Nouvelles de WLCG / FR

D. Bouvet
L. Duflot

Paysage Français et EU

- Nouvelle convention CEA – CC-IN2P3 courant juillet
- CINES, IDRIS en déficit sur les fluides
- RENATER toujours difficultés financières, arrêterait rendez-vous?
- Mesocloud
  - Réunion onboarding
  - 1ere réunion intégration des sites FG (Lille, IPHC, LPNHE, IRFU, Sorina, LLR, Andrei) le 10/6
    - Retours ?
- EGI ne serait pas inscrit dans les ESFRI
-

DOMA / OTF

- US CMS pourrait changer de stratégie et concentrer le stockage vers FNAL avec certains T2 CPU-only (OTF avril @ HEPiX)
  - FNAL ne pourrait pas héberger le CPU nécessaire au T1 pour HL-LHC (limite de puissance) mais devrait pouvoir héberger la totalité du disque (assez d'espace)
  - C'est la conclusion d'un TF pas encore une décision

DOMA / OTF

- FTS4
  - Réécriture complète
  - Départ d'un dev important
  - Ne pourra pas faire la release pour l'été…
  - Les tests ATLAS avec la version dev actuelle sont positifs, prêt a passer une partie de la prod dessus
  - Le message passe a CHEP était moins pessimiste (plus politiquement correct?)
- OTF pourrait re-lancer un sondage cette année sur les sites et leur évolution

Prep DC27

- Au DOMA d'avril ATLAS a présenté ses prévisions de besoins pour DC27 (CMS l'avait fait fin 2025)
  - Utilisation des métriques actuelles, en couvrant 99% des cas
  - Disk ↔️ disk 13000Gbs
  - Tape export 155Gbs
  - Tape → disk 170Gbs
  - "to be more realistic were may need to investigate in DC27 the effect of transferring small files"
  - Google sheet

WLCG OB

- Proposition de statut T1 ALICE pour UPB (Roumanie)
  - Processus en cours depuis ~ 1 an, se passe bien, la dernière barrière était LHCOPN qui est en service depuis ce printemps
  - Avis favorable de l'OB
- Proposition de statut T1 CMS pour un DC en Serbie
  - Re-structurations, plus de contact depuis plusieurs mois
  - Relance
- WLCG et GPU : au vu de l'intérêt Stefano Piano, ancien computing coordinator, a été nommé pour animer le sujet
- Présentation sur le hardware et les coûts (non discutés ici, voir talk demain) par A. Sciaba

Global server market

- TSMC makes most CPUs and GPUs for data Centers
  - Not much competition
- Samsung is the biggest DRAM maker (~38% market share), but SK Hynix leads the HBM market (~58%)
  - HBM much more profitable…
- Processors : AMD gains market share, ARM deployed massively by hyperscalers and AI compagnies
- NVIDIA dominate the GPU market

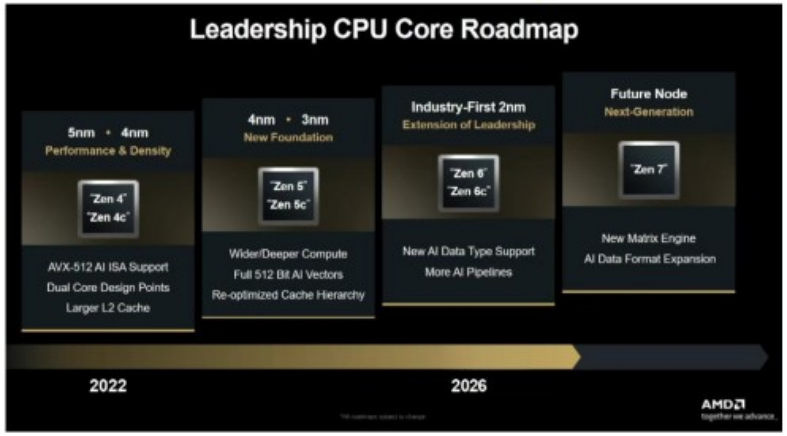
Global server market

- NAND flash market dominated by Samsung, SK Hynix and Kioxia/WD
  - Revenues reaching $170 billion, +140% from 2025
  - NAND wafer prices increased by ~250% in a year and still increasing by ~35%/quarter
  - Production capacity will not increase until 2027-28

