

Gestió energètica eficient per a grans infraestructures TIC (Resum Executiu)



Abril 2022

Víctor Mora Antolino
victor.mora@upc.edu
<https://www.ac.upc.edu>

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

ÍNDEX

- Expectatives i realitats	3
- Visió holística	4
- Eficiència energètica: entorn arquitectura	5
- Eficiència energètica: entorn servei	6
- Eficiència energètica: entorn físic	7
- Estalvi d'energia: Clúster DAC	10
- Estalvi d'energia: /rdlab	11
- Criteris energètics per licitacions TIC	15
- Eficiència energètica: Com ho fem?	16
- Bibliografia	19

Expectatives i realitats



Visió holística

- **Grans infraestructures TIC**

- **Marge operatiu de treball en 3 nivells:**

- Físic: Servidors (Xarxa?)
- Servei: Prestacions als usuaris i als administradors
- Arquitectura: Paradigma d'orquestració

- **Estratègia**

- Estratègia energètica de les TIC $\subset$ Estratègia energètica
- Inversions + gestió transitori = Futur



Queden fora de l'abast d'aquesta presentació els temes energètics del CPD així com les estratègies...

Eficiència energètica: Entorn Arquitectura

- **Paradigma TIC tradicional (standalone)**

- Servidor físic = servei (entorns altament cohesionats)
- Redundància **física** de tots els components (FA, discos locals, xarxa...)
- Qualitat de servei rígida
- Duplicitats de backups externs
- instal.lacions "sobrevingudes":

Servidors no enrackables, aires condicionats domèstics, sense tancaments...

- **Model agregat escalable (clustered)**

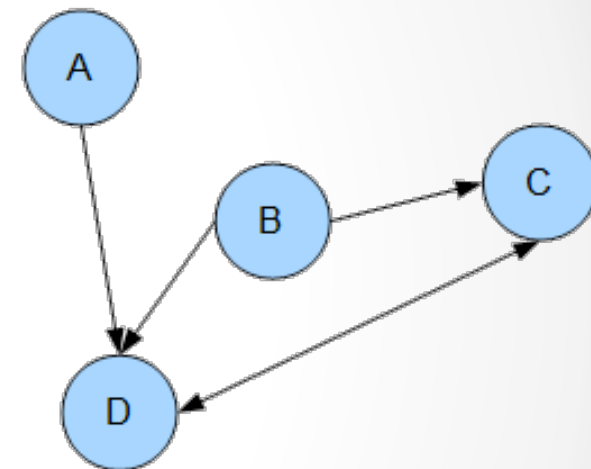
- Desacoblament entre servidors físics i serveis
- Redundància: **Lògica** a nivell de servei i **Física** únicament de components vitals
- Compartició de recursos $\Leftrightarrow$ Escalabilitat/adaptabilitat a la demanda real (Entorns virtuals/Cloud, storage, snapshots...)
- Instal.lacions planificades i certificades (PUE, TIER II+ ANSI/TIA-942)

Eficiència energètica: Entorn Servei

- **(Re)Organització del model TIC**

- *Meta-serveis: Proporcionen suport al sistema/administradors (restringits)
(Orquestradors, backups, monitorització...)*

- *Serveis usuari: Prestacions obertes als usuaris*



- **Plasticitat de serveis:**

- *Interdependència entre serveis fiables (redundància lògica)*

- *Racionalització de recursos i Escalabilitat*

- *Benefici solidari: la millora d'un servei repercuteix a la resta
(Servidors més eficients, millores al emmagatzemament...)*

