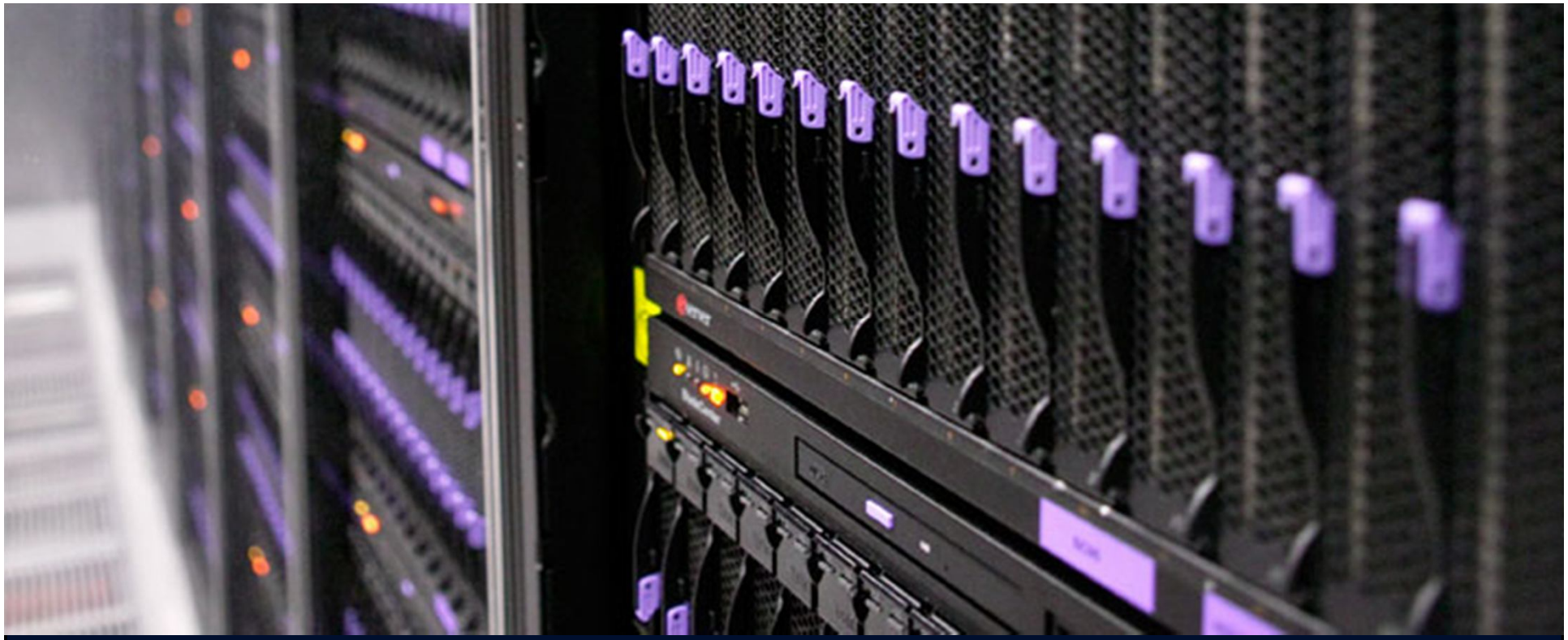


The Unstoppable Transformation (revolution) of IT Sector

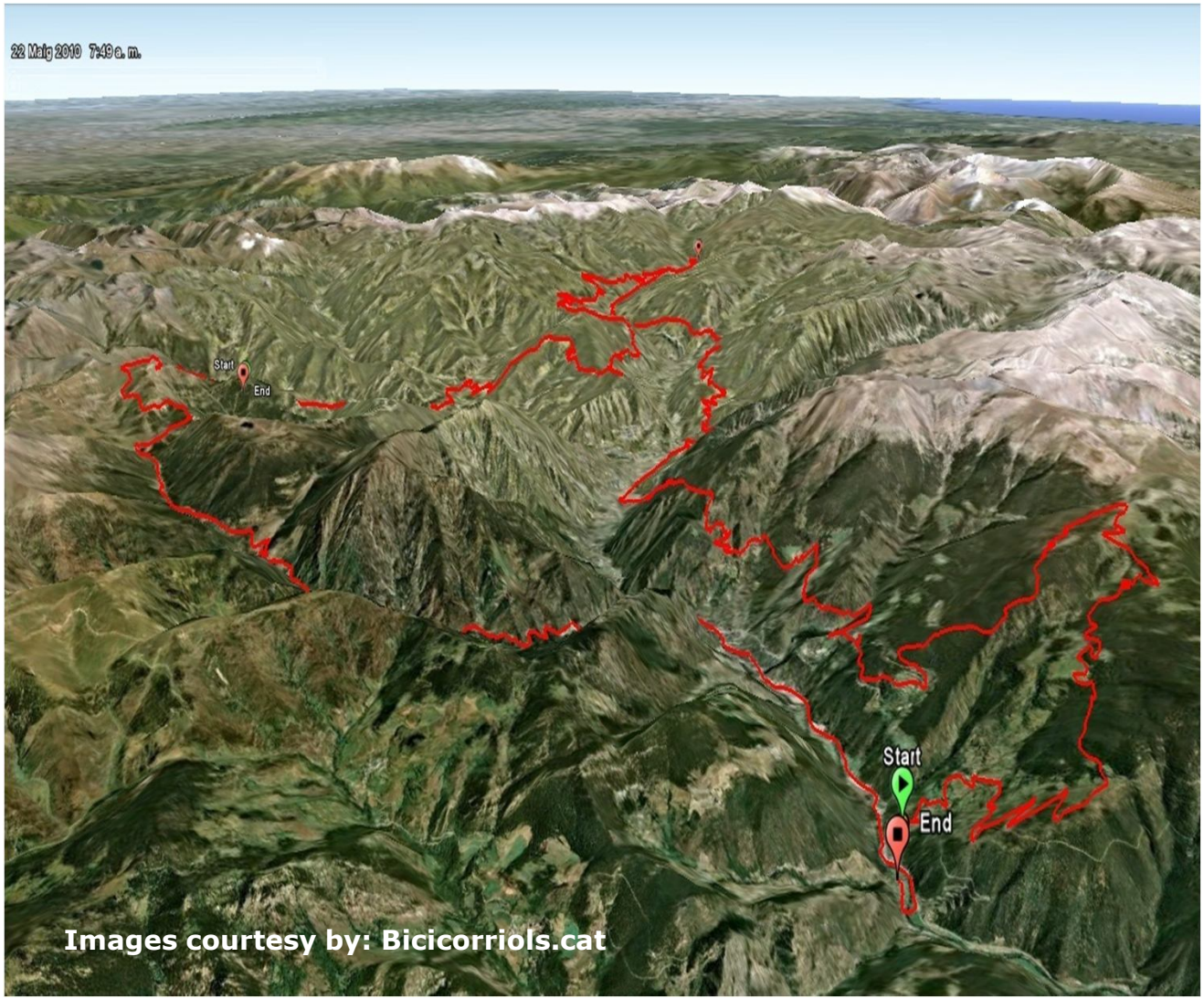
Jordi Torres
UPC Barcelona Tech
Barcelona Supercomputing Center (BSC)

"Aules Empresa" FIB-UPC. January-2011

we records our life … and posts it

Source: Courtesy Campus-Party
(http://www.flickr.com/photos/campusparty)

Capturing every step



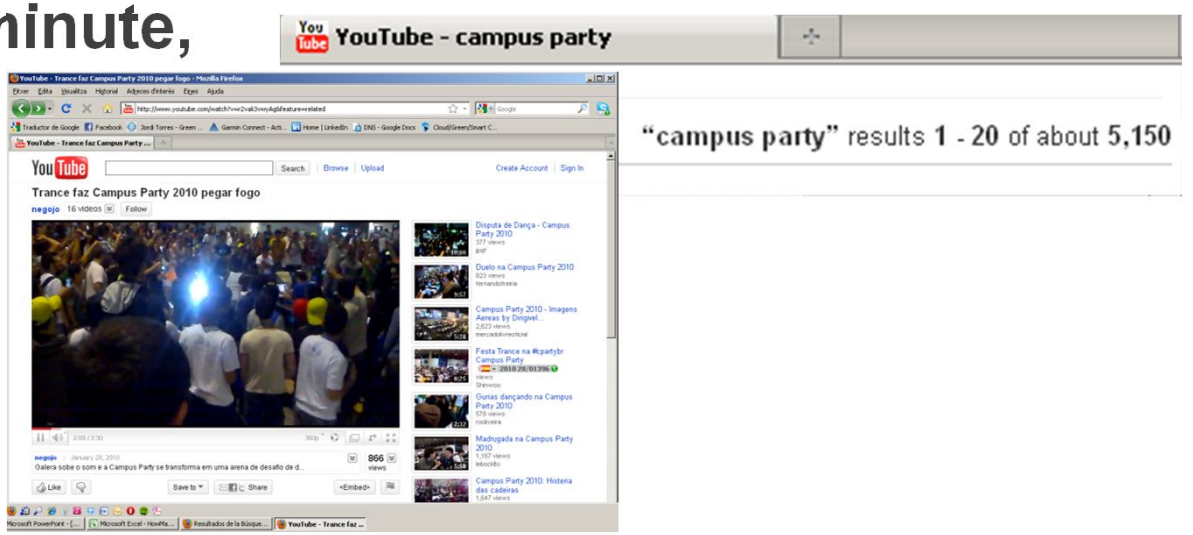
We are posting everyting …

… on the Internet,

Source: Powerpoint ClipArt Gallery

Internet facts:

- **Internet users: There are an estimated 1.9 billion**
- **Activity? E.g: More than 20 hours of YouTube videos are uploaded every minute,**



~~Teraflops, Petaflops?~~

Exabytes, Zettabytes?

Homework:
“La nova taula de multiplicar”

SI decimal prefixes		Binary usage	IEC binary prefixes	
Name (Symbol)	Value		Name (Symbol)	Value
kilobyte (kB)	10^3	2^{10}	kibibyte (KiB)	2^{10}
megabyte (MB)	10^6	2^{20}	mebibyte (MiB)	2^{20}
gigabyte (GB)	10^9	2^{30}	gibibyte (GiB)	2^{30}
terabyte (TB)	10^{12}	2^{40}	tebibyte (TiB)	2^{40}
petabyte (PB)	10^{15}	2^{50}	pebibyte (PiB)	2^{50}
exabyte (EB)	10^{18}	2^{60}	exbibyte (EiB)	2^{60}
zettabyte (ZB)	10^{21}	2^{70}	zebibyte (ZiB)	2^{70}
yottabyte (YB)	10^{24}	2^{80}	yobibyte (YiB)	2^{80}

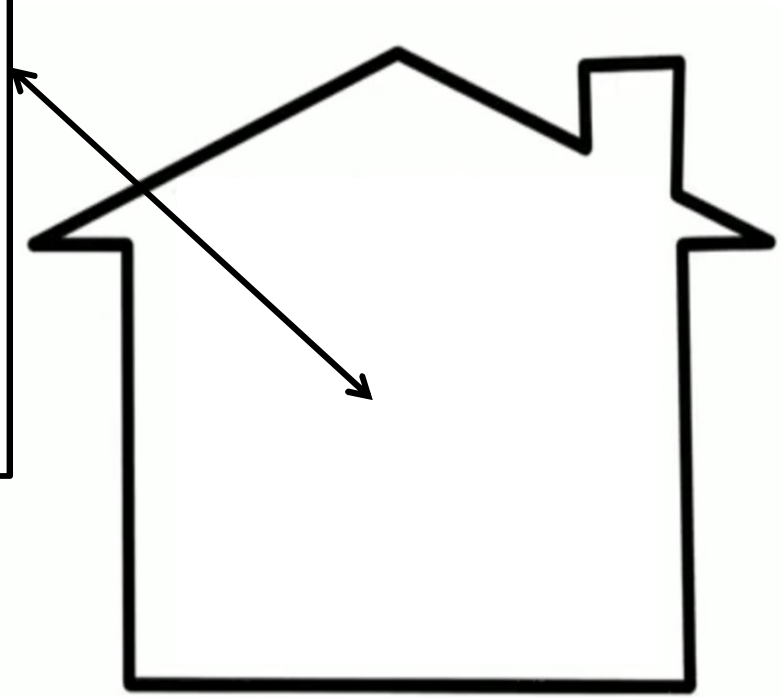
Source: Wikipedia

Where does all this information live?

Where does all this information live?

Where does all this information live?

- **Microsoft+Amazon are estimated to operate over 50,000 servers**
- **Google operates over 1,000,000 servers**



approximately 2% of the world's servers.

Servers in the World?

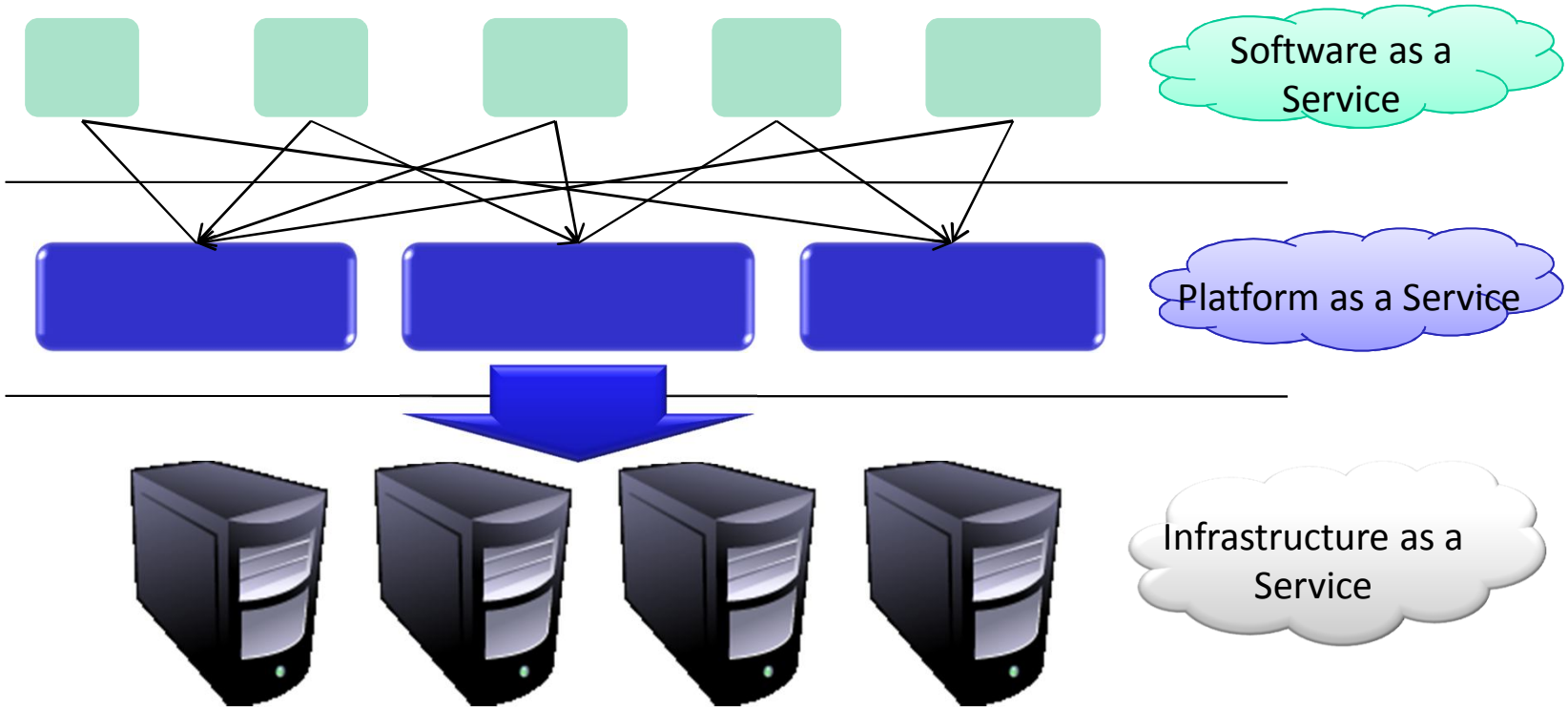
50 millions

the number of physical servers in the world today?

