Le calcul intensif au service de la connaissance

BALANCING CAPEX, OPEX, AND SUSTAINABILITY IN HPC PROCUREMENTS

Context & Challenges

❑ **Supercomputer lifetime duration strategy : A key sustainability and economic lever**

HPC infrastructure ***global efficiency*** depends strongly on environmental, energy, and economic constraints

How to maximize the service provided wrt OPEX+CAPEX and carbon footprint over years ?

- Focus on Perf / Watt (ETS taking over TTS) (procurement & production)

  Energy to Solution   Time to Solution

- Carbon foot print to be considered in a holistic approach

Components Global life cycle, Integration, Energy, Cooling …

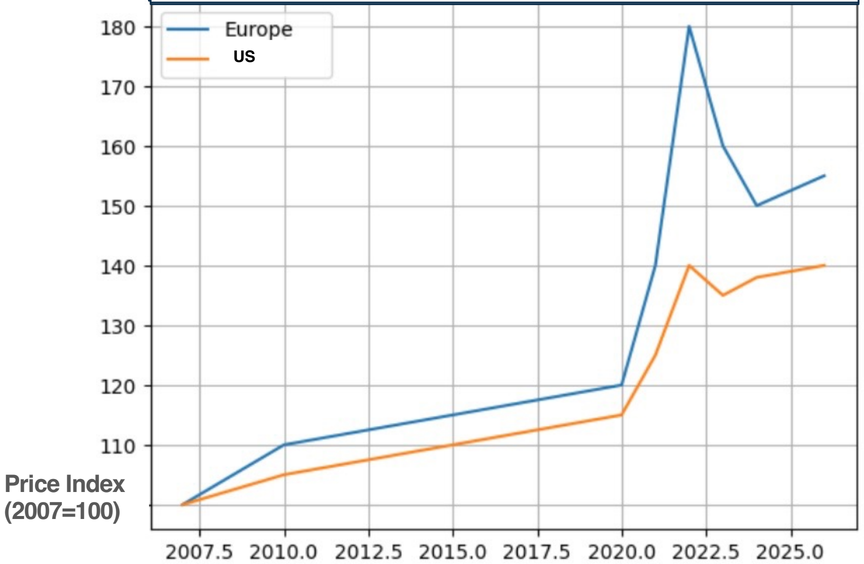
- Consider TCO with the OPEX growth relatively to CAPEX

While Moore's law survives (somehow), Dennard's law is stopped
(dissipation / consumption increases wrt processor surface)

- OPEX strongly impacted by electricity price trend

Let's put it into mathematical models …

Electricity prices in Europe and the US



BALANCING CAPEX, OPEX, AND SUSTAINABILITY IN HPC PROCUREMENTS

Models for actionnable strategies (1/3) :

