Escuela de Ingeniería y Arquitectura
Universidad Zaragoza

Departamento de Informática e Ingeniería de Sistemas
Universidad Zaragoza

Viaje al centro de los datos: ingeniería, arquitectura y ecología

Máster en Ingeniería Informática

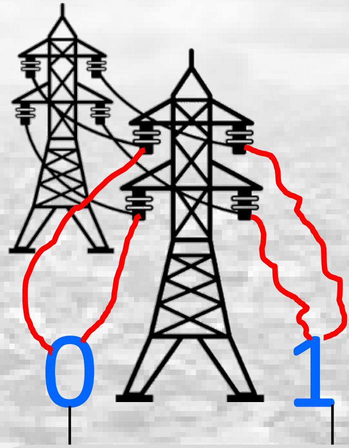
Facultad de Informática. UCM

Madrid, 1 marzo 2018

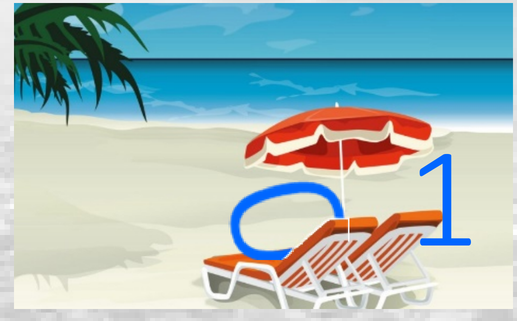
Víctor Viñals Yúfera

Departamento de Informática e Ingeniería de Sistemas
Universidad Zaragoza

instituto de investigación en ingeniería de Aragón
Universidad de Zaragoza



Ing. Eléctrica
Ing. Electrónica



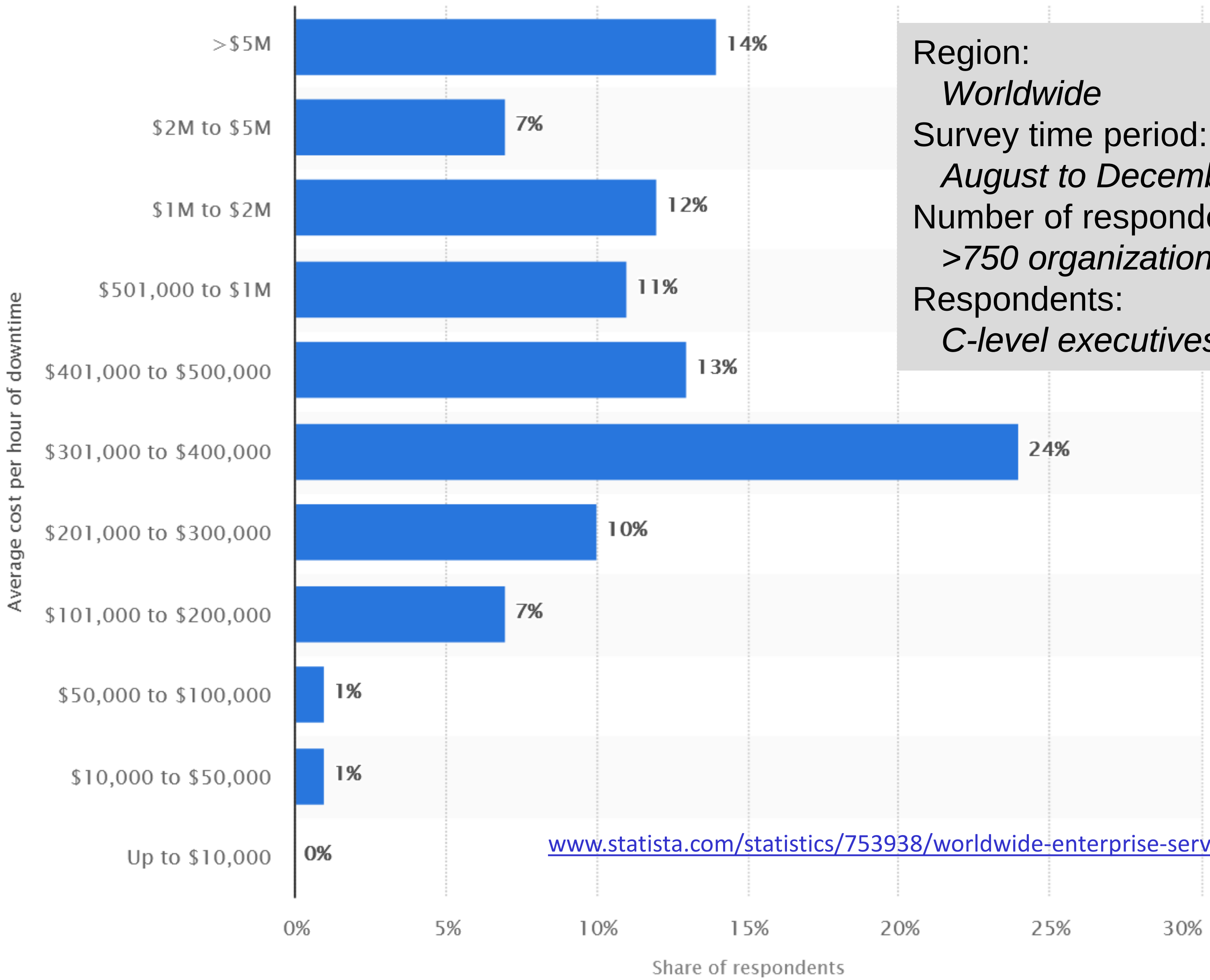
Ing. Mecánica

Arquitectura

Ing. Informática
Ing. Telecomunicación

https://backdropoutlet.com

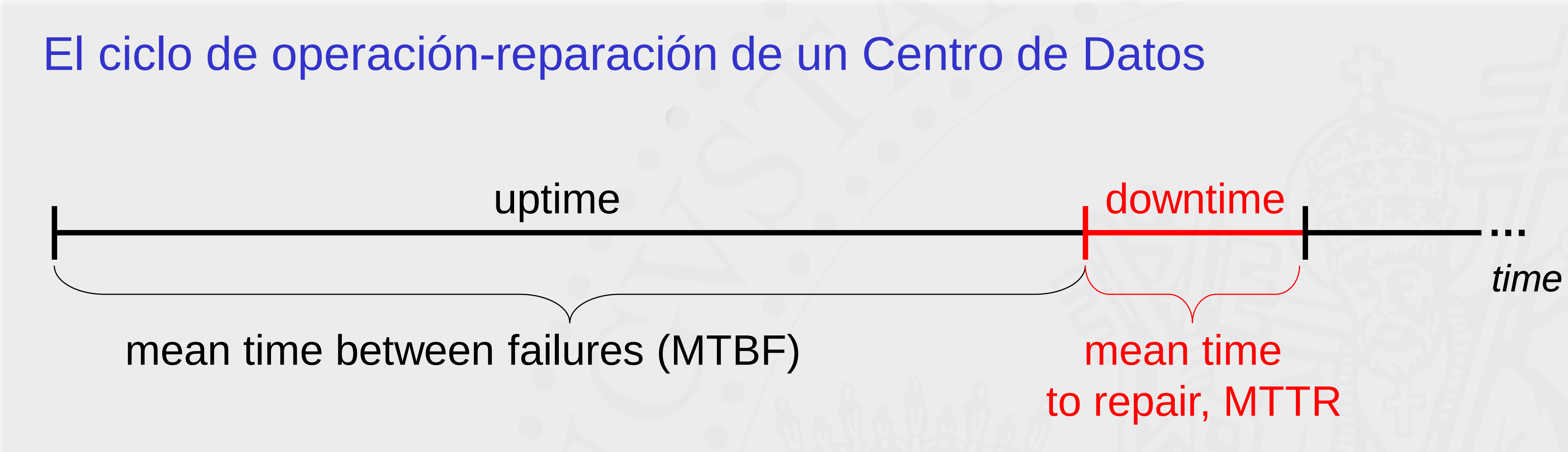
Average cost per hour of enterprise server downtime



Region:
Worldwide
Survey time period:
August to December 2017
Number of respondents:
>750 organizations
Respondents:
C-level executives and IT managers

www.statista.com/statistics/753938/worldwide-enterprise-server-hourly-downtime-cost/

Disponibilidad



- Disponibilidad (%1) = MTBF / (MTBF + MTTR[1])

1. In a service-oriented analysis, what matters is the MTTR (mean time to recover the service). ITIL, for instance, breaks down the MTTR in several sequential times: detection, diagnosis, physical repair and logical recovery.

Cuatro categorías de disponibilidad

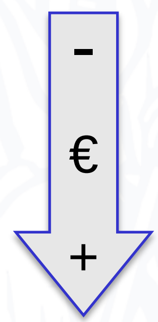
- Categoría = tier, class, level, …

Uptime Institute	tier I	.. tier IV	white papers	www.uptimeinstitute.com
ANSI / TIA 942	tier 1	.. tier 4	compra en www.tiaonline.org	
Bicsi	class F0	.. class F4	compra en www.bicsi.org	
IBM	level 1	.. level 4	público	
TÜV IT	level 1	.. level 4	público ?, europeo[1]	

…

Categorías

1. Rutas únicas sin redundancia
2. Rutas únicas con algunos componentes redundantes
3. Mantenimiento concurrente (dos caminos)
4. Mantenimiento concurrente (dos caminos) con plena tolerancia a los fallos

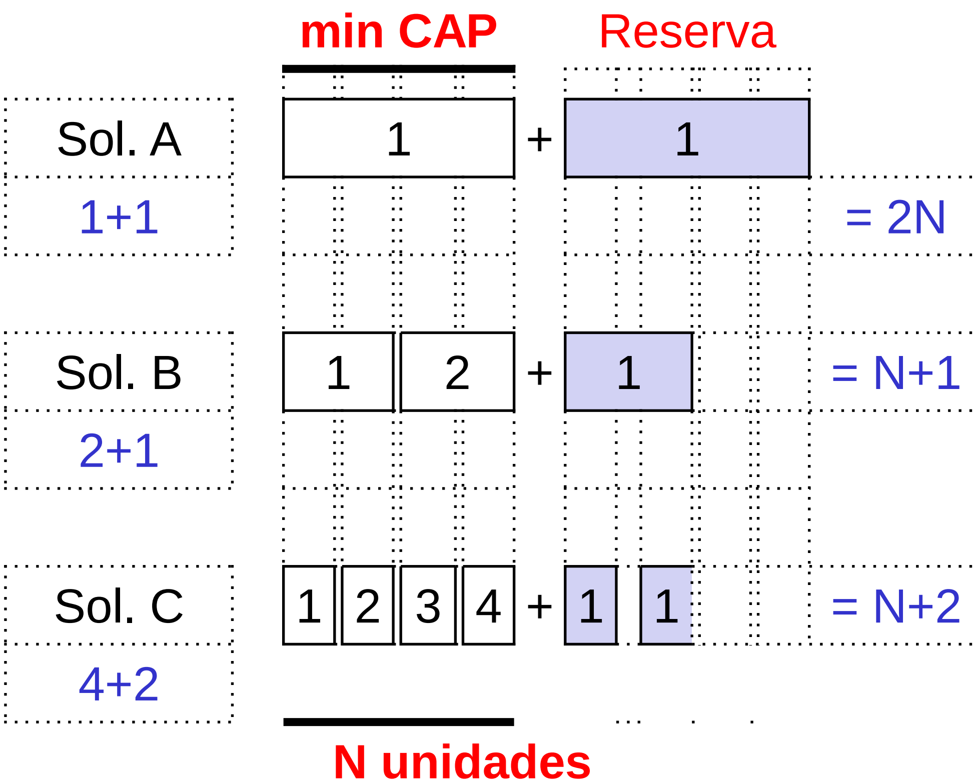


Mantenimiento concurrente

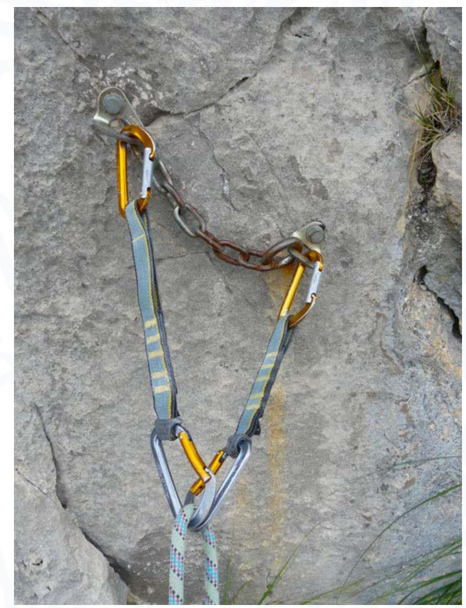
- También llamado *mantenimiento simultáneo*
- La instalación admite actuaciones planificadas relativas a
  - mantenimiento,
  - prueba o
  - ampliación

  … sin cortes de servicio
- Es posible gracias a la redundancia

“Esquema” de Redundancia: 2N, N+1, N+2, …



- El fallo de una unidad de capacidad (+1) es transparente
  - la operación del sistema no se ve alterada -
- Cada solución es diferente:
  - Coste
  - Escalabilidad
  - Eficiencia
  - Mantenimiento



No siempre es fácil determinar la redundancia →

Disponibilidad anual: desde 2 a 6 “nueves”

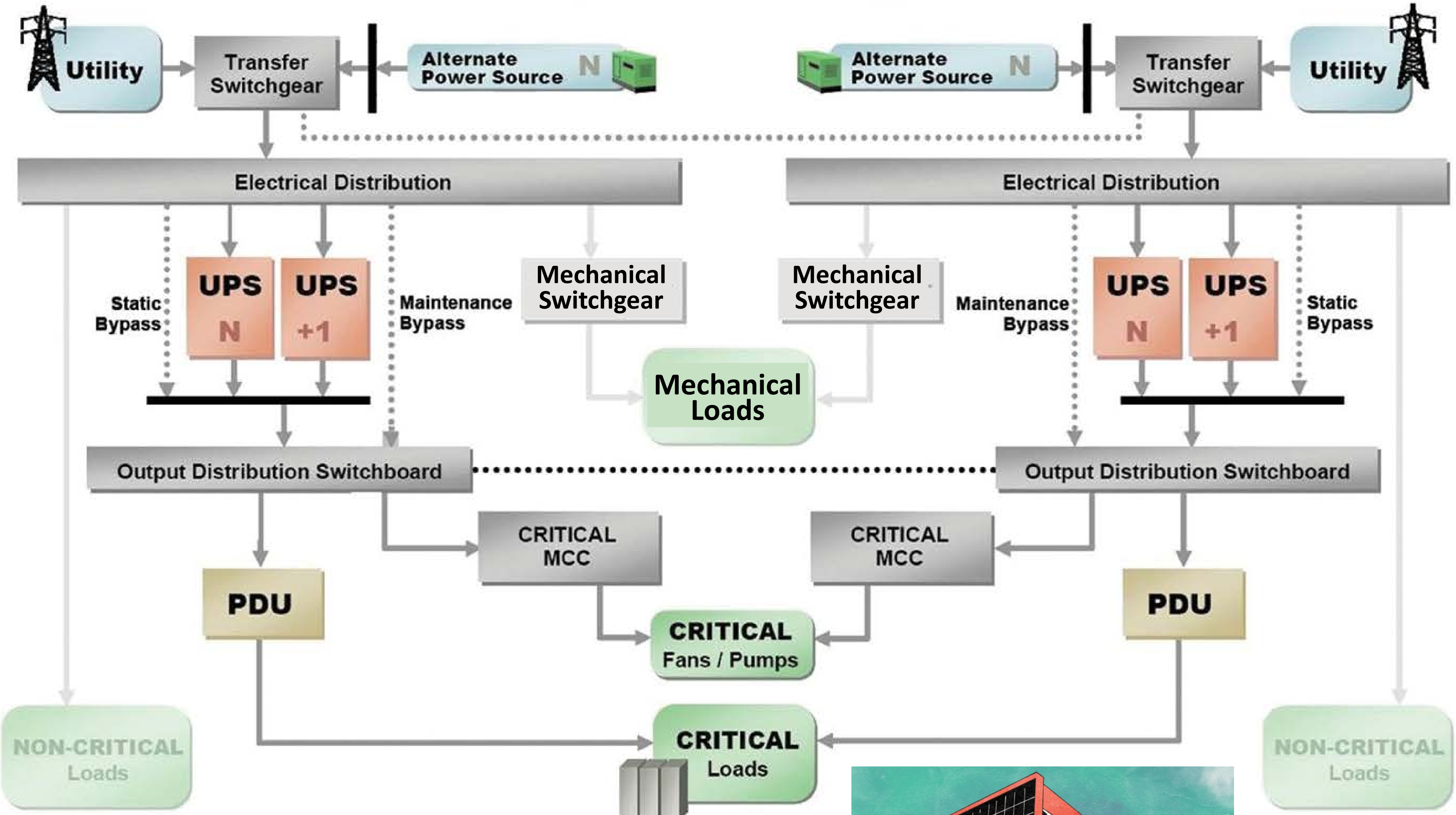
% disponible	Corte anual de servicio	Ejemplo: Uptime Institute[1]	Ejemplos de Negocio
99%	87,6 h	Tier I 99,67% 28,9 h Tier II 99,75% 21,9 h	5,5 días/semana: universidades, ministerios
99,9%	8,76 h	Tier III 99,98% 1,8 h	Coste del corte muy elevado
99,99%	52,56 min	Tier IV 99,995% 26,3 min	Si se cae, fin del bussines …
99,999%	5,26 min		red de cajeros (en teoría !)
99,9999%	31,54 s		Aeroespacial

- El corte de servicio puede ser por error humano (70%), fallo de infraestructura o mantenimiento programado
- No se incluyen desastres naturales ni botones rojos

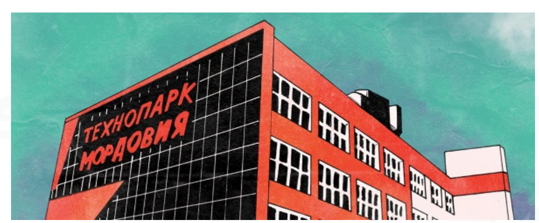
XKCD en español

ENLACE PERMANENTE A ESTA TIRA: HTTP://ES.XKCD.COM/STRIPS/DEVOCION-AL-DEBER/
URL DE LA IMAGEN: HTTP://ES.XKCD.COM/SITE_MEDIA/STRIPS/DEVOTION_TO_DUTY.PNG