Eficiència energètica: Entorn Físic

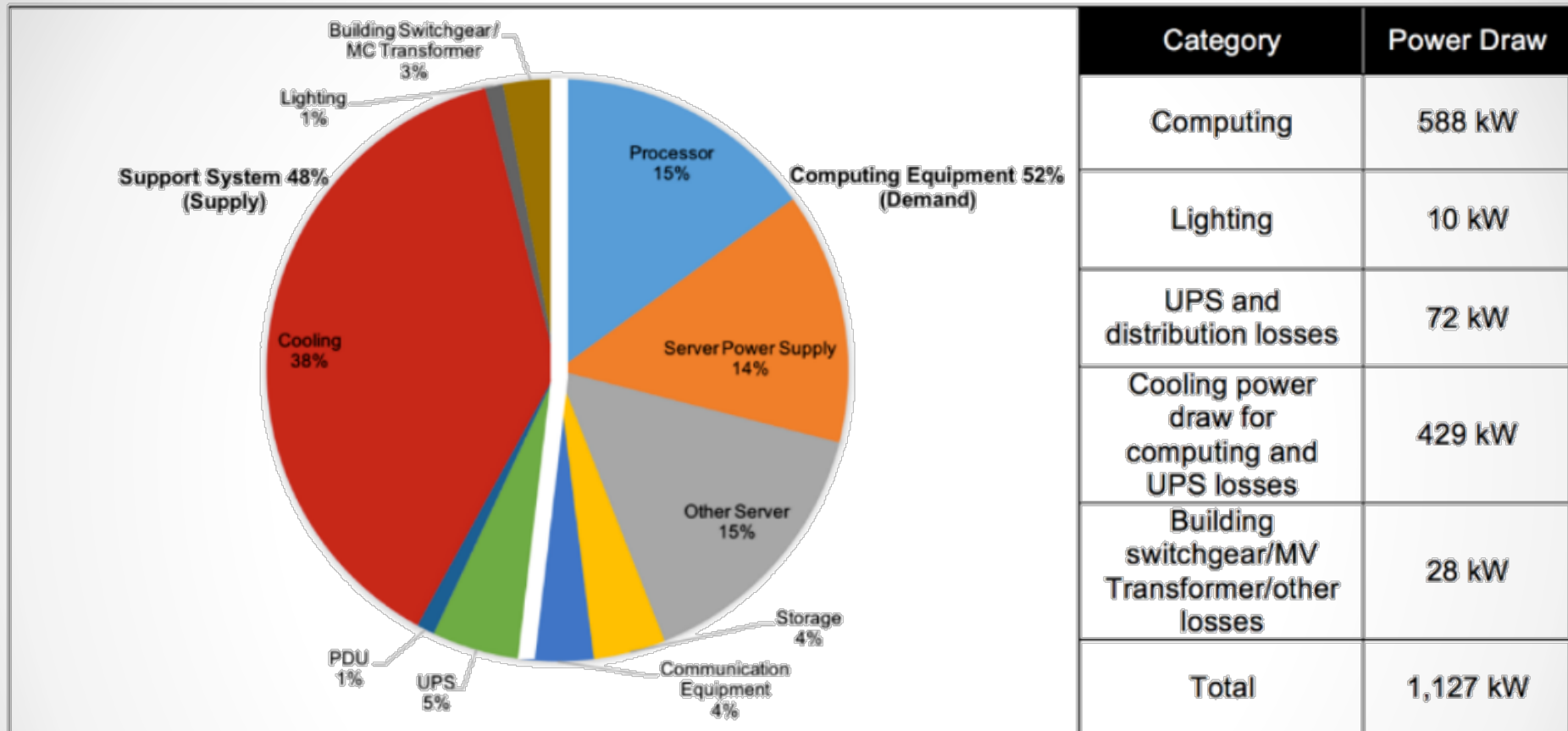


Figure 1. Analysis of a typical 465 m² data center (Emerson 2015).

Apunten a gairebé el 50% de la despesa* energètica del Centre de Processament de Dades!!!

(*) GPUs a part...