On a social level: Cloud Computing is here!

On a social level: Cloud Computing is here!

Cloud Computing: Layers of Abstraction



Source: http://www.slideshare.net/Cumulux/what-is-cloud-computing-presentation?nocache=4789

Cloud Computing: Services

- **Cloud computing services can be classified as:**
  - **Infrastructure as a Service (IaaS)**

    It is at the lowest layer, and it is a wayof delivering basic storage and compute capabilities as standardized services over the network.

  - **Platform as a Service (PaaS)**

    It is in the middle, consisting of higher-layer capabilities that developers can knit into applications.

  - **Software as a Service (SaaS)**

    It is at the highest layer and features a complete application offered as a service, on-demand, via multi-tenancy.

Cloud Computing: next generation of technology

- **History repeats itself**
  - The easiest analogy to explain “the cloud” it that of electricity

- **Pay as you go model**

Outline

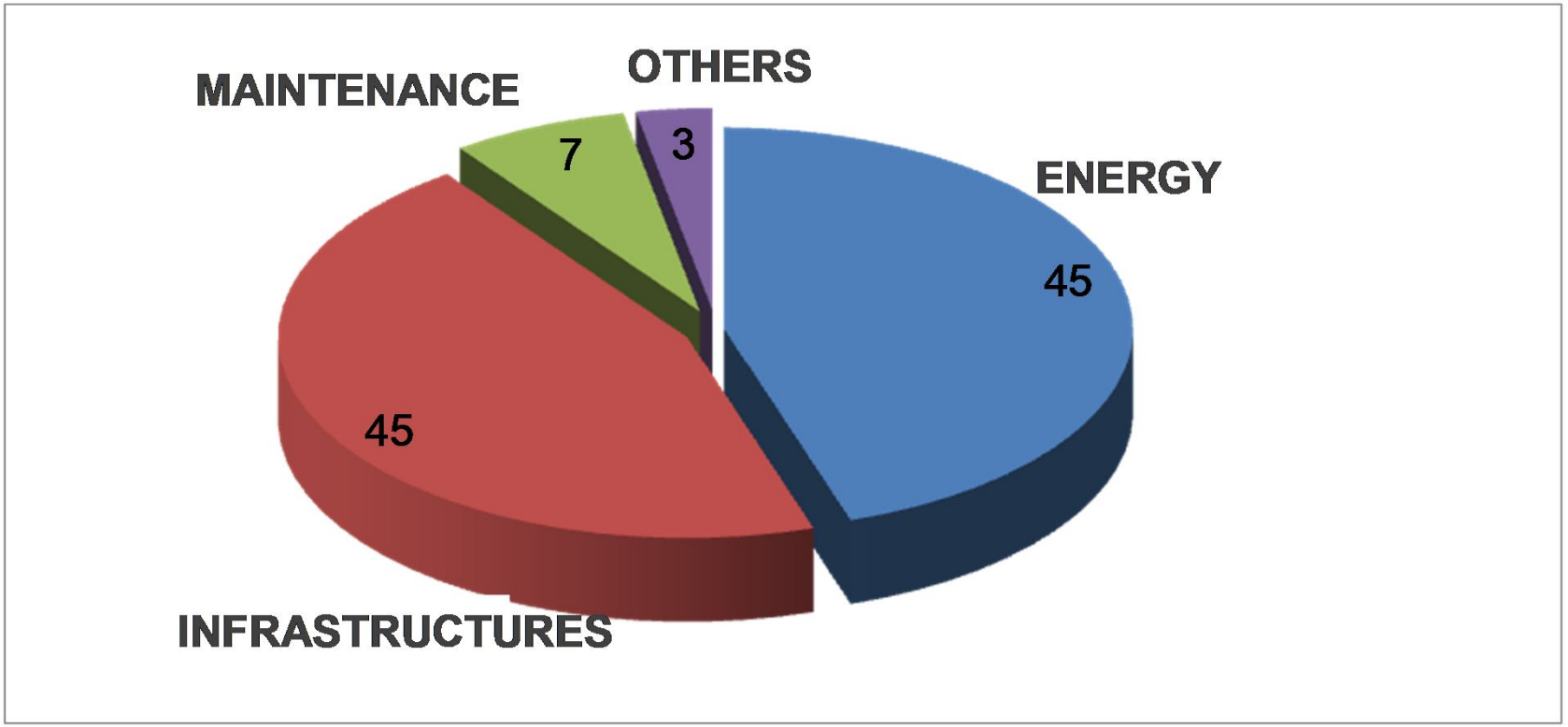
Providers side: The industrialization of ITs

Client side: The Demand for Cloud Services

Cloud Opportunities: Green Computing

OPEX Data Center

- 5 years (%)



Data courtesy of
AST IT Infrastructures
(www.AST-global.com)

Calculation for avg density 7kw/rack
Infrastructure amortization counted as Opex
HW and SW excluded
Data Center Infrastructure valued as Opex

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 CPD edifici Vertex = 1.69 (*)**

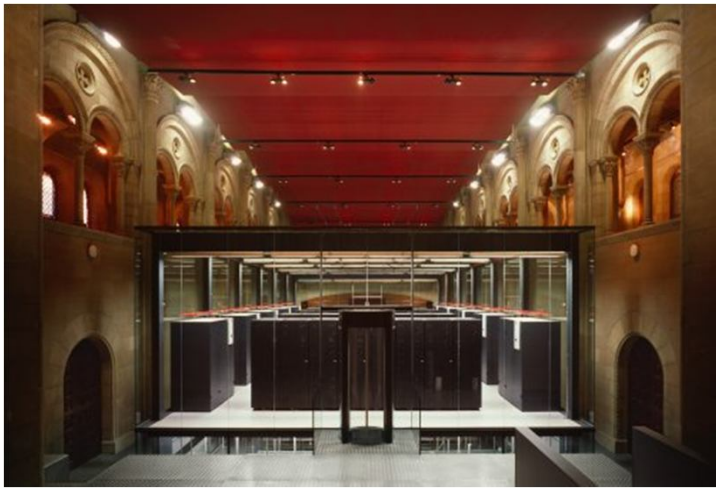
$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

(*) Source: Javier Hidalgo/Jose L Montero

What can we do for DC power efficiency?

- **Can we Minimize losses and thermal/cooling overheads in a Data Center?**

  Example: Reducing cooling costs at Marenostrum

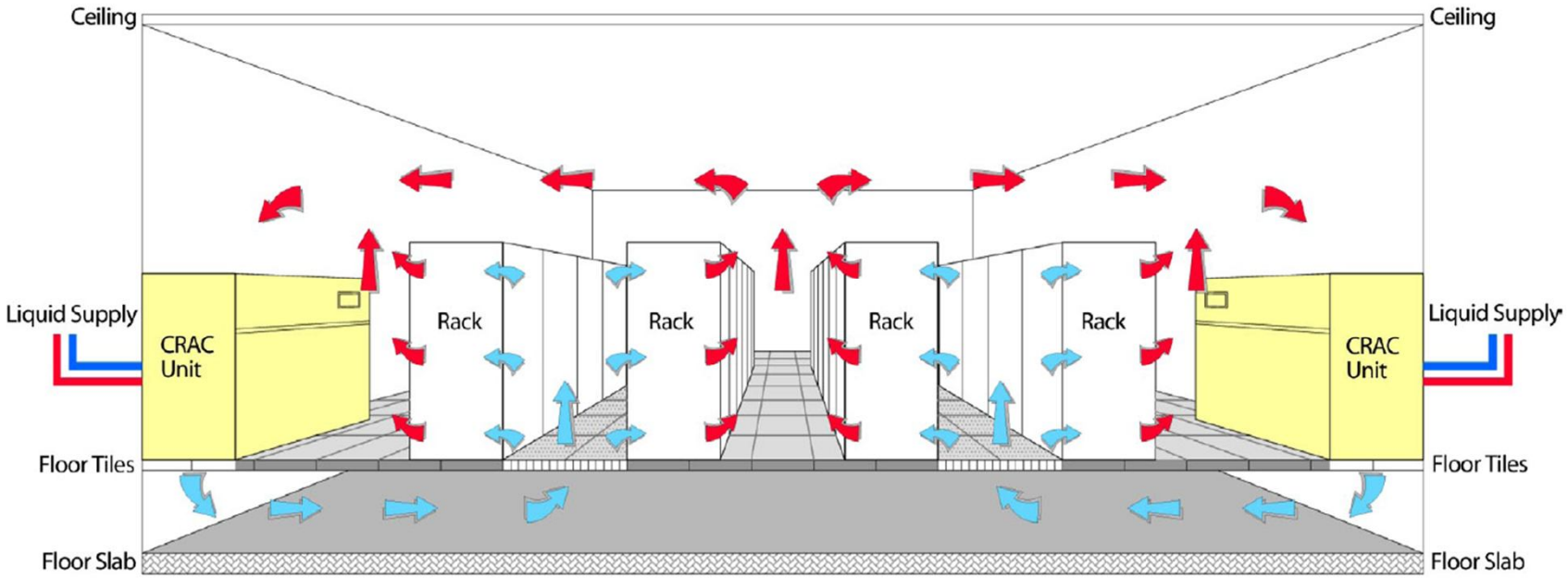


Source: BSC

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

Improving Air Flow Management

- **Usual scenario: Isolate Hot and Cold**



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)

Best practice at Marenostrum

- **Problem measured: to much under-floor pressure**
- **Test: Move some floor tiles**

- **Benefits observed:**
  - Improvement of AC equipment performance
  - Improved the rack-bottom temperature

Best practice at Marenostrum

- **Substituting Floor Tiles**
  - Composite
    - 20% opening

  - Metallic tile
    - 40% opening

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

Source: BSC

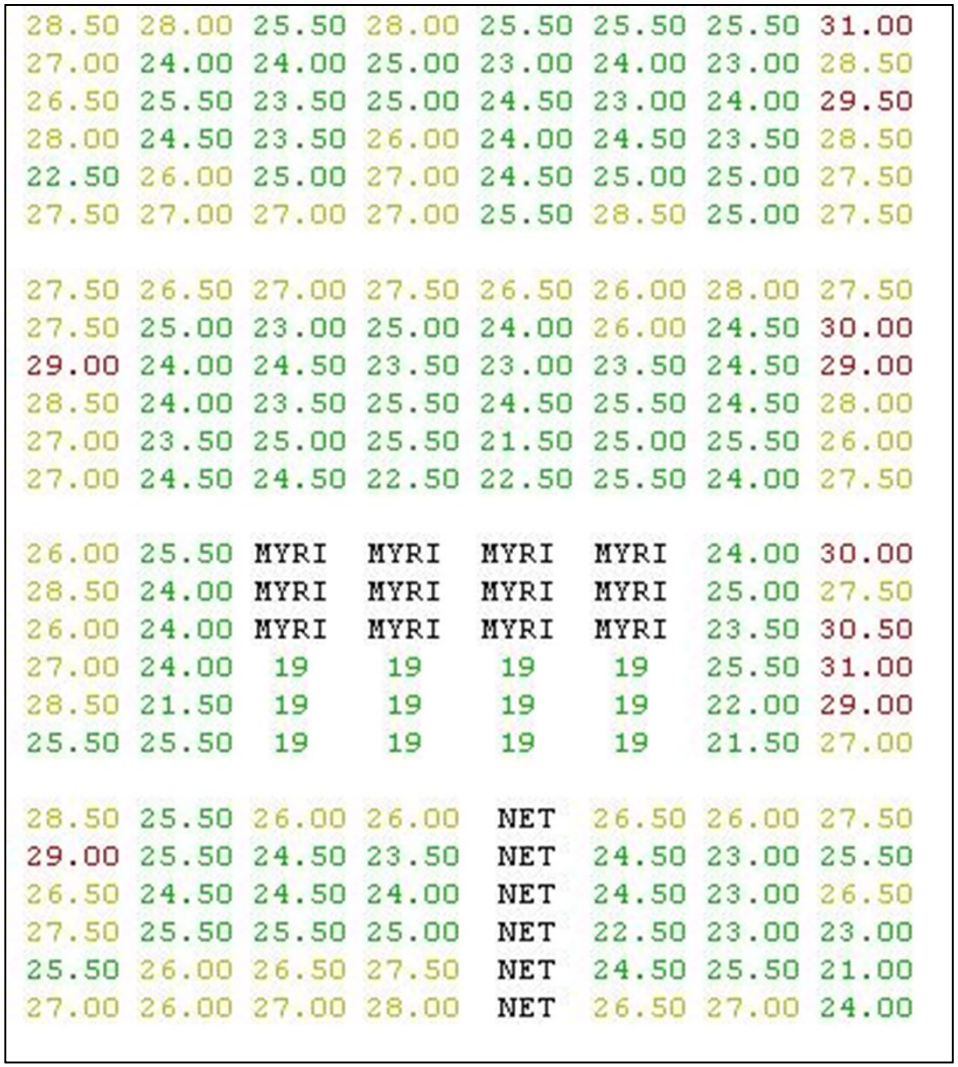
Temperature map

- **Problem measured:**

  not all the Racks have the same temperature

- **Test:**

  Force the air flow



Best practice at Marenostrum

- **Problem measured :**

  not all the Racks have the same temperature

- **Proposal:**

  Methacrylate screens

  installed in front of each rack

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

Source: BSC

Best practice at Marenostrum

- **Results (including other improvements):**

reduction of 10% of power consumption (and CO2)

- **Marenostrum power consumption: approx. 1.2 Mwats**
  - → aprox 1.100.000 €/year