Tier IV: mantenimiento concurrente y tolerancia a fallos



Technopark-Mordovia Data Center
Saransk, Mordovia Republic, Russia

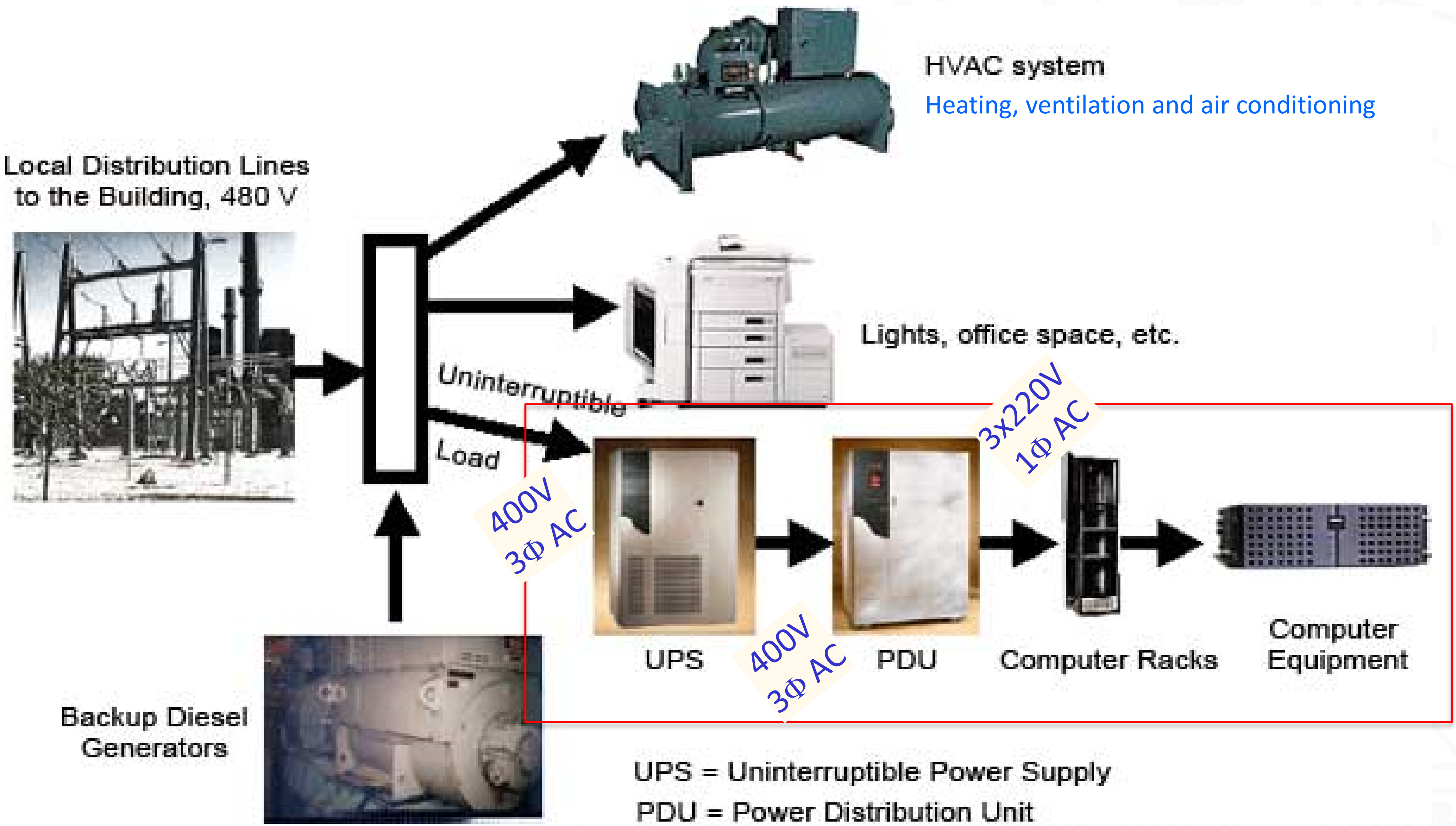


Uptime certifications, Spain, Feb. 2108

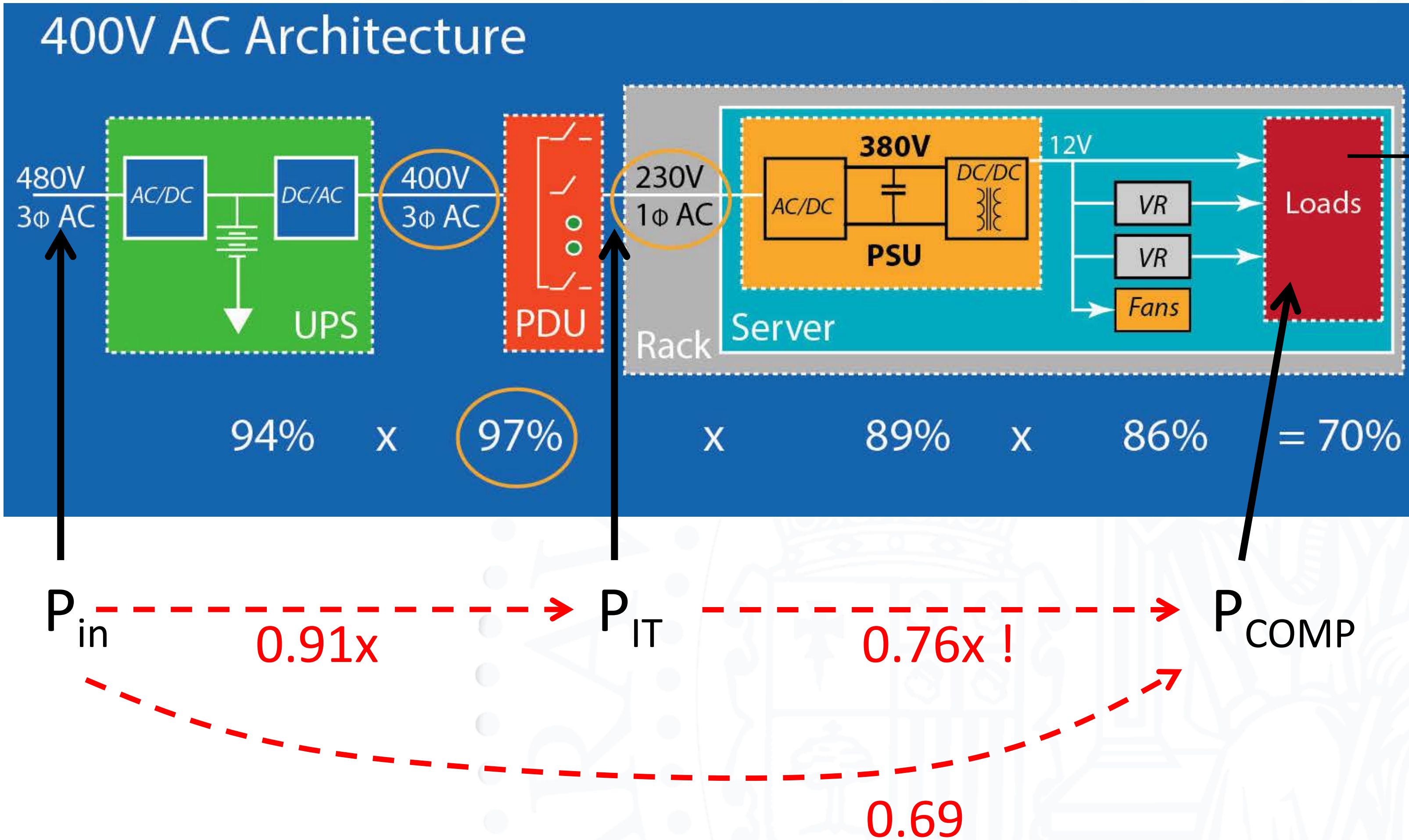
Certification	Tier	Companies	
Operational Sustainability	Tier IV Gold Fault Tolerant	Telefonica Global Technology, BBVAx2	3
Constructed Facility	Tier IV Fault Tolerant	Itconic S.A., BBVAx2, KIO Networks, INDITEX, Telefonica Global Technology	7
	Tier III Concurrently Maintainable	T-Systems ITC Iberia SAU, ASAC Comunicaciones S.L., Silc Immoblesx2, United Nations, T-Systems ITC Iberia SAU, MAPFRE	7
Design Documents	Tier IV Fault Tolerant	Itconic S.A., KIO Networks, INDITEX, BBVAx2, Enagás, Telefonica Global Technology, TISSAT, WALHALLA Data Center	9
	Tier III Concurrently Maintainable	Docalia,S.L., T-Systems ITC Iberia SAU, Silc Immoblesx2, ASAC Comunicaciones S.L., Mediapro, T-Systems ITC Iberia SAU, MAPFRE, United Nations	7
	Tier II Redundant	EDICOM	1
			34


0 1
Ing. Eléctrica
Ing. Electrónica
0 1
Ing. Mecánica
Centro de Datos
Arquitectura
1 0 1 0 0 1 1
Ing. Informática
Ing. Telecomunicación
https://backdropoutlet.com

Electricity Flow in Data Centers



Una cascada de pérdidas !

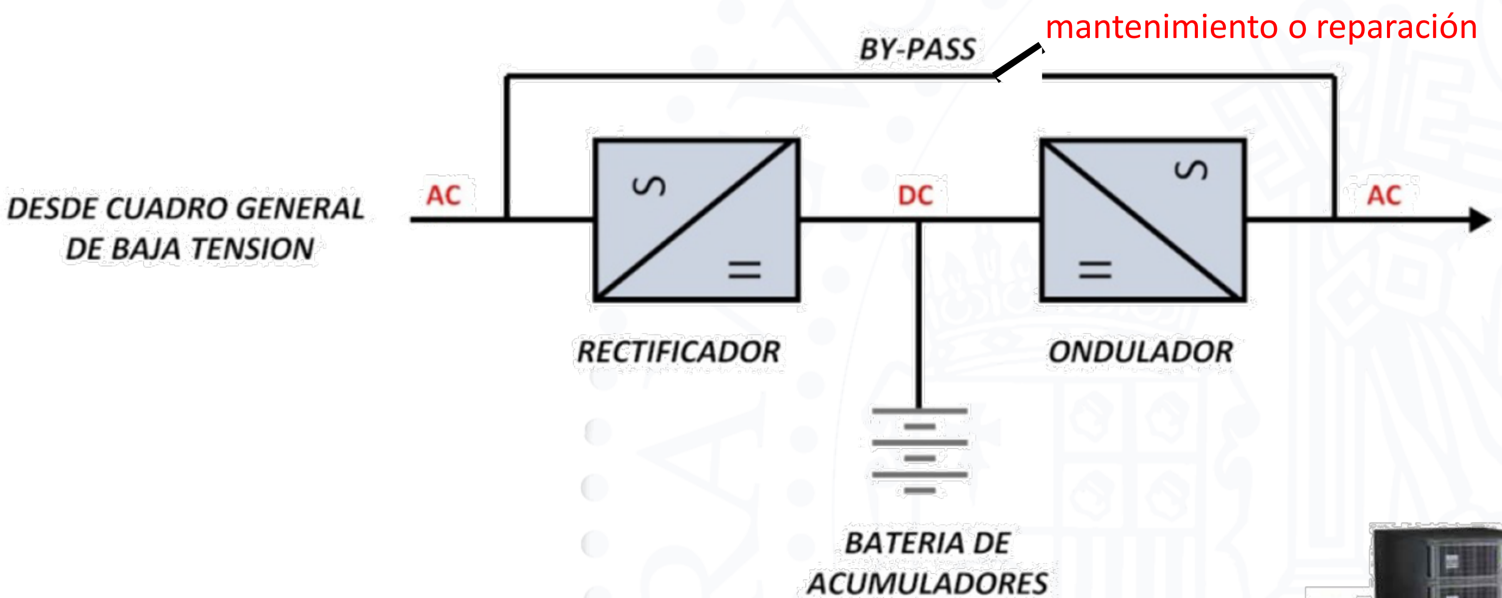


- Processor sockets
- DRAMs
- Chipset
- …

Figure 4.3 of “The Data center as a Computer”, 2nd ed. Efficiency of 400VAC distribution architecture (Source: Pratt, Kumar, Bross, Aldridge, “Powering Compute Platforms in High Effciency Data Centers,” IBM Technology Symposium, 2006).

SAI: Sistema de Alimentación Ininterrumpida

- SAI *estáticos*
  - energía *química*, se guarda en una batería de acumuladores
  - autonomía < 30 minutos
  - *on-line*: el ondulador genera la electricidad alterna para la sala IT, eliminando armónicos, deformaciones de onda, transitorios, etc.

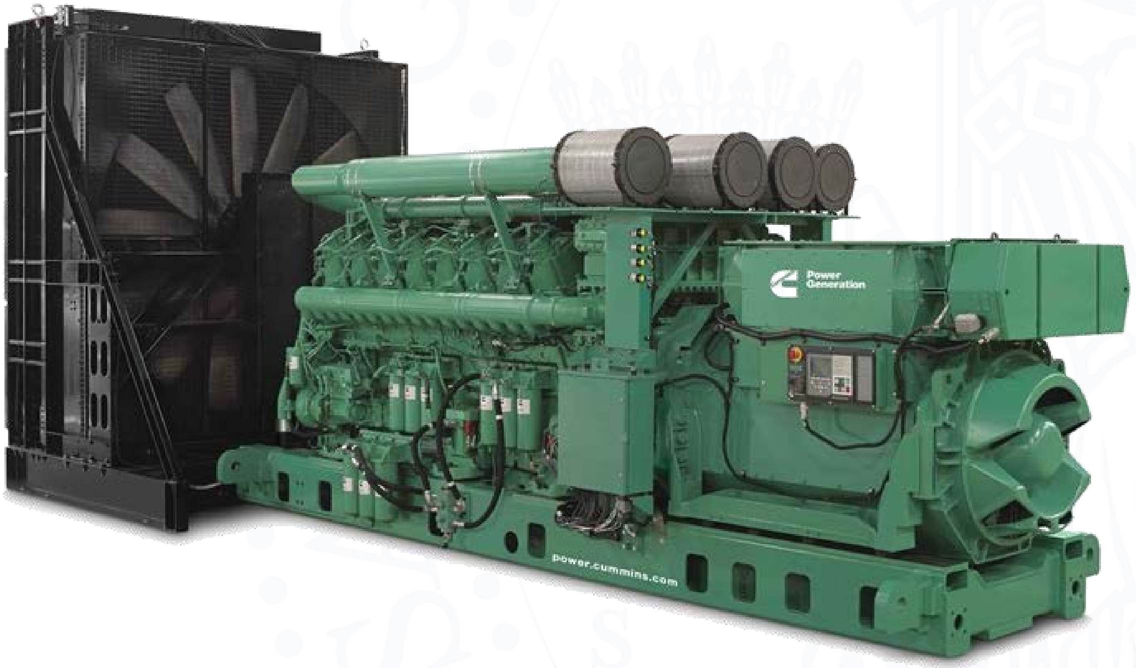


SAI: Sistema de Alimentación Ininterrumpida

- SAI *dinámicos, p.e DRUPS: diesel rotary UPS*
  - energía *cinética*, se almacena en una masa giratoria (volante de inercia)
  - un motor síncrono para mover la masa
  - un grupo electrógeno en línea

Grupos Electrógenos

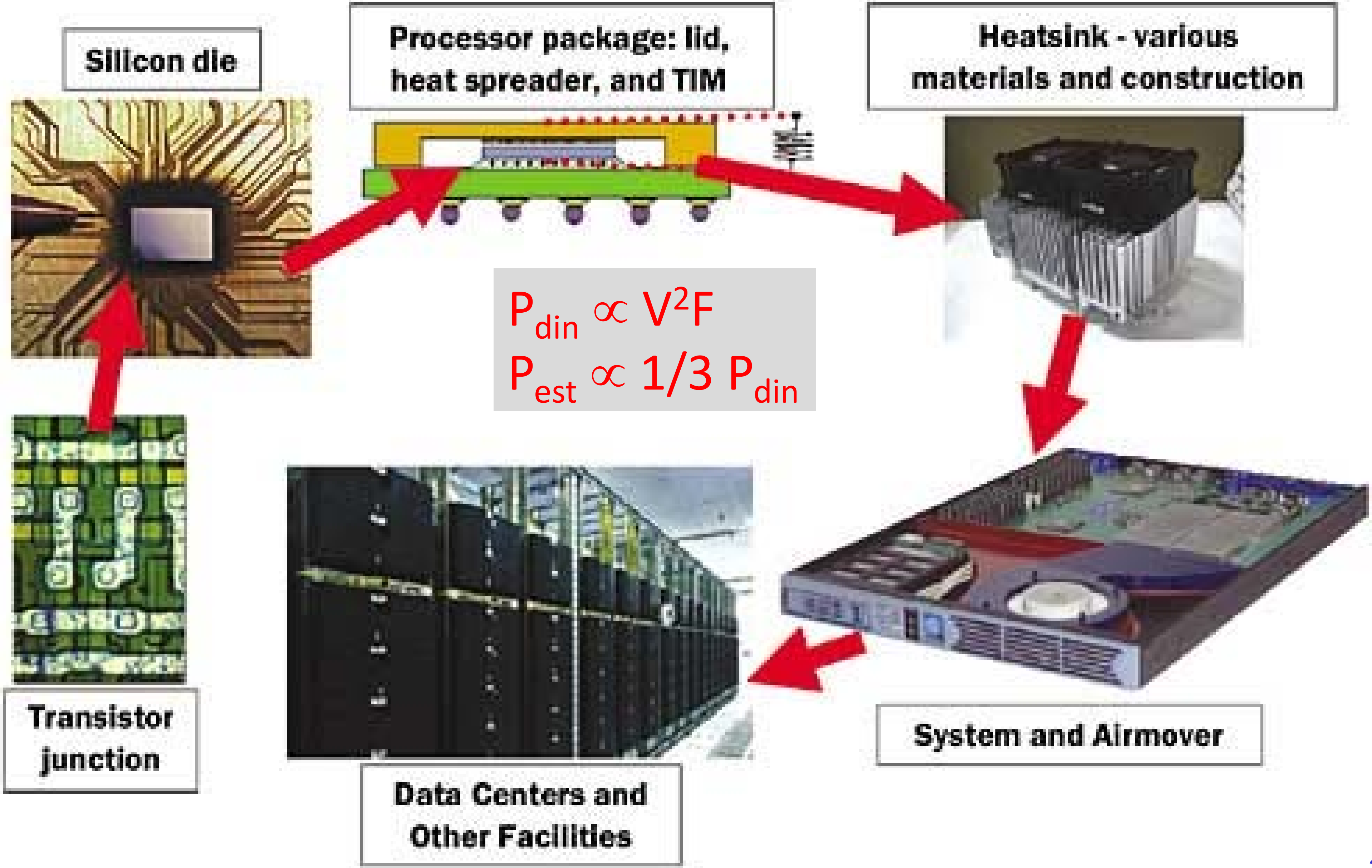
- Motores diésel o gas natural
- Generan energía eléctrica en Baja Tensión (400/230 V, 50Hz)
- Arranque automático en caso de caída de la red eléctrica
  - el Centro de Datos se desacopla de la red exterior
  - el Grupo Electrógeno la sustituye
- Segundos-minutos hasta alcanzar la velocidad de régimen
  - Un SAI (Sistema de Alimentación Ininterrumpida) cubre ese tiempo



Cummins QSK95
2,2 - 3,8 MW


0
1
Ing. Eléctrica
Ing. Electrónica
Ing. Mecánica
Centro de Datos
Arquitectura
Ing. Informática
Ing. Telecomunicación
https://backdropoutlet.com

Generación de calor

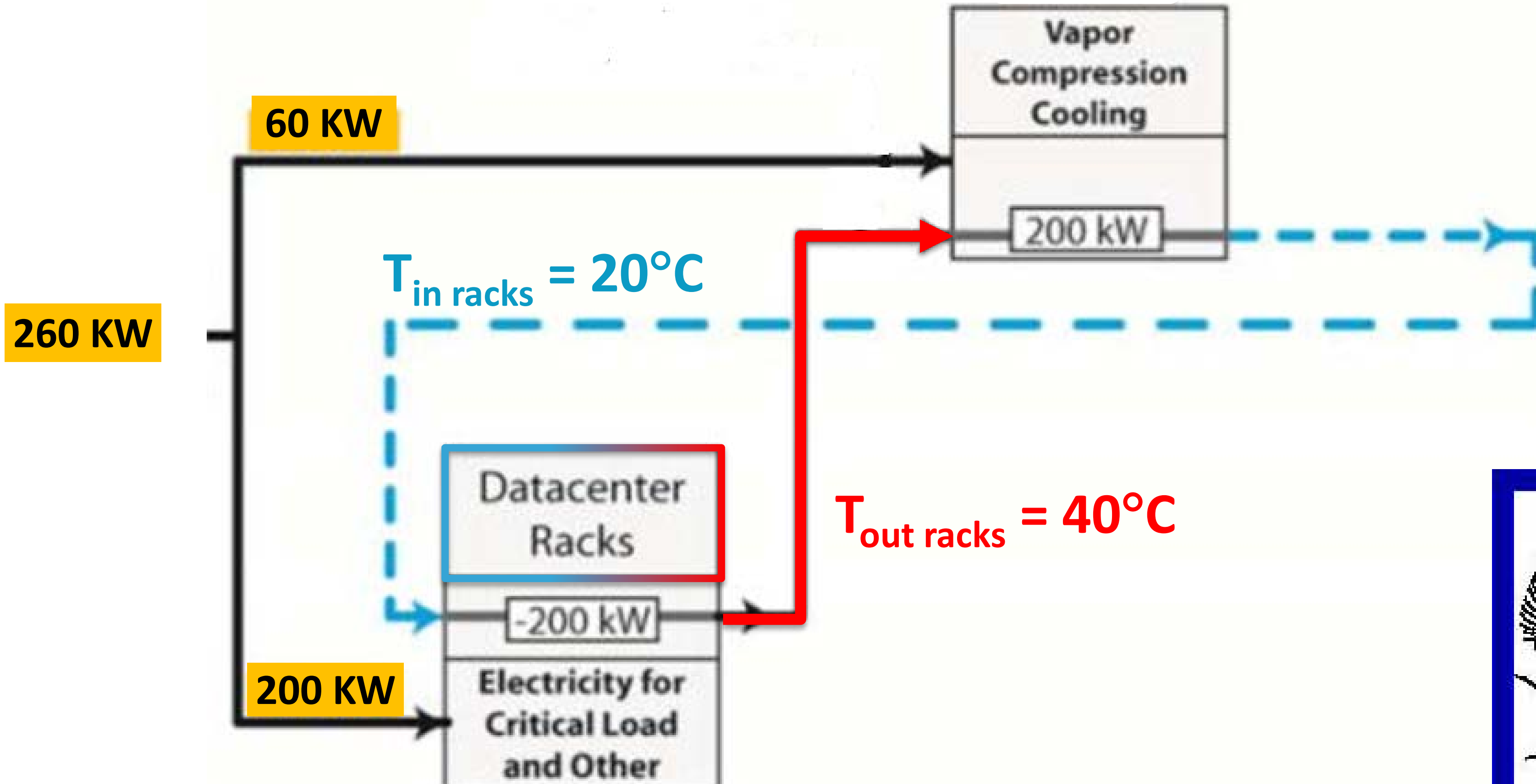


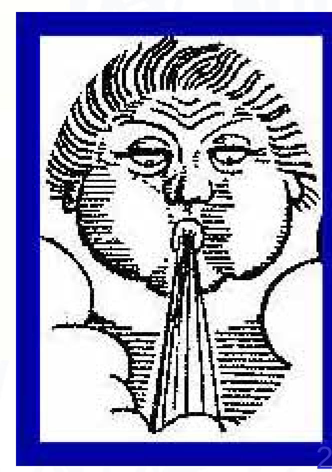
$P_{din} \propto V^2F$

$P_{est} \propto 1/3\ P_{din}$

COP: Coefficient of Performance

COP = 3.3 = 200 KW / 60 KW
= Q removed / Electrical Energy required

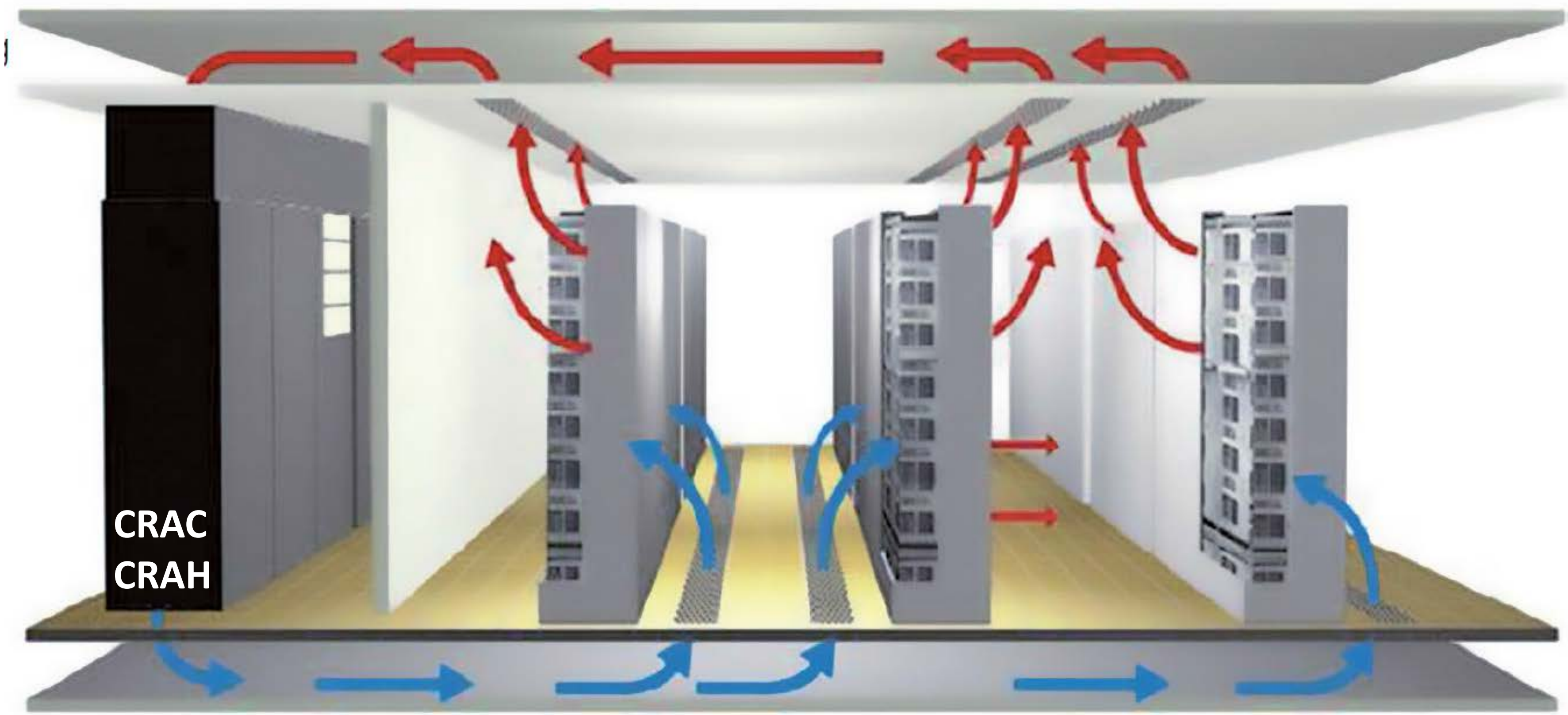




Precedente: hipocausto

Termas en la villa de Vieux-la-Romaine, Normandía, Francia

Pasillo frio - pasillo caliente



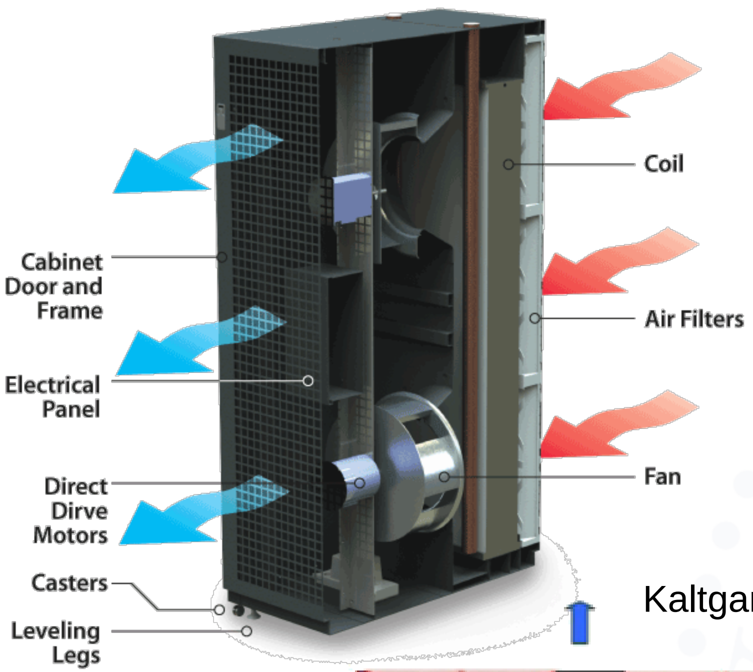
CRAC – Computer Room Air Conditioner (most: ON/OFF)
Condensador en tejado + CRAC con compressor y serpentin con líquido refrigerente