Eficiència energètica: Entorn Físic II

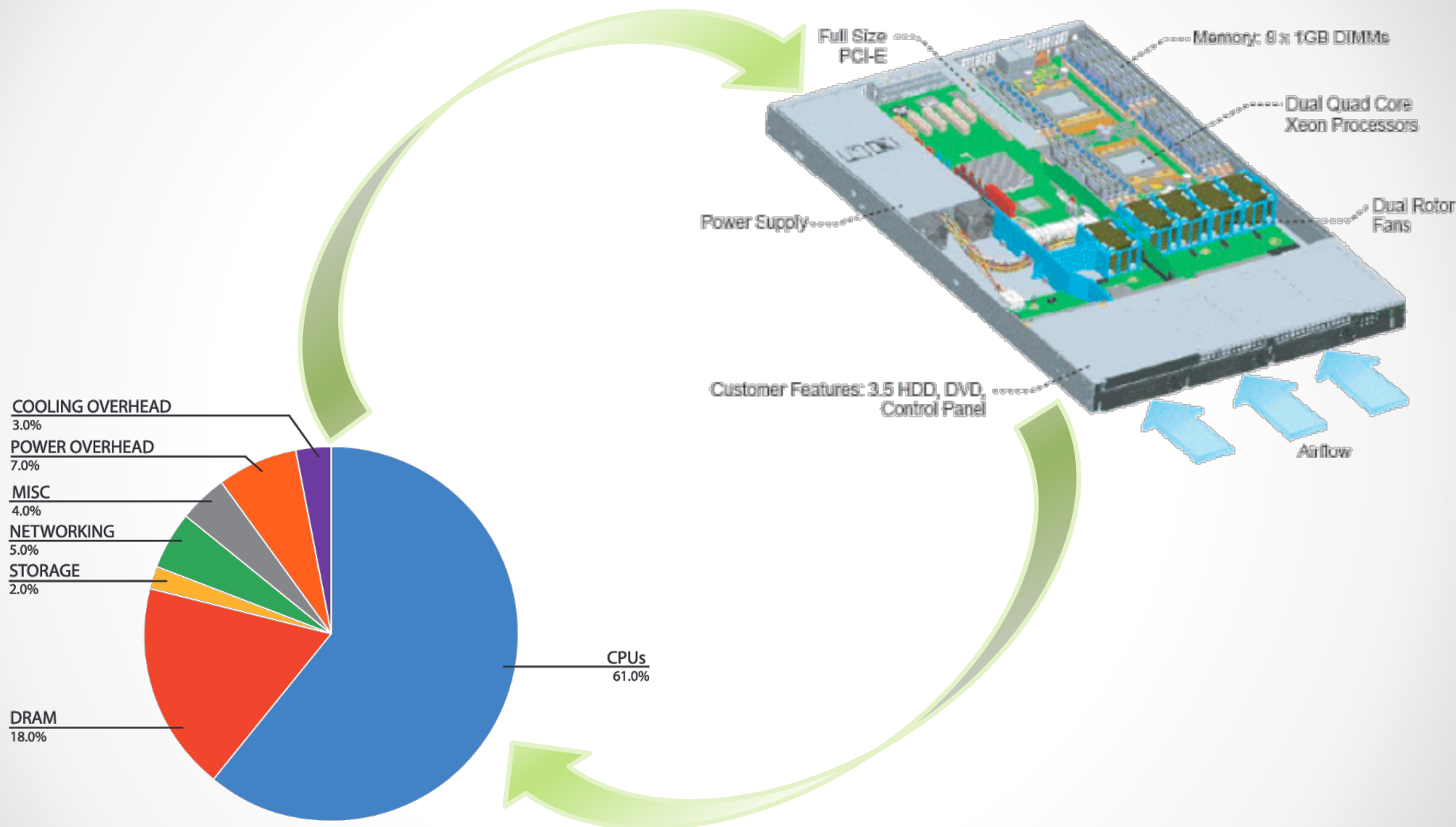