- **PUE**

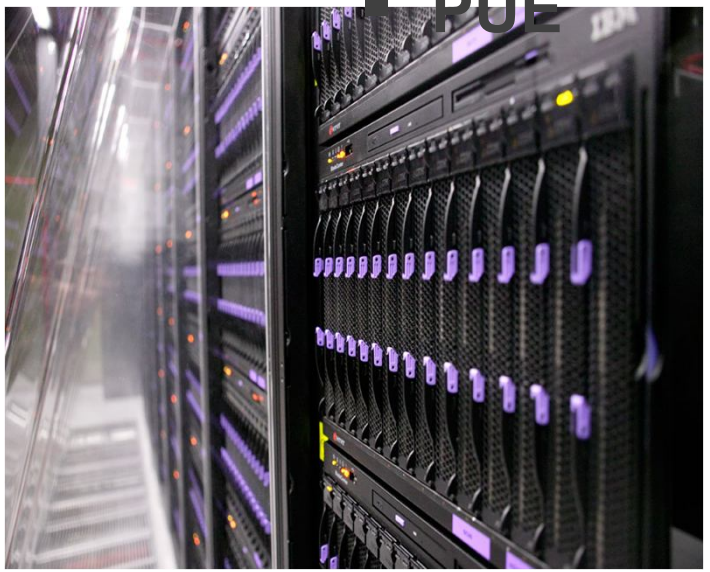


Image courtesy of UPC

PUE ~ 1,3

Time for Containers

PUE ~ 1,3

1

PUE ~ 1,113

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

Estimated ~14% PUE reduction

Decreasing DC power consumption

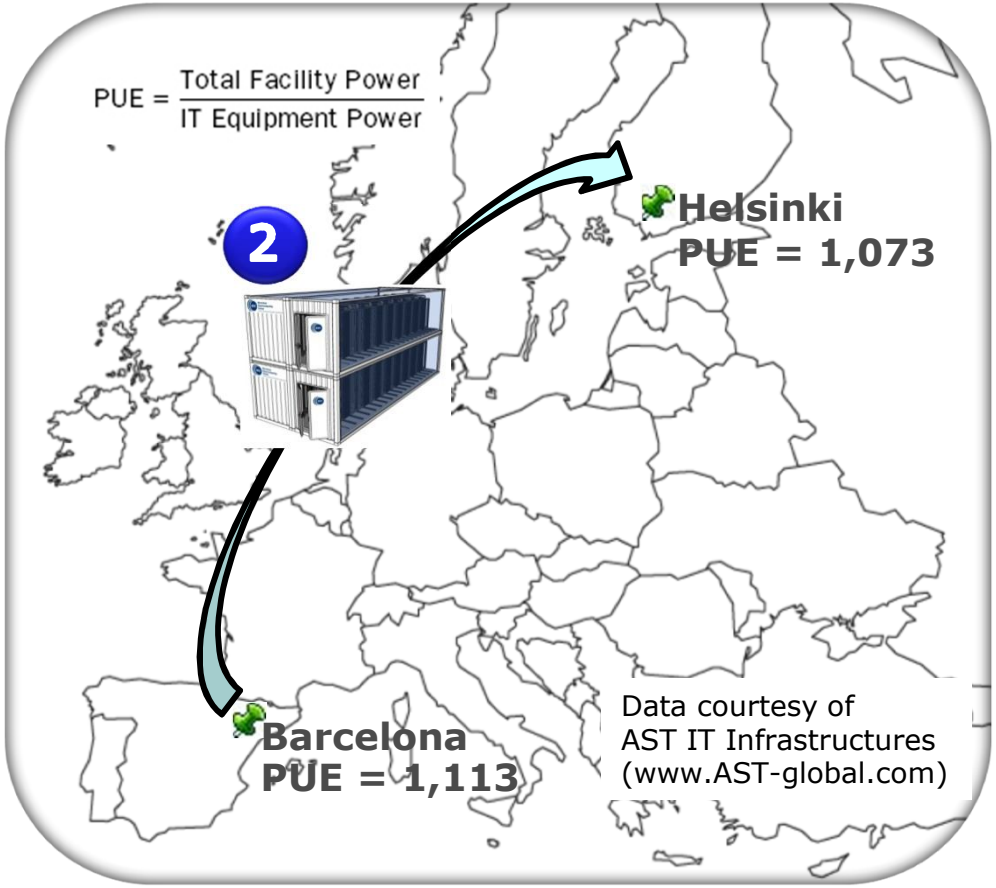
- **How can we decrease DC power consumption without a concomitant decrease in the IT load?**

- **Can we move the Marenostrum in a better location?**

  e.g. more efficient cooling due to climate conditions

Can we put a DC in a better location?

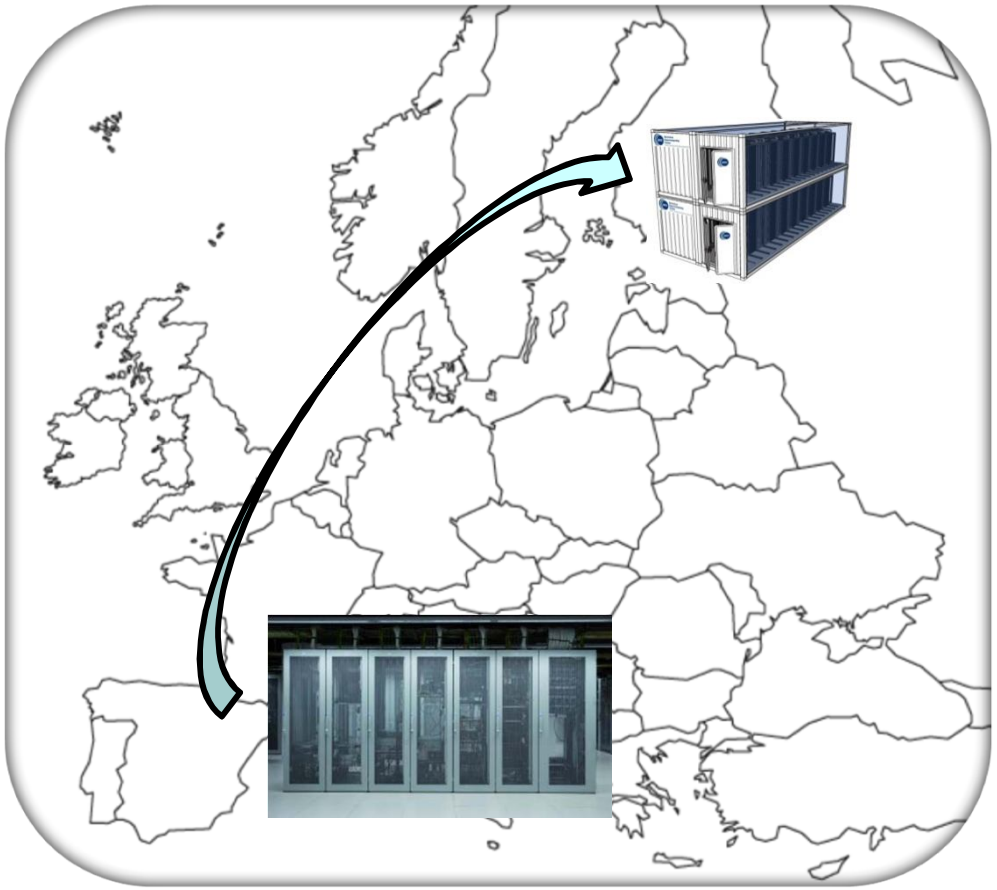
- I containerized marenostrum and it has traveled with me



Estimated ~3,5% of energy could be saved

Differences in cost comparing Cloud with CPD UPC

- **PUE 1.69 → 1,3 → 1,113 → 1,073**



Estimated ~36% of energy could be saved

What else are they doing?

Maximize work done per watt

What else are they doing?

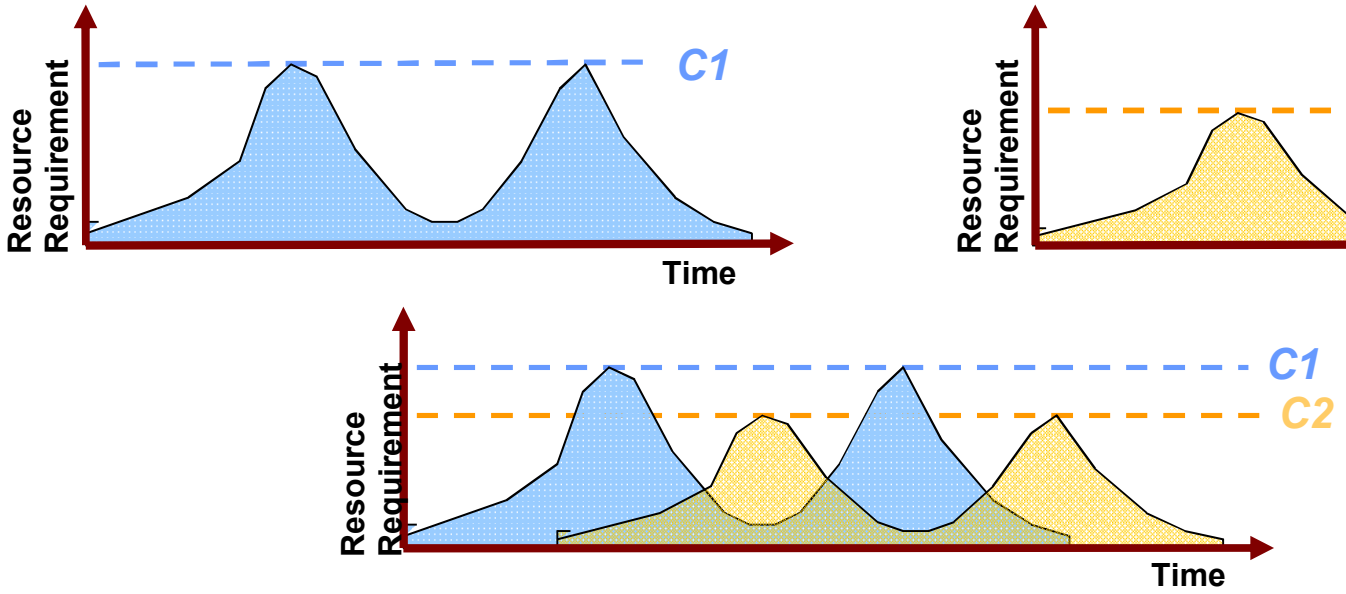
Maximize work done per watt

- **Energy consumption is not only determined by the efficiency of the physical resources**
- **but it is also dependent on the resource management**

Intelligent management of resources may lead to significant reduction of the energy consumption by meeting the performance requirements

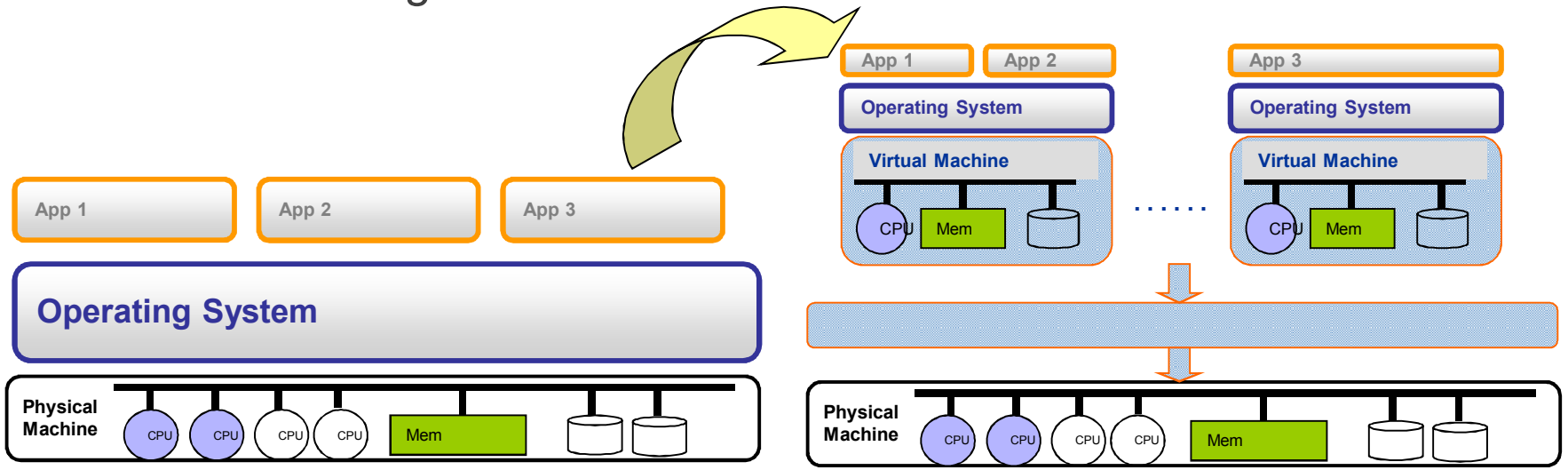
Workload Consolidation

- **Rationalize and reduce operational costs: server consolidation**
  - implies combining workloads from separate applications into a smaller number of systems if there are different peak times
- **IMPORTANT: Heterogeneous workload!**



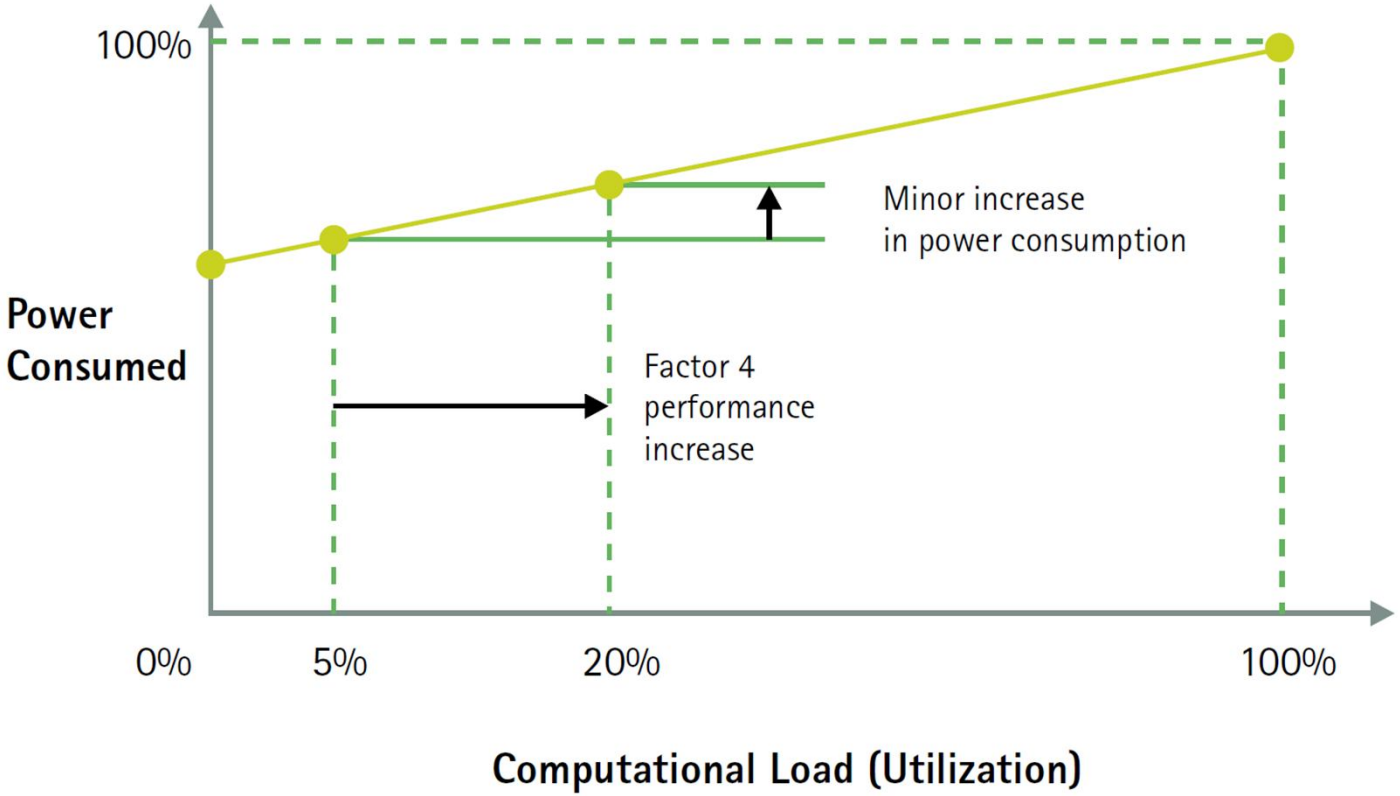
Workload Consolidation

- **Consolidation uses virtualization to share the resources**
  - Virtualization software divides a physical server into isolated virtual environments, enabling organizations to run multiple applications or OS on a single server.



