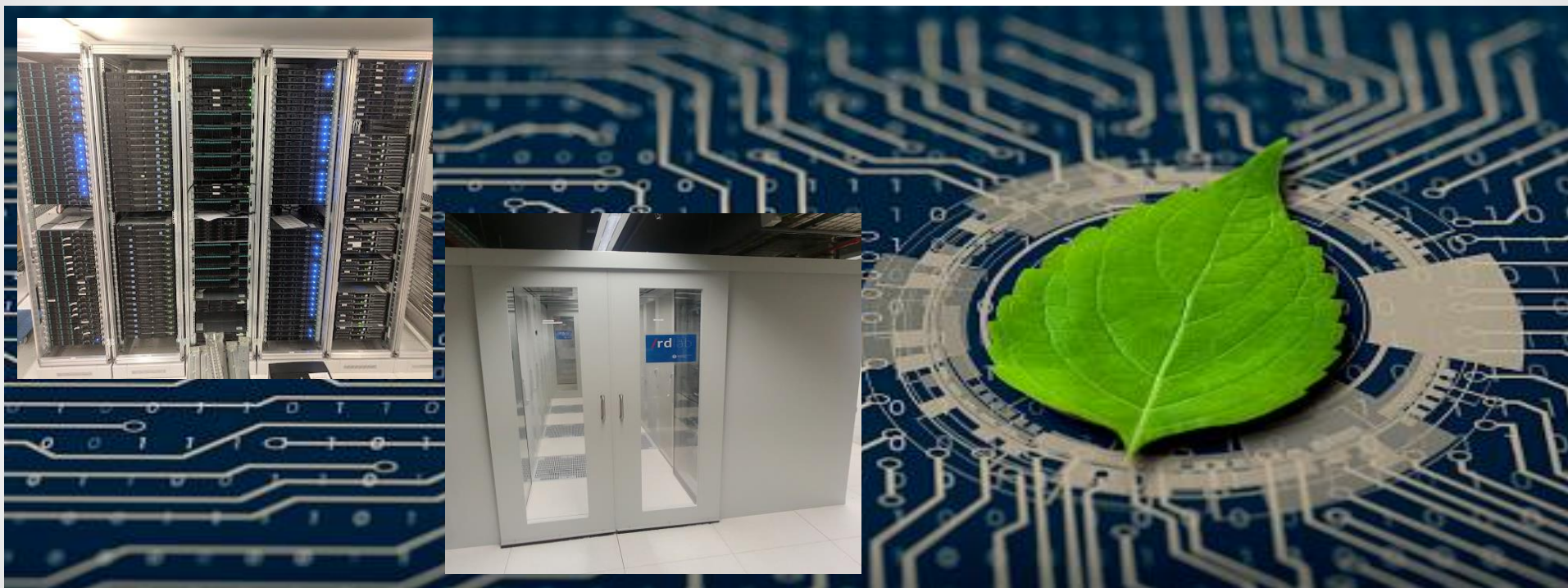


Gestión energética eficiente en grandes infraestructuras TIC



TECNIRIS 81
Octubre 2022



Red IRIS



Gabriel Verdejo Álvarez
gabriel@cs.upc.edu
<https://rdlab.cs.upc.edu>

Índice

- Punto de partida	3
- Visión holística	4
- Eficiencia energética: Arquitectura	5
- Eficiencia energética: Servicio	6
- Eficiencia energética: Entorno físico	7
- Eficiencia energética: Consideraciones	8
- EJ1: Ahorro energético en el clúster del DAC	11
- EJ2: Ahorro energético /rdlab	12
- Eficiencia energética: ¿Como hacerlo?	16
- El reto 10-10-10	19
- Bibliografía	21

Punto de partida

- **Importancia real de la cuestión**

- Ej: Estudio de OVO Energy UK, "Think before you thank"

- Outsourcing: El cloud son "las máquinas de otro", ¡pero el CO2 producido es de todos!

- **Expectativas y realidades**



Visión holística

- **Grandes infraestructuras TIC**
- **Margen operativo de trabajo en 3 niveles**

- Físico: Servidores (¿redes? PoE)
- Servicio: Prestaciones a usuarios y administradores
- Arquitectura: Paradigma de orquestación

- **Estrategia**

- Estrategia energética de les TIC $\subset$ Estrategia energética
- Inversiones + gestión del transitorio = Futuro



Quedan fuera del ámbito de esta presentación los temas referentes a las infraestructuras CPD así como las estrategias

Eficiencia energética: Arquitectura

- **Paradigma TIC tradicional (standalone)**

- Servidor físico = servicio (entornos altamente cohesionados)
- Redundancia **física** de todos los componentes (FA, discos locales, red...)
- Calidad de servicio rígida
- instalaciones “sobrevenidas”

Servidores no enrackables, aires condicionados domésticos, sin cerramientos...