x86 CPUs

- **AMD's Zen 5 (Turin) the latest architecture**
  - 128 cores (192 Zen 5c cores)
  - Zen 6 (Venice) later this year with up to 256 cores and PCIe 6.0 support
  - "Verano" (?) in 2027, it should be a specialized Zen 6 variant with better performance/Watt
- **Intel's Xeon 6+ (Clearwater Forest) this year**
  - 288 E-cores, aimed at cloud, edge AI, telecom
- **Diamond Rapids end of 2026**
  - Up to 192 P-Cores, no SMT
  - Unclear what comes next

AMD CPU roadmap



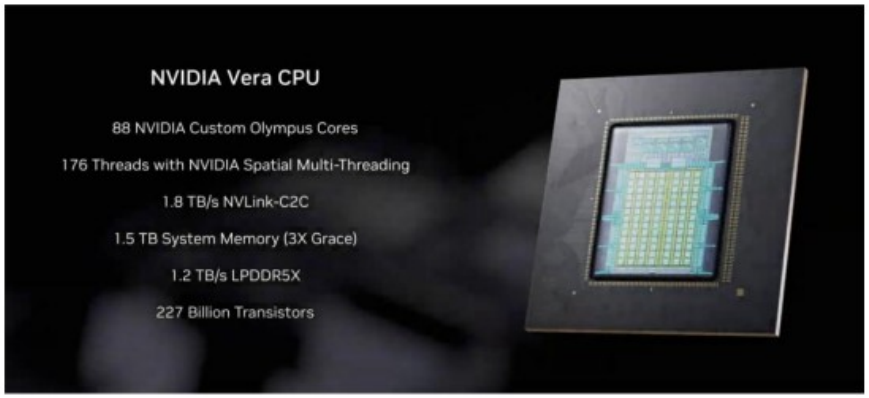
Xeon Roadmap

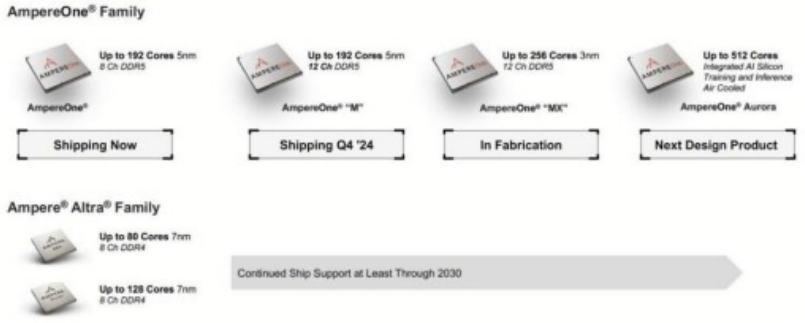
Xeon Family	Availability	Core Type	Process / Packaging
Xeon 6 E-core (Sierra Forest)	June 2024	E-core	Intel 3
Xeon 6 P-core (Granite Rapids)	September 2024	P-core	Intel 3 + EMIB
Xeon 6+ E-core (Clearwater Forest)	1H 2026 (initial target)	E-core	Intel 18A + Foveros Direct 3D / EMIB 3.5D
Diamond Rapids	2H 2026 or later	P-core	Unconfirmed