The energy savings can be up to 20% - 70%

Power consumption after consolidating



Source: Cloud Computing and Sustainability: The Environmental Benefits of Moving to the Cloud Accenture, 2010

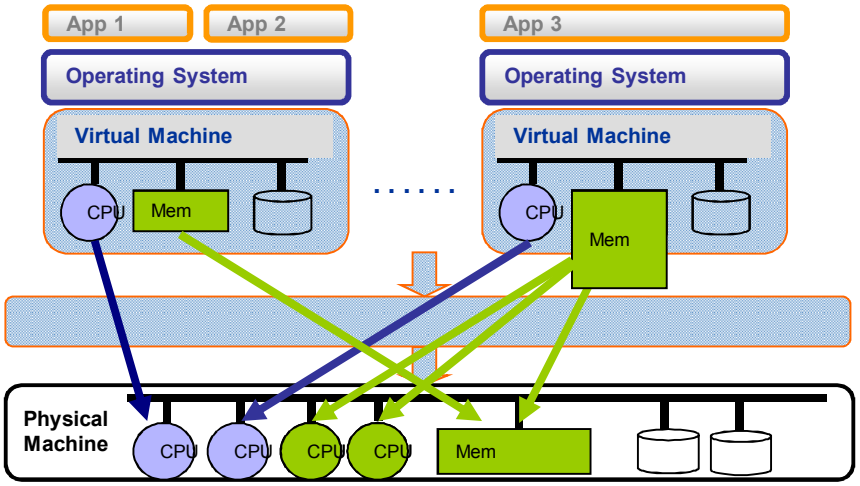
Techniques that virtualization allows

- **In present times, can we go beyond virtualization?**
- **Virtualization allows many techniques:**
  - turning on/off policies,
  - new processors,
  - dynamic voltage and frequency scaling,
  - hybrid architectures on datacenters,
  - federating clouds
  - etc.

Using multi-core processors

- **Middleware allow that applications benefit from new multi-core designs beyond the traditional parallelism**

  Example: Exploitation of Accelerators in memory-bound Applications using memory compression techniques

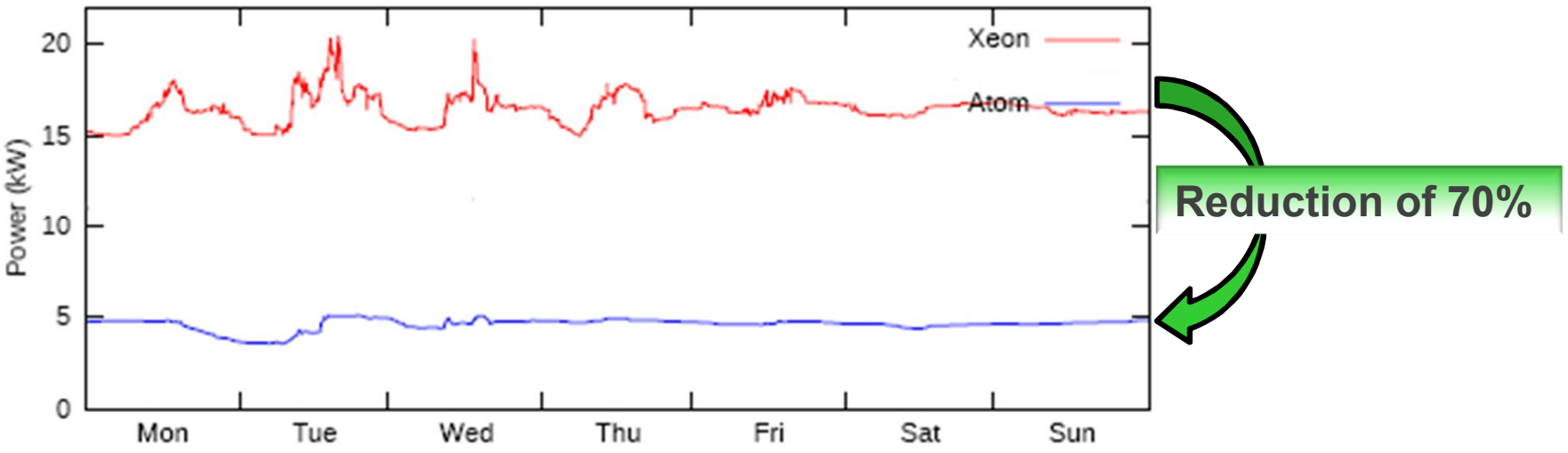


For memory-bound appl. the energy savings can be up to 25%

Use low-power processors

- **Low-power processors vs high-performance processors**

- **Good approach for transactional workloads (webs)**
  - Experiments at BSC:



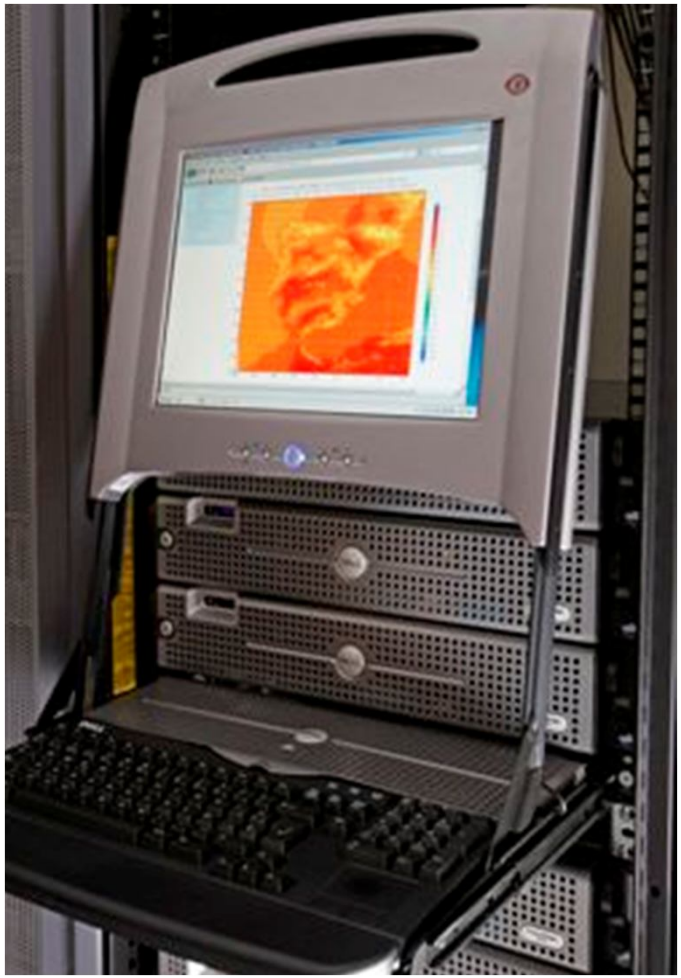
Turn on/off policies

- **However the numerical applications do not behave like this**

  Experiments at BSC with 4 HPC tasks:

  - 1 Xeon * 1 hour → 317 Watts
  - 2 Atom * 5 hours → 398 Watts

  – The most energy-efficient approach in this environment is to run jobs very fast and then **turn off the power of the system**.



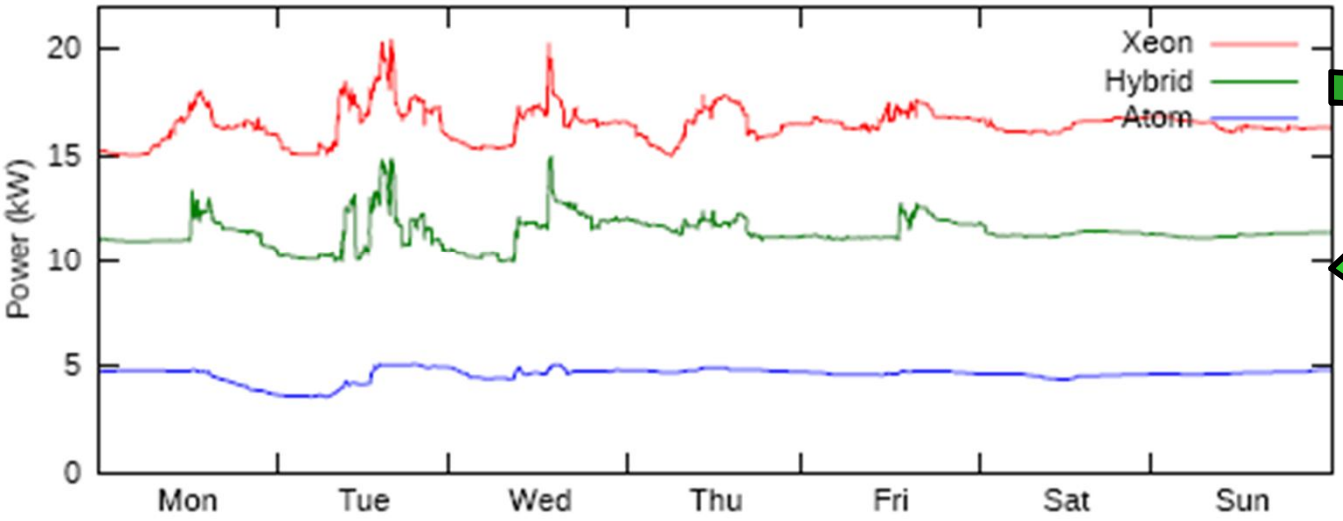
Hybrid hardware

- **But in general the workload in data centers is heterogeneous**
  - For HPC, Xeons work more efficiently than Atoms
  - For Transactional, Atoms work more efficiently than Xeons
- **Therefore, why not build the two together?**

Courtesy of Giovanni Leonardi (Mataró-Sicilia)

Hybrid hardware

- **Better tradeoff in terms of energy and revenue**
  - Experiments at BSC with heterogeneous workload:



Nodes	Work/ON	Power (kW)	DL (%)	R (%)	B (€)
Xeon	25.1 / 27.1	2960.2	59.8	84.2	816.5
Heterogeneous	38.1 / 40.1	2053.7	52.2	83.7	944.7
Atom	52.7 / 54.7	849.9	0.0	83.6	709.6

Cloud Federation

- **Resource Management in a Cloud Federation**
  - Distribution of workload across geographically distributed DC
  - Consolidation at DC level (and turn off a whole DC)
  - Possibility to reallocate the workload to a place where energy or cooling is cheaper
    - e.g. solar energy during daytime across different time zones, efficient cooling due to climate conditions, etc.
  - Etc.

The energy savings can be up to 10%? 30? 50%?

Conclusions

The energy savings can be up to 20% - 70%

Reduction of 25%

Reduction of 70%

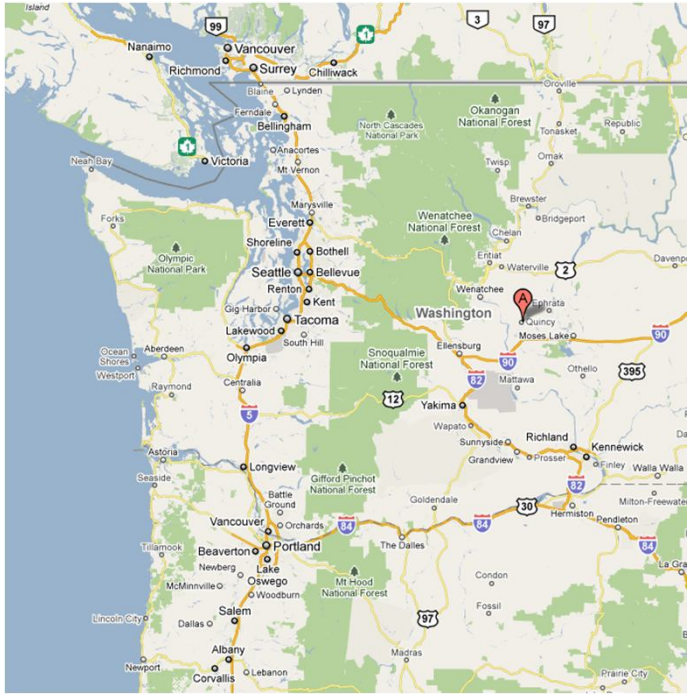
For memory-bound appl. the energy savings can be up to 25%

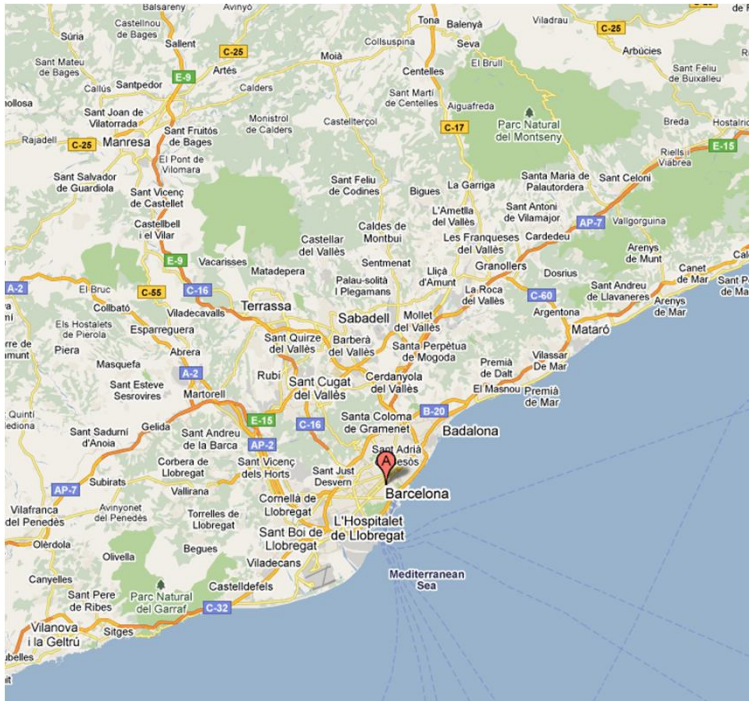
The energy savings can be up to 10%? 30? 50%?