- **Modelo agregado escalable (clustered)**

- Desacoplamiento entre servidores físicos y servicios
- Redundancia: **Lógica** a nivel de servicio y **Física** de componentes vitales
- Compartición de recursos $\Leftrightarrow$ Escalabilidad/adaptabilidad a la demanda real (Entornos virtuales/Cloud, storage, snapshots vs backup tradicional, federación)

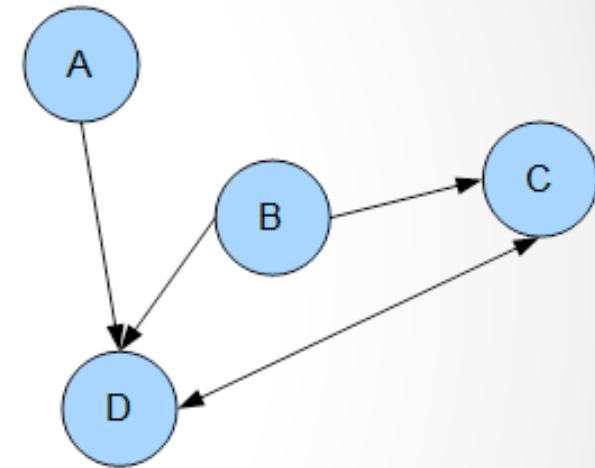
Instalaciones planificadas y certificadas (PUE, TIER II+ ANSI/TIA-942)

Eficiencia energética: Servicio

- **(Re)Organización del modelo TIC**

- *Meta-servicios: Proporcionan soporte al sistema/administradores (restringidos) (Orquestadores, backups, autenticación, monitorización...)*

- *Servicios usuario: Prestaciones abiertas a los usuarios*



- **Plasticidad de servicios**

- *Interdependencia entre servicios fiables (redundancia lógica)*

- *Racionalización de recursos y escalabilidad*

- *Beneficio solidario: la mejora en un servicio repercute en el resto (Servidores más eficientes, mejoras en de almacenamiento...)*