CPU : ARM

Arm CPUs

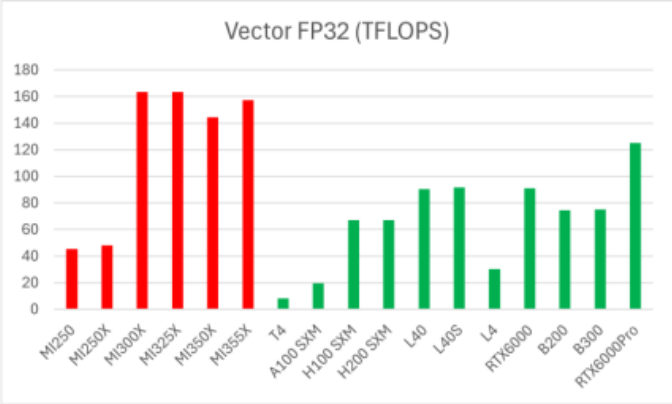
- **Nvidia Vera CPU, 2x performance over Grace**
  - 88 cores, 3 nm, SMT, +40% L3 (162 MB), 1.5 TB of SOCAMM
  - Offered also as a standalone product
- **Arm just announced their own AGI CPU for 2026H2**
  - Targets AI infrastructure, marketed as "agentic CPU"
  - 136 Neoverse V3 cores, 3.7 GHz, PCIe6, DDR5-8800, TDP 300W
- **Ampere**
  - Ampere Altra will carry on for low power compute (up to 128 cores)
  - AmpereOne (192 cores) → AmpereOne M (192 cores, more DDR5 channels and PCIe lanes) → AmpereOne MX (256 cores, 3nm, 512 cores)
- **Fujitsu set to release Monaka in 2027**
  - 2nm, 144 cores, 12-channel DDR5, PCI 6.0 with CXL,
  - Positioned for AI inference, simulation and data processing
  - Followed by Monaka X (2029), Monaka X with NPU (2030), Monaka XX (2031)
- **Arm is being considered as a power-efficient alternative to x86**

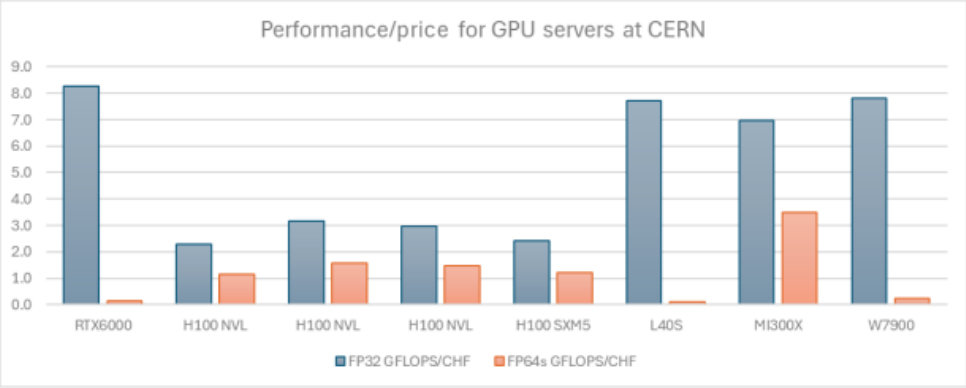


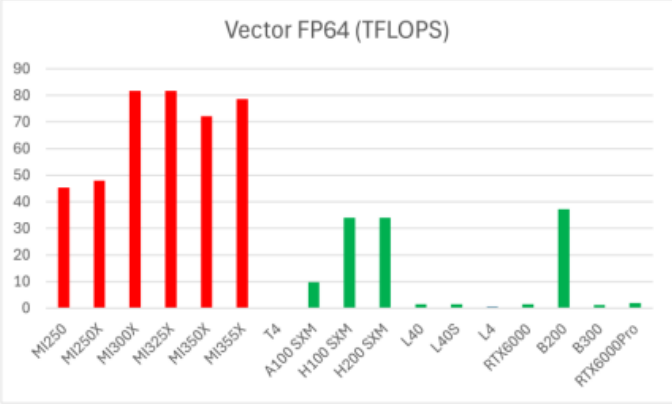


Takeaway message for GPUs

- **Generation-to-generation improvements focus on AI**
  - low precision performance, memory size and bandwidth
- **AMD seems keener on having HPC-oriented products than Nvidia**
  - Nvidia pushing for “emulated” FP32/FP64
- **Current gen AMD GPUs are the FP32 champions**
  - But the RTX 6000 Pro leads on FP32 TFLOPS/$
- **Current gen AMD GPUs are the FP64 champions**