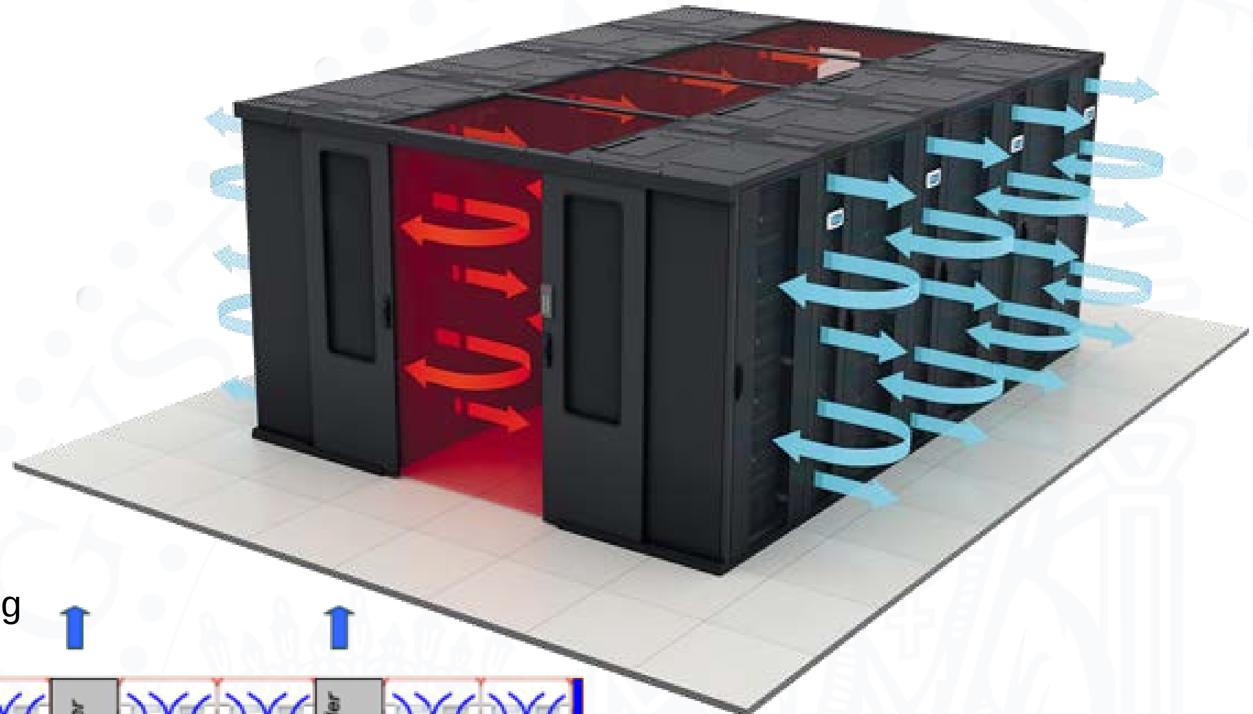
CRAH – Computer Room Air Handler (can have Variable Speed Fans)
Sistema externo para enfriar agua + CRAH con válvula y serpentín para el agua enfriada

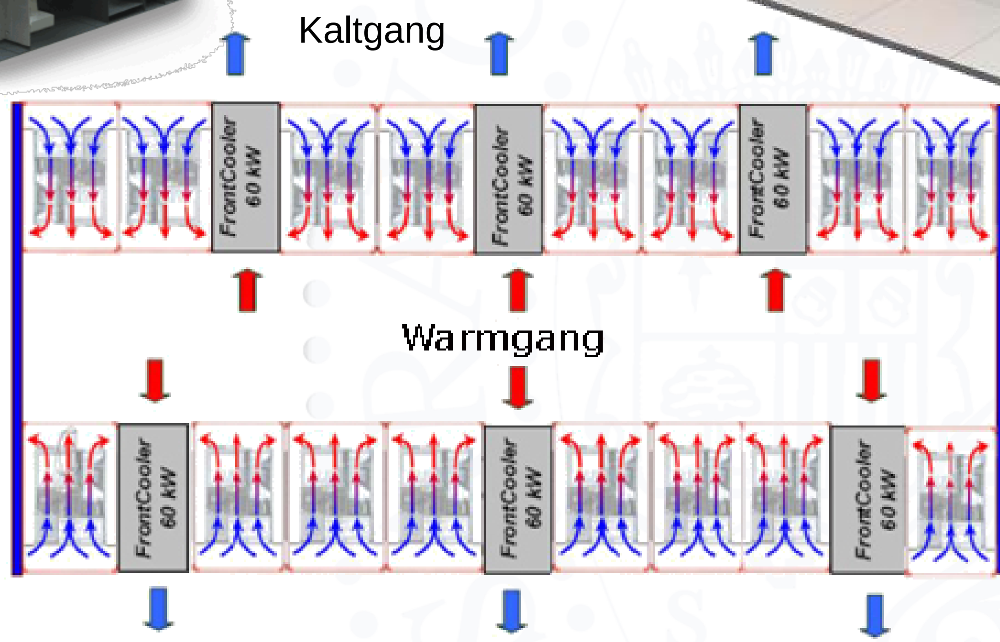
El Centro de Proceso de Datos. PQC.
Ilustrado por Forges. 2012

Antonio Fraguas de Pablo. Madrid.
17-1-1942 22-2-2018

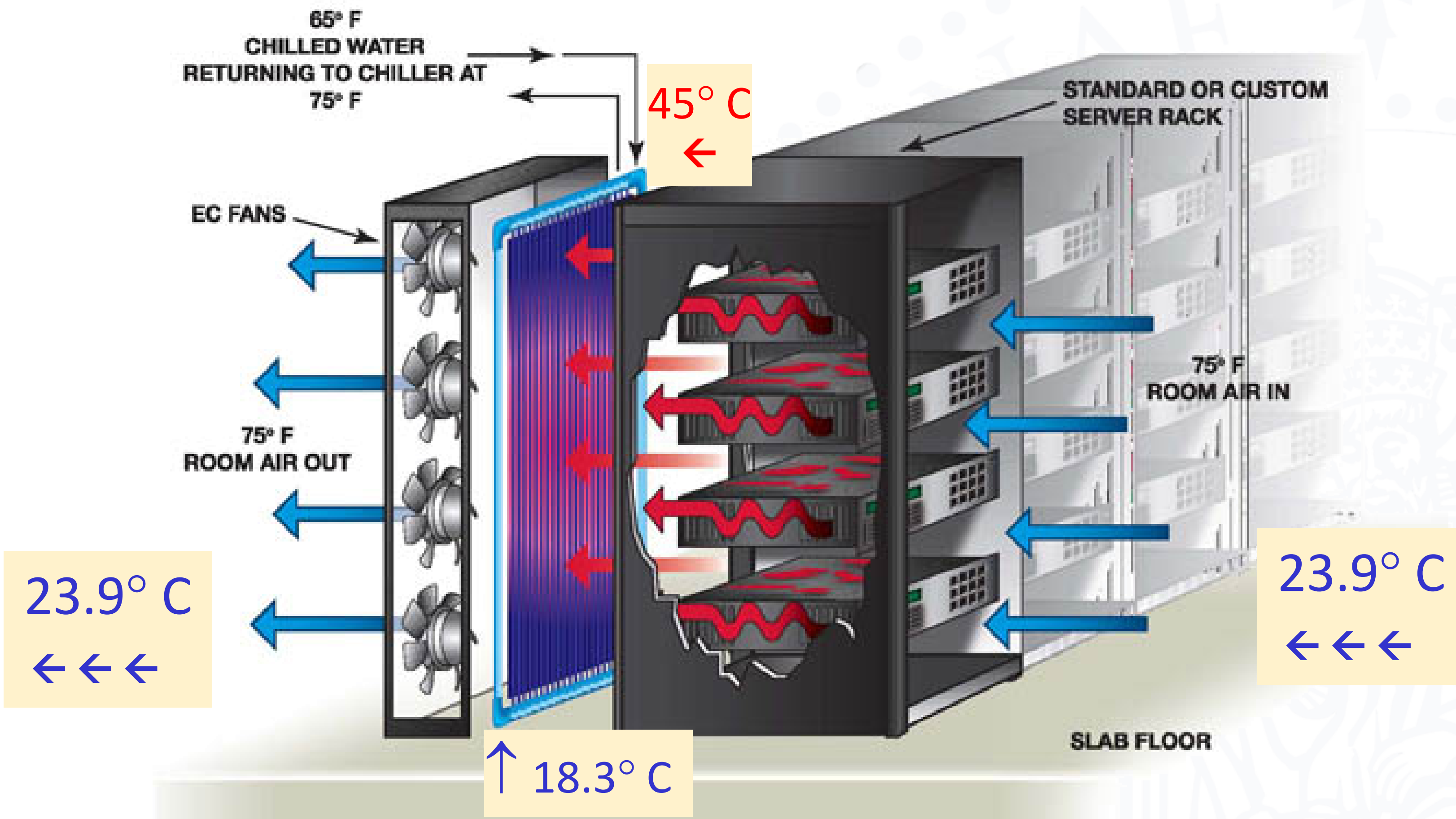
In-row cooling (con contención aire caliente)







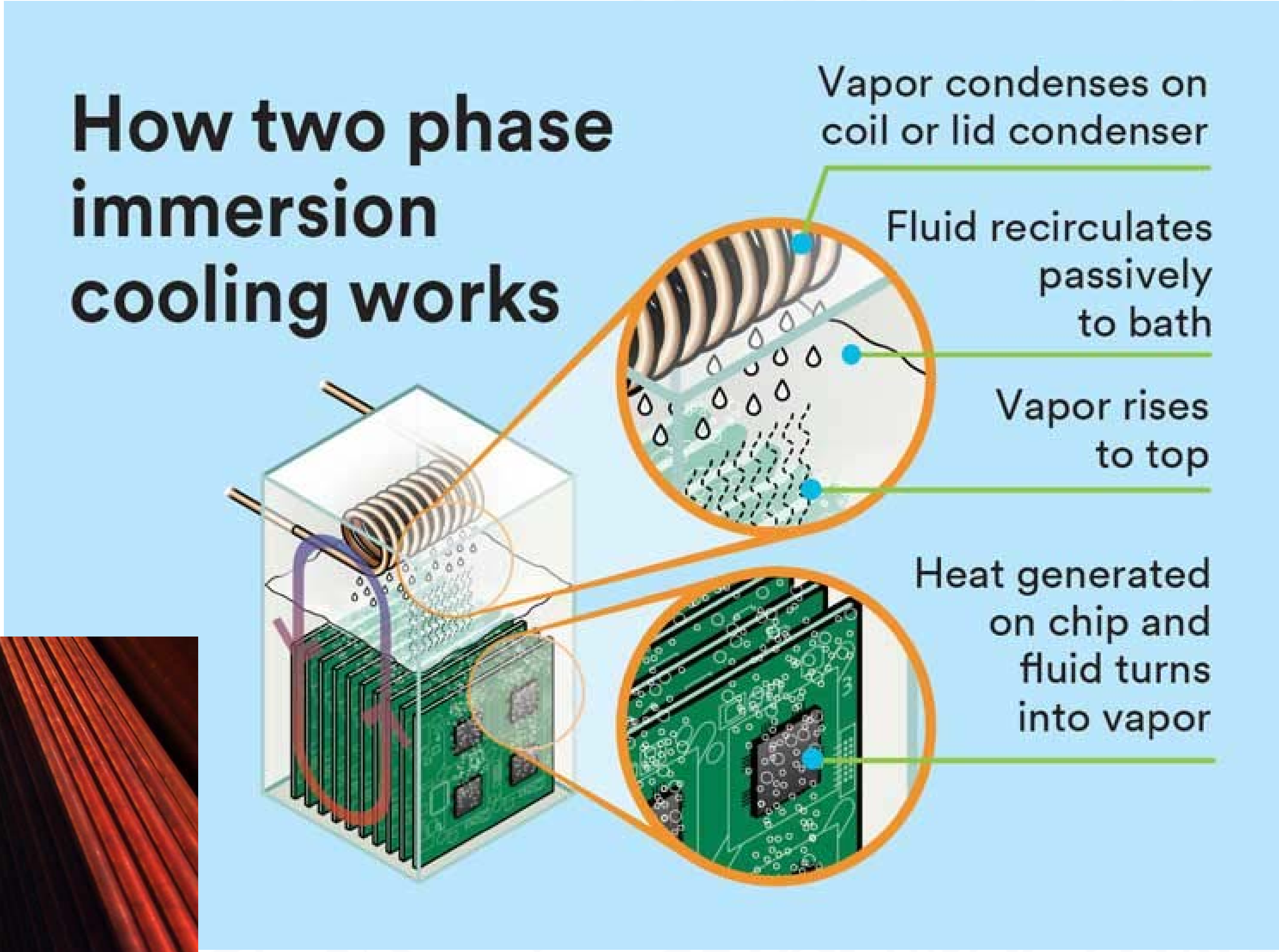
In-rack cooling (puertas refrigeradoras)



- Chilled water from the University of Oregon's central chiller plant will be piped beneath a raised floor in Allen Hall's new data center to cool high-density racks able to safely receive up to 30 kilowatts of power without overheating. Lower-density racks will use the same water system via in-row coolers. (Rendering courtesy of Motivair). Read more: http://djcoregon.com/news/2012/04/09/university-of-oregon-data-center-to-give-researchers-room-to-think/#ixzz3f6jGkB19

Inmersión: 3M™ Novec™ Engineered Fluids

- Novec: electrical insulator
- Potential to save up to 97% in energy usage compared to traditional air-cooling methods




0 1
Ing. Eléctrica
Ing. Electrónica
0 1
Ing. Mecánica
Centro de Datos
Arquitectura
1 0 1 0 0 1 1
Ing. Informática
Ing. Telecomunicación
https://backdropoutlet.com

Racks

Capacidad:
42 U = 186,9 cm
(1U = 4,45 cm)

92,75 cm

60,95 cm

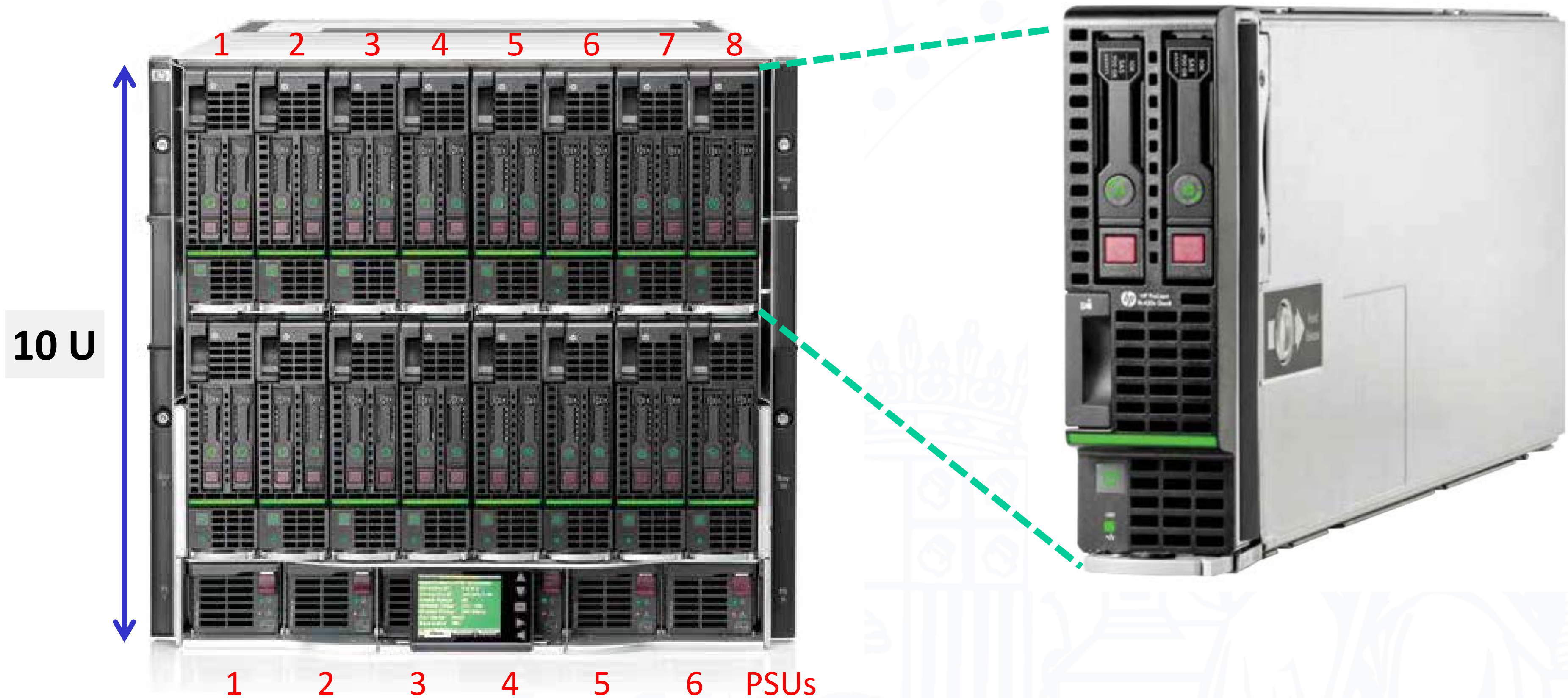
Servidor de Cálculo formato horizontal: 1U, 2U y 3U

- > 2 sockets-chips de gama alta
  - > 10 cores/chip
  - > 4 GB/core de DRAM DDR3-4 (> 80 GB en total)
- 2 fuentes de alimentación
- muchos mini-ventiladores
- disco duro o de estado sólido
- varios puertos Ethernet (cobre o fibra óptica)

2U Supermicro Superserver SYS-6029TP-HTR
8 Intel Skylake-SP 5120 14C/28T 2.2G 19.25M 10.4GT
384 GB = 48x8GB DDR4-2666 ECC REG RDIMM
3.5" SATA3 2TB 6Gb/s 7.2K 128MB Makara BP 512
4xSIOM 2-port GbE 1U bracket, Intel i350-AM2
17.985 € + IVA = **21.761,85 €**

Ejemplo: HP Proliant BL460c Gen8 Blade Server

- http://h18004.www1.hp.com/products/blades/info/what-is-a-blade.html

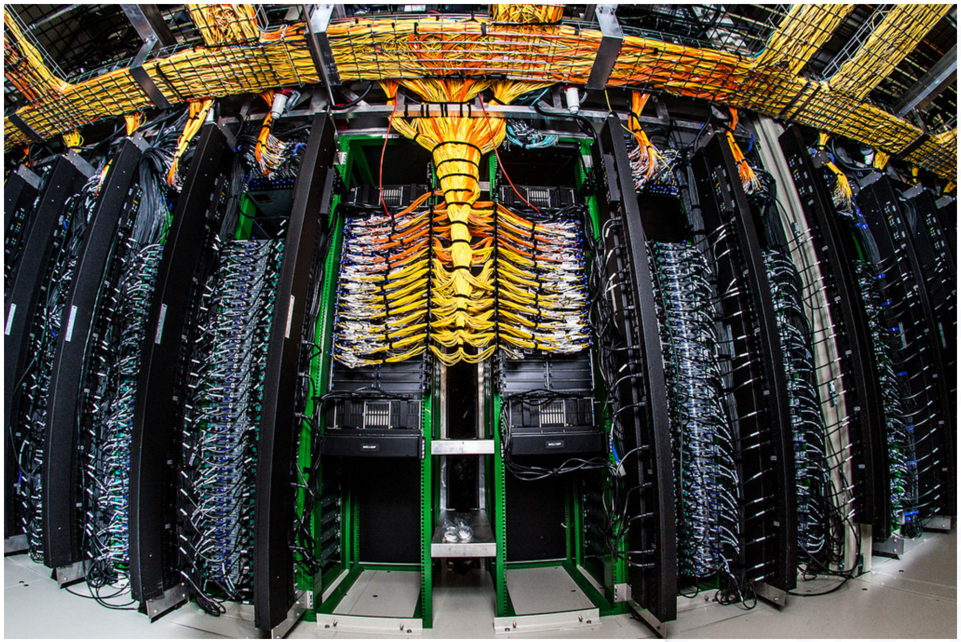


Servidores de Disco

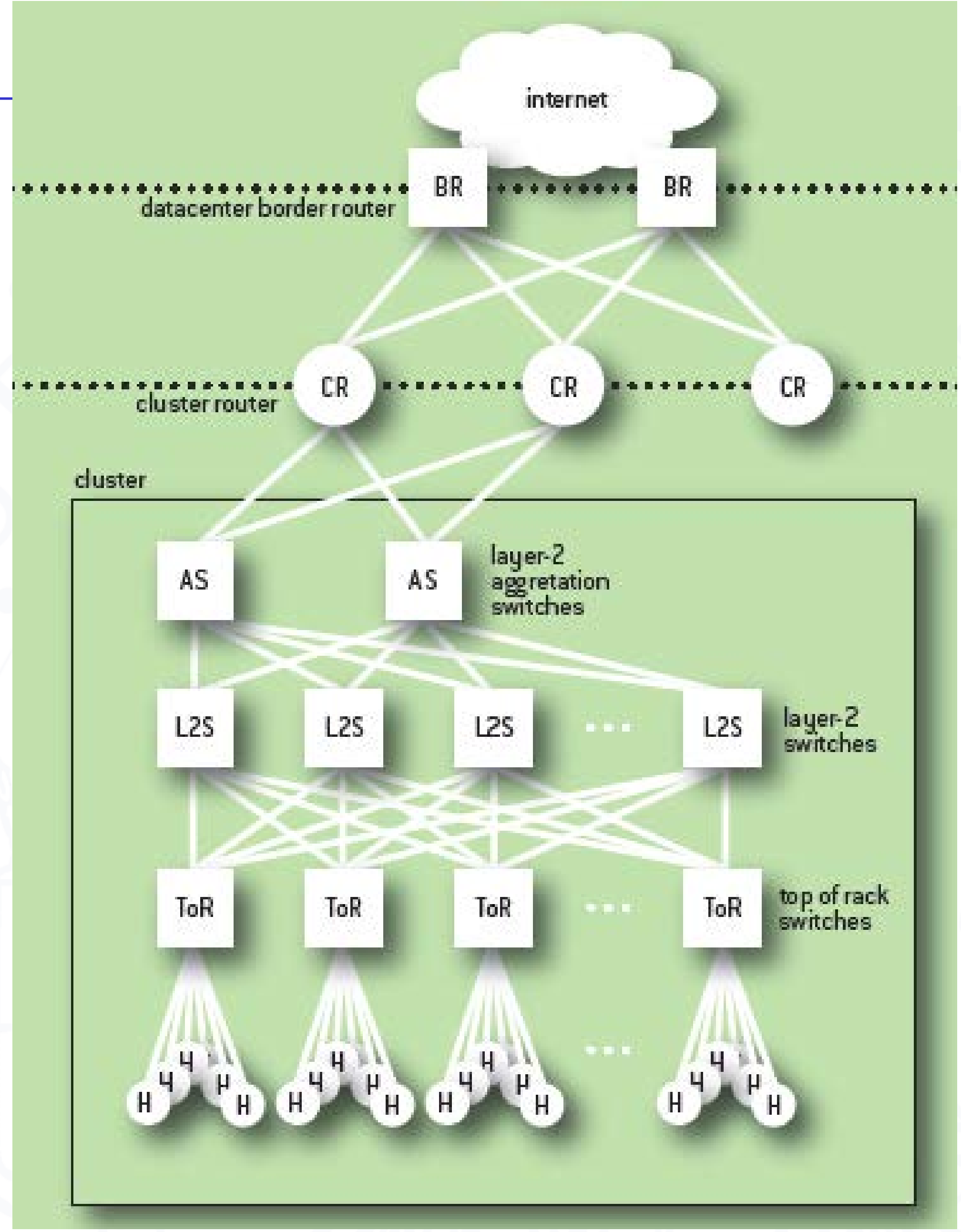
- Ejemplo: Cisco UCS S3260 Storage Server
  - Dual-node Intel® Xeon® processor E5-2600 v4 series
  - Up to **600 TB** of local storage
  - Can be configured with enterprise-class Redundant Array of Independent Disks (RAID)
  - 2 x 40 Gbps Ethernet port connectivity