Eficiencia energética: Entorno físico

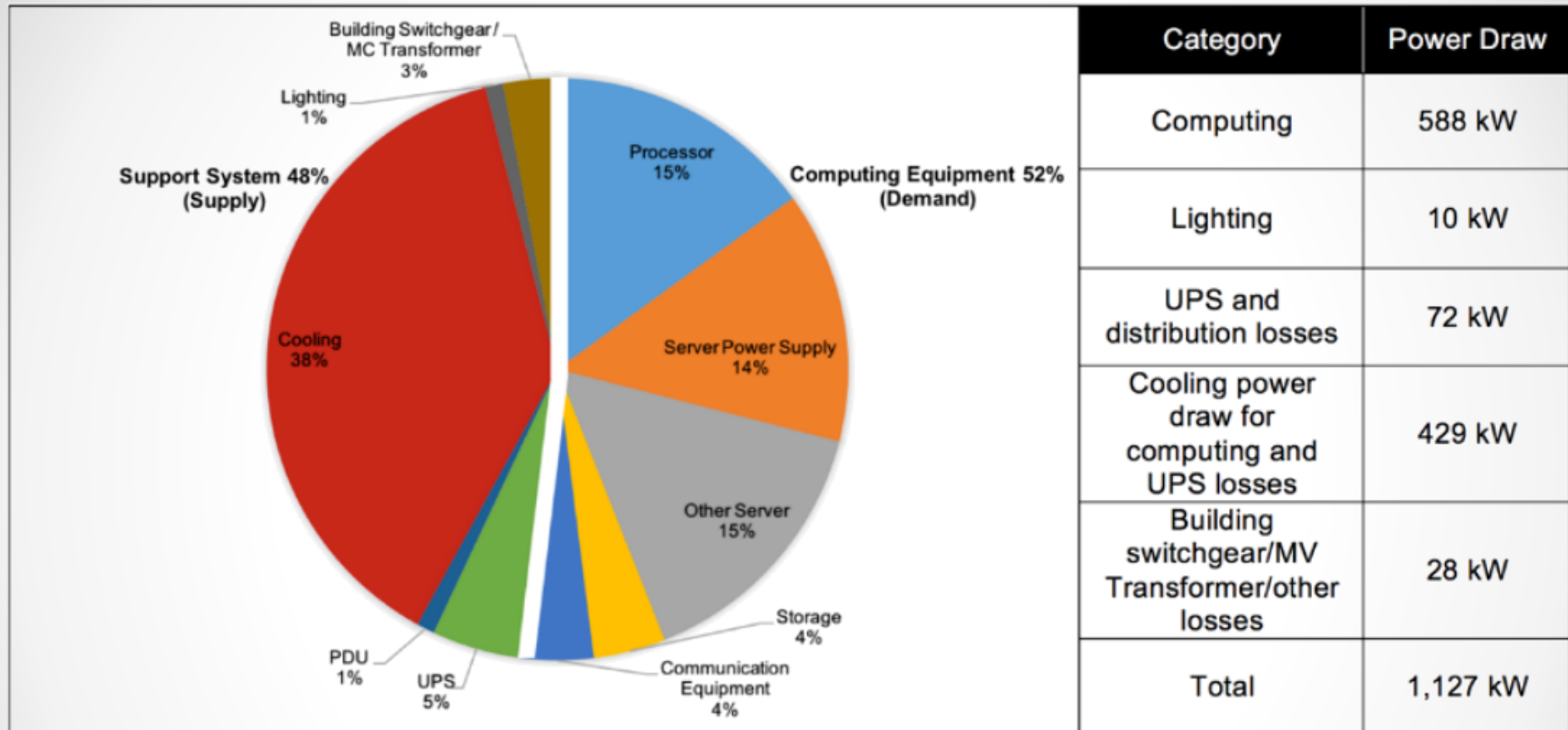


Figure 1. Analysis of a typical 465 m² data center (Emerson 2015).

!!!Apuntamos al 50% del gasto* energético del Centro de Procesamiento de Datos!!!

(*) GPUs a parte...