Energy cost depends on location

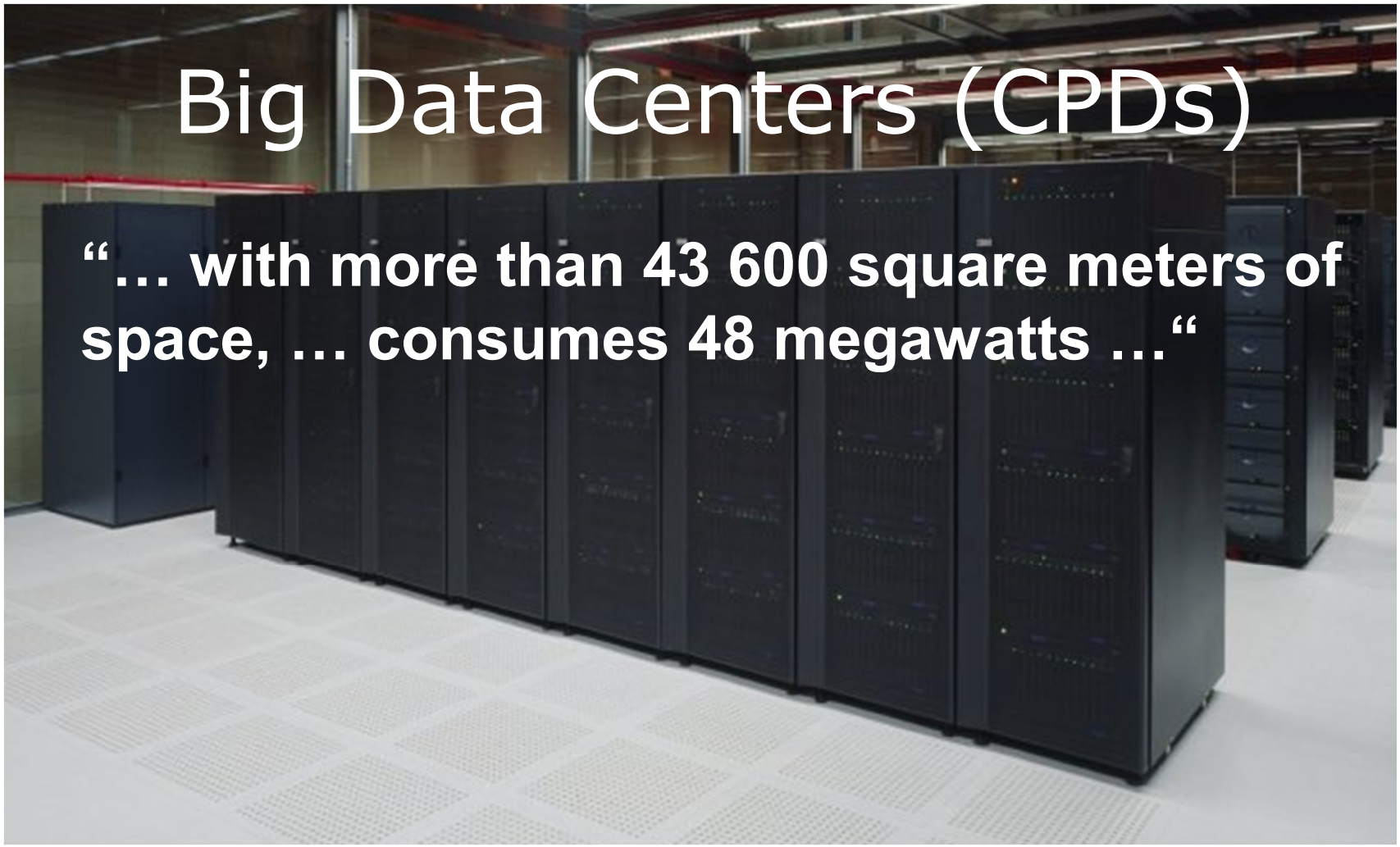
- **Examples:**
  - Quincy (Washington, USA) → ~ 0,015€ (0.019$) kW/h
  - Barcelona → ~ **0,12** € kW/h

  → It is 8 times less





What is the size of a DC?

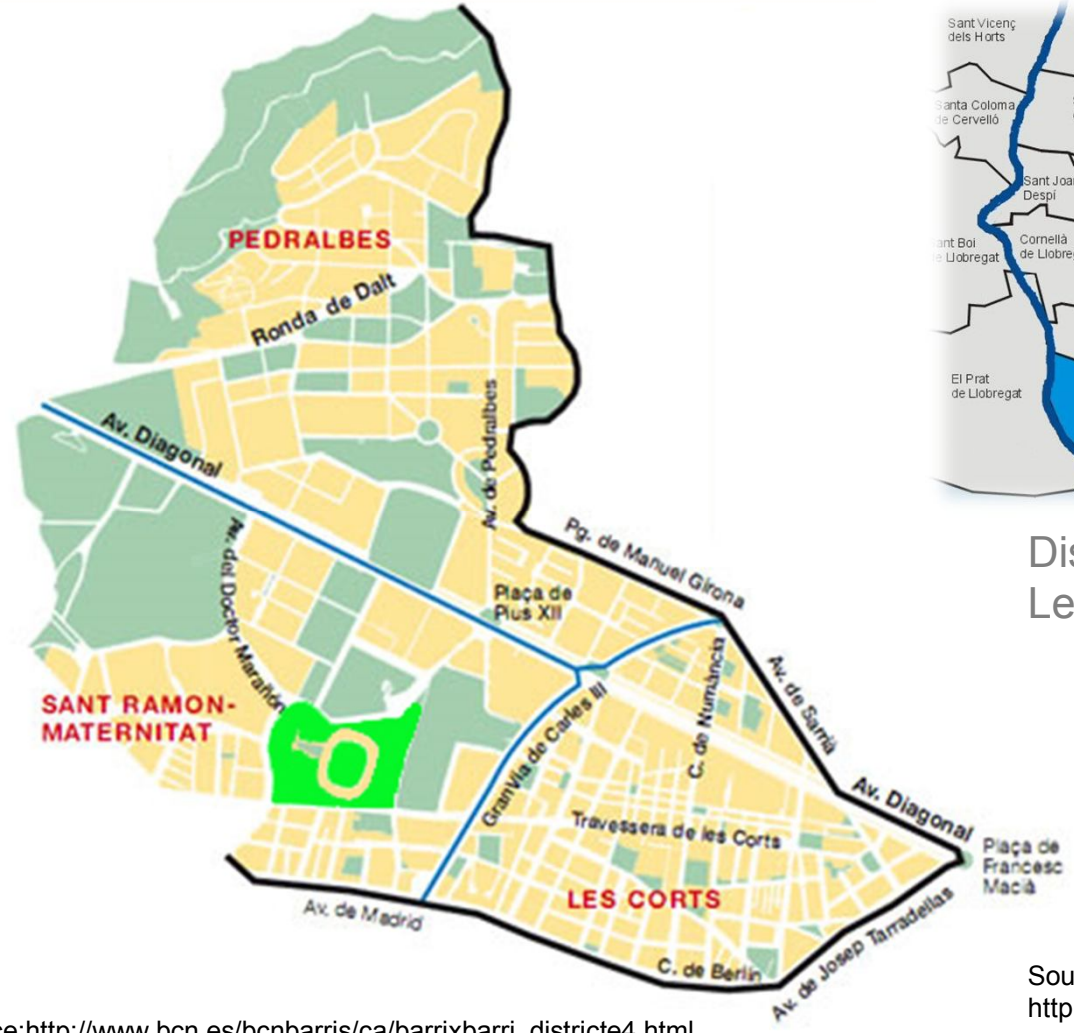


Source: *Tech Titans Building Boom By Randy H. Katz. IEEE Spectrum, February 2009*

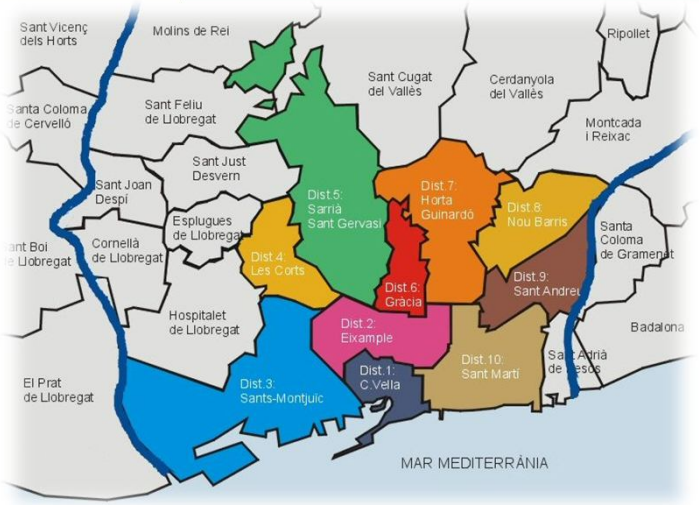
Does anyone know what this means?

foto: Google

Does anyone know what this means?



Source:http://www.bcn.es/bcnbarris/ca/barrixbarri_districte4.html



Districts of Barcelona,
Les Corts coloured yellow.

Source:
http://en.wikipedia.org/wiki/Les_Corts,_Barcelona

Summary: what are providers doing to provide cheaper services?

- **unprecedented economies of scale reduce overall cost and increase efficiencies**
  - Commodity hardware (servers, networks, storage)
- **All servers virtualized with many concurrent loads**
  - Achieves high average utilization
- **Standardization and automation (self-managed)**
  - Minimize labor-policy, not human decide on workload placements
- **Good location for cooling and energy**
- **Etc.**

Outline

Providers side: The industrialization of ITs

Client side: The Demand for Cloud Services

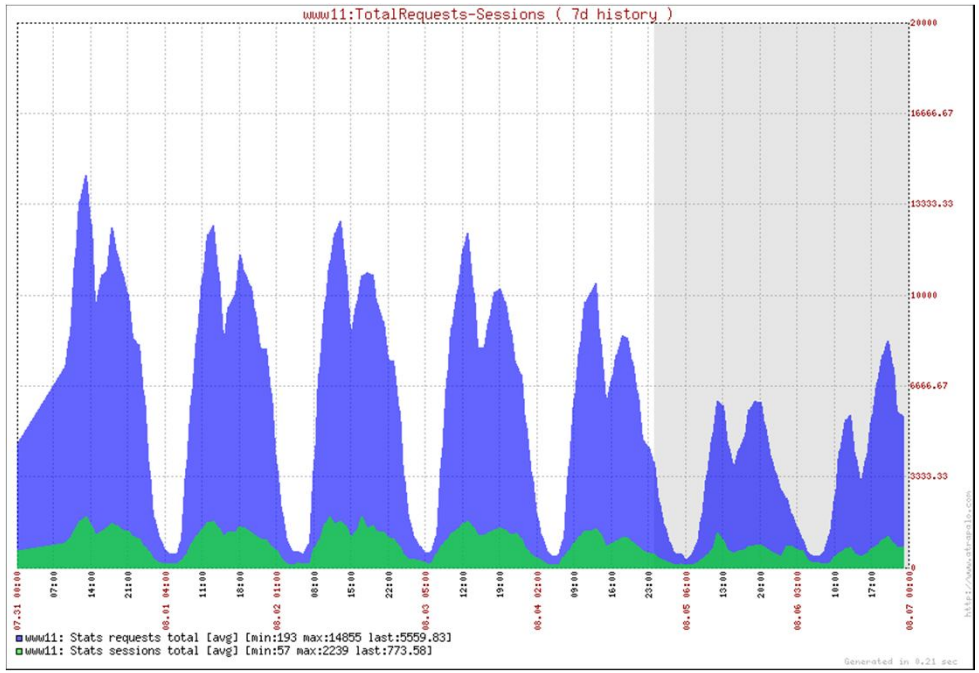
Cloud Opportunities: Green Computing

Reasons for adoption

- **The business community has begun to embrace cloud computing as a viable option**
  - to reduce costs (opex & capex)
  - Scalability
  - business agility

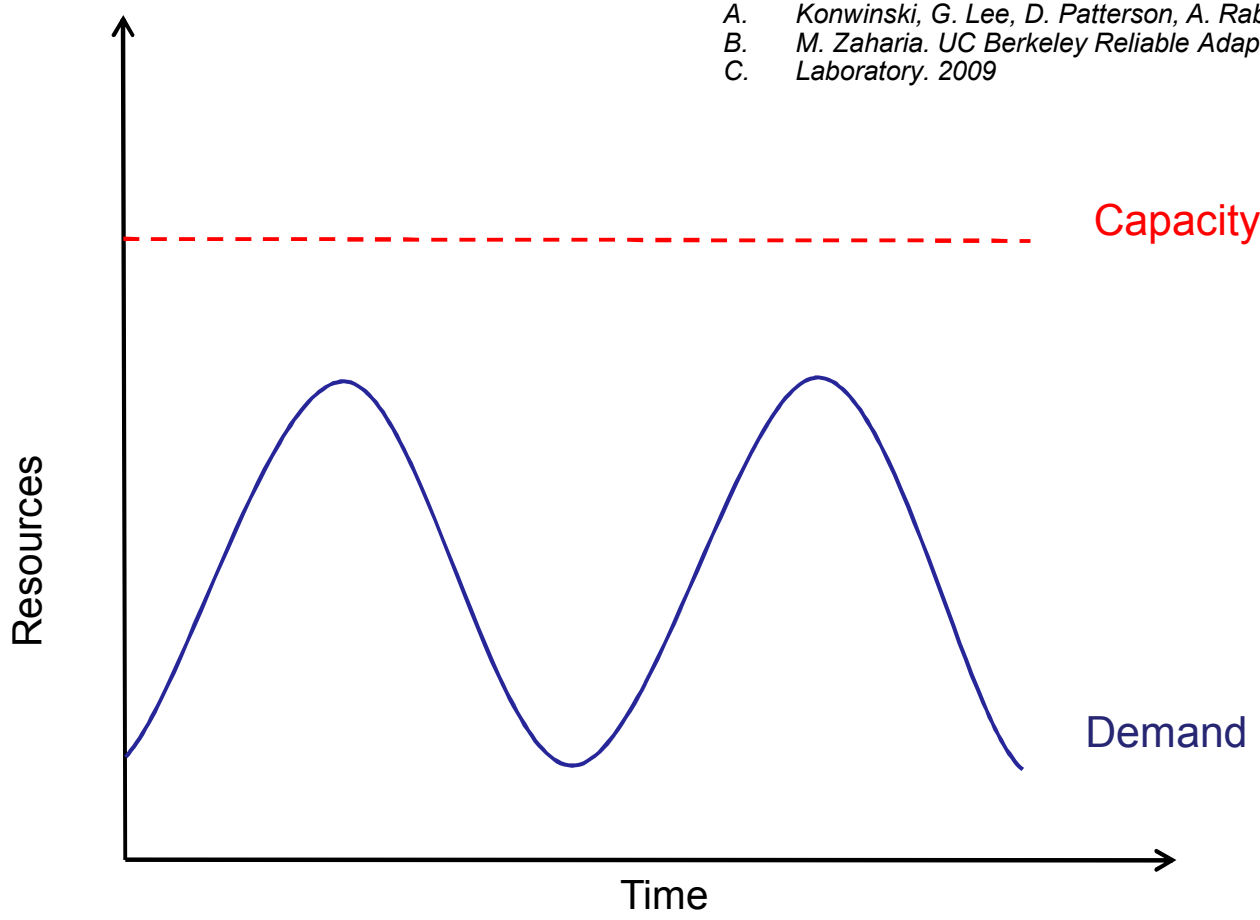
- **Target Exemple:**

  (website weekendly workload)