1 TB = 2^{40} B ≈ 10^{12} B

Red de Comunicaciones



http://cableporn.tumblr.com/image/140538130922



Ejemplo: Cisco 3548

- Switch orientado a:
  - High-Frequency Trading (HFC, inversión en bolsa automática)
  - Big Data (análisis de enormes cantidades de datos en tiempo real)
- 48 puertos 10 Gigabit Ethernet
- Rendimiento
  - Todos los puertos pueden transmitir simultáneamente a su velocidad de línea (non-blocking)
  - Level 2: 480 GBps (max, tramas)
  - Level 3: 720 Mpps (max, paquetes)

1 U

Figura: cisco.com

http://www.tel2u.com/es/redes/472336-nexo-3548-48-puertos-sfp-0882658506017.html

El negocio

Tipos de CDs según el grado de propiedad de IT

- **Tres opciones de Centros de Datos** según el grado de propiedad de las infraestructuras
  - Añadimos una cuarta opción que aclara la posición del **cloud**

Licencias de software ♦

		IT (servidores, almacenamiento, red)		
		Propiedad	*Alquiler*	*No tengo*
Instalaciones (frio, electricidad, internet, seguridad)	*Propiedad*	**Enterprise** **(o Corporative)** ♦		
	Alquiler	**Colocation** **(o Housing)** ♦	**Managed Hosting** • Leasing de mis servidores ♦ • Cuenta en servidores compartidos	
	No tengo			[Closed Application] SaaS ♦ [Development] PaaS ♦ [Virtual Machine] IaaS ♦

aumenta la externalización (outsourcing)

Enterprise data centers

1. Traditional
   - e.g. Banking, energy, government departments,
2. Warehouse Scale Computer (WSC)
   - e.g. Google, Amazon, Facebook, Apple, Microsoft's online internet services
3. Supercomputer
   - Scientific applications,
     recently including Big Data analytics and Cloud Computing

		Servidores		
		Propiedad	Alquiler	No tengo
Instalaciones	Propiedad	**Enterprise**		
	Alquiler	**Colocation**	**Managed Hosting** • Leasing • Cuentas	
	No tengo			Cloud

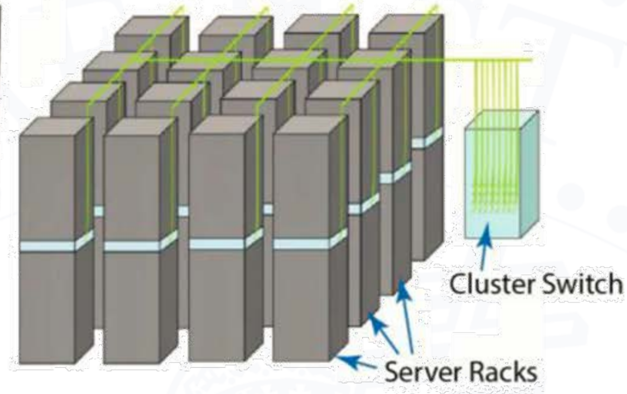
Traditional Enterprise Datacenter

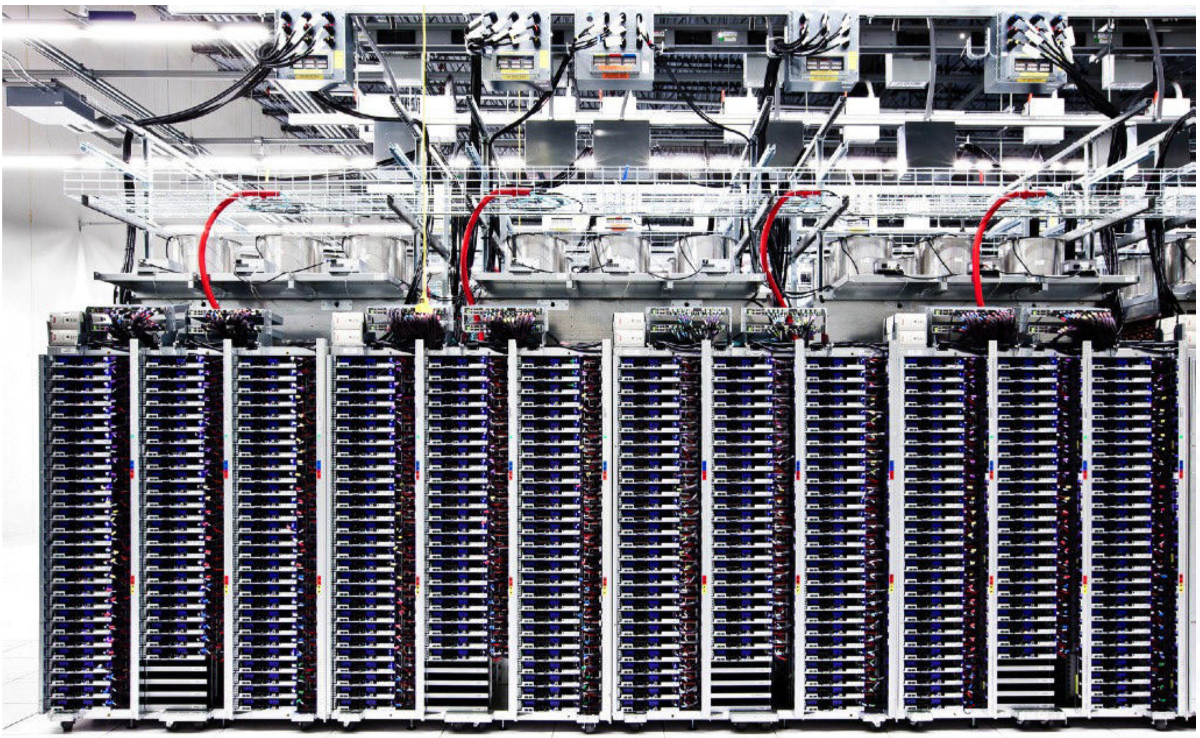
- Centro de proceso de datos de Cantabria, del Banco Santander (2011)
  - 100 millones de clientes
  - picos de 10.000 tps
  - dos centros independientes
  - 120 cubos cerrados con 14, 16 o 18 racks
  - Hw heterogéneo: servidores pizza, blades y mainframes
  - aspira a certificación TIER IV (Tier Certification of Constructed Facility)
  - PUE 1,25
  - 20 MW
    (30% del consumo
    de la ciudad de Santander)

		Servidores		
		Propiedad	Alquiler	No tengo
Instalaciones	Propiedad	**Enterprise**		
	Alquiler	**Colocation**	**Managed Hosting** • Leasing • Cuentas	
	No tengo			Cloud

Warehouse Scale Computer

Figure 1.1: Sketch of the typical elements in warehouse-scale systems: 1U server (left), 7’ rack with Ethernet switch (middle), and diagram of a small cluster with a cluster-level Ethernet switch/router (right).





		Servidores		
		Propiedad	Alquiler	No tengo
Instalaciones	Propiedad	**Enterprise**		
	Alquiler	**Colocation**	**Managed Hosting** • Leasing • Cuentas	
	No tengo			Cloud

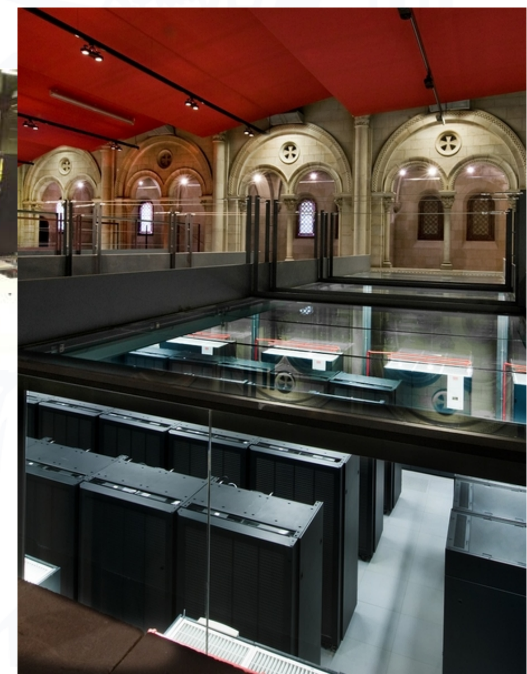
Figure 1.2:

Picture of a row of servers in a **Google** WSC, 2012

Supercomputer

- BSC Mare Nostrum **IV 2017** (public, UPC-Barcelona)
  - Peak Performance of 11.15 PetaFLOPS
  - 384.75 TB of main memory
  - 3,456 Main Compute Nodes (165,888 cores in 48 racks)
    - 2x Intel Xeon Skylake Platinum 8160 24C a 2.1 GHz
    - 216 nodes, 12x8 GB DDR4-2667 DIMMS (8GB/core)
    - 3240 nodes, 12x32 GB DDR4-2667 DIMMS (2GB/core)
  - 14 PB of disk storage, 3 racks
  - Interconnection networks:
    - 100Gb Intel Omni-Path Full-Fat Tree
    - 10 Gb Ethernet
  - SUSE Linux Enterprise Server 12 SP2
  - 33.7 kW/rack, 1.62 MW, > 1,4 M€/year,
    Rear Door Cooling by means of chilled water

		Servidores		
		Propiedad	Alquiler	No tengo
Instalaciones	Propiedad	**Enterprise**		
	Alquiler	**Colocation**	**Managed Hosting** • Leasing • Cuentas	
	No tengo			Cloud



Colocation

http://www.pulsant.com/colocation/

- **Pulsant**: over 3,000 customers and 15 years of experience
- Owns and operates 10 data centres
  - Milton Keynes Colocation Data Centre
    - Within one hour of London
    - Tier 4 enterprise class facility
    - 3400 m^2 (< 60x60m)
    - Allows high rack power densities, up to 15Kw
    - Connected by a resilient,
      multi-carrier high speed network with 2ms of network latency
    - ISO/IEC 27001:2005 compliant.
      Management system intended to bring information security
      under explicit management control
    - PCI compliant (Payment Card Industry Data Security Standard)

		Servidores		
		Propiedad	Alquiler	No tengo
Instalaciones	Propiedad	**Enterprise**		
	Alquiler	**Colocation**	**Managed Hosting** • Leasing • Cuentas	
	No tengo			Cloud

Options for Colocation

http://www.pulsant.com/colocation/

Capacity available from single racks, cages and private suites to entire data halls

Private Cage Features

A caged area isolates your equipment from the rest.
Cages are built on demand to satisfy the floor space, power, network connectivity and cooling requirements.

- Starting at 3 cabinet minimum
- Customized security options

Locked Rack Cabinets

Rack cabinets available with individual combination.

- 44U rack cabinets
- Full rack, half and quarter options
- 24/7 intelligent hands available
- 24/7 onsite access

Dedicated Private Suites

Private suites are ideal where a client requires a large amount of floor space combined with the maximum level of access control.

- Hard-walled private area
- 230 m^2 floor space available
- Customised security options

		Servidores		
		Propiedad	Alquiler	No tengo
Instalaciones	Propiedad	**Enterprise**		
	Alquiler	**Colocation**	**Managed Hosting** • Leasing • Cuentas	
	No tengo			Cloud

Colocation, IP transit, cloud or hosting services

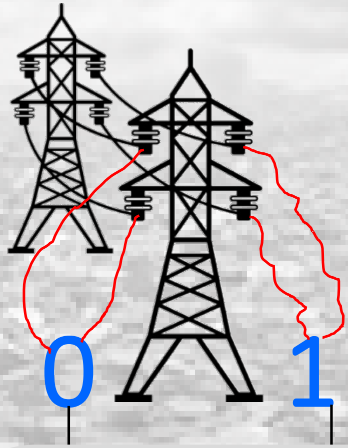
50 centros de datos
http://www.datacentermap.com/

Non-neutral DC's Carrier neutral DC's Building with multiple DC's

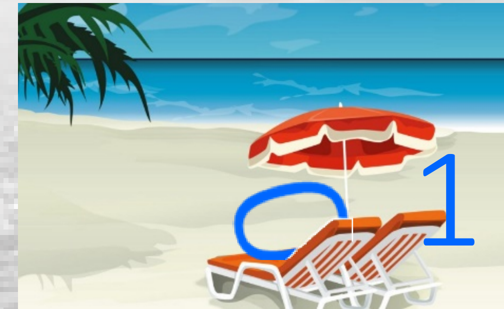
15 centros de datos

http://www.datacentermap.com/

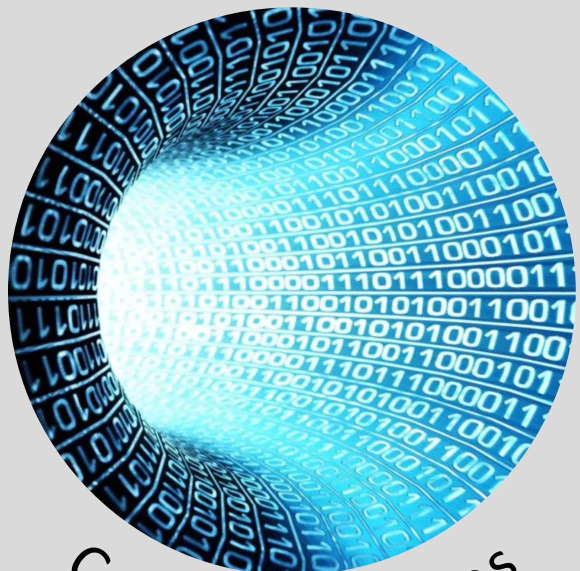
Non-neutral DC's   Carrier neutral DC's   Building with multiple DC's



Ing. Eléctrica
Ing. Electrónica



Ing. Mecánica



Centro de Datos

Arquitectura

https://backdropoutlet.com

Ing. Informática
Ing. Telecomunicación

ACT

ESSENTIAL DETAILS

ORGANIZATION:	ACT, Inc.
LOCATION:	Iowa City, Iowa
ONLINE:	March 2008
NOTABLE FEATURES:	Closed-loop geothermal cooling, HVAC dry coolers, tornado-resistant structure, hot and cold aisles. Active-active configuration with a second Data Center 5 miles (8 [illegible] away, linked by ACT's own dark fiber. LEED-Platinum certified.
TIME TO DESIGN AND BUILD:	13 months
SIZE:	6,700 sq. ft. (622.5 sq. m) total, with 4,000 sq. ft. (371.6 sq. m) of hosting space.

"People Get Ready"
Jeff Beck and Rod Stewart
1985

FIGURE 1-3 Tornado-resistant glass allows daylight into parts of the building regularly occupied by people. Image provided courtesy of Neumann Monson Architects.

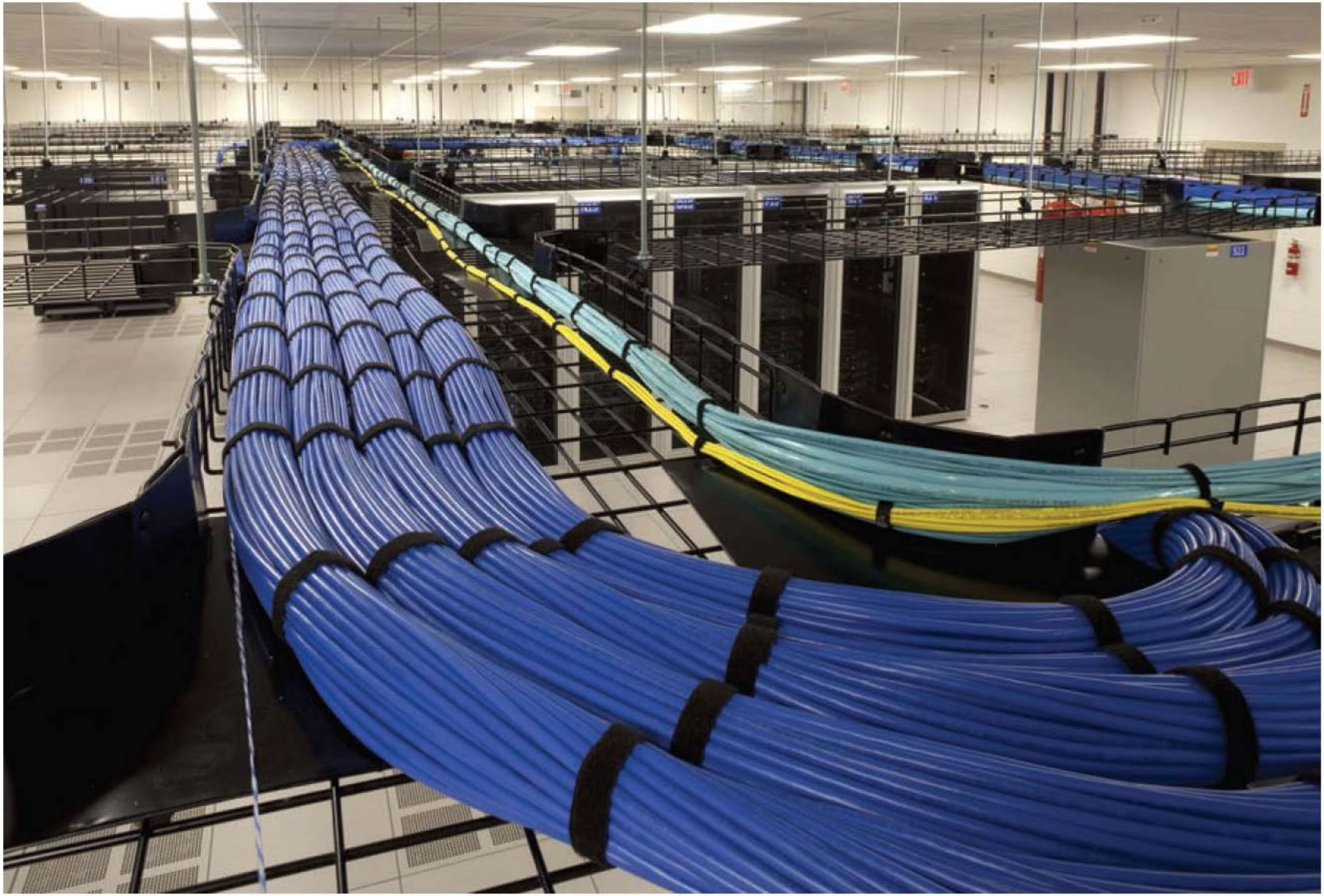


FIGURE 1-4 Structured cabling is routed overhead. Electrical conduits and cooling are provided below a raised floor. Image provided courtesy of Neumann Monson Architects.

ESSENTIAL DETAILS

ORGANIZATION: Bahnhof	
LOCATION: Stockholm, Sweden	
ONLINE: September 2008	
NOTABLE FEATURES: Artificial daylight, greenhouses, waterfalls, and a 687 gallon (2,600 liter) saltwater fish tank. Two Maybach diesel submarine engines used for standby power.	
TIME TO DESIGN AND BUILD: 20 months	
SIZE: 10,764 sq. ft. (1,000 sq. m) total, with 5,382 sq. ft. (500 sq. m) of hosting space and 2,153 sq. ft. (200 sq. m) for back-of-house systems. Remaining space is for office and personnel areas.	
IT-USABLE CAPACITY: 800 kW	

The above-ground entrance to Bahnhof's Data Center in Stockholm, Sweden.
Embedded 30 m underground in a former nuclear bunker, the Data Center retains its wartime codename, **Pionen White Mountains**. Images provided courtesy of Bahnhof.

FIGURE 3-6 Artificial waterfalls within the Bahnhof Data Center.