Eficiencia energética: Entorno físico II

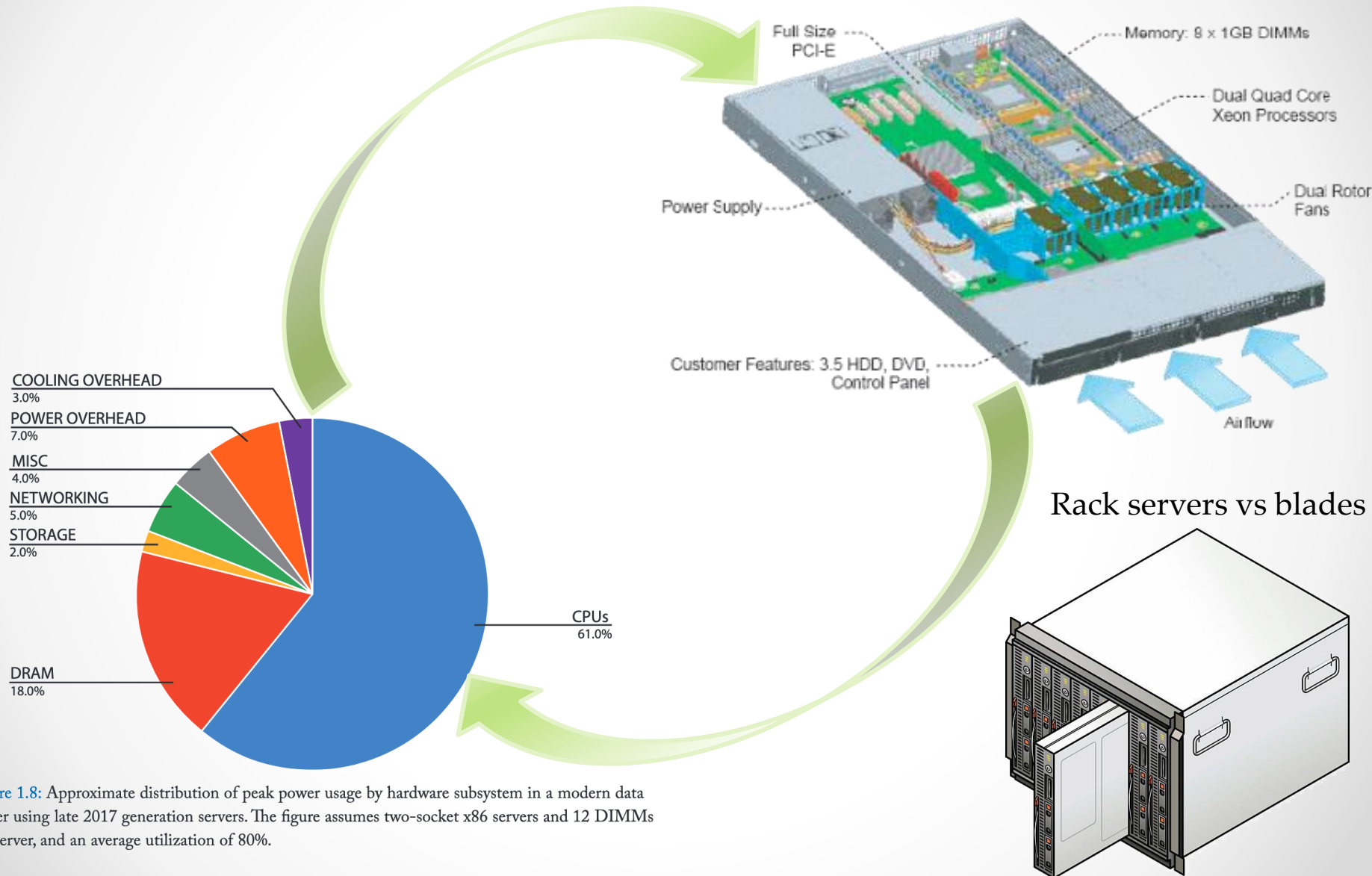


Figure 1.8: Approximate distribution of peak power usage by hardware subsystem in a modern data center using late 2017 generation servers. The figure assumes two-socket x86 servers and 12 DIMMs per server, and an average utilization of 80%.

Eficiencia energética: Entorno físico III

- CPU + Memoria + Disco + Fuentes de Alimentación**

➤ **Thermal Design Power (TDP):** A metric that is expressed in watts. It refers to the amount of power/heat a cooling system (i.e. heatsink, fan) is expected to dissipate to prevent overheating.

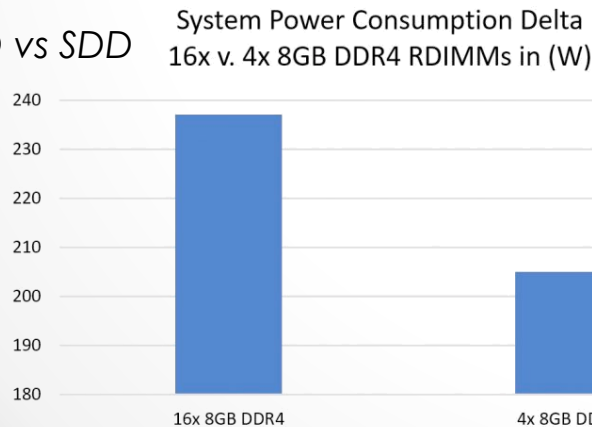
Power Per Watt: A measure of the energy efficiency of a particular computer architecture. It measures the rate of computation delivered by a computer for every watt of power consumed.

➤ CPU **cores** vs CPU **cores + threads**: Temperatura – rendimiento – Uso medio

➤ RAM: ECC vs non-ECC

➤ FA 1000w ≠ consume 1000w

➤ HDD vs SSD



Efficiency level certifications

80 Plus test type	Icon	115 V internal non-redundant				230 V internal redundant				230 V EU internal non-redundant			
		10%	20%	50%	100%	10%	20%	50%	100%	10%	20%	50%	100%
80 Plus			80%	80%	80%						82%	85%	82%
80 Plus Bronze			82%	85%	82%		81%	85%	81%		85%	88%	85%
80 Plus Silver			85%	88%	85%		85%	89%	85%		87%	90%	87%
80 Plus Gold			87%	90%	87%		88%	92%	88%		90%	92%	89%
80 Plus Platinum			90%	92%	89%		90%	94%	91%		92%	94%	90%
80 Plus Titanium		90%	92%	94%	90%	90%	94%	96%	91%	90%	94%	96%	94%