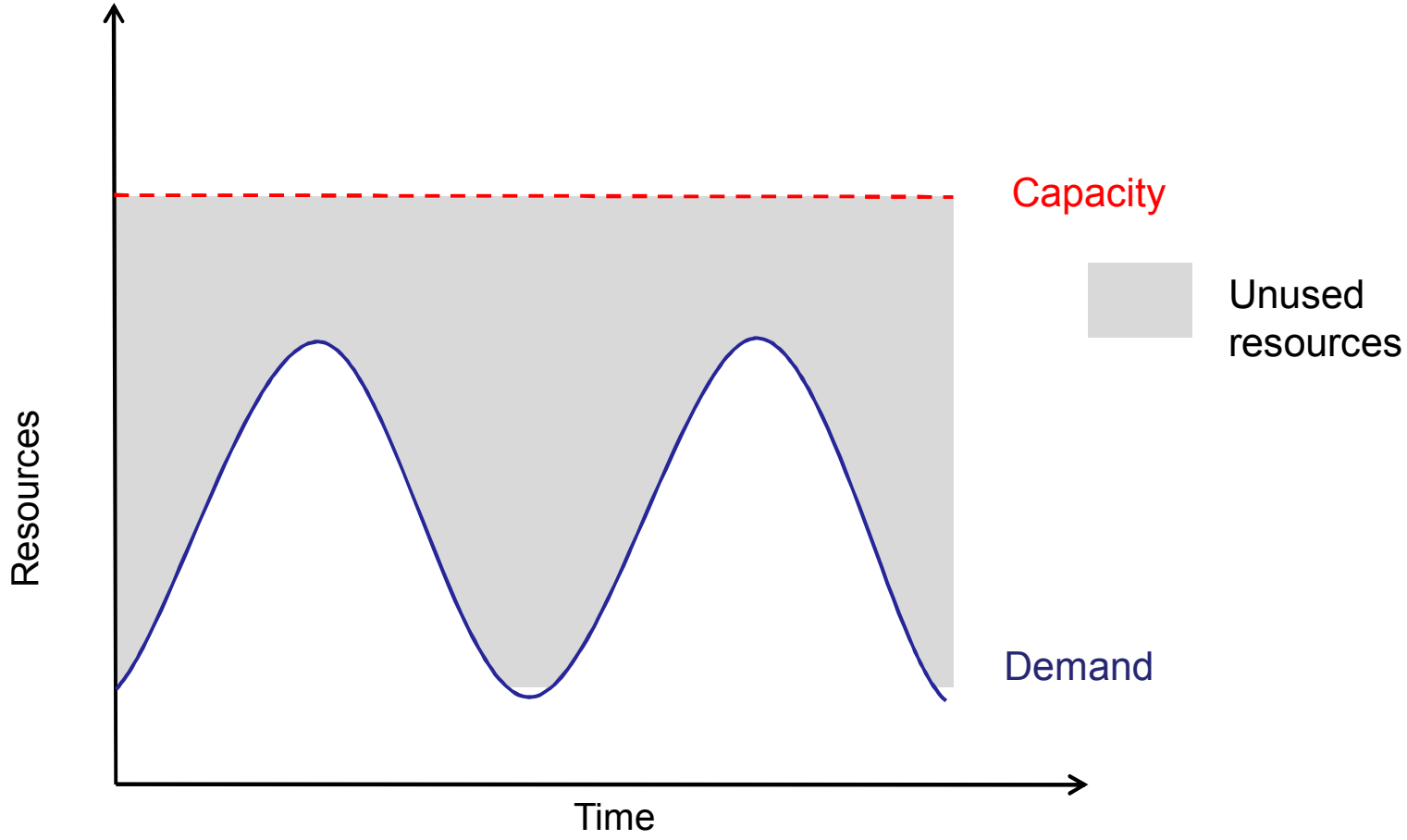


Capacity versys Demand

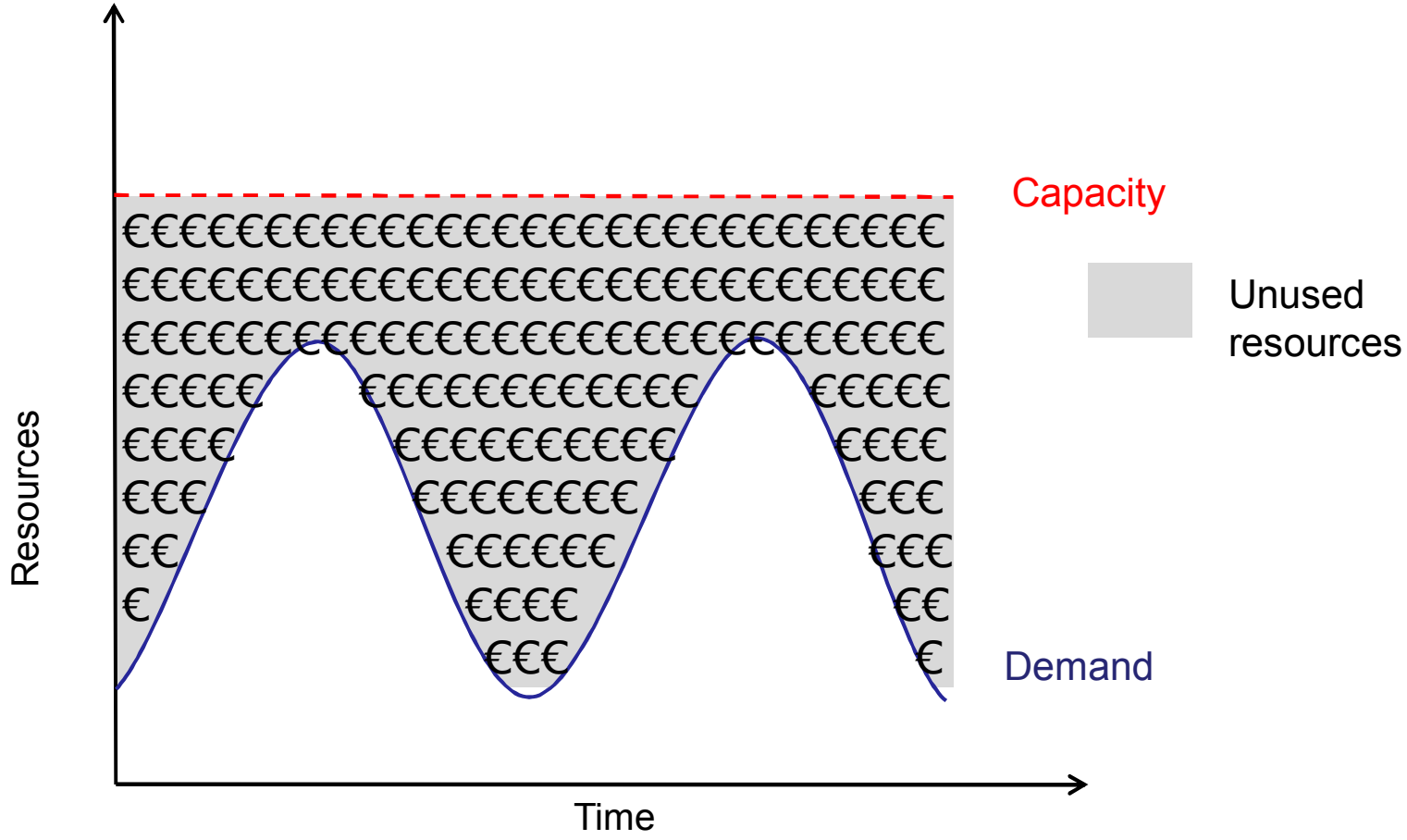
Source: Above the Clouds: A Berkeley View of Cloud Computing.
M. Ambrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz,
A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stooica,
B. M. Zaharia. UC Berkeley Reliable Adaptive Distributed Systems
C. Laboratory. 2009



Lost of resources

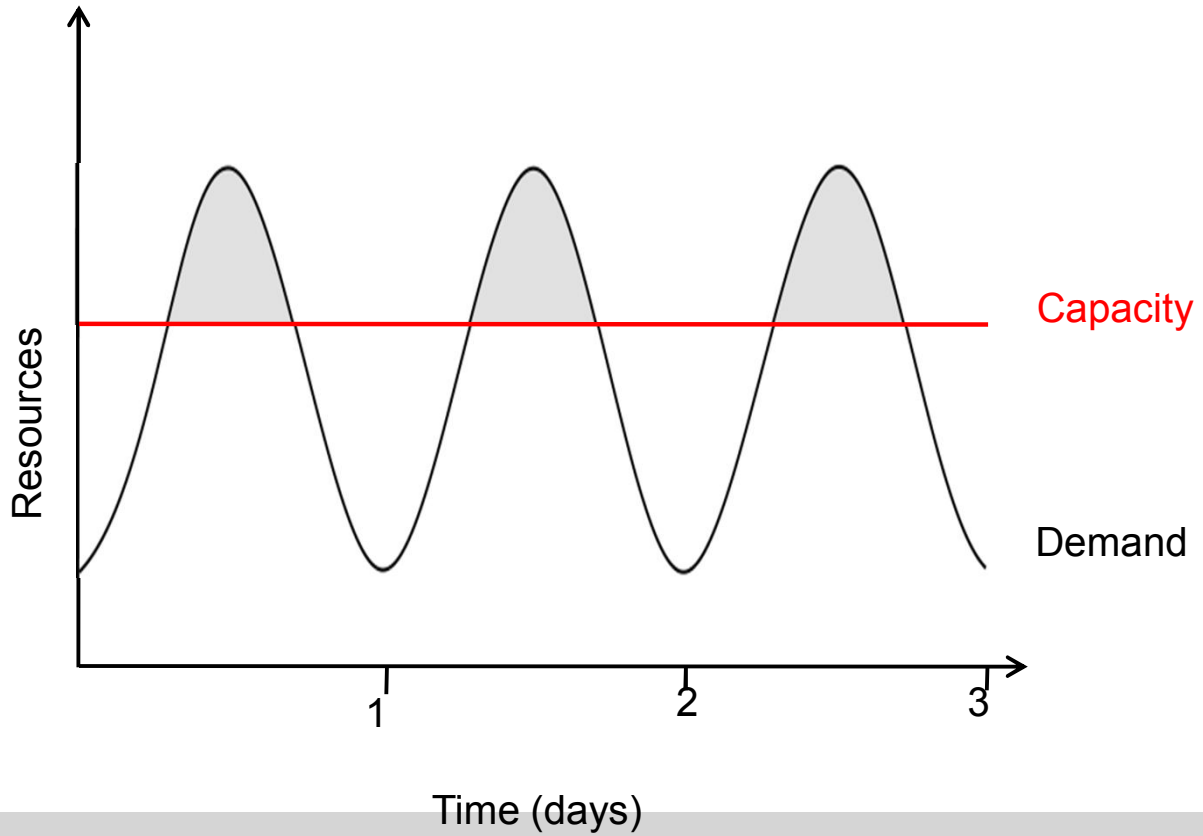


Lost of resources



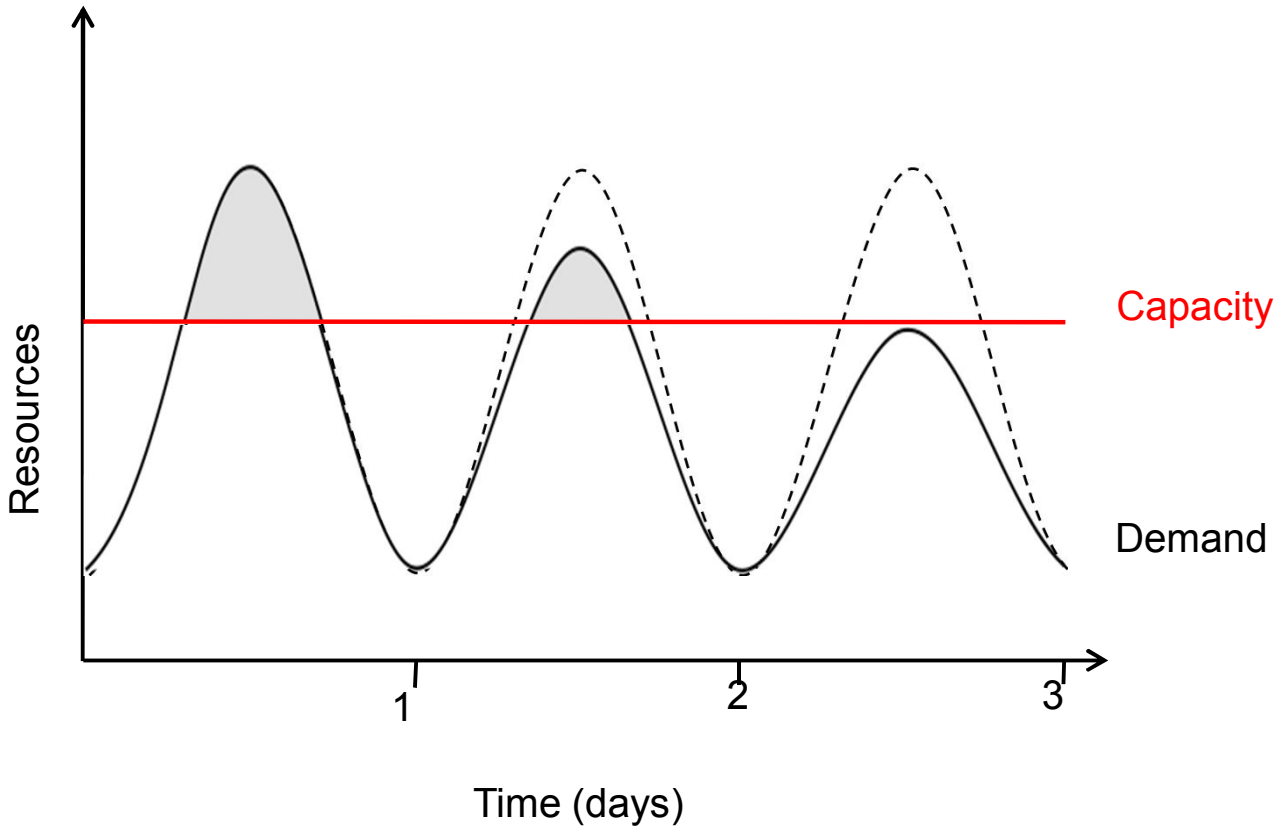
Can we make a careful approach?