Pionen White Mountains

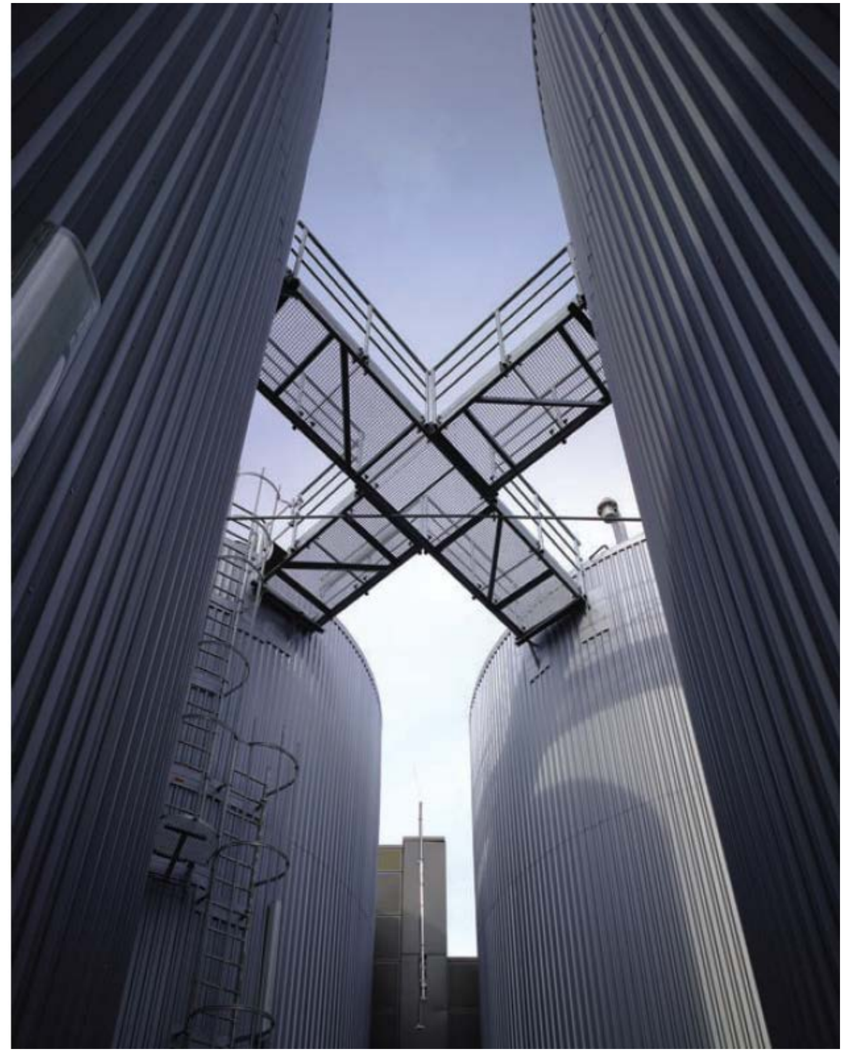
Two Maybach diesel
submarine engines,
complete with accompanying
sound-horn,
provide standby power for the
Data Center.

ESSENTIAL DETAILS

ORGANIZATION: Citi	
LOCATION: Frankfurt, Germany	
ONLINE: May 2008	
NOTABLE FEATURES: Green roof and green wall features. Reverse osmosis water treatment for cooling. Harvested rainwater used for irrigation. Rotary UPS. First Data Center certified LEED-Platinum.	
TIME TO DESIGN AND BUILD: 20 months	
SIZE: 228,000 sq. ft. (21,181.9 sq. m) total, with 100,000 sq. ft. (9,290 sq. m) of hosting space	
POWER: 14.4 MW from two separate feeds	
TIER: III	

FIGURE 7-3 Citi's Data Center includes a green wall and a green roof that reduce the heat load on the building and last longer than conventional materials. Image by Christian Richters Photography.

Citi uses reverse osmosis to reduce sediment buildup in its cooling towers and saves about 13.2 million gallons (50 million liters) of water per year. Image by Christian Richters Photography.



Implementing diesel rotary UPS systems helped Citi eliminate batteries and other building materials. Image by Christian Richters Photography.

ESSENTIAL DETAILS

ORGANIZATION: Digital Realty Trust	
LOCATION: Chicago, Illinois	
ONLINE: First constructed from 1912 to 1914. Converted to telecom use between 1998 and 2000. Acquired by Digital Realty Trust in 2005.	
NOTABLE FEATURES: Designated historic landmark building. More power capacity than Chicago O'Hare International Airport. One of the world's largest carrier hotels. An 8.5 million gallon (32.2 million liter) brine tank for thermal storage. Includes a 20,000 sq. ft. (1,858.1 sq. m) data hall that was first in the United States to be certified LEED-Gold.	
TIME TO DESIGN AND BUILD: Varies by data hall	
SIZE: 1,133,391 sq. ft. (105,295.5 sq. m)	
POWER: 102 MW	

350 E. Cermak Road, Chicago, Illinois. Colocation Data and Telco Center. 109 MW
First constructed from 1912 to 1914. Converted to telecom use between 1998 and 2000. Acquired by Digital Realty Trust in 2005

FIGURE 8-2 Built in the early 1900s as a printing plant, the building's strong floors and large vertical shafts made it an excellent candidate for conversion to hosting space.

FIGURE 8-3 Digital Realty Trust installed a second layer of windows behind the originals, providing more physical security while allowing natural light in to certain areas.

FIGURE 8-4 Architectural details from the building feature symbols from the history of printing.

ESSENTIAL DETAILS

ORGANIZATION:	eBay
LOCATION:	Phoenix, Arizona
ONLINE:	October 2011
NOTABLE FEATURES:	In-building data hall, containerized Data Centers on the roof. Air and liquid cooling to cabinets, outside air cooling, adjustable computing capability.
TIME TO DESIGN AND BUILD:	22 months
SIZE:	42,000 sq. ft. (3,901.9 sq. m) total, with 14,000 sq. ft. (1,300.6 sq. m) of hosting space in the building and rooftop containers.
POWER:	4MW for IT initially, expandable to 6 MW for the in-room data hall and 6 MW for containers
TIER:	II

Project Mercury, Phoenix, Arizona. eBay. 2011. Tier II. 4-6 MW for in-house IT + 6 MW for containers on the roof

Greetings program. Visitors first encounter the eBay Data Center’s Tron theme at its entrances

Illuminated panels at the end of each row reinforce the eBay Data Center's futuristic appearance

FIGURE 9-4 Liquid cooling piping is routed below a 6 ft. (1.8 m.) raised floor while power, cooling, and cabling are distributed overhead in eBay's Project Mercury Data Center.

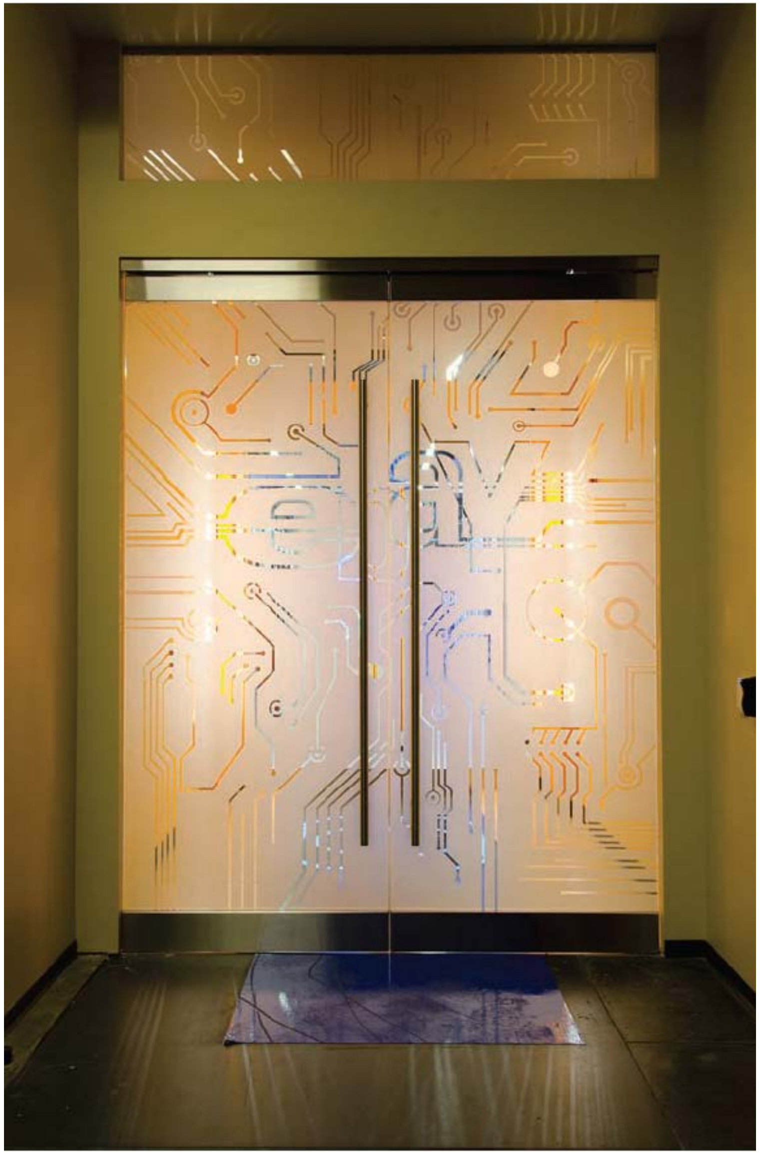


FIGURE 9-10 Power strips glow in a darkened Data Center container.

ESSENTIAL DETAILS

ORGANIZATION: Facebook	
LOCATION: Prineville, Oregon	
ONLINE: May 2011	
NOTABLE FEATURES: 277 volt power distribution. Evaporative cooling system. Power over Ethernet (PoE) LED lighting. Custom servers. LEED-Gold certified.	
TIME TO DESIGN AND BUILD: 15 months for 147,000 sq ft. (13,656.7 sq. m) first phase	
SIZE: 307,000 sq. ft. (28,521.2 sq. m) total, with about 261,000 sq. ft. (24,247.7 sq. m) of hosting area and 46,000 sq. ft. (4,273.5 sq. m) of electrical room space. Mechanical systems are located in an enclosed area on roof and not included in size figures.	
POWER: 30 MW available for IT	
TIER: III	

Prineville, Facebook, Oregon. 2011. Tier III. 30 MW IT. Starts “Open Compute Project”. PUE = 1.07 at full load

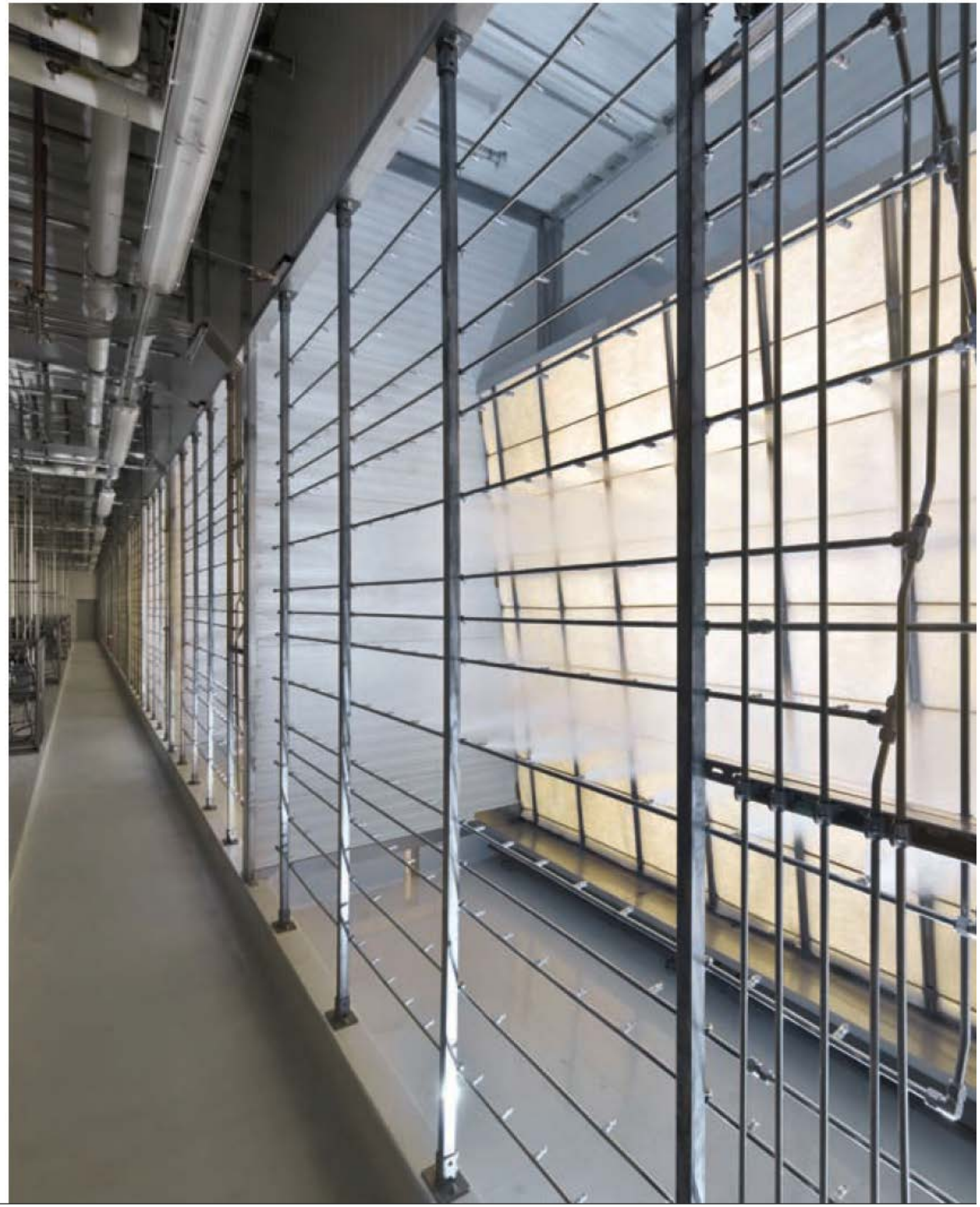
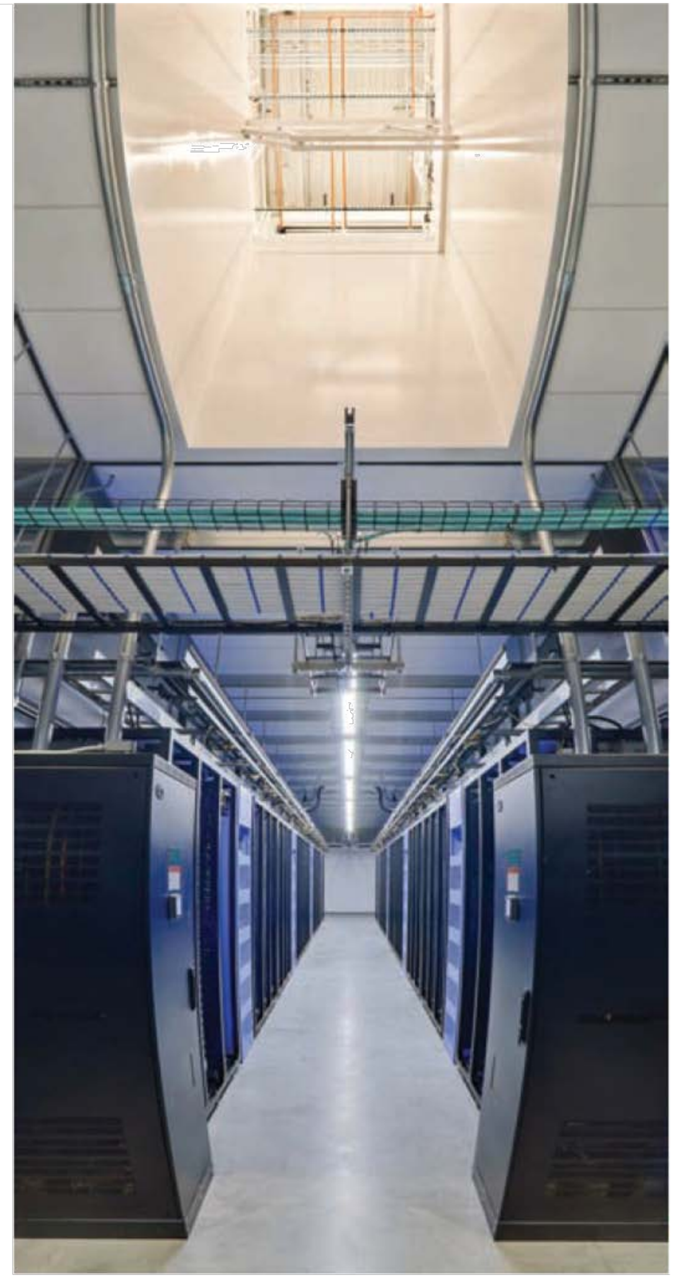
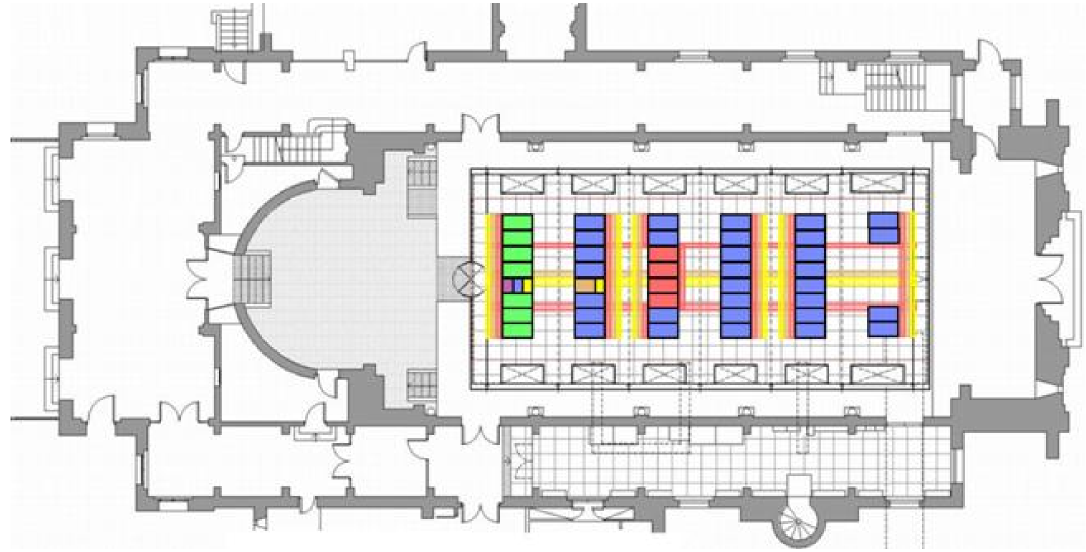


FIGURE 10-4
When the outside air used to cool the Data Center is too dry, moisture is added by way of a misting system. A mist eliminator prevents water that has not been absorbed by the air from proceeding into the Data Center.

FIGURE 10-8
Air passes through mist eliminators, at left, and propelled by variable speed fans, at right.



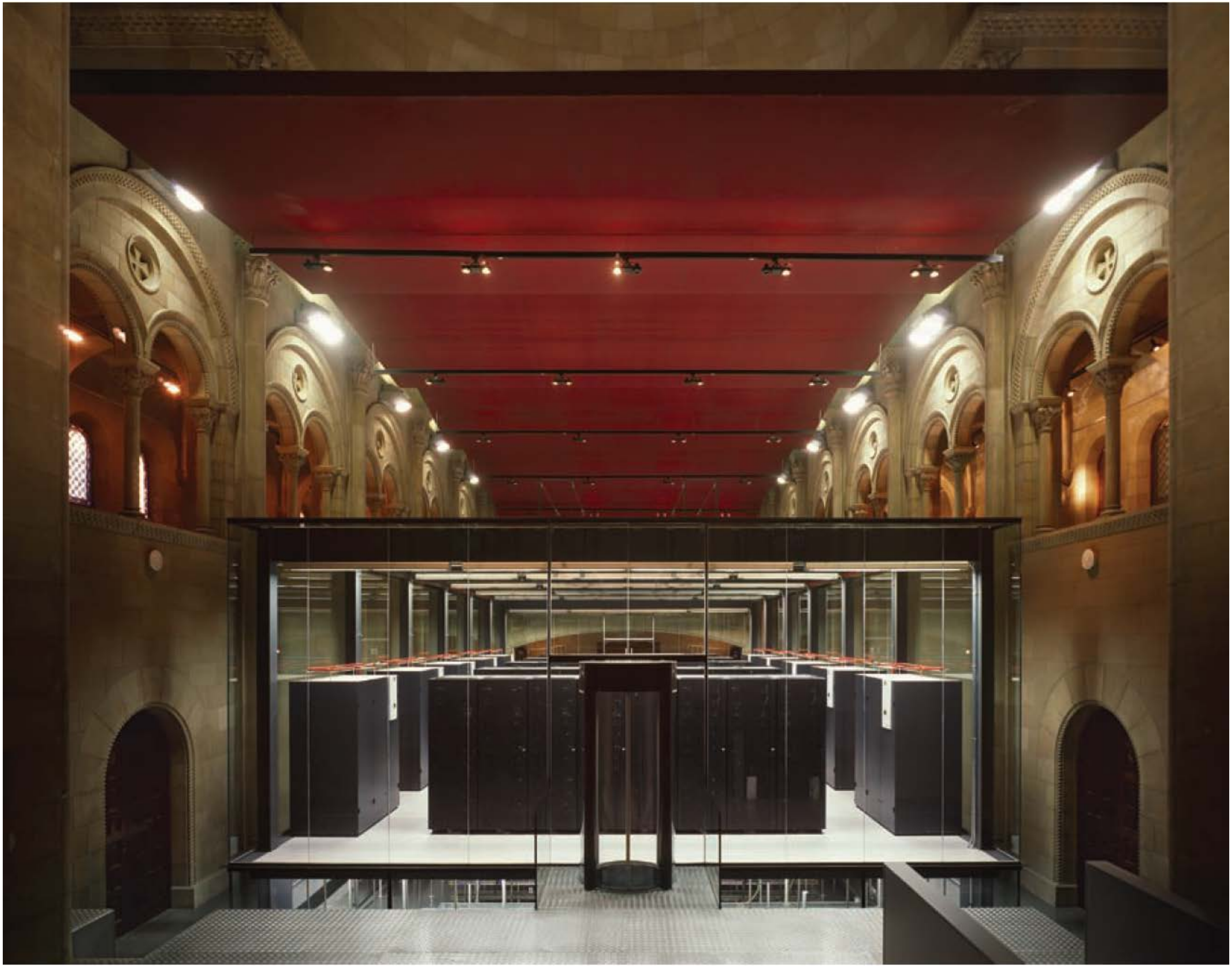
Cool air descends into the cold aisles of Facebook's Data Center.