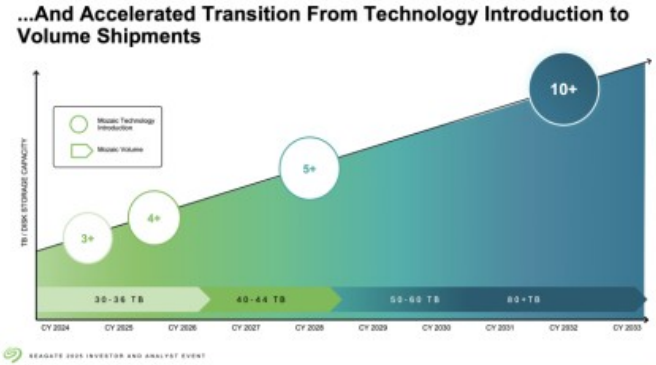


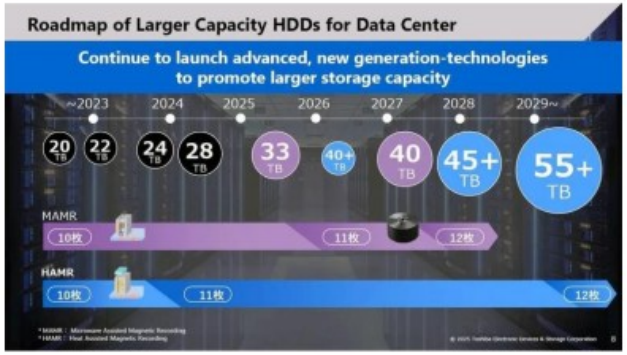




Disk storage

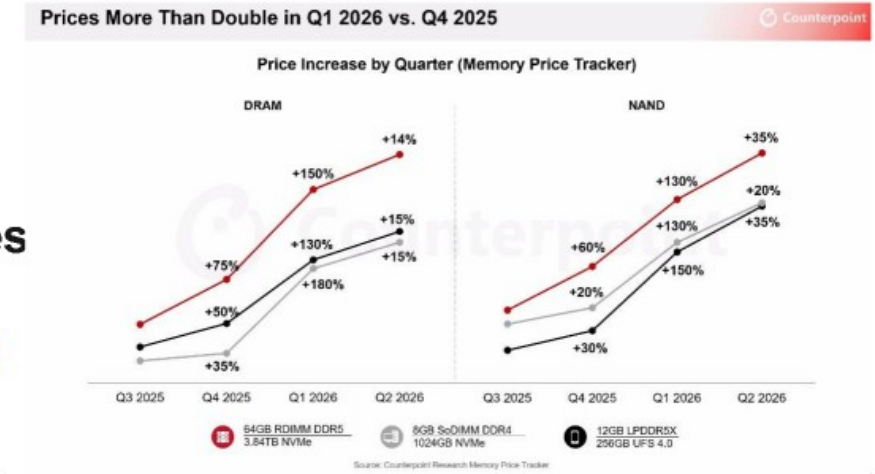
- **HDD capacity is growing at a rate similar to storage demand (thus number of shipped units remains ~constant)**
  - Today's "sweet spot" for price is around 24-26 TB
- **All HDD makers use some sort of energy-assisted magnetic recording to increase capacity per platter**
  - Seagate is the technology leader with 44 TB HAMR drives
  - Toshiba had a more cautious roadmap, using more platters and microwave-assisted MR (MAMR), with HAMR adopted after 2027
  - WD chose energy-assisted perpendicular MR (ePMR) + SMR until 60 TB, then fully adopt HAMR
  - All makers expected to reach 80-100 TB capacity around 2030-2031
- **Finally, progress on performance is planned**
  - Seagate intends to increase the number of actuators
  - WD will introduce high performance drives aiming at 8x bandwidth and 4x I/O using Dual-Pivot and High Bandwidth technologies