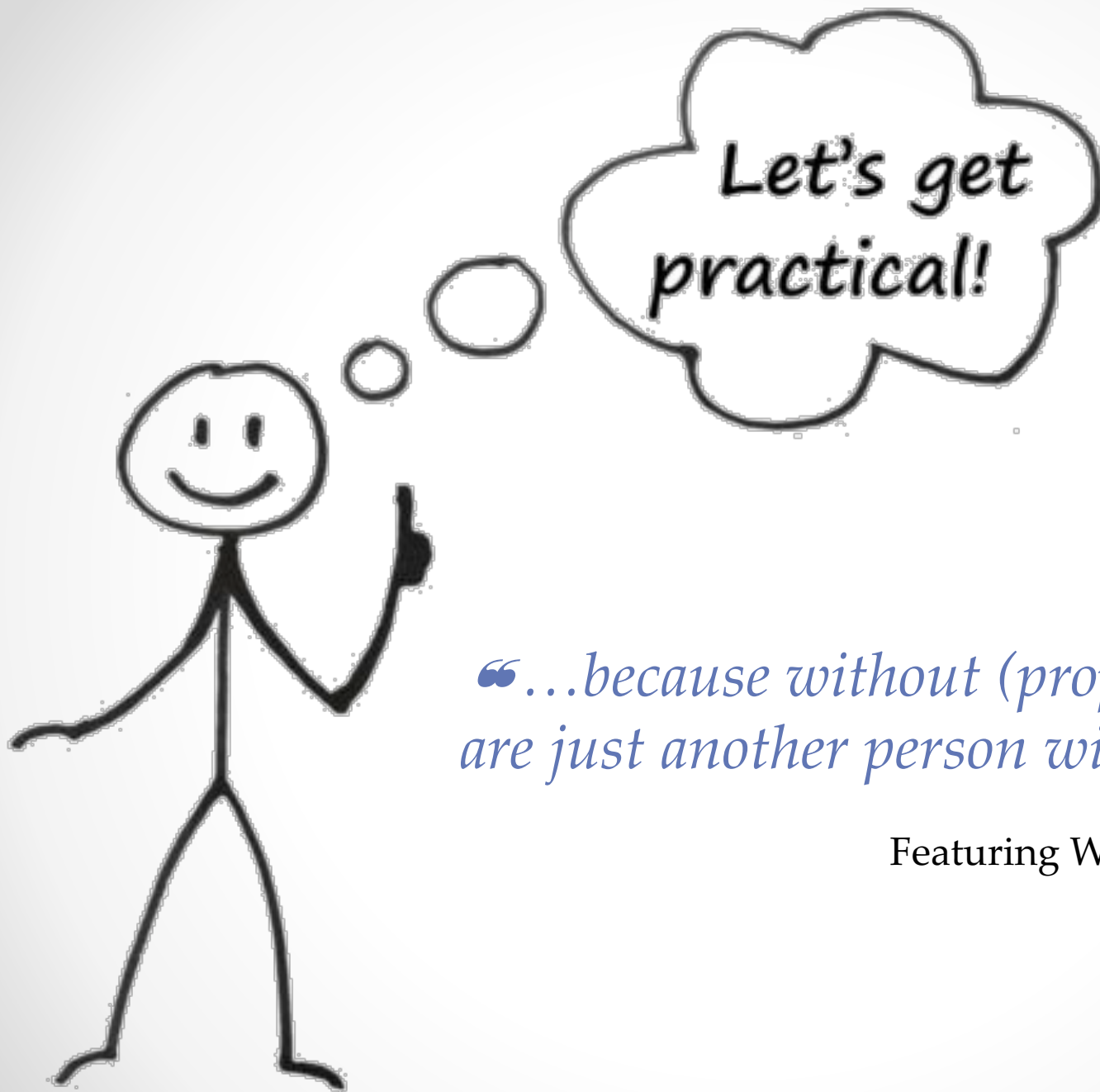