ESSENTIAL DETAILS

ORGANIZATION:	Barcelona Supercomputing Center - Centro Nacional de Supercomputación
LOCATION:	Barcelona, Spain
ONLINE:	April 2005
NOTABLE FEATURES:	Server environment is located in 1920s chapel and houses the MareNostrum supercomputer
TIME TO DESIGN AND BUILD:	7 months
SIZE:	Overall chapel floor space undetermined. Hosting area is 1,722 sq. ft. (160 sq. m)
POWER:	1.4 MW used by the entire facility, with about 850 kW consumed by computing hardware
TIER:	0
CABINET LOCATIONS:	48
POWER DENSITY:	Varies by cabinet, up to 22 kW maximum per cabinet maximum

"Song of Gulls". Islands. King Crimson. 1971



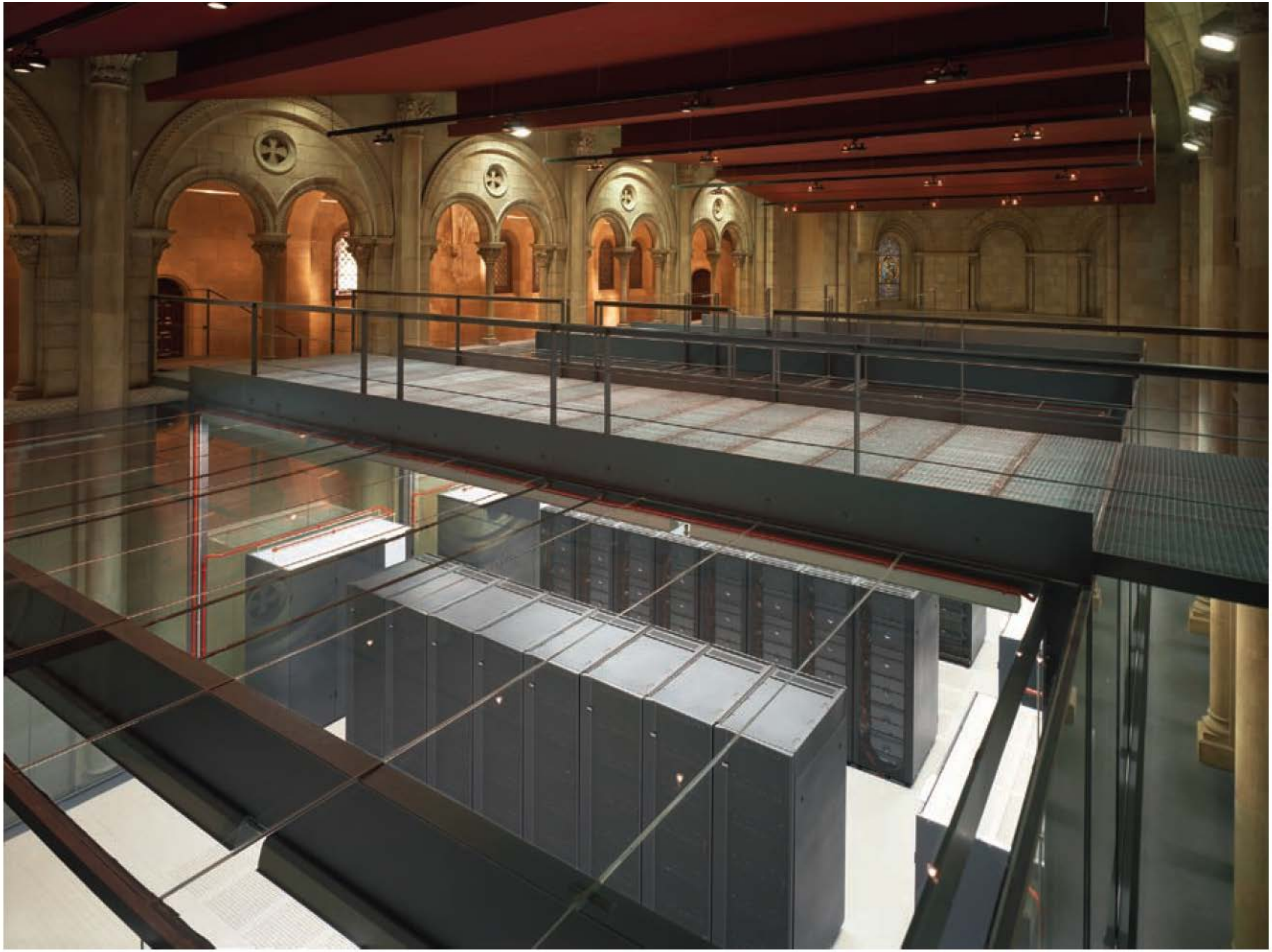
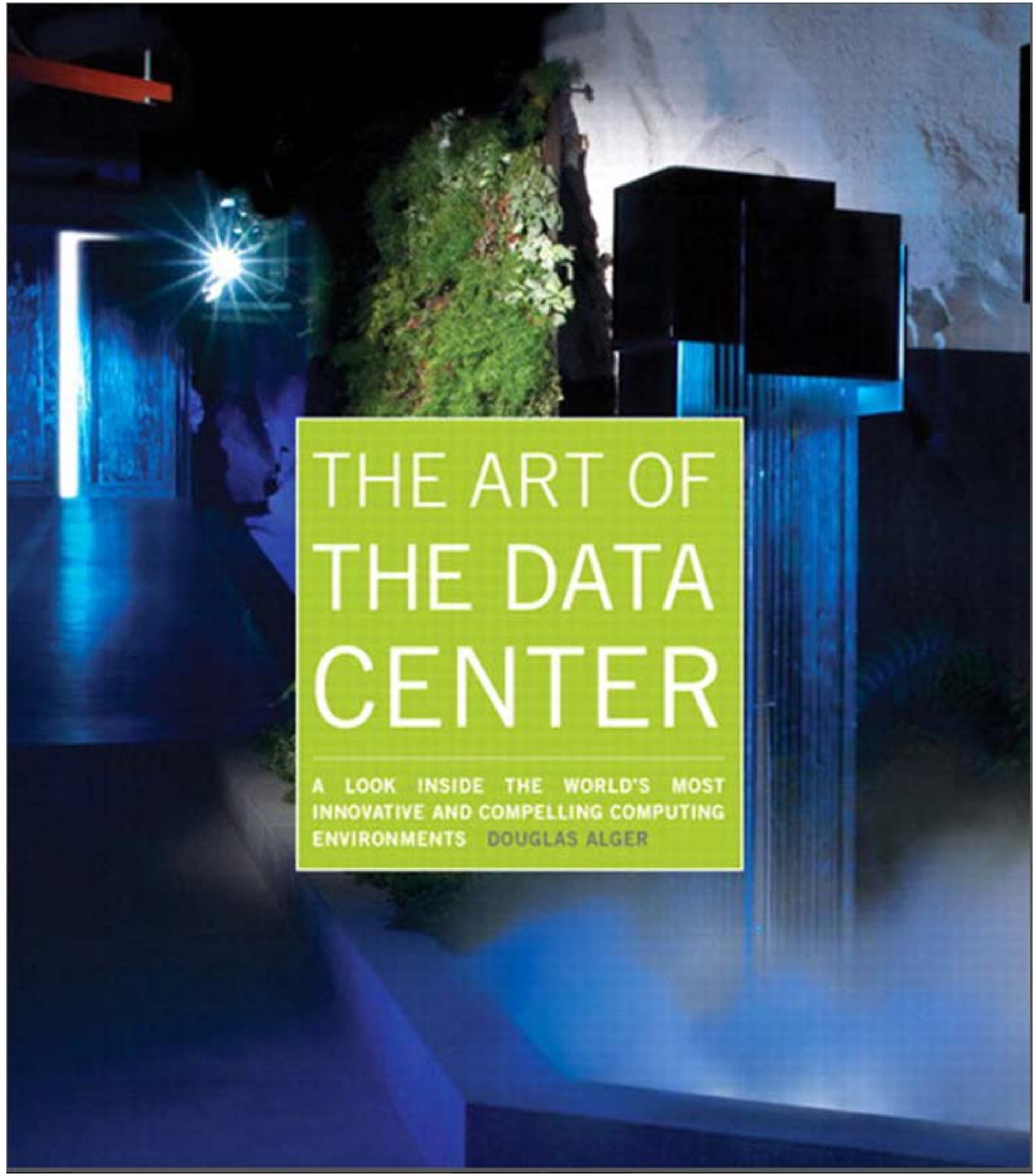


FIGURE 4-4 The height of the supercomputing center's glass walls were chosen to optimize airflow and align with the physical dimensions of the chapel.


IBM



Douglas Alger

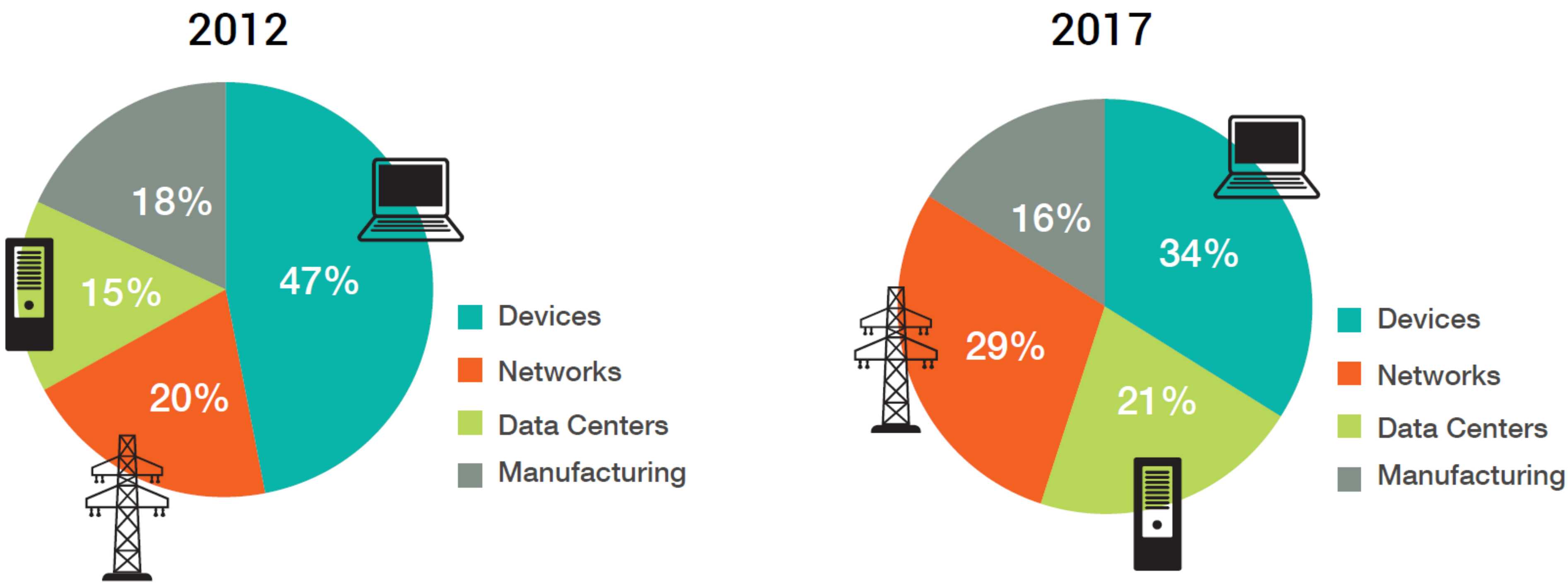
First printing September 2012


Ecología
(poca energía, muy renovable)


"Across the tech sector we need to recognize that data centers will rank by the middle of the next decade among the large users of electrical power on the planet." Brad Smith, President-Microsoft

Consumo de Energía IT

Main components of electricity consumption for the IT sector

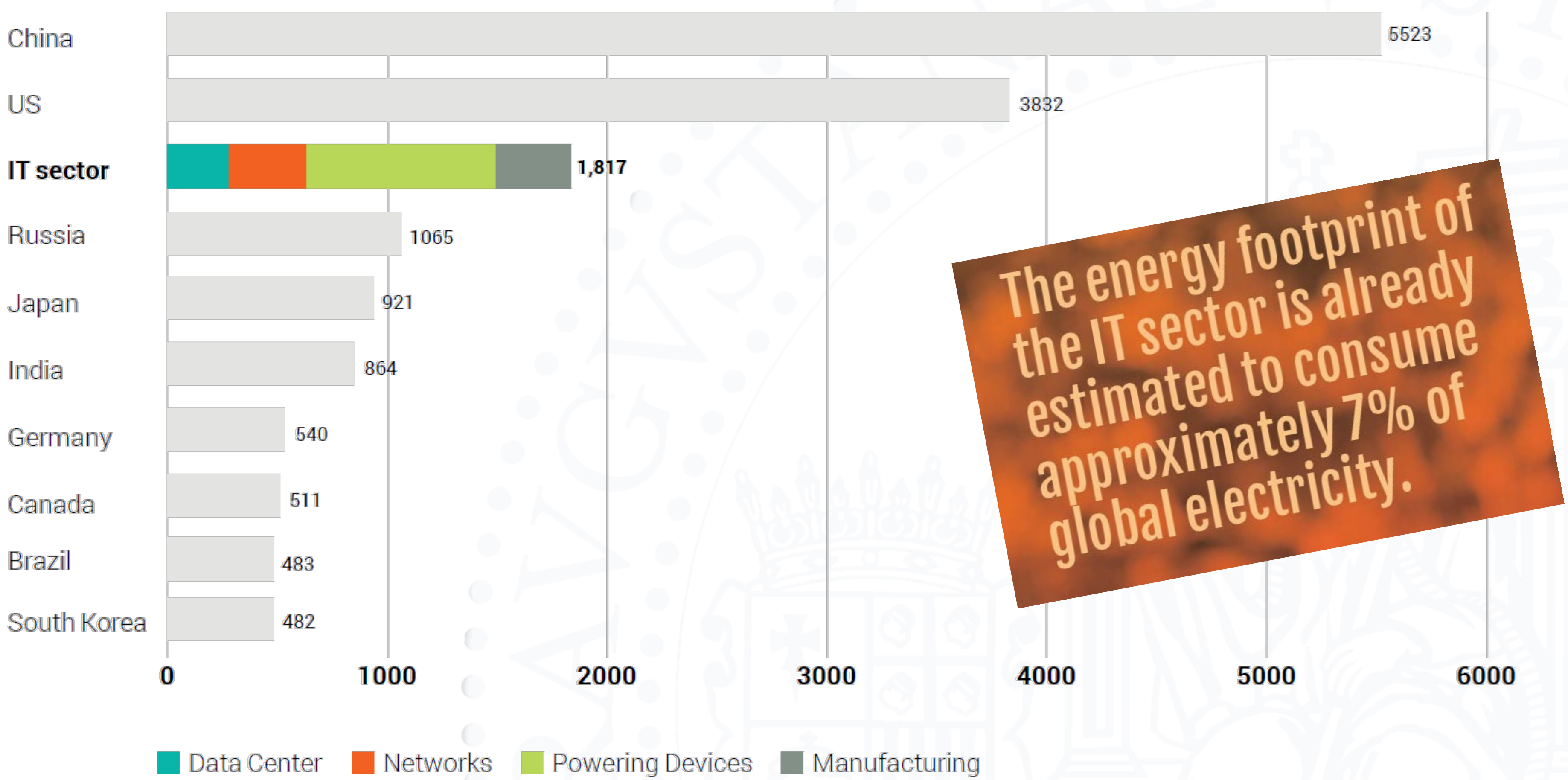


Main components of electricity consumption for the IT sector, 2012. From “Emerging Trends in Electricity Consumption for Consumer ICT”

Clicking clean: who is winning the race to build a green internet?
Green Peace Inc., January 2017.

Consumo de Energía IT

2012 Electricity Consumption; Countries Compared to IT Sector in billion kWh



Source: Emerging Trends in Electricity Consumption for Consumer ICT, Peter Corcoran and Andres Andrae (2013) and CIA World Factbook. China/Russia/Canada figures are from 2014.

Clicking clean: who is winning the race to build a green internet?
Green Peace Inc., January 2017.

2020 IT Carbon Footprint by sub-sector

820 M tons CO2

360 M tons CO2

260 M tons CO_2

2007 Worldwide IT carbon footprint:
2% = 830 M tons CO_2

Comparable to the global aviation industry

Expected to grow to
- **4%** by 2020
- **13%** by 2030

SMART 2020: Enabling the Low Carbon Economy in the Information Age, The Climate Group, 2008.

¿ Qué futuro energético tienen los Centros de Datos ?

- Bajar costes operativos es una prioridad
  para TODOS los tipos de Centros de Datos,
  o sea,
  hay que calcular lo mismo
  consumiendo menos energía
  - Mejorar la EFICIENCIA → mejorar el “PUE”
  - Construir fuentes de energía propias
    - con permiso de las compañías eléctricas, o con abogados –

- Los negocios relacionados con Internet
  Google, Apple, Facebook, Amazon, eBay, Microsoft, YouTube, …
  saben que sus clientes y sus trabajadores
  son proclives a las fuentes renovables de energía,
  o sea,
  construir por construir, construyamos 100% renovables !

PUE

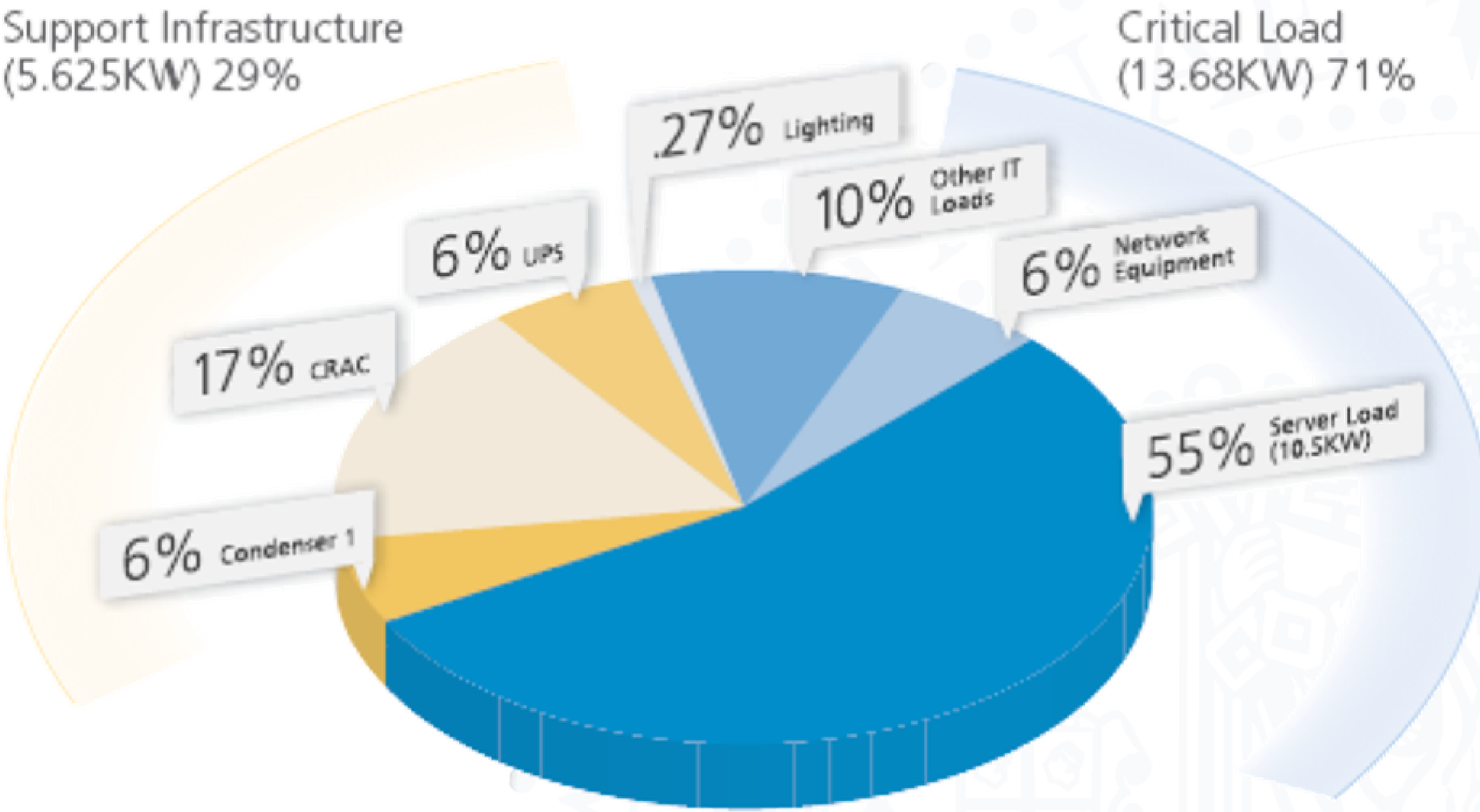
- Power Usage Efectiveness

$$PUE = \frac{P_{TOT}}{P_{IT}} = \frac{P_{IT} + P_{noIT}}{P_{IT}}$$

- P_{IT}
  - Servidores de cálculo
  - Almacenamiento: discos, cintas
  - Redes de datos: entre servidores, conexión a Internet
- P_{noIT}
  - Climatización
  - Pérdidas SAI
  - Pérdidas en transformación AC-AC
  - Iluminación, seguridad, sistema anti-incendios
  - Pérdidas distribución eléctrica

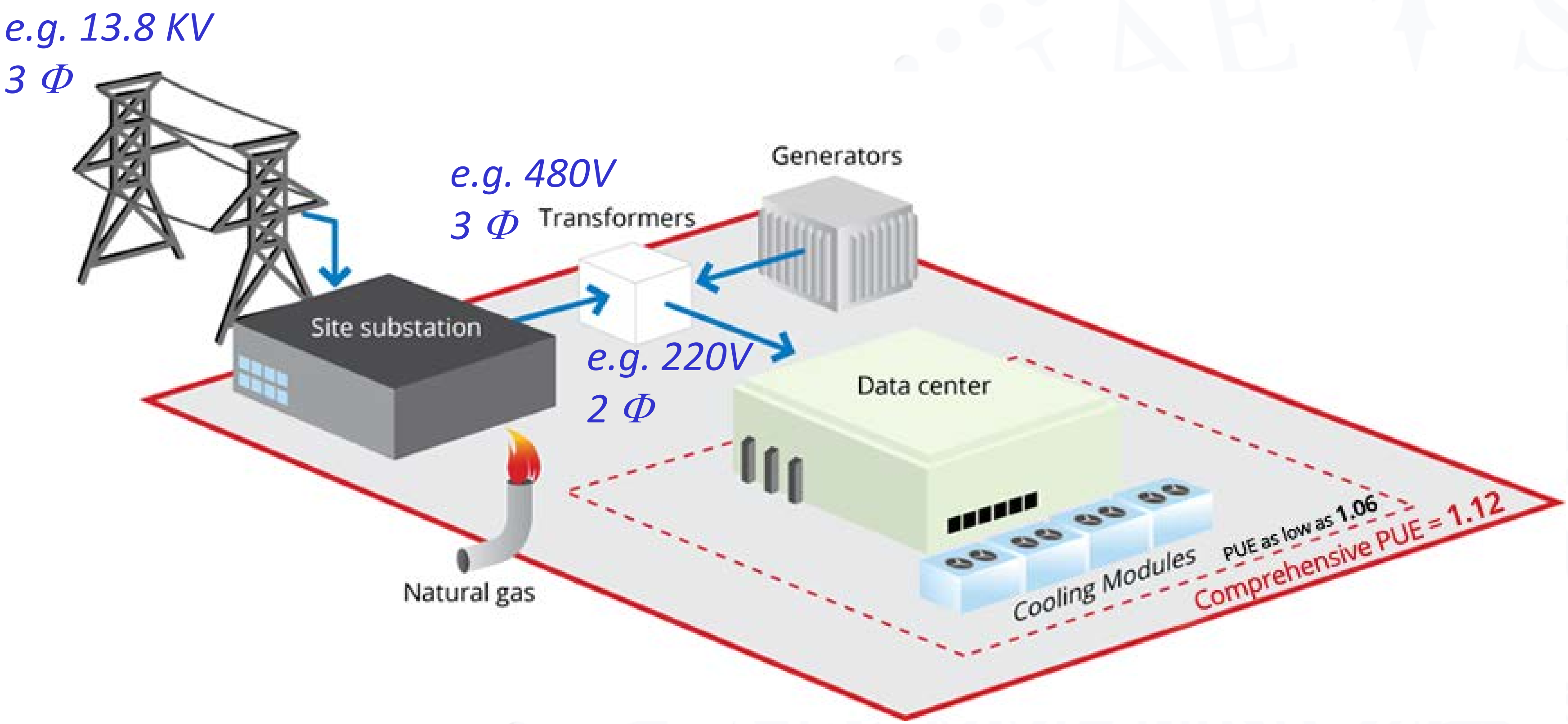
The Green Grid Data Center Power Efficiency Metrics: PUE and DCiE, White Paper #6, 2007