- **Underestimation is even worse, user dissatisfaction is difficult to measure.**

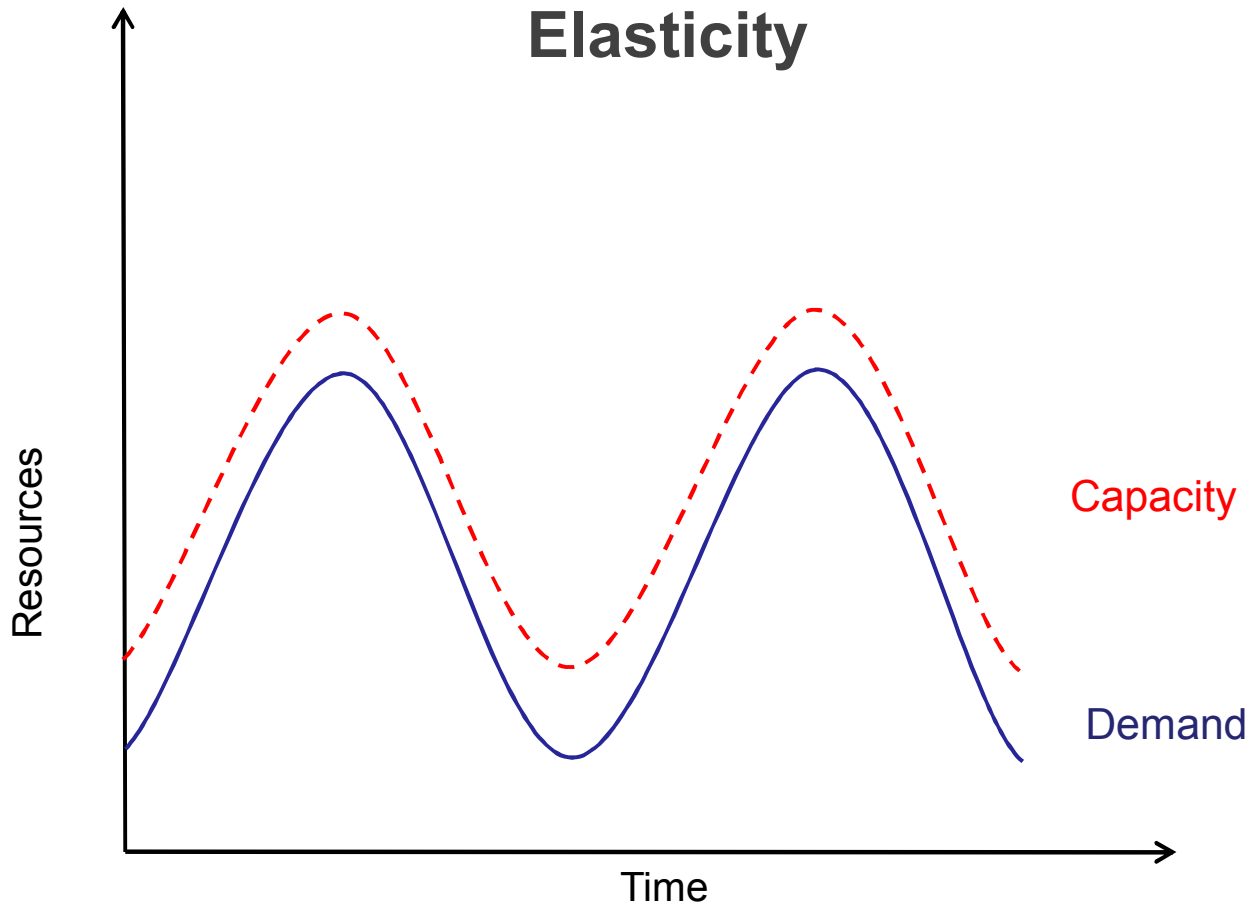


The demand self-adjusts to the supply

- **Clients are lost**



The Cloud allows …

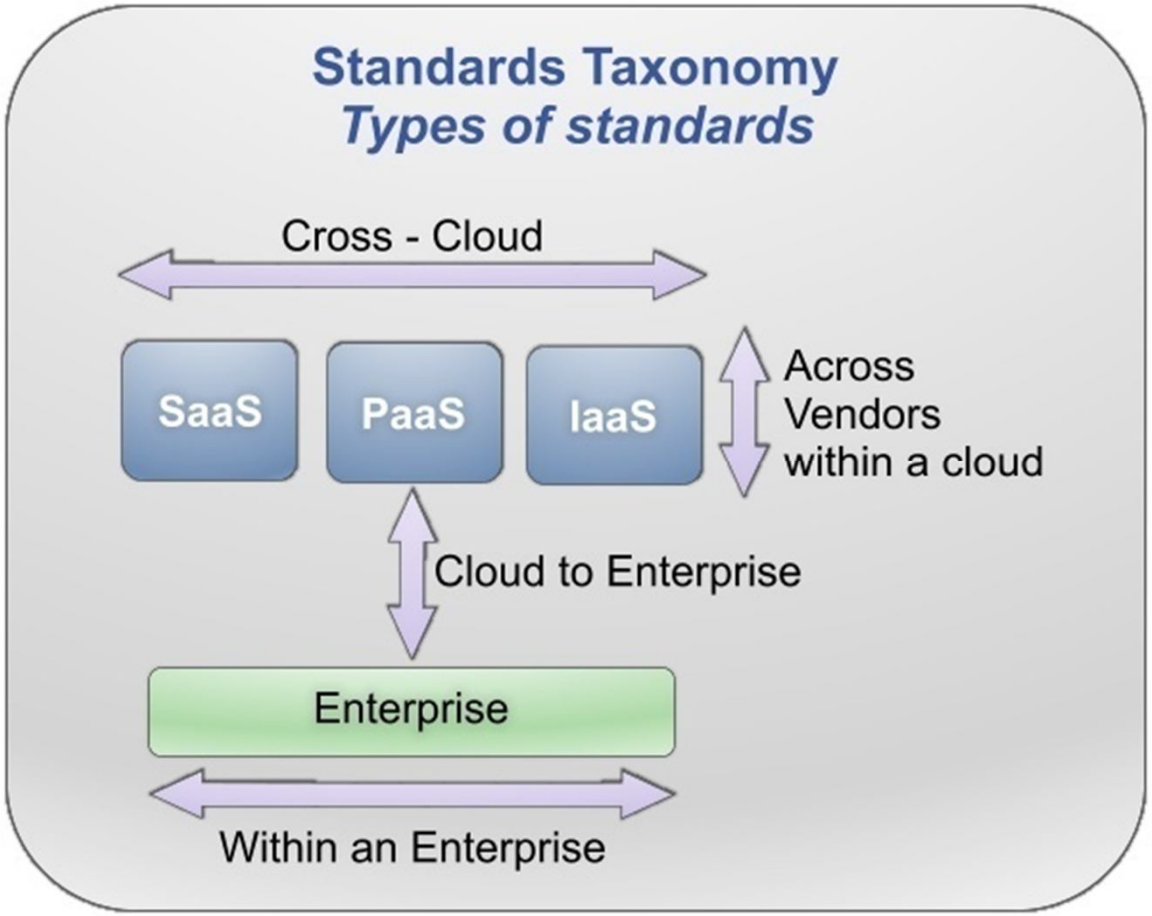


Reasons for adoption

- **“Hours to provision versus weeks**
  - While most organizations adopt cloud computing strategies for **cost reasons**, the biggest reported benefits turn out to be **agility**

Difficulties in adopting the Cloud

- **Lack of standards:**



Source: http://groups.google.com/group/cloud-computing-use-cases

Risks of Cloud Computing

- **Availability**
- **Privacy**
- **Legislation**
- **Data theft and Loss**
- **…**

→Vendors are addressing these by offering:

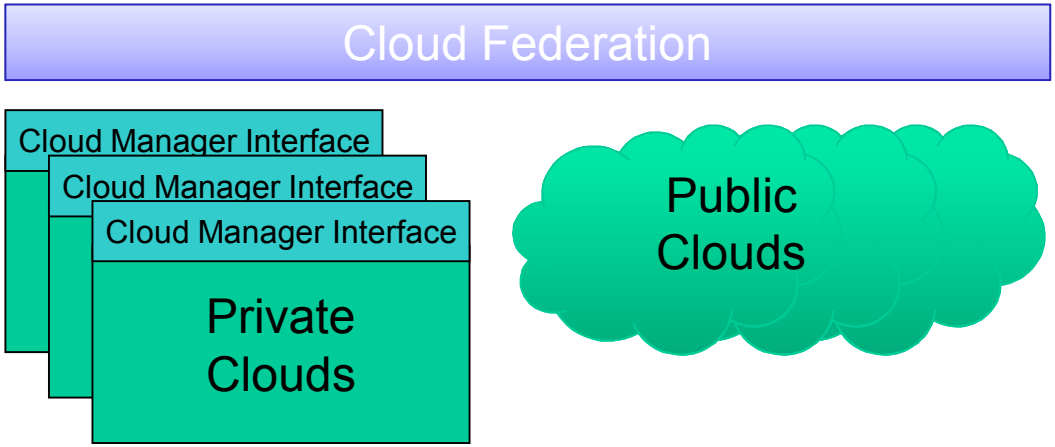
- In continent / in country hosting
- **Private clouds** 
- Secured IaaS (e.g. using VPN)
- …

→Customers are addressing these risks by:

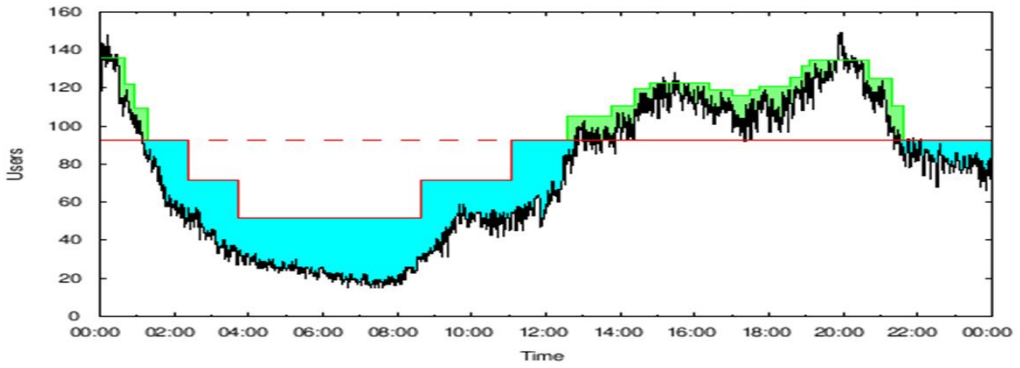
- including cloud computing in backup, recovery and data protection (encryption).

Next: Federació de Clouds

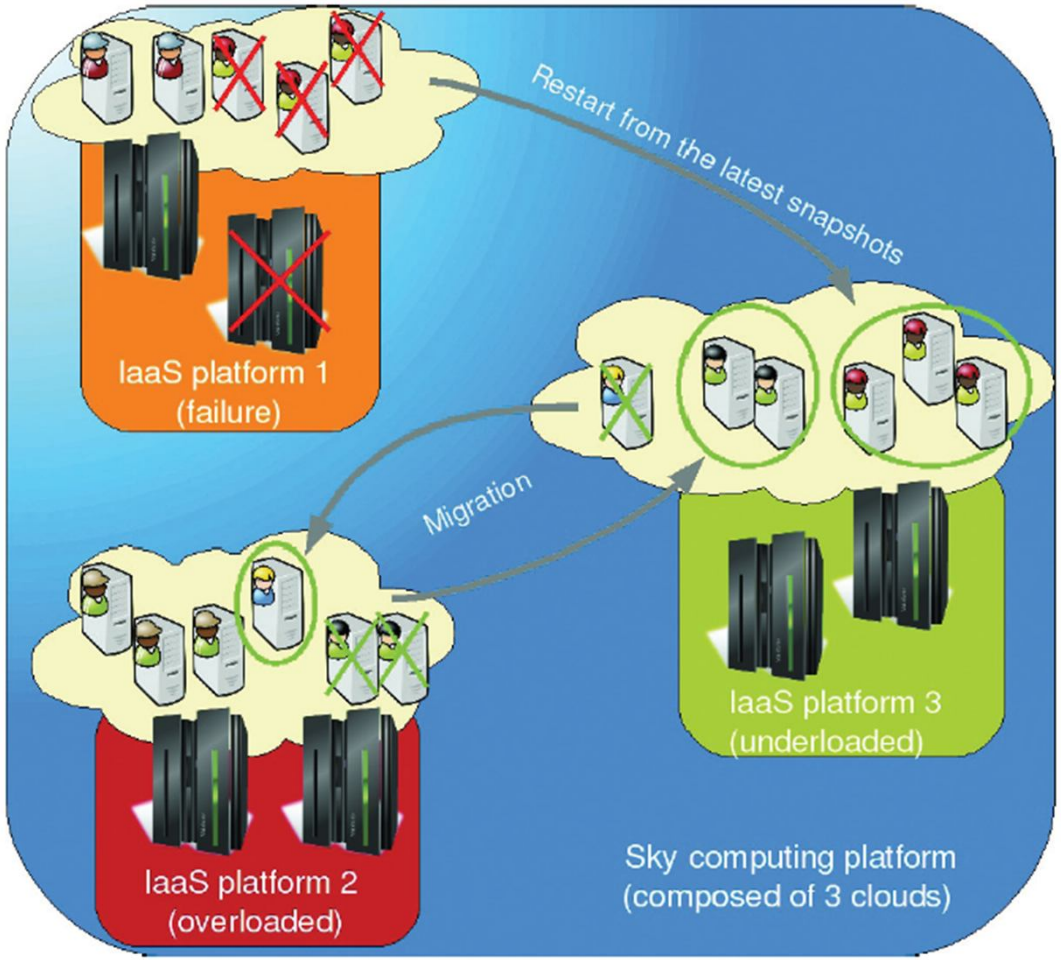
- **Federació de Clouds interns+intern, intern+externs, extern+extern.**



- **Els proveïdors poden fer outsource I insource dde recursos dins de la federació**



Next: Federació de Clouds



Source: "Managing Virtual Resources: Fly through the Sky", ERCIM News, 83

Outline

Provides side: The industrialization of ITs

Client side: The Demand for Cloud Services

Cloud Opportunities: Green Computing

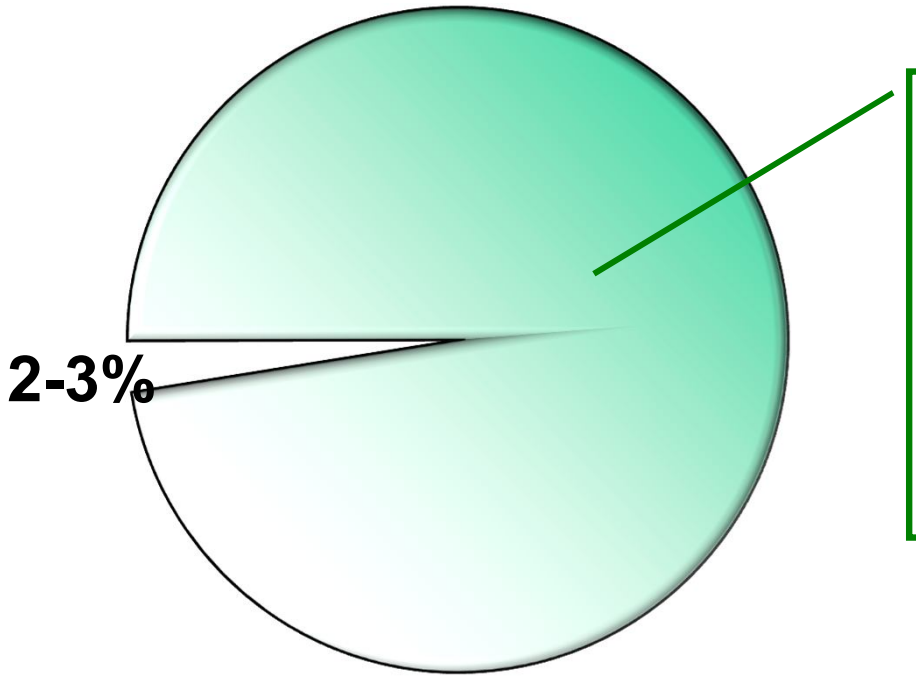
Impact of the computing on to the environment:

Foto: J.T.