Figure 1.8: Approximate distribution of peak power usage by hardware subsystem in a modern data center using late 2017 generation servers. The figure assumes two-socket x86 servers and 12 DIMMs per server, and an average utilization of 80%.



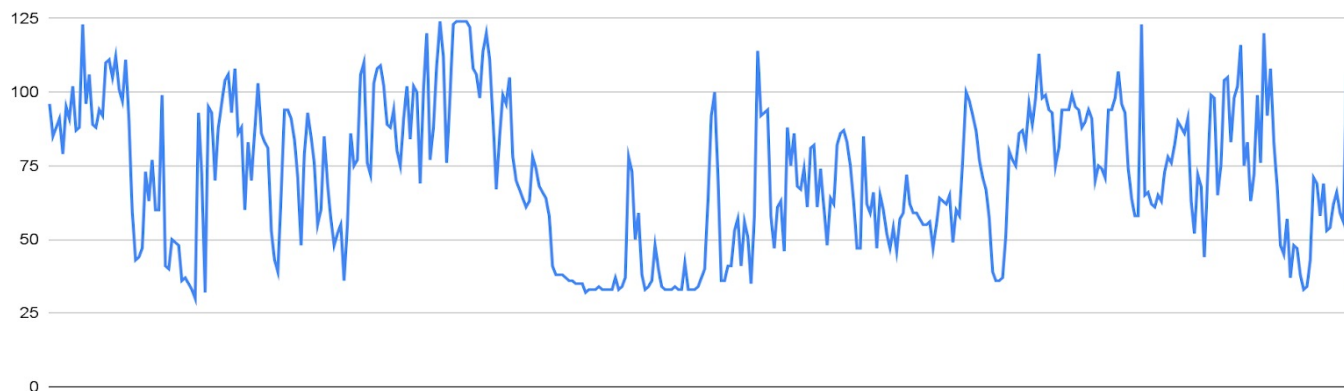
“...because without (proper) data, you are just another person with an opinion.”