Example1: Raritan Production DC–New Jersey (≈2009)



$$\text{PUE} = \frac{\text{Total Power} = 19.3\ \text{kw}}{\text{Total IT Equipment Power} = 13.7\ \text{kw}} = 1.4$$

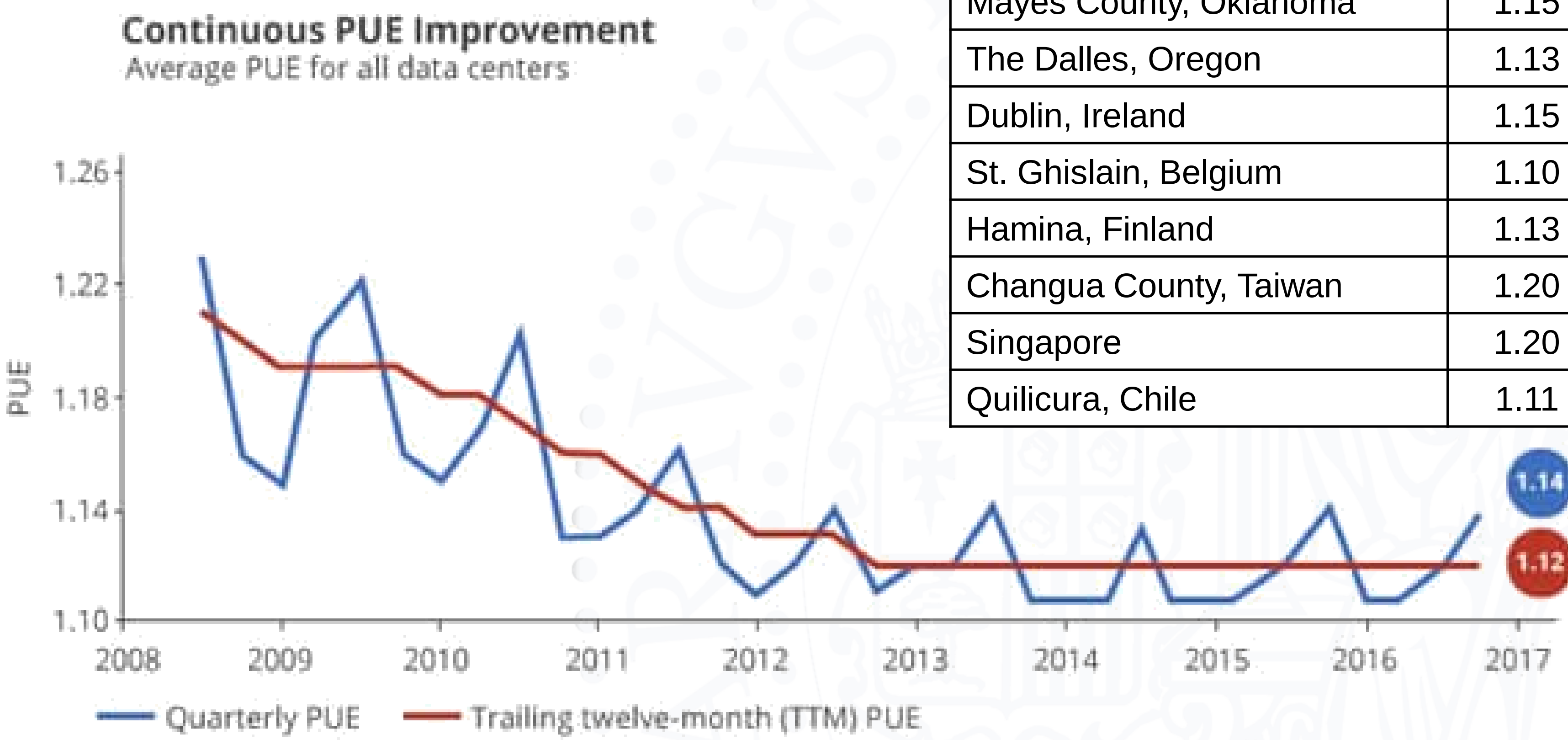
Google Data Centers PUE measurement



“The average PUE for all Google data centers is 1.12, although we could achieve a 1.06 PUE with narrower boundaries”

Google Data Centers

- PUE data
  for all large-scale
  Google data centers



Q3 2016 performance	PUE	
Campuses	3-month	12-month
Douglas County, Georgia	1.16	1.12
Lenoir, North Carolina	1.12	1.11
Berkeley County, South Carol.	1.16	1.13
Council Bluffs, Iowa	1.14	1.10
Council Bluffs, Iowa-2	1.14	1.12
Mayes County, Oklahoma	1.15	1.12
The Dalles, Oregon	1.13	1.13
Dublin, Ireland	1.15	1.15
St. Ghislain, Belgium	1.10	1.09
Hamina, Finland	1.13	1.12
Changua County, Taiwan	1.20	1.18
Singapore	1.20	1.21
Quilicura, Chile	1.11	1.12

100% Renewable Energy Commitments

2011

- December: Facebook

2012

- May: Apple
- September: Google

2013

- February: rackspace (the open cloud company)
- March: salesforce, box

2014

- November: amazon web services
- October: Microsoft

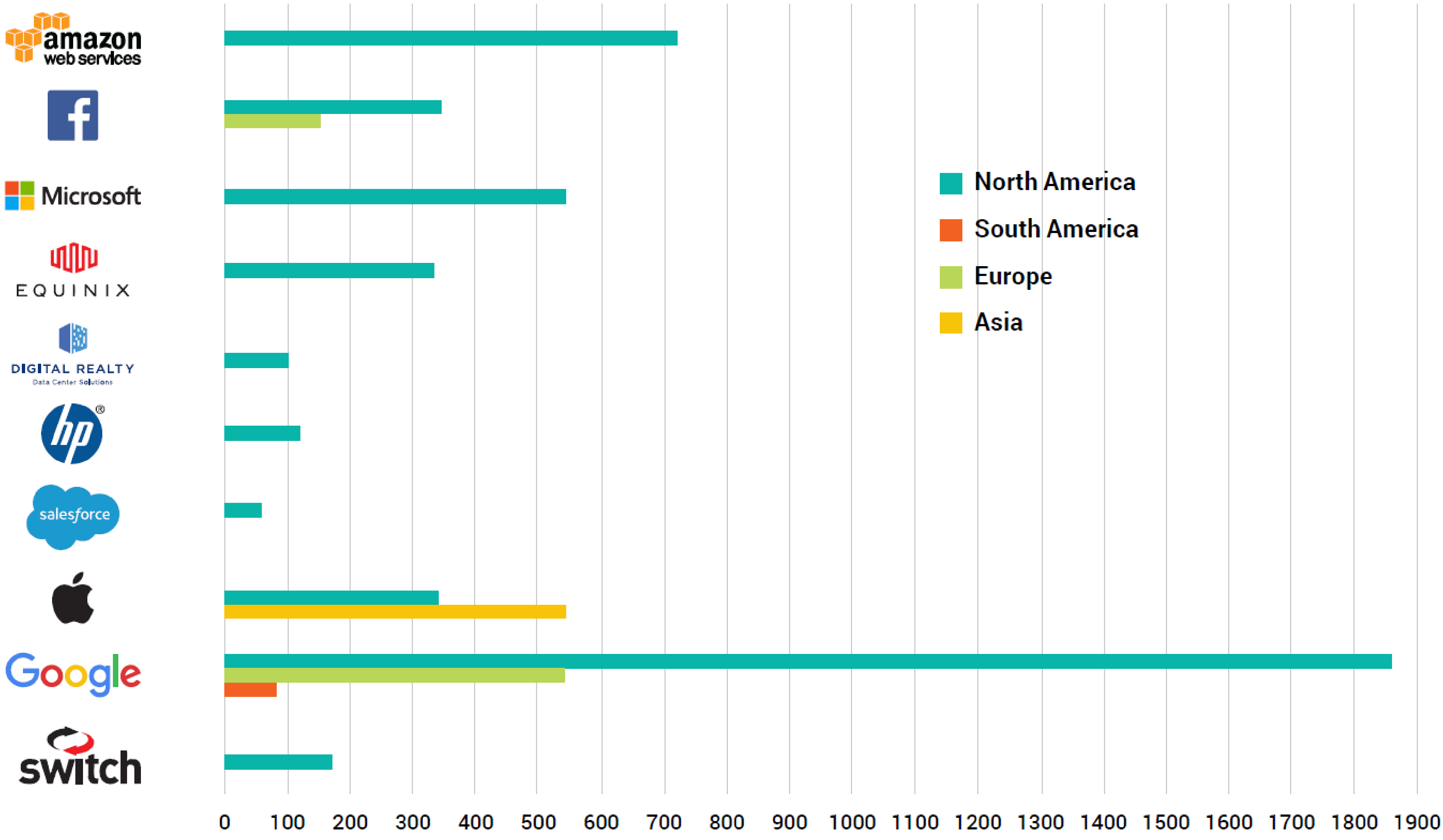
2015

- March: EQUINIX
- May: DIGITAL REALTY (Data Center Solutions)
- June: NAVER
- November: switch
- December: Adobe

2016

- March: hp
- April: Etsy
- September: Hewlett Packard Enterprise

IT Sector Renewable Contracts 2010-16 (MW)



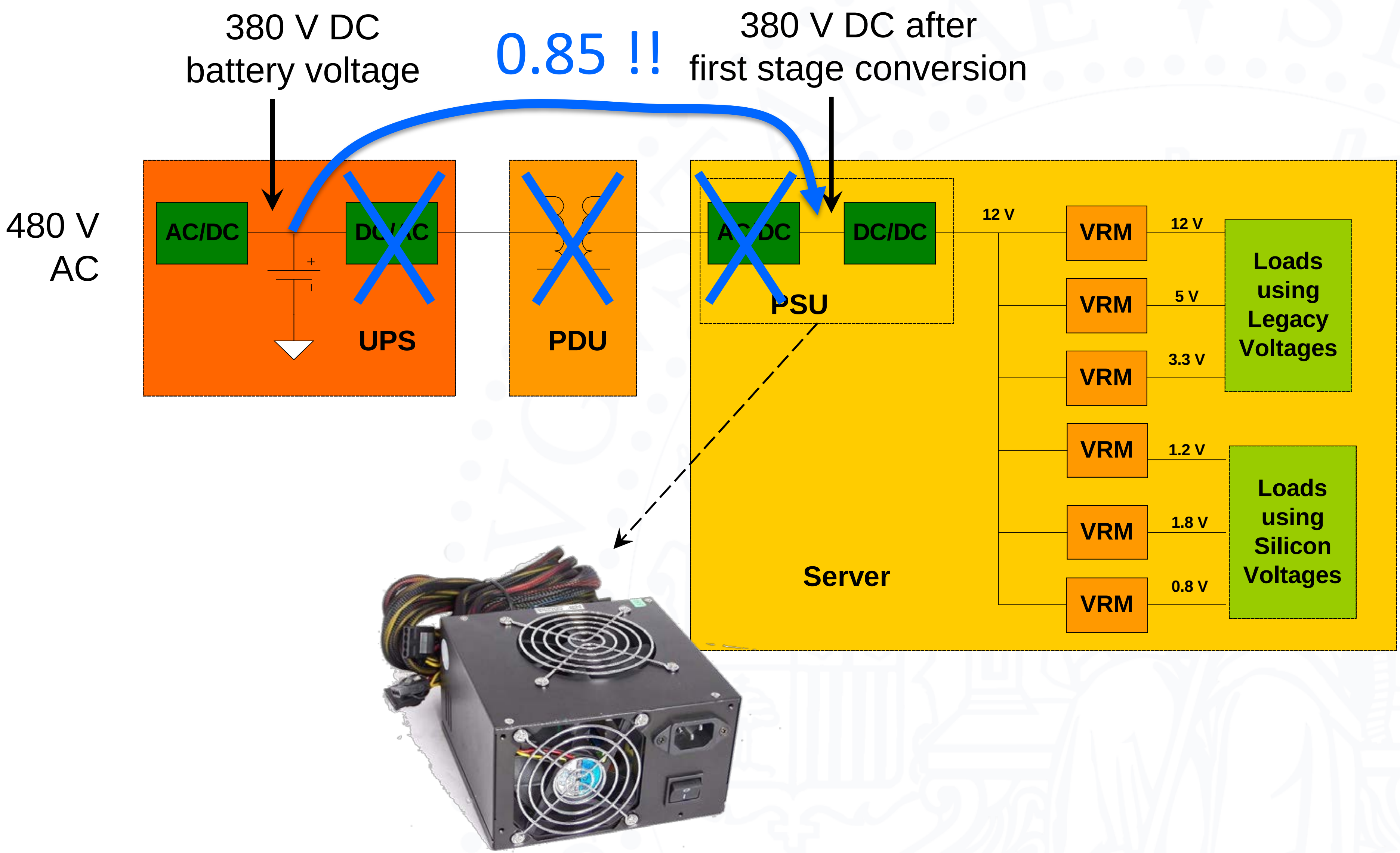
- www.reviewjournal.com/business/energy/switch-reaches-agreement-nevada-regulators-leave-nv-energy
- December 9, 2016 - 4:25pm. **Switch reaches agreement with Nevada regulators to leave NV Energy.** By SEAN WHALEY
- CARSON CITY — Data storage company Switch and state regulators have reached an agreement to allow the company to leave as a retail customer of NV Energy. The agreement, which must be approved by the Nevada Public Utilities Commission (PUC), would essentially reverse the PUC's 2015 denial of Switch's original "exit application." The company could pay as much as $27 million in an "exit fee" to leave as a customer of the utility, which does business in Southern Nevada as Nevada Power. The fee is intended to protect remaining utility customers for investments already made in electricity infrastructure to meet demand.
- The company would leave by Aug. 1 at the latest. Switch made no comment on the proposed agreement, which is expected to come up for a vote this month. The agreement also allows Switch to secure its own energy supplies for its massive data storage facility being built at the Tahoe Reno Industrial Center near the Tesla gigafactory east of Reno. Switch's Northern Nevada data center is expected to cost $3 billion when fully built and will house seven buildings totaling 6.49 million square feet. The project broke ground in September 2015. www.datacenterdynamics.com/app-cloud/switch-plans-largest-data-center-in-the-world-in-reno/94871.article
- Switch in 2014 became the first company in years to file an exit application with the PUC, relying on a 2001 law passed by the Nevada Legislature. But the PUC denied the application on a 2-1 vote, citing concerns about whether the amount of the exit fee was appropriate and would protect remaining ratepayers. Since then, however, three large gaming companies — Wynn Resorts Ltd., MGM Resorts International and Las Vegas Sands Inc. — were approved to leave as retail customers of the utility. Wynn and MGM did so Oct. 1, while Sands opted not to pursue a departure. MGM was required to pay an $86.9 million exit fee and Wynn had to pay $15.7 million. Just last month Caesars Entertainment Corp. also applied to withdraw as a retail customer of NV Energy and purchase power on the wholesale market. The intent of the 2001 law was to lessen pressure on electricity rates during an energy crisis that was occurring at the time. Those circumstances no longer exist. Instead, large companies that exit are able to negotiate more favorable rates for power.
- Currently only large utility customers can leave under the 2001 law. But Nevada voters in 2018 will be asked to weigh in on a measure that would clear the way for all utility customers large or small to seek out cheaper electricity prices through competition. Question 3 passed in November but must win approval a second time before the Legislature could enact a law opening the process to competition for everyone.
- After Switch was denied, it reached an agreement with Nevada Power to stay with the utility and get all of its electricity from renewable energy at two Southern Nevada large-scale solar facilities called Switch 1 and Switch 2. The stipulation filed with the PUC dated Thursday says Switch will continue to take power from both of these facilities.
- Despite this agreement, Switch in July filed suit in U.S. District Court against the PUC and NV Energy over the denial of its exit application, arguing that others were allowed to leave. The lawsuit asserted that approvals of other exit applications constituted a violation of Switch's 14th Amendment right to equal protection under the law. The stipulation does not address this lawsuit.
- Founded in 2000, Switch says its patented data centers are globally recognized for their revolutionary high-density design, superscale cloud campus, unparalleled telecommunications services and industry-leading sustainability.

Muchas Gracias !!

« Ce n’est qu’une forêt de champignons », dit-il.
Voyage au centre de la Terre
1864
Jules Verne

Backup

Typical AC Distribution Today



Centros de datos, potencia y refrigeración[1]

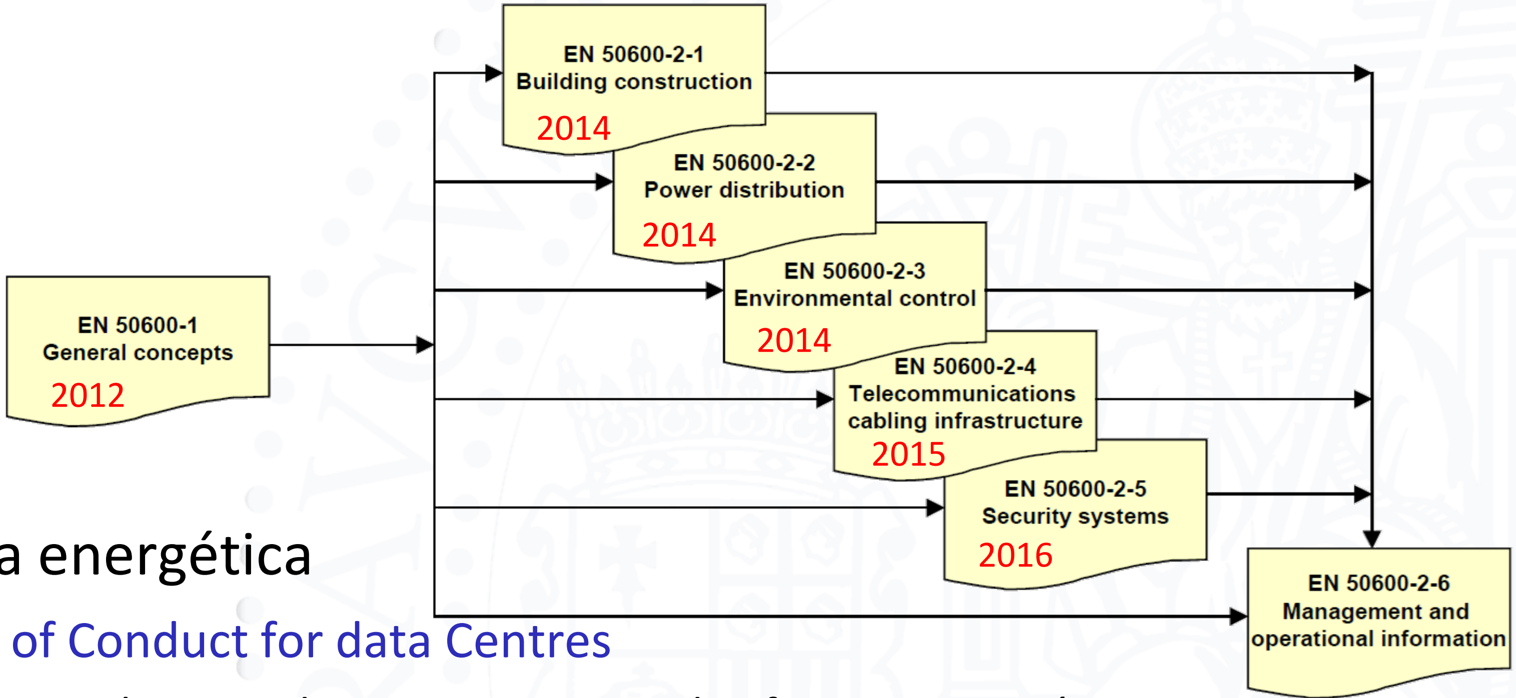
Potencia por rack de 42U (Kw)	Empresa convencional CDs propios o *collocation* Horizonte de 10 años	Gran volumen y dependencia de IT	Proveedores de servicios IT y Supercomput.	Tipo de refrigeración	
21			- Racks completamente llenos		in-rack
20					
19					
18					
17					
16				in-row	
15					
14		p.e. - 2 Blade de 6U/rack: 6Kw - 20 servidores/rack de 1U-500w: 10Kw			
13					
12					
11					
10	equipos nuevos				
9					
8				Gestión de aire convencional: pasillo frio - pasillo caliente	
7					
6 (media max)	equipos viejos				
5 (media hoy)					
4 (media min)					
3					

Organismos y sus normas, fuera de Europa

- Pautas de diseño (best practices)
  - The **Uptime Institute** (1993, USA)
    - Uptime Institute Tier Ratings for Data Centres → referencia en **certificación** ($)
  - Telecommunications Industry Association (**TIA**, 1988, USA)
    - ANSI/TIA 942 - Infrastructure Standard for Data Centres → referencia en **diseño**
  - Building Industry Consulting Services International (**BICSI**, 1974, USA) → referencia en **disponibilidad**
    - ANSI/BICSI 002-2014 - Data Centre Design and Implementation Best Practices
  - American Society of Heating, Refrigeration and Air Conditioning Engineers (**ASHRAE**, 1894, USA)
    - TC9.9 Guidelines for Mission Critical Facilities) → referencia en **condiciones ambientales**
- Códigos de conducta en operación
  - Association for Computer Operations Management (**AFCOM**, 1980, USA)
    - best practices, insights, tools and resources to successfully manage today's data center infrastructure
- Eficiencia energética
  - The **Green Grid** (2007)
  - US Departament of Energy (**DOE**)
  - US Environmental Protection Agency (**EPA**)

Organismos y sus normas en Europa

- Pautas de diseño y operación: EN 50600 series (60-70€/documento)
  - Desarrollado por el CENELEC:
    European Committee for Electrotechnical Standardization



- Eficiencia energética
  - EU Code of Conduct for data Centres
    - programa voluntario de mejora continua de eficiencia energética, partiendo de lo que hay…

Recordemos, definiciones de Centro de Datos

- ANSI/TIA 942-2005
  Telecommunications Infrastructure Standard for Data Centers
  - This Standard specifies the minimum requirements for telecommunications infrastructure of data centers and computer rooms including single tenant enterprise data centers and multi-tenant Internet hosting data centers

    ***data center**: a building or portion of a building whose primary function is to house a computer room and its support areas*

- Code of Conduct on Data Centres Energy Efficiency v. 2.0-2009
  European Commission, … , Institute for Energy, Renewable Energies Unit

    *the term "**data centres**" includes all buildings, facilities and rooms which contain enterprise servers, server communication equipment, cooling equipment and power equipment, and provide some form of data service (e.g. large scale mission critical facilities all the way down to small server rooms located in office buildings)*

About PUE™ Invention and Development

- The **Green Grid Association** is a non-profit, open industry consortium of end users, policy makers, technology providers, facility architects, and utility companies working to improve the resource efficiency of information technology and data centers throughout the world
- PUE is a Trade Mark belonging to the Green Grid Association, explained in:
  The Green Grid Data Center Power Efficiency Metrics: PUE and DCiE, White Paper #6, **2007**
  www.thegreengrid.org/Global/Content/white-papers/The-Green-Grid-Data-Center-PowerEfficiency-Metrics-PUE-and-DCiE
- *PUE™: A Comprehensive Examination of the Metric*, White Paper #49, **2012**
  - "Since its original publication in 2007, PUE has been globally adopted by the industry. Over the past years, The Green Grid has continued to refine the metric measurement methodology with collaborative industry feedback. This collective work has been brought together here to simplify the absorption and use of the PUE metric. To produce this document, The Green Grid consolidated all its previously published material related to PUE and included new material as well.
    …
    This document allows executives to gain a high level of understanding of the concepts surrounding PUE, while providing in-depth application knowledge and resources to those implementing and reporting data center metrics."