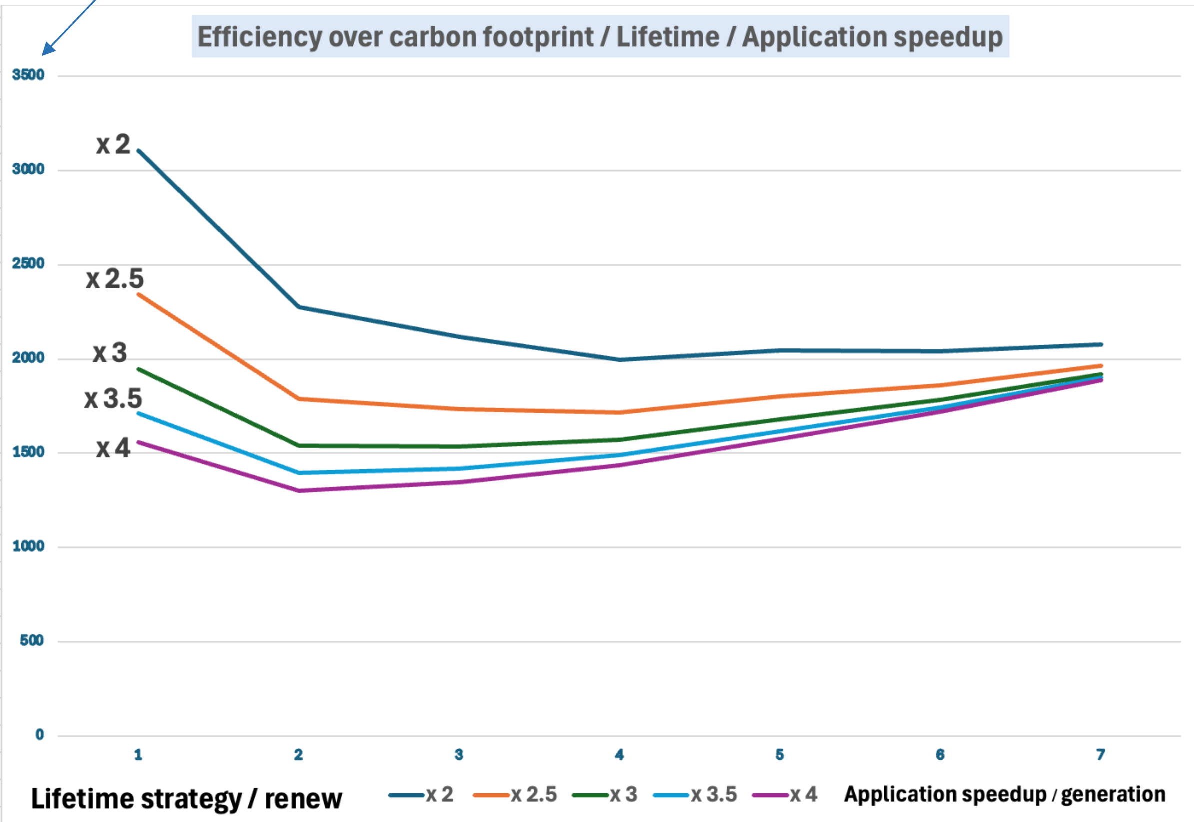
Efficiency related to carbon footprint

❑ Variables in use (summary)

$$Ay = A^{\left(\frac{1}{g}\right)} = e^{(\ln A)/g}$$

- g : years between techno generation
- A : App acceleration / per generation
- Ay : A on a per year basis
- Fab-CO2 : Carbon emissions due to fabrication (gpus, cpus, memories, racks …)
- Prod-CO2 : Carbon emission due to production phase (energy consumption)
- PUE : Power Usage Effiency - Data center
- Usage : Percentage usage of facility

Carbon emissions related to app performance



in years

Results are in carbon emissions related to application production

For a range of application speedups :
Carbon emissions / application speedup / based on conservation duration
(sum up over 10 years scenario)

GENCI
Le calcul intensif au service de la connaissance

BALANCING CAPEX, OPEX, AND SUSTAINABILITY IN HPC PROCUREMENTS

Models for actionnable strategies (2/3) :

Efficiency related to global budget i.e. TCO

For a fixed budget over n years :

What is the optimal conservation duration depending on Applications acceleration ?

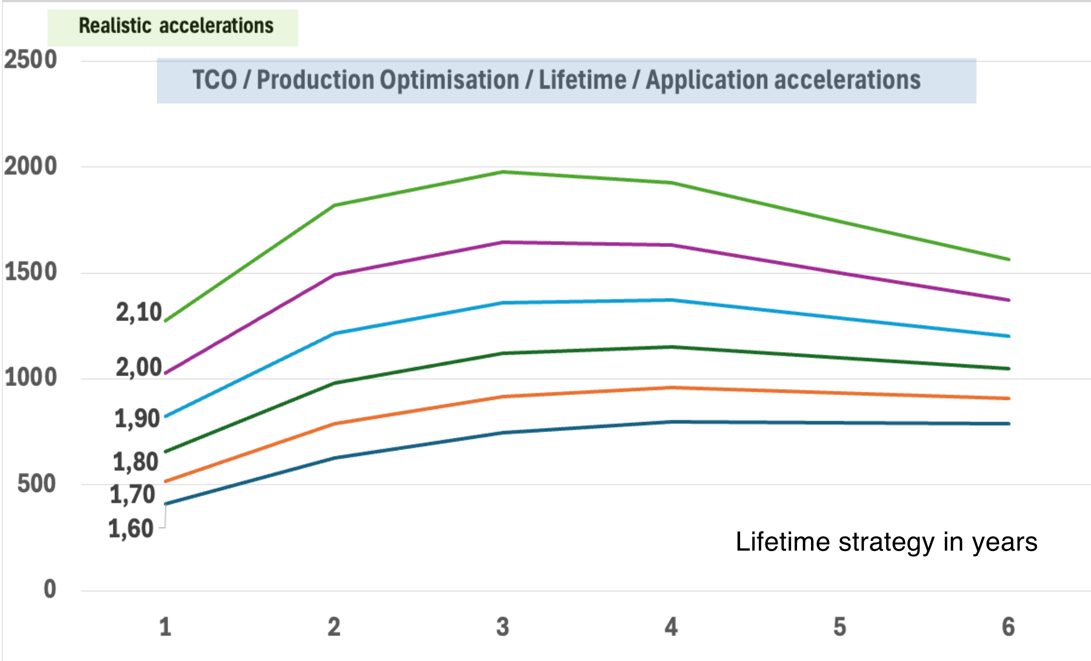
❑ Variables in use (including some of the Carbon Foot Print model)