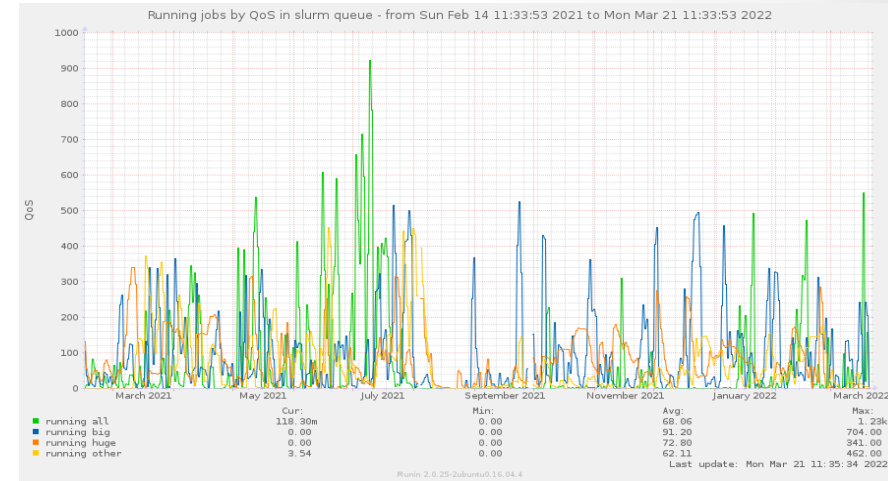
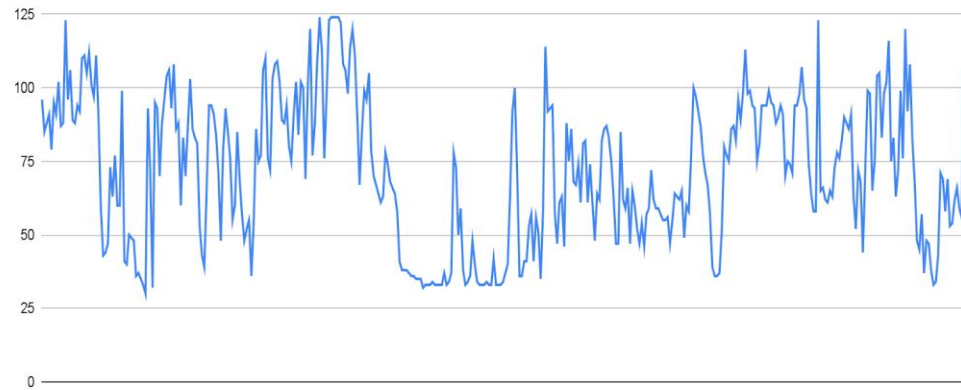
“...because without (proper) data, you are just another person with an opinion.”

Featuring William Edwards Deming

EJ1: Ahorro energético en el clúster del DAC

- **122 nodos de cómputo ~ 2.500.000 jobs anuales**
- **Uso irregular** (Slurm power managenet)