Load levels in data centers (2)

Typical shared WSC, which mixes several types of workloads and includes online services.

Google's most highly utilized WSCs, where large continuous batch workloads execute.

WSC = Warehouse-Scale data Centers

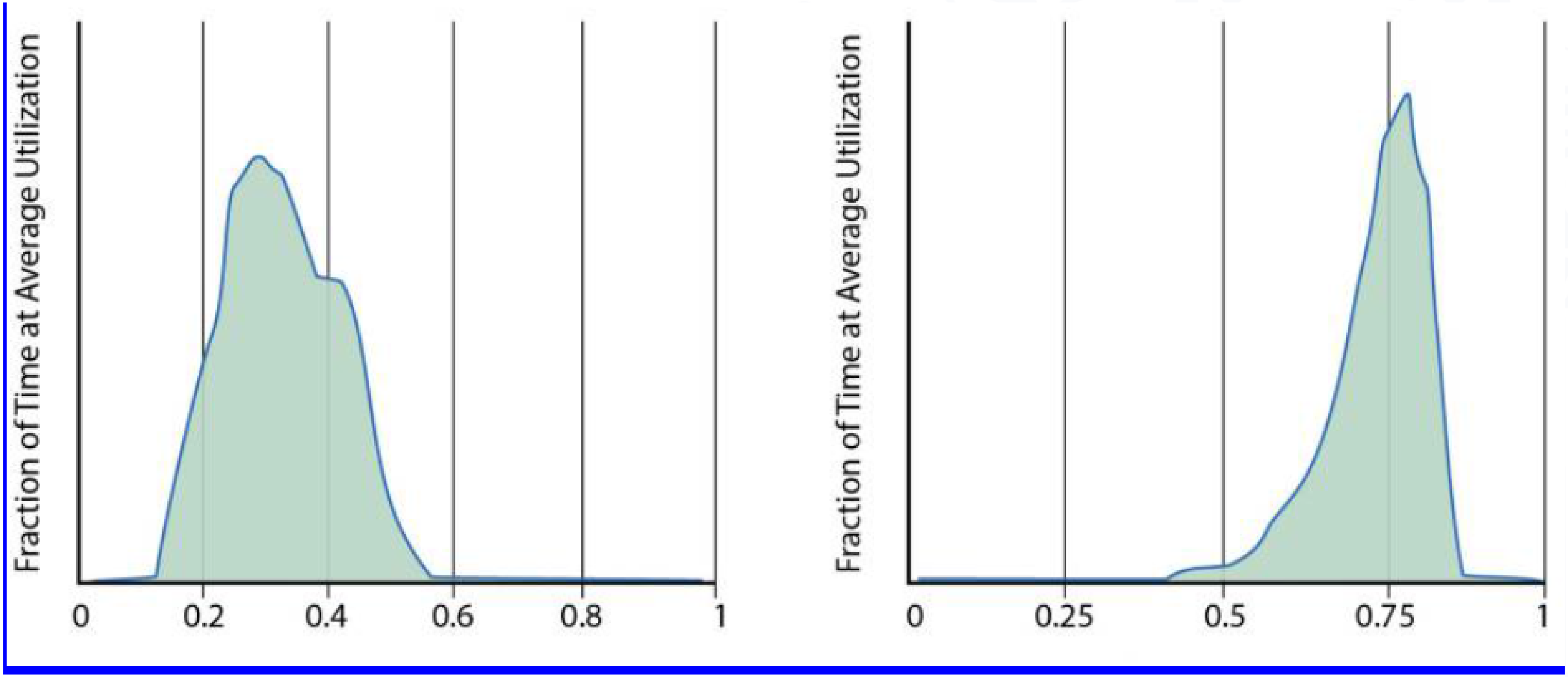
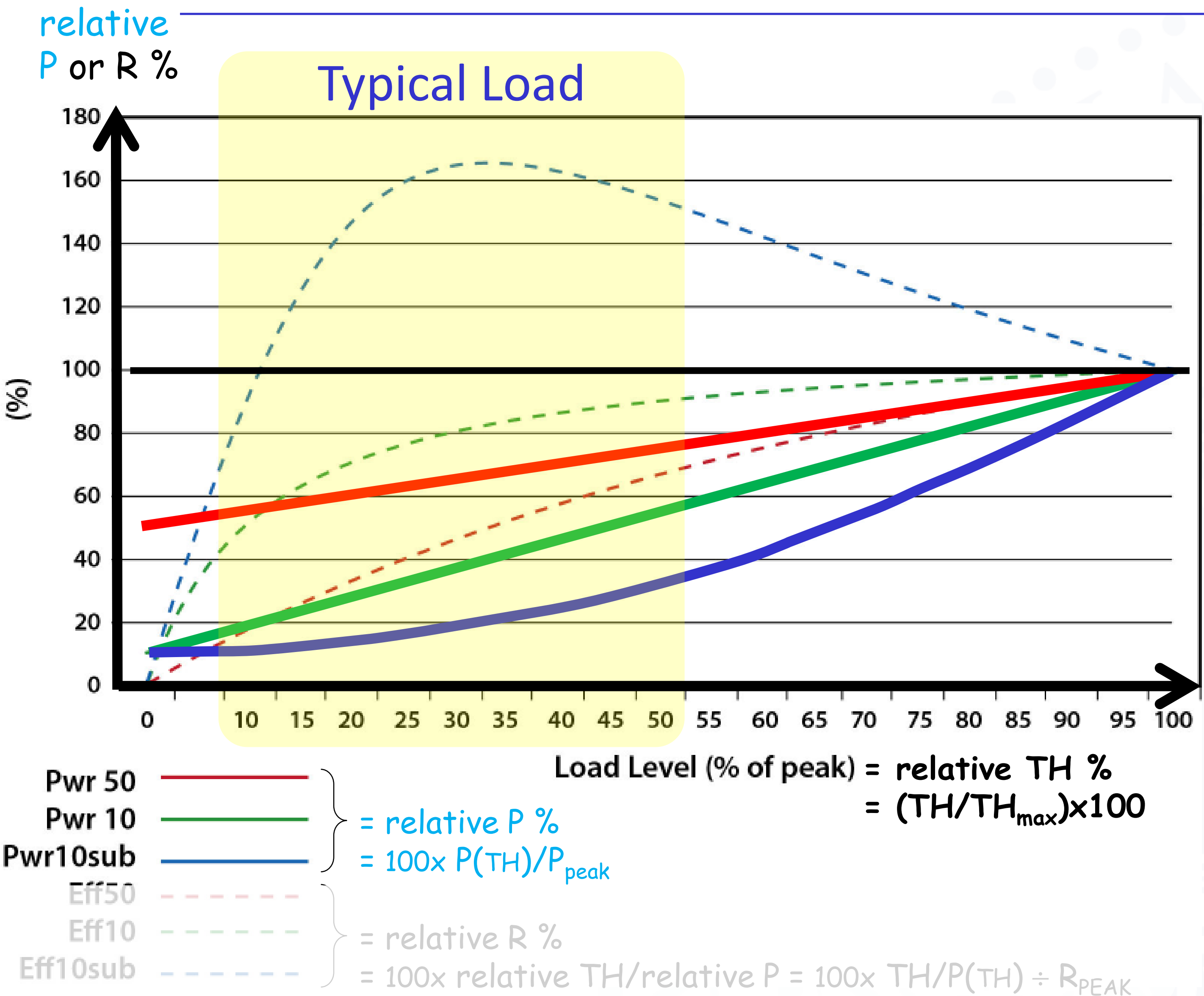


Figure 5.4: Average activity distribution of a sample of two Google clusters, each containing over 20,000 servers, over a period of 3 months (January-March 2013).

Figure 5.4 of "The Data center as a Computer", 2nd ed

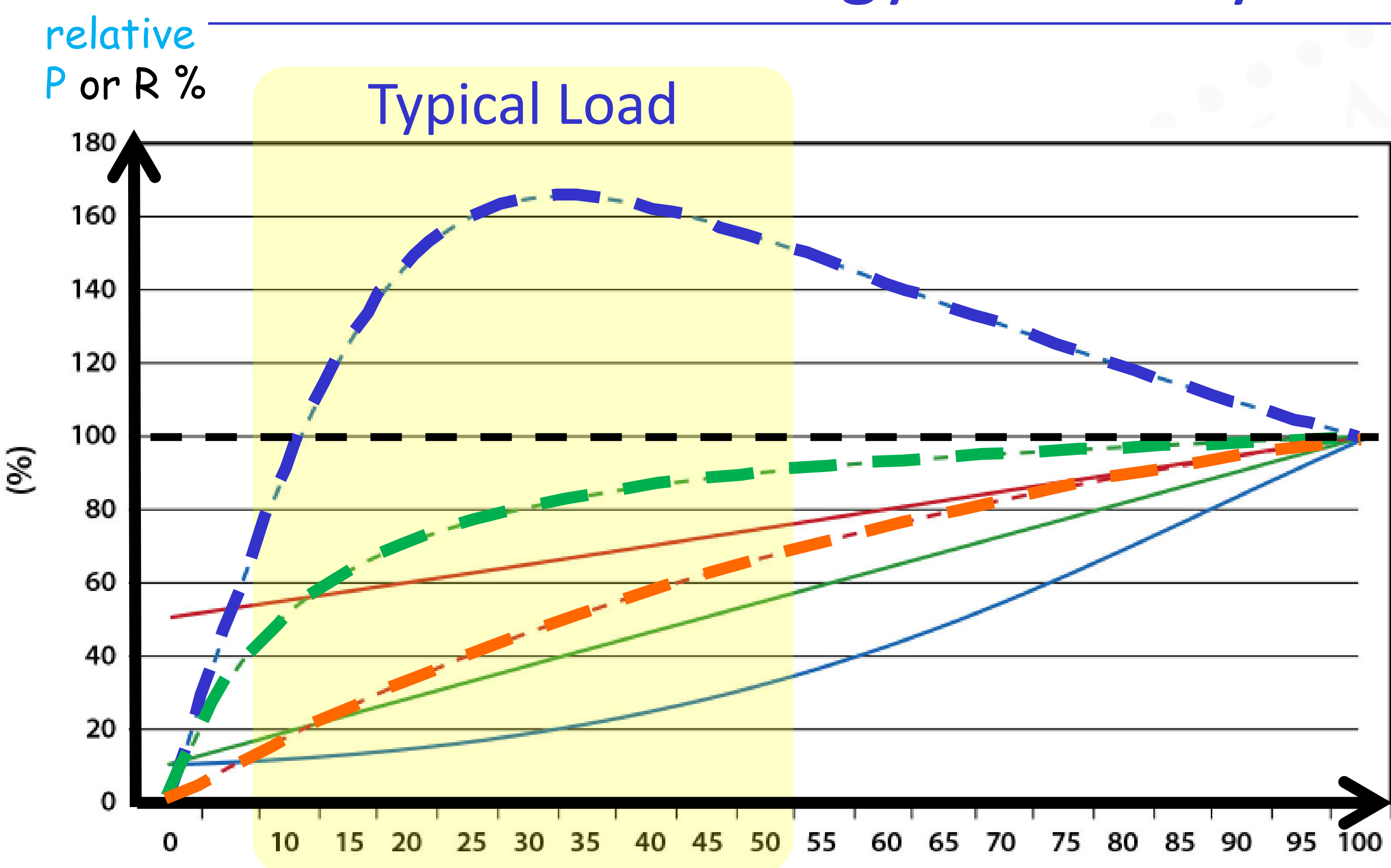
Power and energy efficiency vs. load level



- Three systems with the same power and throughput at peak: | | |
- Solid lines are *relative* power **P**: 100x P(TH)/Ppeak
- Pwr50: a typical server with idle power at 50% of peak
- Pwr10: a more energy-proportional server with idle power at 10% of peak
- Pwr10sub: a sublinearly energy-proportional server with idle power at 10% of peak

Figure 5.5 of "The Data center as a Computer", 2nd ed

Power and energy efficiency vs. load level



Load Level (% of peak) = relative TH %

= $(TH/TH_{max})\times 100$

Pwr 50

Pwr 10

Pwr10sub

= relative P %

= 100x P(TH)/P_{peak}

Eff50

Eff10

Eff10sub

= relative R %

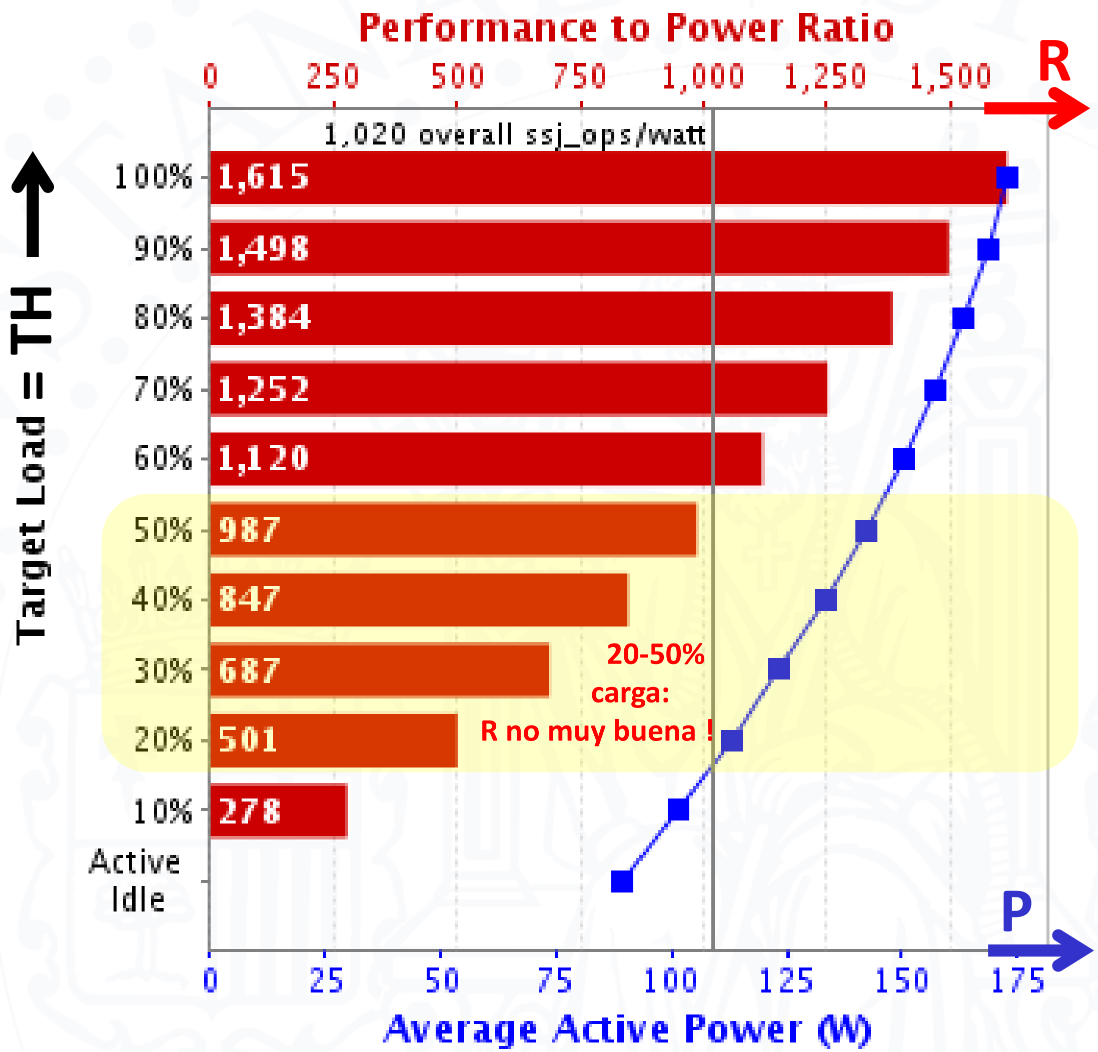
= 100x relative TH/relative P = 100x TH/P(TH) ÷ R_{PEAK}

- Three systems with the same power and throughput at peak: | | |
- Solid lines are *relative* power **P**: 100x P(TH)/Ppeak
- Dashed lines are *relative* efficiency **R**: 100x TH/P(TH) ÷ R_{PEAK}
- Pwr50: a typical server with idle power at 50% of peak
- Pwr10: a more energy-proportional server with idle power at 10% of peak
- Pwr10sub: a sublinearly energy-proportional server with idle power at 10% of peak

Figure 5.5 of "The Data center as a Computer", 2nd ed

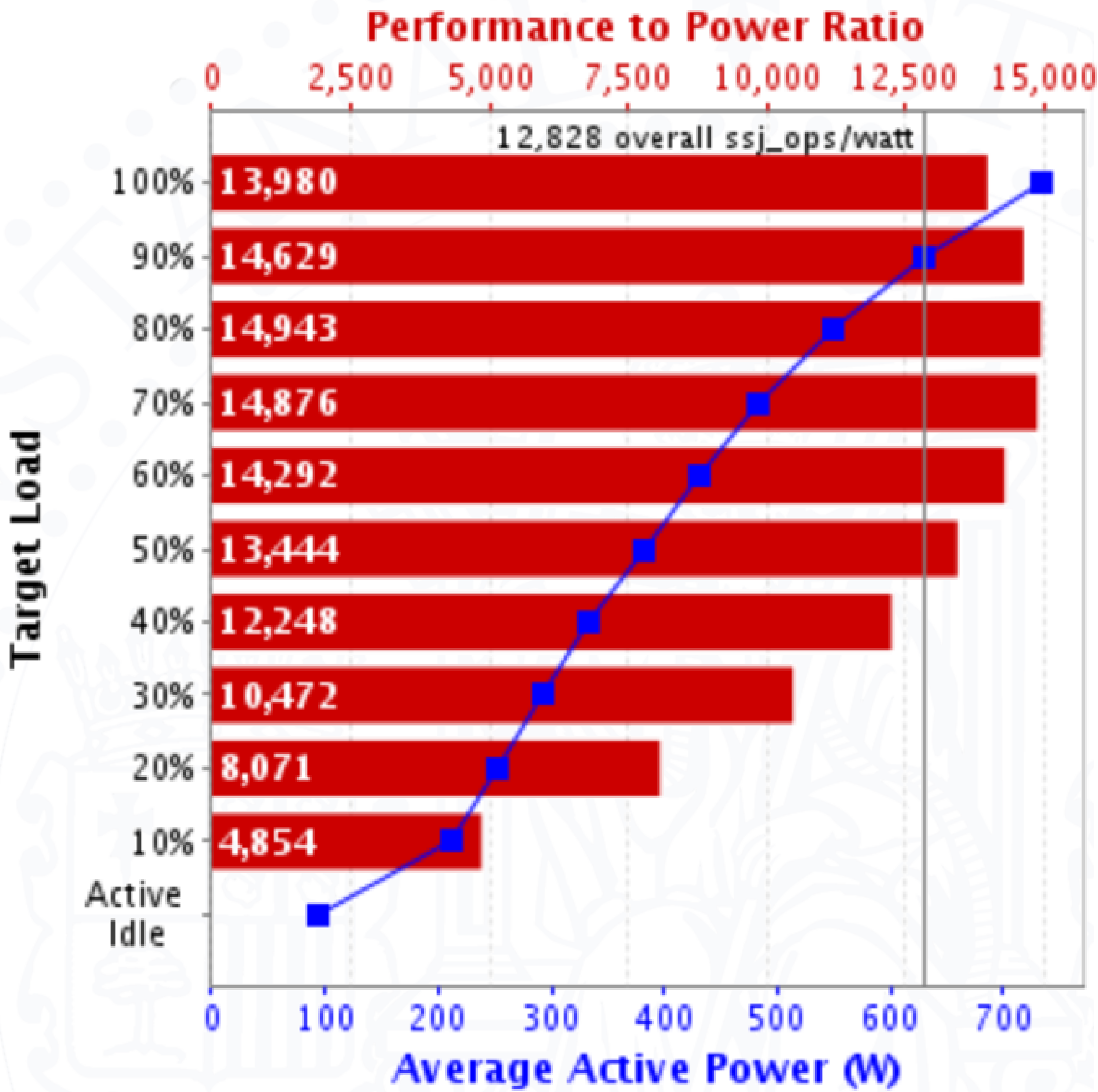
SPECpower_ssj® ASUS RS160-E5 (2008)

Performance			Power	Performance to Power Ratio
Target Load	Actual Load	ssj_ops	Average Active Power (W)	
100%	99.2%	278,927	173	1,615
90%	89.7%	252,327	168	1,498
80%	80.4%	226,095	163	1,384
70%	70.0%	196,907	157	1,252
60%	59.8%	168,211	150	1,120
50%	49.9%	140,335	142	987
40%	40.1%	112,784	133	847
30%	30.1%	84,723	123	687
20%	20.1%	56,512	113	501
10%	10.0%	28,193	102	278
Active Idle		0	89.4	0
			∑ssj_ops / ∑power =	1,020.9



SPECpower_ssj2008 - FUJITSU Server PRIMERGY RX4770 M4 (2018)

Performance			Power	Performance to Power Ratio
Target Load	Actual Load	ssj_ops	Average Active Power (W)	
100%	99.9%	10,249,832	733	13,980
90%	90.0%	9,232,029	631	14,629
80%	80.0%	8,208,471	549	14,943
70%	70.0%	7,180,090	483	14,876
60%	59.9%	6,151,365	430	14,292
50%	50.0%	5,128,907	382	13,444
40%	39.9%	4,096,830	334	12,248
30%	30.0%	3,074,915	294	10,472
20%	20.0%	2,047,755	254	8,071
10%	10.0%	1,026,271	211	4,854
Active Idle		0	94.9	0
			∑ssj_ops / ∑power =	12,828



Form Factor:	2U
CPU Name:	Intel Xeon Platinum 8176M
CPU Characteristics:	28-Core, 38.5MB L3 Cache
CPU Frequency (MHz):	2.1 GHz
CPU(s) Enabled:	112 cores, 4 chips, 28 cores/chip
Memory:	24 x 8 GB DDR4