Featuring William Edwards Deming

Estalvi d'energia al clúster del DAC

- **122 nodes~2.500.000 jobs anuals**
- **Hi ha moments d'ús intens i altres on molts nodes no fan res**

Nodes actius 2021/2022



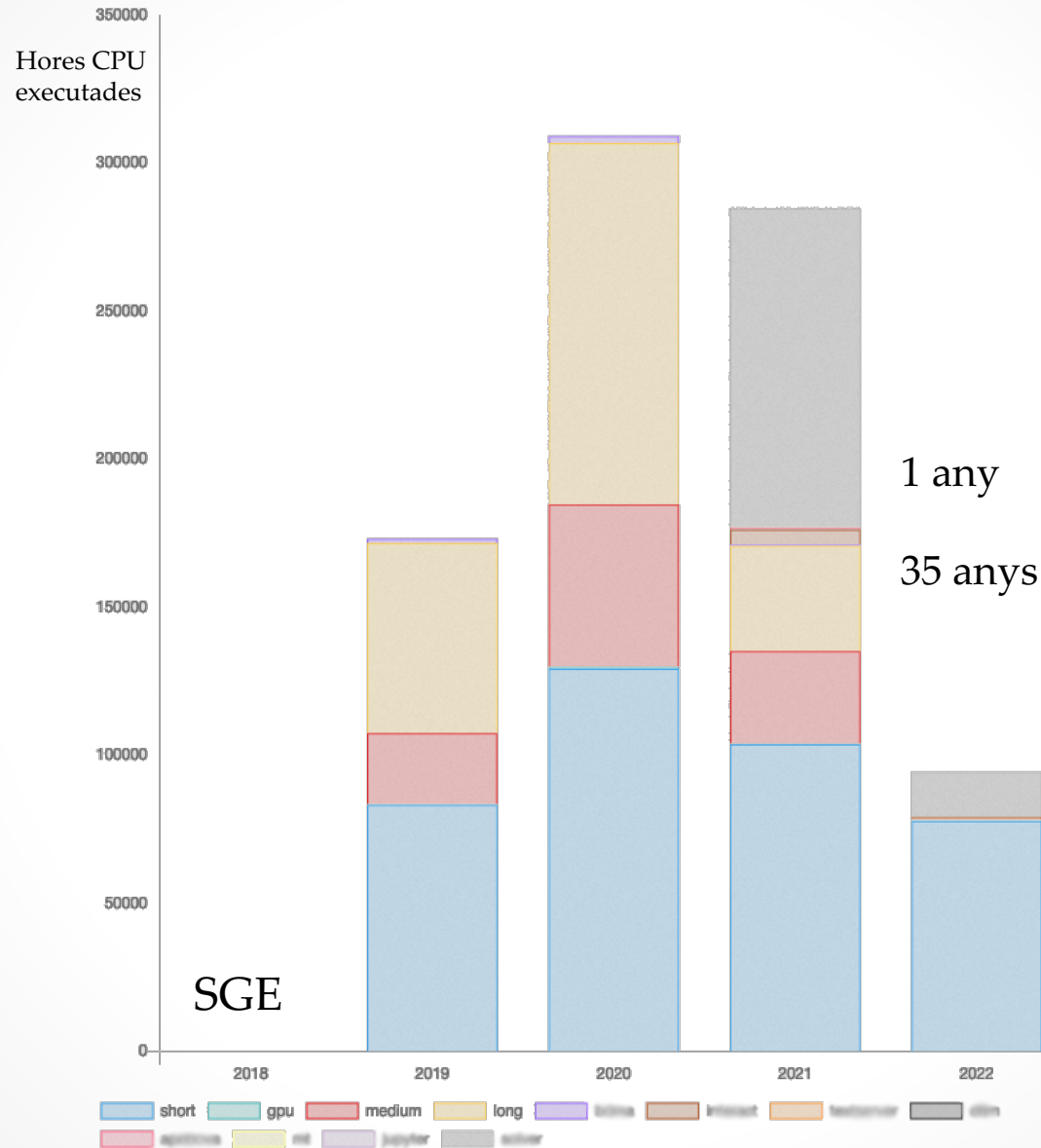
- **Node típic**

- 1 Font de 460 W
- 2 CPUs físiques x 6C/12T
- 128 GB (16 mòduls de 8 GB)
- 1 Disc SATA 1TByte

Estat	Potència	Intensitat	KWh	€ / hora*	€ / dia 100 nodes	€ / any 100 nodes
Desconnectat	0 W	0 A	0	0	0	0
Standby	10 W	0,143 A	0,01	0,00142705	3,4	1241
Ences sense 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 24 x CPU + IO + HDD + VM	182 W	0,803 A	0,18	0,0256869	61,6	22484

(*) preu de la llum agafat segons una mitjana del preu actual 0,142 € KWh

Estalvi d'energia al /rdlab: Servei HPC



1 any = 365d * 24 h = 8.760h

35 anys = 306.600h

Estalvi d'energia al /rdlab: Cas real discos R320



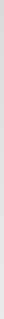
Equip del 2015:

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



# Discos	Tecnologia	Estat	Model	MAX Consum elèctric
0		repòs		55.3w
1	HDD	repòs	Seagate Constellation 1Tbyte 7.2k rpm	62.2w
1	HDD	Estressat E/S	Seagate Constellation 1Tbyte 7.2k rpm	71.5w
1	SDD	repòs	Samsung 860 EVO 1TB	58.0w
1	SDD	Estressat E/S	Samsung 860 EVO 1TB	66.3w

Estalvi d'energia al /rdlab: Cas real Dell 1950



Equip del 2010:

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



# Discos	Tecnologia	Estat	Model	MAX Consum elèctric
2		repòs	Western Digital 500GB RE3 7.2k rpm	169w
2	HDD	Estressat E/S	Western Digital 500GB RE3 7.2k rpm	171.8w

Consum energètic +2.5 x R320!

Estalvi d'energia al /rdlab: Cas real Dell R410



Equip del 2020:

2 x Intel Xeon 4214R

2 x FA 500w (A500E-S0)

1 x SATA 1TBytes

64 GB RAM



# FA1	FA2	Estat	Consum elèctric
connectada	connectada	apagat	23w (+24%)
connectada	DES connectada	apagar	17.5w
connectada	connectada	repòs	110.5w (+7.5%)
connectada	DES connectada	repòs	102.1w

Redundància física de components
vs redundància servei/Arquitectura

Criteris de consum energètic de servidors TIC

1. Font d'alimentació

Seleccionar una font d'alimentació que proporcioni una potència inferior al 125% del consum màxim del sumatori de components del servidor (CPU+RAM+Discos+Placa base) i que tingui eficiència dintre de la norma 80Plus. Eviteu si és possible fonts d'alimentació redundants, és millor buscar flexibilitat a nivell de servei i arquitectura TIC, excepte pels casos de sistemes d'alta disponibilitat (HA) o equips crítics.

Puntuació: 0 per les 80 Plus i donar 2 punt addicionals per millora de categoria energètica.

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

2. Processador

Servidors genèrics : Seleccionar la CPU dintre de la família compatible amb el socket de la placa base que tingui menor consum i que proporcioni els cores/threads requerits. Són preferibles els sistemes amb refrigeració passiva vers els tradicionals amb ventiladors. Recordeu que generalment cal configurar/habilitar els estats de baix consum energètic a la BIOS.

Servidors HPC/específics: Segons necessitats.