Nodes actius 2021/2022



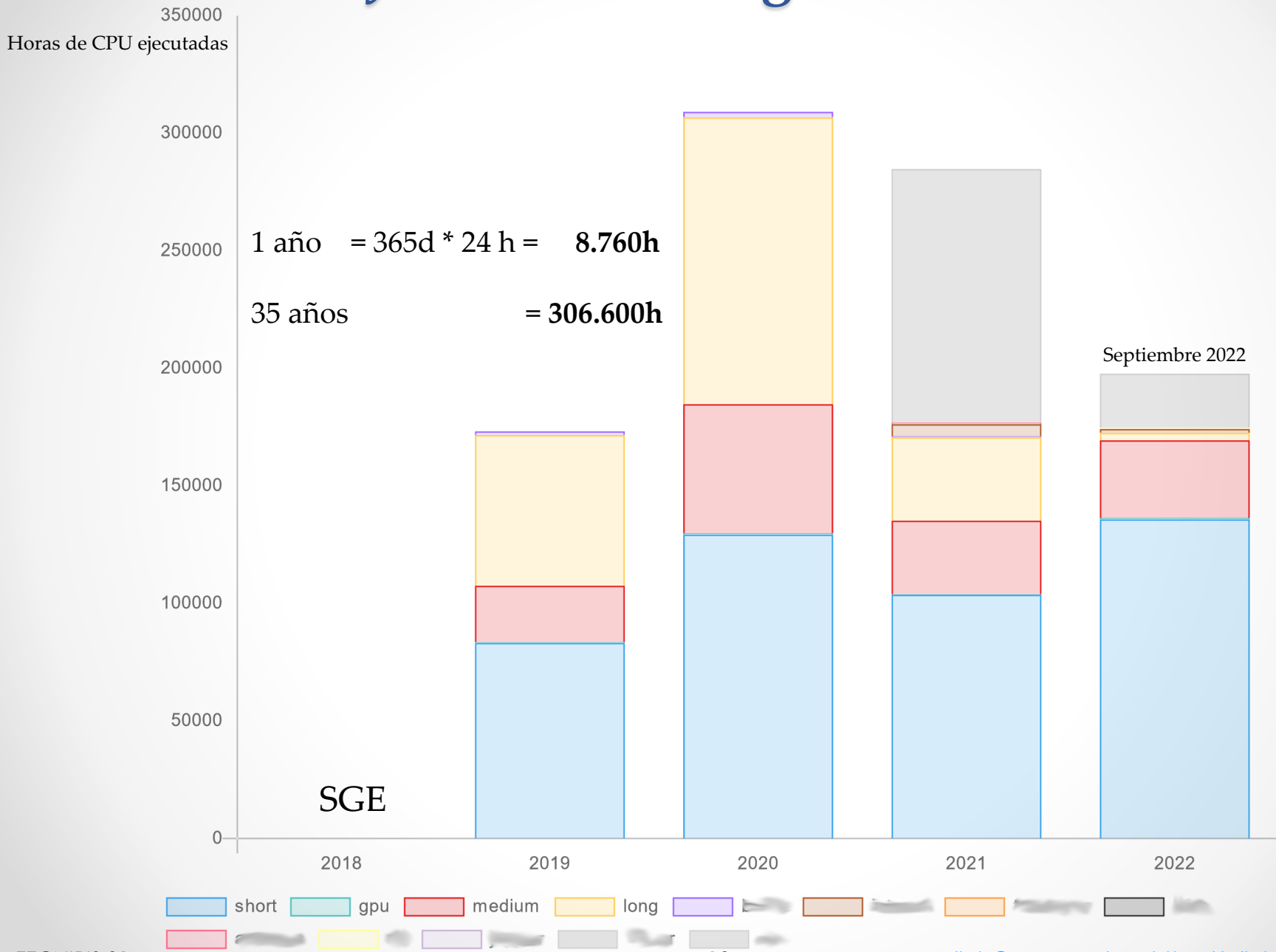
• Consumo nodo típico

- 1 FA de 460 W
- 2 CPUs físicas x 6C/12T
- 128 GB (16 x 8 GB ECC)
- 1 Disco SATA 1TByte

(*) precio de la luz calculado como media del precio actual 0,142 € KWh

Estado	Potencia	Intensidad	KWh	€ / hora*	€ / día 100 nodos	€ / año 100 nodos
Desconectado eléctricamente	0 W	0 A	0	0	0	0
Standby	10 W	0,143 A	0,01	0,00142705	3,4	1241
Encendido sin carga de jobs	81 W	0,368 A	0,08	0,0114164	27,4	10001
Stress 1 CPU	103 W	0,457 A	0,1	0,0142705	34,2	12483
Stress 1 IO	91 W	0,405 A	0,09	0,01284345	30,8	11242
Stress 10 CPU	151 W	0,654 A	0,15	0,02140575	51,4	18761
Stress 10 IO	118 W	0,489 A	0,12	0,0171246	41,1	15002
Stress 24 CPU	160 W	0,708 A	0,16	0,0228328	54,8	20002
Stress 24xCPU + IO + HDD + VM	182 W	0,803 A	0,18	0,0256869	61,6	22484

EJ2: Ahorro energético /rdlab



EJ2: Ahorro energético /rdlab II



Equipo del 2015:

2 x Intel E5-2430 (12 cores)
1 x SATA 1TBytes
32 GB RAM



# Discos	Tecnología	Estado	Modelo	Consumo (MAX)
0		Reposo		55.3w
1	HDD	Reposo	Seagate Constellation 1Tbyte 7.2k rpm	62.2w (+12%)
1	HDD	Stress E/S	Seagate Constellation 1Tbyte 7.2k rpm	71.5w (+29%)
1	SSD	Reposo	Samsung 860 EVO 1TB	58.0w (+5%)
1	SSD	Stress E/S	Samsung 860 EVO 1TB	66.3w (+19%)

EJ2: Ahorro energético /rdlab III



Equipo del 2010:

2 x Intel Xeon 5120 (4 cores)
2 x SATA 500Gbytes
4 GB RAM
(controladora RAID*, 2FA..)



# Discos	Tecnología	Estado	Modelo	Consum (MAX)
2		Reposo	Western Digital 500GB RE3 7.2k rpm	169w
2	HDD	Stress E/S	Western Digital 500GB RE3 7.2k rpm	171.8w (+2%)*

¡¡Consumo energético de +2.5 x R320!!