ICT industry, which includes the Internet, produces roughly 2 to 3 percent of global GHG emissions.

Source: Climate Group's Smart2020 report (www.smart2020.org)"

ICT reduces emissions of other activites



ICT can significantly contribute to control and reduce the 98% of emissions caused by other activities and industries.

ICT has already achieved the increase in business efficiency

Source**:** (1) http://www.archives.gov/research/american-cities/images/american-cities-040.jpg
(2) http://www.noticiasdealava.com/ediciones/2007/01/16/economia/alava/fotos/3014023.jpg

Do you have these at home?

Or …

- Do you read newspapers from the newsagents?

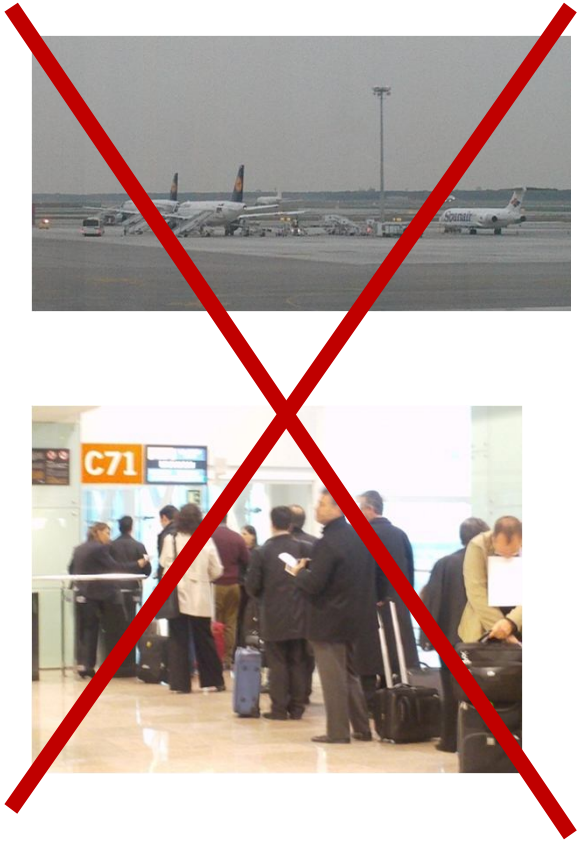
ICT reduces waste of resources

- Enable dematerialization
- …

ICT reduces waste of resources

ICT reduces waste of resources

- Reduce the need to travel



Courtesy: Dani Urgell – CISCO

The future is digital

- **world is becoming instrumented**

  Sensors are being embedded across entire ecosystems

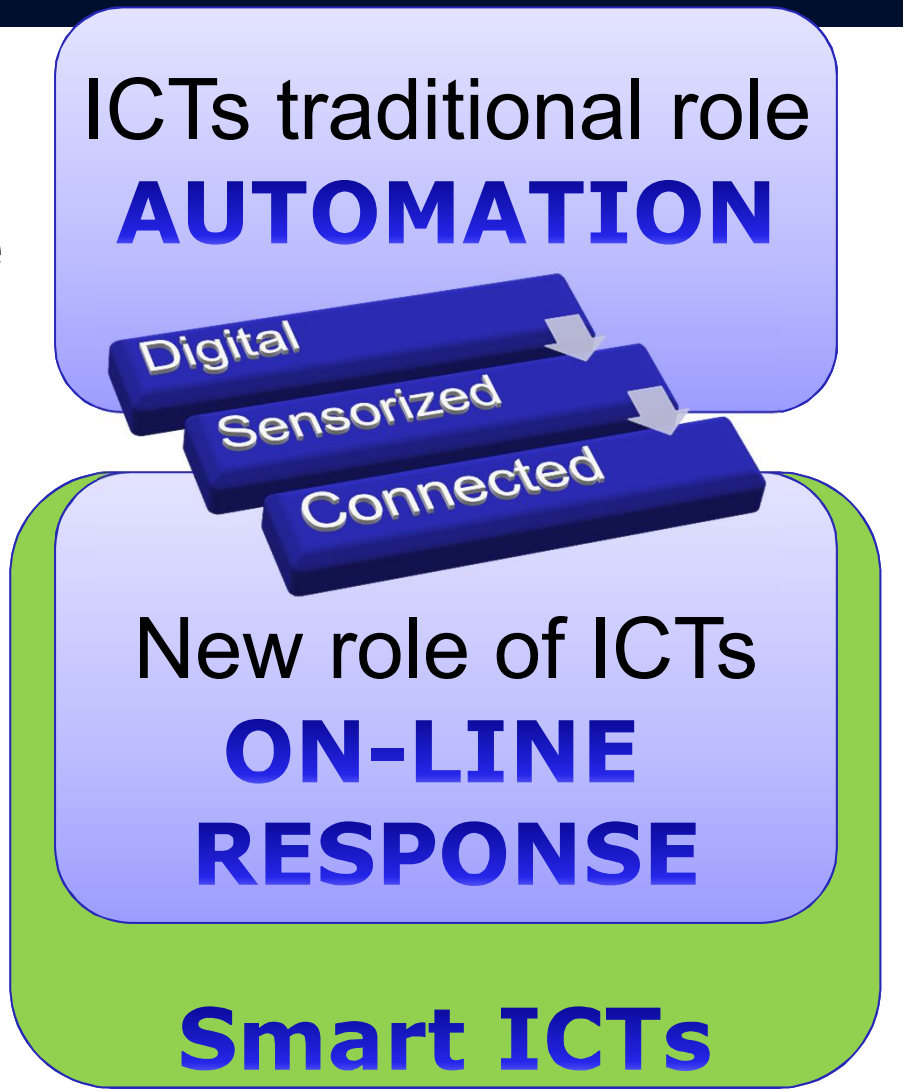
  Mobile phones, cameras, cars, shipping containers, intelligent appliances, RFID tags by the hundreds of millions

- **All becoming interconnected.**

The role of ICTs

- The resulting volume of data promises insight and intelligence to solve some of our biggest problems

- **but only if we can process and make sense of it in real time.**



Future sustainability thanks to smart ICTs

- **smart ICTs?**

DEBATE. Las TIC, una nueva oportunidad / **Jordi Torres**

Tecnología 'smart' para una Europa sostenible

Europa está preocupada por su crecimiento sostenible. Los atascos de tráfico en las principales ciudades europeas tienen un coste estimado de 135 millones de euros al año, lo cual supone que los conductores pierdan un total de cinco días al año al volante. El 40% de la energía producida se desperdicia hasta llegar al consumidor por la ineficiencia en la gestión. Estamos malgastando energía y emitiendo a la atmósfera dióxido de carbono (CO_2), lo que dificulta que Europa pueda cumplir con sus ambiciosos objetivos para luchar contra el cambio climático.

El diagnóstico es claro y la solución no pasa sólo por aumentar las infraestructuras, sino por un uso mucho más eficiente de estas. Por ello, la UE apoya firmemente las nuevas tecnologías capaces de mejorar la eficiencia energética y hacer que, en general, los sistemas de transporte sean más inteligentes o que los edificios consuman menos recursos de energía, por ejemplo. Y por ello, las tecnologías de la información y comunicación (TIC) son la piedra angular para la Europa del futuro, basada en una economía de alta eficiencia energética y bajos índices de emisiones de carbono. Las emisiones de CO_2 derivadas de las TIC supondrán un 2,8% del total de emisiones globales para el 2020. Estas mismas TIC ayudarán a evitar un 15% de las emisiones globales procedentes de otra gran parte de la industria. Un balance extremadamente positivo a favor de las TIC.

Numerosos ejemplos evidencian el potencial de las denominadas *smart* TIC o TIC inteligentes para aportar soluciones a la sostenibilidad europea, y las instituciones europeas así lo han reconocido. Pero queda pendiente un problema fundamental previo: el déficit de estudiantes en ingenierías TIC. Los gobiernos deben abordar también esta problemática y apoyar decididamente a las universidades técnicas, si realmente desean alcanzar la sostenibilidad a través de innovaciones tecnológicas.

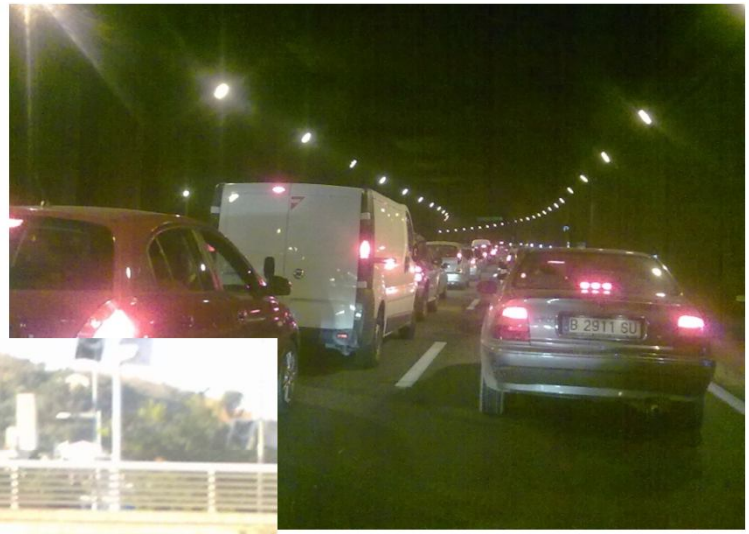
¿Y nosotros, la ciudadanía, qué podemos hacer? Si usted conoce a un o una joven que está sopesando elegir una carrera centrada en estudios técnicos, puede, cuando menos, alentarle a ello, puesto que no sólo le permitirá estudiar una carrera con un gran futuro, sino que, además, contribuirá indudablemente a labrar una Europa más sostenible.●

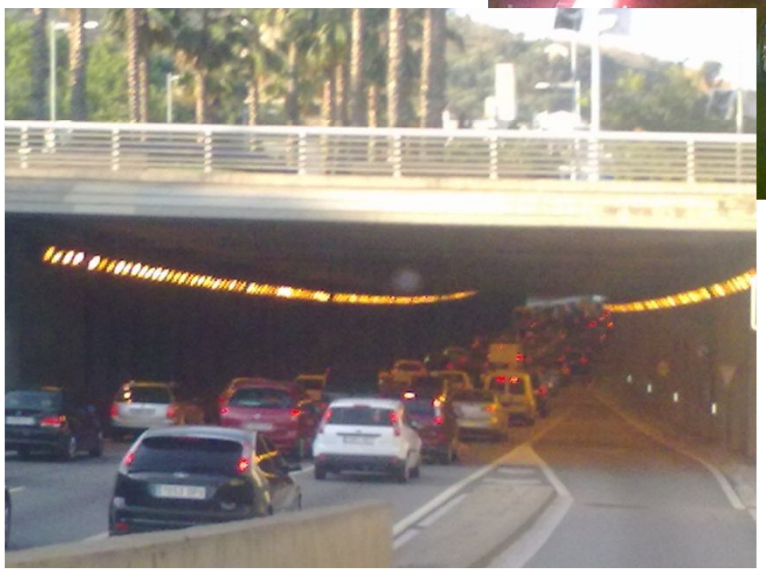
J. TORRES, *catedrático UPC, Barcelona Supercomputing Center-Centre Nacional Supercomputació*

Green and Sustainable EU Growth

- **Traffic jams**

  cost Europe €135 billion a year, and drivers lose five days per year while sitting in traffic.





Green and Sustainable EU Growth

- **Smart Transport Systems**
  - Making roads and cars “smarter” with intelligent transport systems (ITS) such as
    - sensor networks,
    - RF tags,
    - positioning systems
    - …
  - The Internet can interconnect diverse technologies and make mobility more efficient through the real-time management of public and private transport resources.

Foto: David Carrera

Green and Sustainable EU Growth

- **Smart Energy Grids**
  - Today, up to 40 percent of the energy produced might be lost on its way to the consumer,
  - but Internet connectivity, computing power, digital sensors, and remote control of the transmission and distribution system will help make grids smarter, greener, and more efficient.

Foto: J.T.

Foto: SMART 2020 Report

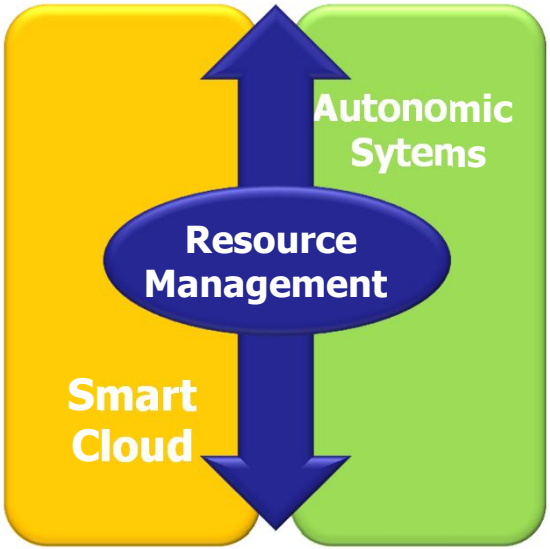
Autonomic Systems and eBusiness Platforms research line

Research objective

Autonomic resource allocation and heterogeneous workload management

for Internet-scale virtualized data centers (cloud computing)

comprising heterogeneous clusters of hybrid hardware.



Thank you for your attention

www.JordiTorres.org

facebook.com/jordi.torres.vinyals

Jordi.Torres@bsc.es

http://es.linkedin.com/in/jorditorresvinyals

JordiTorresBCN

About the author:

Jordi Torres has a Master degree in Computer Science from the Technical University of Catalonia (UPC Barcelona Tech, 1988) and also holds a Ph. D. from the same institution (Best UPC Computer Science Thesis Award, 1993).

Currently he is a full-time professor in the Computer Architecture Department at UPC and is a manager for the *Autonomic Systems and eBusiness Platforms* research line at **Barcelona Supercomputing Center** (BSC).

His current principal interest as a researcher is to make IT resources more efficient in order to obtain more sustainable IT and to focus on the resource management needs of modern distributed and parallel cloud computing environments. He has worked on a number of EU and industrial research and development projects. He has more than a hundred publications in the area and was involved in several conferences in the field of distributed and parallel systems.

He has been Vice-dean of Institutional Relations at the Computer Science School, and currently he is a member of the University Senate and member of the Board of Governors.

He can be reached at **www.JordiTorres.org**