CM = Consum Màxim = Màxim(consum en watts de tots els processadors ofertats)

Cm = Consum mínim = Mínim(consum en watts de tots els processadors ofertats)

VE = Variància energètica = CM - Cm

Puntuació: $10 * (CM - \text{Consum processador}) / VE$

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

Referència per CPUs AMD: <https://www.amd.com/en/processors>

NOTA: Són preferibles els sistemes amb 2 o més processadors físics a un mateix equip que servidors separats per l'estalvi energètic sempre que no interfereixin amb serveis d'alta disponibilitat (HA).

3. RAM:

Servidors amb processadors d'equips portàtils o baix consum: 2 DIMMs per processador físic. Servidors amb processadors de servidors genèrics: 4 DIMMs per processador físic. Servidors amb processadors equips HPC/específics: Segons el fabricant

Puntuació: $10 * (\#DIMMs \text{ recomanats per la gamma del processador} / \#DIMMs \text{ ofertats})$

NOTA: En cas de que s'ofrenin configuracions amb menys DIMMs dels recomanats per processador, es puntuarà amb 0 punts per ineficient.

4. Disc

Servidors genèrics: Seleccionar preferentment discos SSD per capacitats de fins a 2TBytes. Recordeu configurar correctament al sistema operatiu la gestió SMART dels discos i TRIM si s'escau. Servidors de disc/HPC/específics: Avaluar segons el cas i necessitats.

Puntuació: $10 * (\#discos \text{ SSD ofertats} / \#discos \text{ totals ofertats})$

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

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

Formula Puntuació energètica = FA * 0.3 + Processador * 0.4 + RAM * 0.1 + DISC * 0.2

Eficiència energètica: Com ho fem?



Despesa vs **Inversió**

Eficiència energètica: Com ho fem? II

- **Proposta: Repte 10 – 10 – 10**

➤ **10%** de reducció de consum elèctric als servidors dels **CPDs** (Omega + Vx)

Mesurable (benchmarking)

Impacte real (cream skimming)

➤ **10%** d'inversió econòmica del consum elèctric anual previst de servidors als CPD

Autofinançat

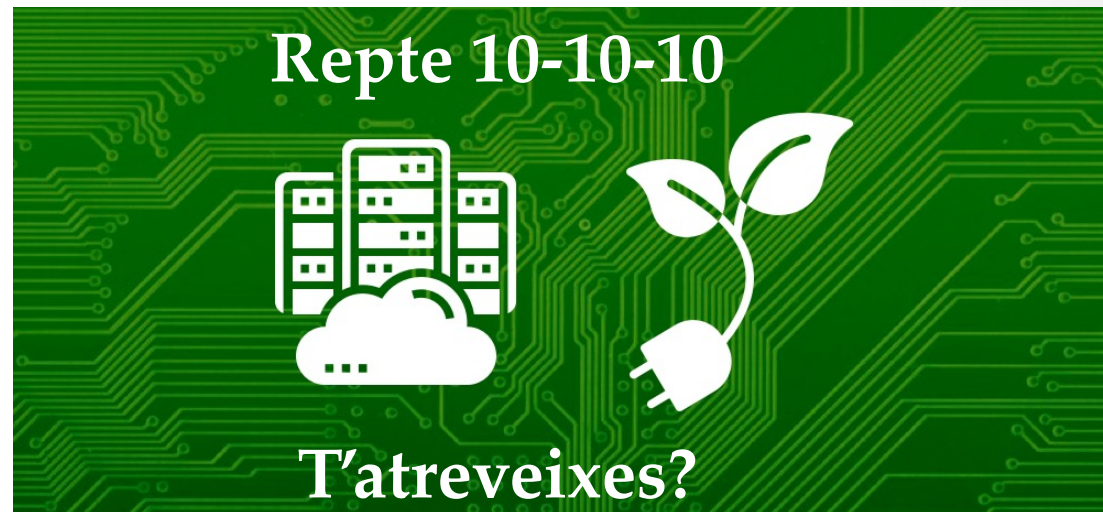
Retorn de la inversió continu

➤ **10** Mesos

Finit al temps

Termini d'aprovisionament

Prototipus vs MVP



...i recordeu

*Els GRANS canvis
tenen impacte!*

*Els petits canvis, son
poderosos...*

Bibliografia

[www1] The Problem of Power Consumption in Servers

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

[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/

[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)

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

[www8] RAM power consumption

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

Bibliografia 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>

Bibliografia 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/>