EJ2: Ahorro energético /rdlab IV



Equipo del 2020:

2 x Intel Xeon 4214R

2 x FA 500w (A500E-S0)

1 x SATA 1TBytes

64 GB RAM



FA1	FA2	Estado	Consumo eléctrico (MAX)
Conectada	DESconectada	Apagado	17.5w
Conectada	Conectada	Apagado	23w (+31%)
Conectada	DESconectada	Reposo	102.1w
Conectada	Conectada	Reposo	110.5w (+8%)

Redundancia física de componentes vs redundancia del servicio (Arquitectura)

Eficiencia energética: ¿Cómo hacerlo?

- **Punto de partida**



- Resistencia a los cambios: ¿por qué cambiar?
- Cultura institucional: Siempre lo hemos hecho así, tengo ya mucho trabajo...
- Servicios e instalaciones descentralizadas: Bolets(Setas) y reinos de taifas
- Trabas legales y administrativas:
 - Ley de contratación pública
 - Distintas fuentes de financiación (subvenciones, TT, proyectos competitivos)
 - Gastos elegibles
 - Facturación por servicio vs. Facturación por inventariable (hardware)

➤ ...

Eficiencia energética: ¿Cómo hacerlo? II

1. Fuente de alimentación

Seleccionad una fuente de alimentación que proporcione una potencia inferior al 125% del consumo máximo del sumatorio de componentes del servidor (CPU+RAM+Discos+Placa base) y que tenga eficiencia dentro de la norma 80Plus. Evitad, si es posible, fuentes de alimentación redundantes. Se recomienda buscar la flexibilidad a nivel de servicio y arquitectura TIC, a excepción de casos de sistemas de alta disponibilidad (HA) o equipos críticos.

Puntuación: 0 para las 80 Plus y añadid 2 puntos adicionales para cada mejora de la categoría energética.

Referencia: https://en.wikipedia.org/wiki/80_Plus

2. Procesador

- Servidores genéricos : Seleccionad la CPU dentro de la familia compatible con el socket de la placa base que tenga menor consumo y que proporcione los cores/threads requeridos. Son preferibles los sistemas con refrigeración pasiva versus los tradicionales con ventiladores. Recordad que generalmente es necesario configurar/habilitar los estados de bajo consumo energético a la BIOS.

- Servidores HPC o específicos (GPUs...): Según necesidades.

CM = Consumo Máximo = Máximo(consumo en watts de todos los procesadores ofertados)

Cm = Consumo mínimo = mínimo(consumo en watts de todos los procesadores ofertados)

VE = Varianza Energética = CM - Cm

Puntuación: $10 * ((CM - \text{Consumo procesador}) / VE)$

Referencia para CPUs Intel: <https://ark.intel.com/content/www/us/en/ark.html>

Referencia para CPUs AMD: <https://www.amd.com/en/processors>

NOTA: Son preferibles sistemas con 2 o más procesadores físicos en un mismo equipo, que servidores múltiples desde el punto de vista de ahorro energético para propósito general.

3. RAM

- Servidores con procesadores de equipos portátiles o bajo consumo: 2 DIMMs por procesador físico. Servidores con procesadores para servicios genéricos: 4 DIMMs por procesador físico.

- Servidores con procesadores para equipos HPC/específicos: Según el fabricante, arquitectura del procesador (dual channel, quad channel...) y número de CPU sockets.

Puntuación: $10 - \text{Valor Absoluto} (\# \text{DIMMs recomendados para la gama del procesador} - \# \text{DIMMs ofertados})$

NOTA: En el caso de ofertar configuraciones con menos DIMMs de los recomendados por el fabricante del procesador, se puntuará con 0 puntos por ineficiente.

4. Disco

- Servidores genéricos: Seleccionad preferentemente discos SSD hasta los 2TBytes. Tened en cuenta la activación en el sistema operativo de la gestión SMART de los discos y TRIM si es necesario.

- Servidores de disco/HPC/específicos: Según necesidades.

Puntuación: $10 * (\# \text{discos SSD ofertados} / \# \text{discos totales ofertados})$

Referencia: <https://en.wikipedia.org/wiki/S.M.A.R.T.>

Referencia: [https://en.wikipedia.org/wiki/Trim_\(computing\)](https://en.wikipedia.org/wiki/Trim_(computing))

Fórmula Puntuación Energética = FA * 0.3 + Procesador * 0.4 + RAM * 0.1 + DISC * 0.2