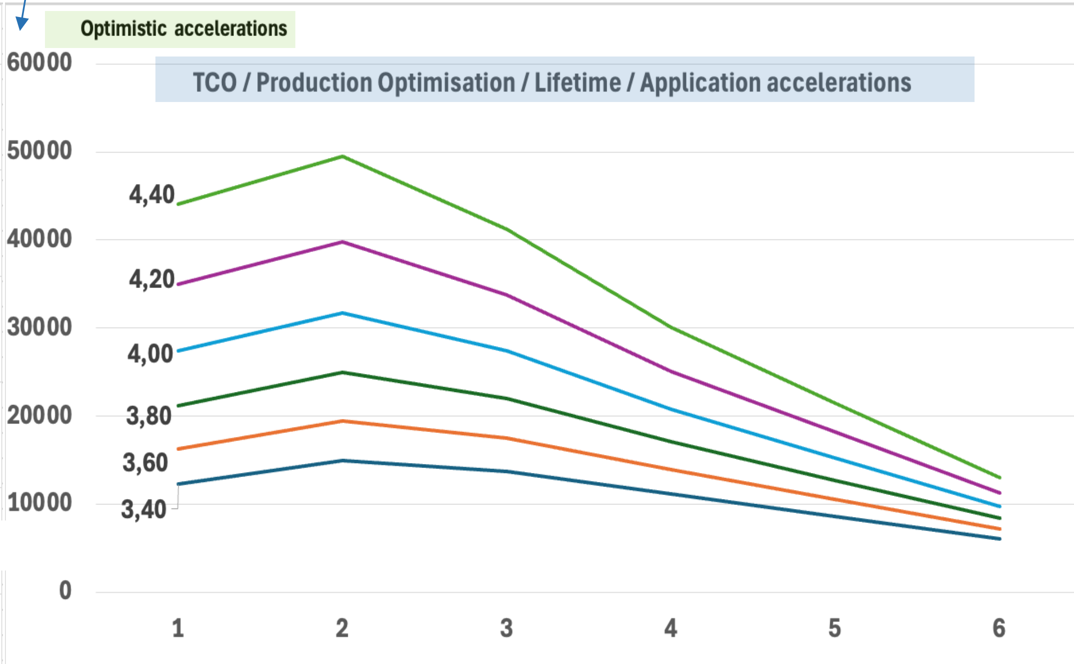
- Annual Budget = CAPEX/n + OPEX)

With OPEX/CAPEX increasing over generations (energy part ~10% ratio / year)

- Energy price / maintenance

Application production related to total budget





GENCI
Le calcul intensif au service de la connaissance

BALANCING CAPEX, OPEX, AND SUSTAINABILITY IN HPC PROCUREMENTS

Models for actionnable strategies (3/3) :

Conclusions - Limits & considerations

Conclusions :

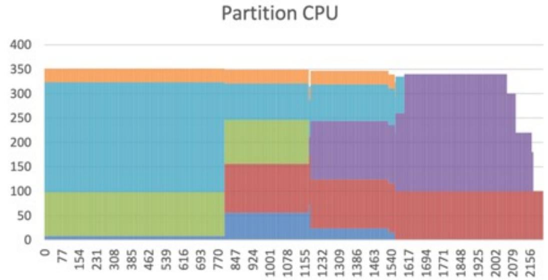
- Similar results either on TCO or Carbon footprint
- Acceleration of application is key.
  - Below x1.8, flat after 4 years and more ➔ don’t change !
  - Reasonable speedup (x2 – x3) optimal lifetime is between 4 and 3 years
  - Optimistic / breakthrough speedup (> x3) could tend to 2 years
  - From x2, longer lifetime (after optimum) leads to :
    - ➢ non optimal use of budget
    - ➢ carbon footprint inefficiency