Price evolution

- **CPUs have been mostly spared by price increases**
  - AI data centers are fully GPU-centered
  - This might change soon, due to the explosion of agentic AI
- **Memory prices increased by ~5x**
  - Production will increase by just ~20%/year and won't meet demand until 2030
  - Shortages may happen
- **Storage costs maybe the biggest concern**
  - HDD became ~30-50% more expensive, SSDs 5x more!
  - Tape is not immune, tape drives price increased by ~30% and by about ~5-10% for tape media
- **AI not the only reason but the main one**
  - Helium shortage may impact the entire semiconductor and storage industry



Impact on resource acquisition

- **Computing servers**
  - Increase of memory prices is enough to increase $/HS23 by 2-3x
  - Unlikely to get back to the “flat budget” curve within the next few years
- **GPU servers**
  - Needs still marginal, but sites will have less money to spare for GPUs
- **Disk storage**
  - Currently, a big source of concern, given the experiments’ needs
  - Flash vs. HDD price gap has become enormous
- **Tape storage**
  - Probably price reductions will not happen as previously expected

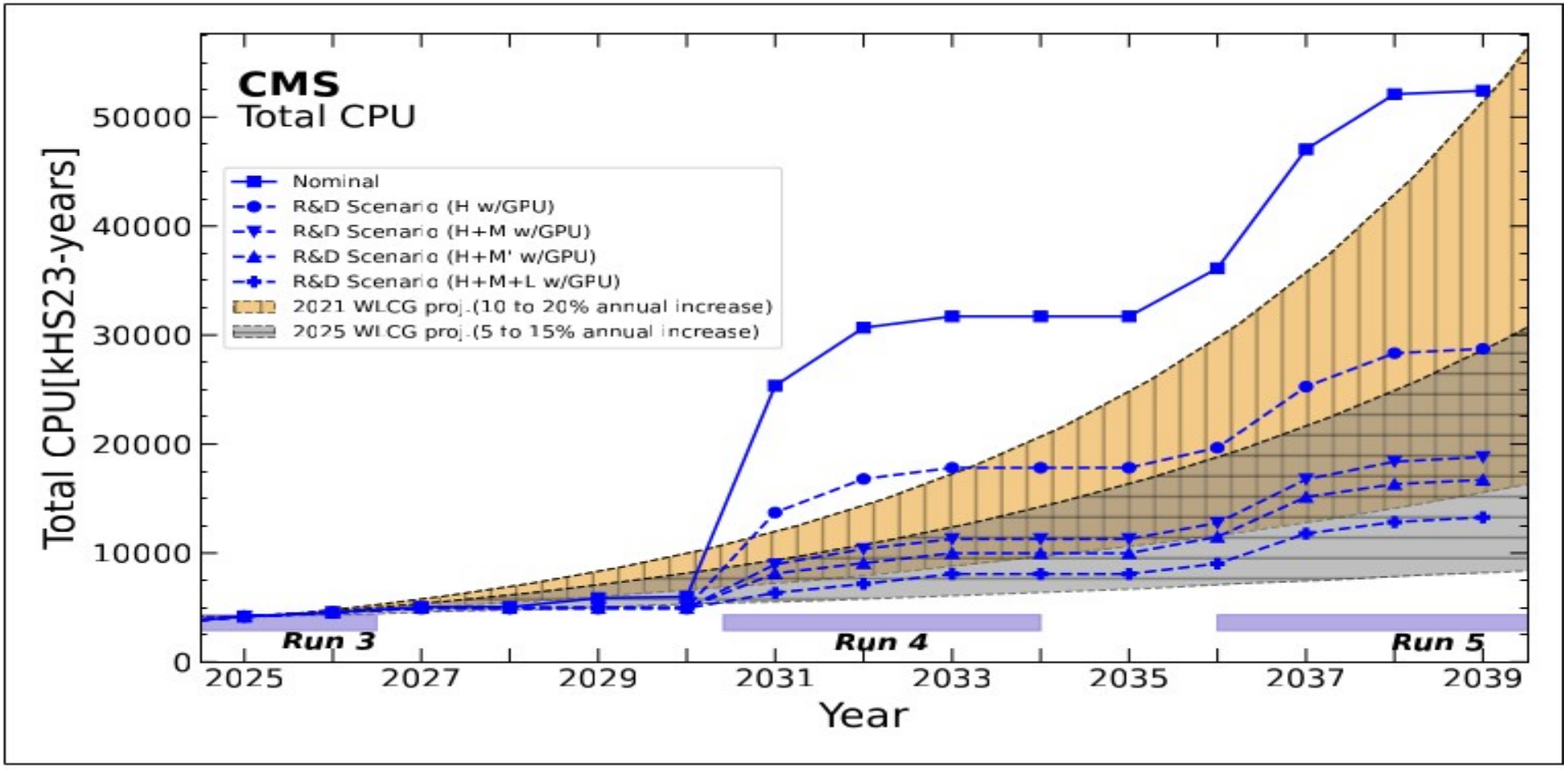
CMS Conceptual Computing Model

- Sorti en mars CERN CDS mais pas beaucoup de pub autour
  - Ne ferme aucune porte de R&D donc difficile de se projeter
  - Propose une discussion dédiée plus tard avec aussi le CM ATLAS, ici quelques points saillants
- Taux de trigger grands, plus de parking → besoins de stockage très importants. Primary processing hors site, en particulier si GPU devient important
- *Coût total CPU pour le début du Run 4 si on l'achetait dans un seul DC: 25 MHS23-years 170 racks de 20 dual-proc AMD Epyc 9754 CPU 200m2 5MW ~100MCHF et 2-5MCHF coût électrique annuel*
-

CMS Conceptual Computing Model

- *Specifically for the CMS HLT application, benchmarks show that one NVIDIA Tesla T4 is roughly equivalent to 1400 HS23 [113], and a newer NVIDIA L4 to 3200 HS23*
- Le scenario minimal ne rentre pas du tout dans les coûts !