Eficiencia energética: ¿Cómo hacerlo? III



Gasto **vs** Inversión
Compra **vs** Uso
Propiedad **vs** Consorciado público

El Reto 10-10-10

- **Propuesta:**

- **10%** de reducción del consumo eléctrico de **servidores** en **CPDs** (Omega)
 - Medible (benchmarking)
 - Impacto real (cream skimming)
- **10%** inversión económica del consumo eléctrico **anual** previsto CPD (servidores)
 - Autofinanciado
 - Retorno de la inversión continuo
- **10** Meses
 - Finito en tiempo
 - Periodo de aprovisionamiento



... y recordad

*...pero los
GRANDES cambios,
¡tienen impacto!*

*Los pequeños
cambios, son
poderosos...*

Bibliografía

[www1] The Problem of Power Consumption in Servers

<https://www.infoq.com/articles/power-consumption-servers/>

<https://www.ovoenergy.com/ovo-newsroom/press-releases/2019/november/think-before-you-thank-if-every-brit-sent-one-less-thank-you-email-a-day-we-would-save-16433-tonnes-of-carbon-a-year-the-same-as-81152-flights-to-madrid>

[www2] The Datacenter as a computer, 3rd edition

<https://www.morganclaypool.com/doi/pdfplus/10.2200/S00874ED3V01Y201809CAC046>

[www3] Energy efficiency insight into small and medium data centres

https://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2017/7-appliances-products-lighting-and-ict/energy-efficiency-insight-into-small-and-medium-data-centers-a-comparative-analysis-based-on-a-survey/2017/7-086-17_Vasques.pdf/

<https://www.serverstack.in/2019/01/19/difference-between-rack-servers-and-blade-servers/>

[www4]. The Datacenter of the future (Dell)

https://education.dellemc.com/content/dam/dell-emc/documents/en-us/2021KS_Ganesh-The_Data_center_of_the_Future.pdf

[www5] Breakdown of power consumption in servers (Emerson 2015)

https://www.researchgate.net/figure/Breakdown-of-power-consumption-in-servers-Emerson-2015_fig4_329278608

[www6] Thermal design power (TDP)

https://en.wikipedia.org/wiki/Thermal_design_power

[www7]. Performance per watt (PPW) + RAM Consumption

https://en.wikipedia.org/wiki/Performance_per_watt

<https://www.servethehome.com/ddr4-dimms-system-power-consumption-tested/>

Bibliografía II

[www9] PSU efficiency

https://en.wikipedia.org/wiki/80_Plus

[www10] Storage Area Network (SAN)

https://en.wikipedia.org/wiki/Storage_area_network

[www11] Network-Attached Storage (NAS)

https://en.wikipedia.org/wiki/Network-attached_storage

[www12] HDD vs SSD

<http://www.amplicon.es/docs/SSD-vs-HDD-white-paper.pdf>

[www13] HDD vs SSD power consumption

<https://ssdsphere.com/ssd-power-consumption-comparison/>

[www14] HDD vs SSD power consumption 2

<https://www.anandtech.com/show/7682/the-wd-black2-review/9>

[www15] HDD vs SSD lifetime and reliability

<https://www.backblaze.com/blog/are-ssds-really-more-reliable-than-hard-drives/>

[www16] Nvidia RTX 3090 power consumption analysis – GPU power consumption

<https://www.profesionalreview.com/2020/10/23/nvidia-rtx-3090-cuatro-nvlink/>

<https://www.tomshardware.com/features/graphics-card-power-consumption-tested>

Bibliografía II

[www17] GPUs power consumption

<https://www.igorslab.de/en/nvidia-geforce-rtx-3090-founders-edition-review-between-even-between-evaluate-and-common-decadence-if-price-is-not-all/16/>

[www18] Slurm HPC Green guide

https://slurm.schedmd.com/power_save.html

[www19] CPU cores vs CPU cores + threads (AMD)

<https://www.anandtech.com/show/16261/investigating-performance-of-multithreading-on-zen-3-and-amd-ryzen-5000/5>

[www20] CPU cores vs CPU cores + threads (Intel)

<https://www.intel.com/content/www/us/en/gaming/resources/hyper-threading.html>

[www21] Hyperthread vs Multithread

<https://www.pcpartguide.com/is-hyperthreading-worth-it/>

[www22] Hyperthread vs Multithread 2

<https://www.cgdirector.com/cpu-cores-vs-logical-processors-threads/>

[www23] Martin Lindstrom – The ministry of common sense

<https://www.martinlindstrom.com/books-by-martin-lindstrom/>

[www24] Cream skimming – Prototype vs Minimum Viable Product

https://en.wikipedia.org/wiki/Cream_skimming

<https://www.boldare.com/blog/first-version-of-your-app-mvp-vs-prototype/>