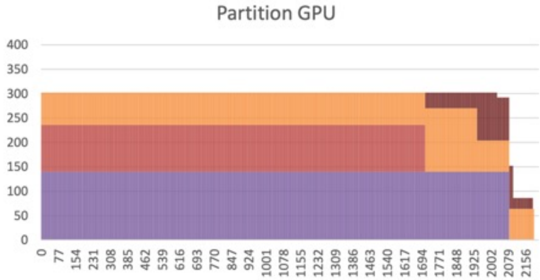
Limits & Considerations :

- *Pace and increase of consumption between technology generation are assumed constant*
- *CAPEX sky-rocketing due to memory crisis not included in the model !*
- *Carbon footprint is based on French energy mix*
- *Holistic Carbon footprint is complex (different lifetime contributors to consider / decision target)*
- *Scope is homogenous configuration*
- *Assumption that applications are pre-existing for comparisons*
- *Applications acceleration must be considered for <u>global throughput</u>*
- *Warning for high lifetime : maintenance (5 – 10% of CAPEX per year) may rise or become not available*

BACKUP SLIDE

TCO Method





Variable definition :

t : wallclock time of the workload
ECt : Energy consumption by computational nodes
EDt : Energy consumption from disks and service nodes
EUt : Energy consumption from internal cooling devices
c : total core count of the proposed configuration
cwl : core count used for the selected workload
T : Number of second in 5 years
N : Number of runs of the workload during 5 years
ET : Total energy consumption during 5 years
EC : Energy of computational part during 5 years
ED : Energy of storage and service nodes during 5 years
EU : Energy of internal cooling device during 5 years
TCOE : TCO due to energy consumption during 5 ans
PUEair : Power Usage Effectiveness for air cooled devices
PUEcoldw : Power Usage Effectiveness for "Cold water" cooled devices
PUEhotw : Power Usage Effectiveness for "Hot water" cooled devices

Then :

$$E_T = E_C + E_D + E_U$$

$$E_T = \left[\left(E_{Ct} * \frac{c}{c_{wl}} * PUE_{compute}\right) + \left(E_{Dt} * PUE_{disk}\right) + \left(E_{Ut} * PUE_{cdu}\right)\right] * \frac{T}{t}$$

$$TCO_E = E_T * EnergyCost$$

$$TCO = Purchase + Maintenance + Infrastructure + Staff + TCO_E$$

Where:

- *Purchase* is the acquisition cost of the machine
- *Maintenance* is the maintenance cost over 5 years
- *Infrastructure* is the specific cost of infrastructure for the machine
- *Staff* is the cost for complementary staff for machine exploitation over 5 years

Cost of the workload (TCO_b) takes into account the number of time (N) it twill be possible to run it over 5 years

$$N = \frac{c}{c_{wl}} * \frac{T}{t}$$

$$TCO_b = {TCO}/_{N}$$

Energy calculations :

$$E_C = E_{Ct} * \frac{c}{c_{wl}} * PUE_{compute} * \frac{T}{t}$$

$$E_D = E_{Dt} * PUE_{disk} * \frac{T}{t}$$

$$E_U = E_{Ut} * PUE_{cdu} * \frac{T}{t}$$

With

$$PUE_{compute} = 1 + \left[(PUE_{air} - 1) * \%_{air-for-compute}\right] + \left[(PUE_{coldw} - 1) * \%_{coldw-for-compute}\right] + \left[(PUE_{hotw} - 1) * \%_{hotw-for-compute}\right]$$

$$PUE_{disk} = 1 + \left[(PUE_{air} - 1) * \%_{air-for-disk}\right] + \left[(PUE_{coldw} - 1) * \%_{coldw-for-disk}\right] + \left[(PUE_{hotw} - 1) * \%_{hotw-for-disk}\right]$$

$$PUE_{cdu} = 1 + \left[(PUE_{air} - 1) * \%_{aircdu}\right] + \left[(PUE\_coldw - 1) * \%_{holdw-for-cdu}\right] + \left[(PUE_{hotw} - 1) * \%_{hotw-for-cdu}\right]$$