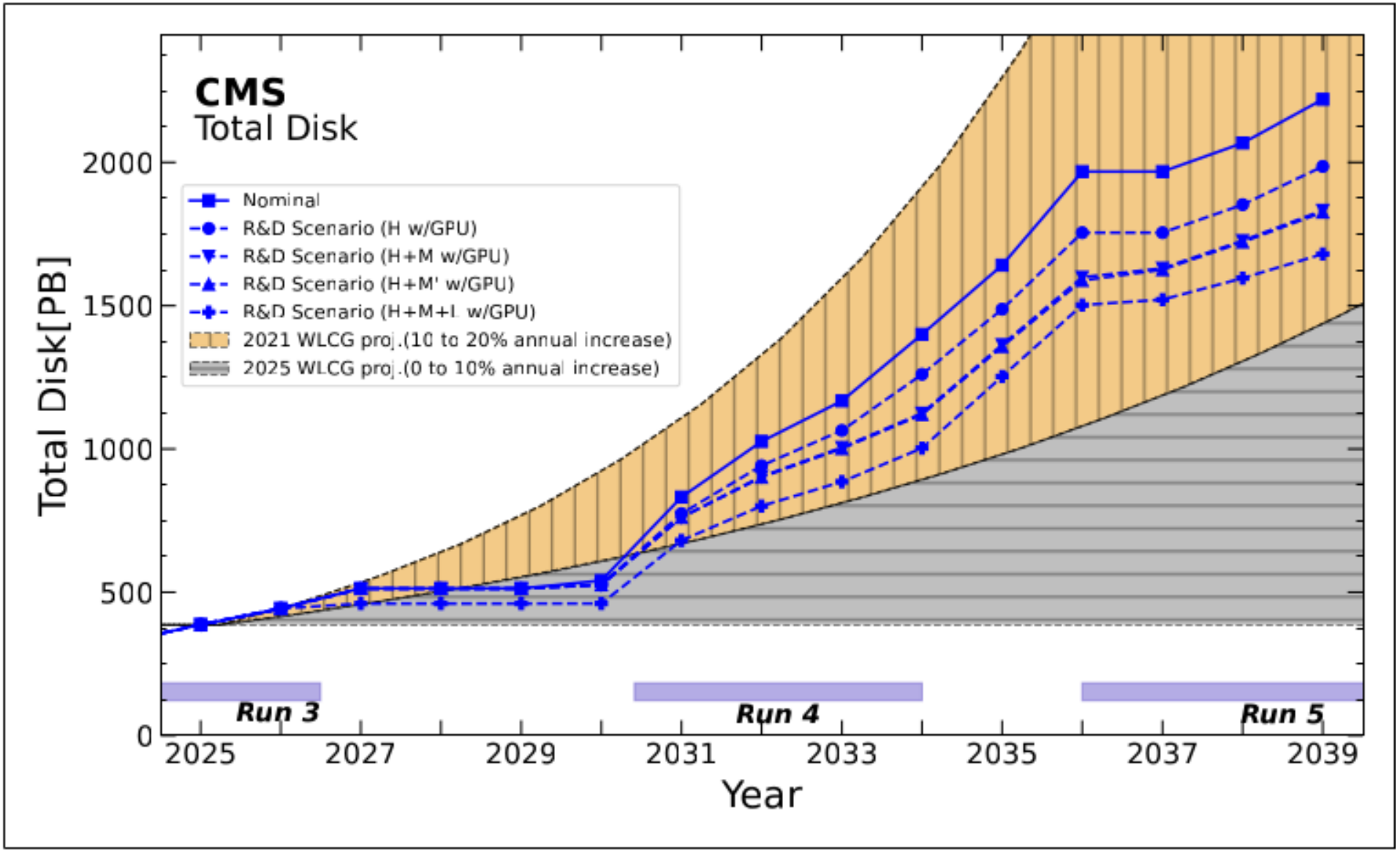
CMS CPU



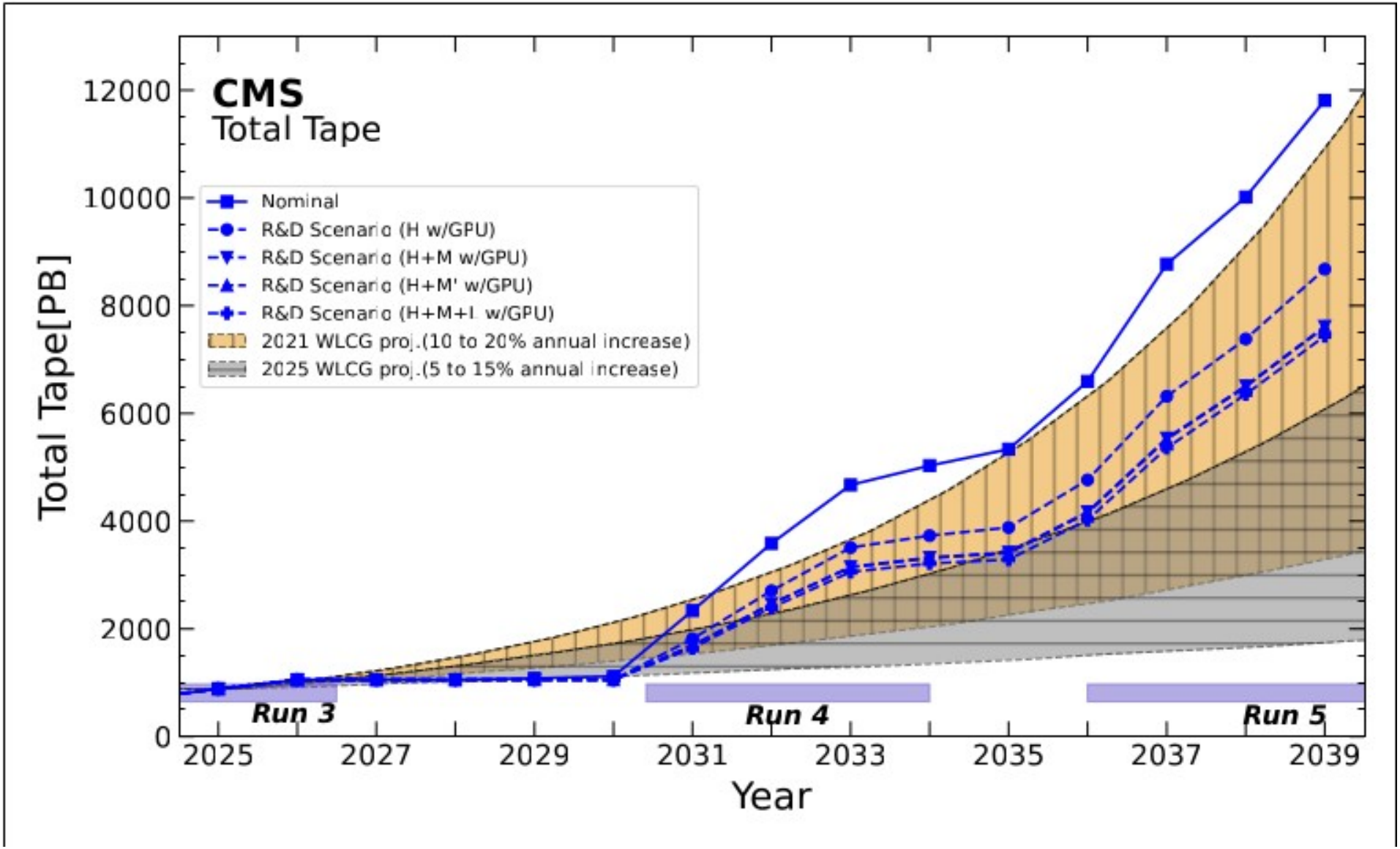
Zone grisée : 5-15 % baisse de coût

Tous les scenarii sauf « carrés » incluent l'utilisation de GPU !

CMS disk


CMS
Total Disk
Nominal
R&D Scenario (H w/GPU)
R&D Scenario (H+M w/GPU)
R&D Scenario (H+M' w/GPU)
R&D Scenario (H+M+I. w/GPU)
2021 WLCG proj.(10 to 20% annual increase)
2025 WLCG proj.(0 to 10% annual increase)
Total Disk[PB]
0
500
1000
1500
2000
2025
2027
2029
2031
2033
2035
2037
2039
Year
Run 3
Run 4
Run 5

CMS tape


CMS
Total Tape
Nominal
R&D Scenario (H w/GPU)
R&D Scenario (H+M w/GPU)
R&D Scenario (H+M' w/GPU)
R&D Scenario (H+M+I. w/GPU)
2021 WLCG proj.(10 to 20% annual increase)
2025 WLCG proj.(5 to 15% annual increase)
Total Tape[PB]
0
2000
4000
6000
8000
10000
12000
2025
2027
2029
2031
2033
2035
2037
2039
Year
Run 3
Run 4
Run 5