



HAL
open science

Optimizing the Impact of Upgrading Computer Equipment

Gabriel Breuil, Olivier Hermant, Renaud Pawlak

► **To cite this version:**

Gabriel Breuil, Olivier Hermant, Renaud Pawlak. Optimizing the Impact of Upgrading Computer Equipment. ICT4S, Jun 2024, Stockholm (Suède), Sweden. 10.1109/ICT4S64576.2024.00033. hal-04886183

HAL Id: hal-04886183

<https://minesparis-psl.hal.science/hal-04886183v1>

Submitted on 14 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimizing the Impact of Upgrading Computer Equipment

Gabriel Breuil, *Constellation* Olivier Hermant, *Mines Paris, PSL University* Renaud Pawlak, *Cincheo*

Abstract—Upgrading hardware platforms and infrastructure, like personal computer, servers, or supercomputers suffers from a tension when it comes to carbon neutrality. Increasing performances are an incentive for fast renewal, while manufacturing environmental costs urge us to extend as much as possible the lifetime of devices. Thus, carbon neutrality seems to be difficult to reach. In this article, we try to answer quantitatively to the question “what is the optimal point for which we may tend to the carbon neutrality of infrastructures”? Our approach is based on a model of the greenhouse gas emissions, the primary energy and the monetary cost of two parameters that vary country by country: the use of a microprocessor, and its fabrication. We have investigated three different renewal scenarios: the maintain of the workload capacity, the maintain of the number of processor cores, and the maintain of the number of microprocessors. Finally, we apply our model to two concrete CPUs. The most efficient solution (*i.e* with the lowest impacts) is to replace the entire batch of CPUs with the newer batch as soon as possible. This optimal-replacement strategy depends on the specifications of CPUs and allows to determine the optimal scenario.

Index Terms—Data centers, hardware upgrading rate, carbon-neutral system, impact models, life cycle assessment, resource management.



1 INTRODUCTION

In 2019, 4.1 billions users of Information and Communication Technologies (ICT) have been identified, and the demand is still growing. This increase is partly driven by the desire of having more and more efficient ICT infrastructures [1]. The energy cost of the fabrication of the equipment (end-user devices, Data Centers and network) corresponds to 44% of the total ICT energy consumption. Nicola Jones explains in [2] that by 2030 the ICT will increase up to 21% of the total electricity demand. These power consumption expectations split into a constant factor for the production of ICT and consumer devices, and an exponential increase for networks and Data Centers. The demand of Data Center services is indeed expected to expand within the next few years, driving an effort towards more efficiency. To do so, one has to enhance their workload capacity, which can be done for example by virtualization and increasing the performance of the electronic components used in Data Centers. As a consequence, each year, new generation of processors, hard disks, RAM, optical fibre, and so on are made available in the market [3]–[8] and replace older equipment. Moreover, even if the workload of Data Centers is not saturated such that an upgrade of electronic components is not strictly necessary, hardware replacement still may occur. Indeed, components have generally a three-year warranty that coerces customers to a three-year life-cycle [9], [10].

Carbon-neutral systems and infrastructure are a key issue for the green transition. However, carbon neutrality in the ICT is not an easy task and may be difficult to reached. Moreover, it exists only few investigations on carbon-neutral data-centers, which are quite recent [11]–[13]. These works suggest new policy instruments, technological methodology, and investigation on the overall

energy consumption of Data Centers.

The huge amount of data transferred through the network, the periodic upgrade of electronic components in Data Centers and the energy consumption of the infrastructure are the main reasons for an increase of energy consumption in Data Centers. The energy consumption of the infrastructure can be divided in two almost equal parts: computing equipment (processor, service power supply, storage, communication equipment, and so on) and support systems (mainly cooling, building switch-gear, PDU, and UPS). Even though the support systems are composed of various elements, cooling is predominant and represents the majority of the energy consumption of the infrastructure [14], [15]. The energy consumption of the computing equipment is balanced among various categories [2], [15]–[18].

One of the challenges of the digital 21st century is to reach carbon-neutral systems, infrastructures, architectures etc. Thus one needs to evaluate and reduce the energy consumption and the environmental impact of Data Centers. Over the last decades, to evaluate the performance of Data Centers, metrics such as the Power Usage Effectiveness (PUE), Thermal Design Power (TDP), Power Consumption (PC), Power Effectiveness (PE), and so on have been used. Scientists have worked on the enhancement of Data Centers performance by optimizing these metrics [16], [19]–[21]. Among those, TDP and PUE are the most commonly used. TDP gives the maximal power value of a CPU when it is used at 100% and PUE gives the ratio of the total facility energy load (the support system and the computing equipment) divided by the ICT energy load (the energy load of the computing equipment). During the past decade, the main focus has been to get a PUE as close as possible to 1. Indeed, it

would mean that all of the energy load of the facility is dedicated to the ICT equipment. Doing so, it has led to optimizing the efficiency of the cooling and power infrastructure [22], [23]. Nevertheless, optimizing the PUE has no direct benefit on the efficiency of ICT nor on the actual energy consumption and environmental impact of Data Centers and of their electronic components.

To go further on the improvement of Data Centers, one should also take into account their global impact on the environment through a Life Cycle Assessment (LCA) including the LCA of its components [24]–[26]. The LCA allows to consider the environmental impacts of the manufacturing, transport, usage and end-of-life of the components of Data Centers.

Rabih Bashroush and co-authors have focused their work on the upgrade of hardware in Data Centers [27]–[30]. In the study published in 2018 [27], Bashroush has put in the spotlight the energy saving and the impact of the hardware upgrade. The model is based on the replacement of one server by another one, which performances are assessed by Koomey’s law. The payback point is then determined as the time at which the upgrade starts showing energy benefits. Moreover, the model is validated by using real-life dataset, in particular a lifecycle impact analysis [31] of servers shows that the use phase of a Data Center has the strongest impact on the total consumption. In addition to that, Doyle *et al.* have studied in [30] the economical impact – and so the money saving – of the hardware renewal with Bashroush’s model. They established that it is possible to obtain a return on investment in the case of reducing the server population. Then, in [29], Bashroush *et al.* suggested that a renewal with re-manufactured servers is an efficient alternative for 5-6 years old servers.

It is essential to detail the hardware renewal policy and choose to replace CPUs, drive bay, etc. only when it has a positive impact globally: immediate or amortized money, energy and CO₂ emissions gains, while reducing useless wastes. The review of Jin *et al.* [32] analyses the servers’ energy consumption and classifies the different components in Data Centers. The ranking for energy consumption is as follows: the central processing unit (CPU) is the higher energy consumer (32% of a server), the power supply is 15% of a server and the memory is 14%. At the exception of cooling, the dominating factor is therefore CPU. Moreover, the energy consumption of CPUs has been widely studied in order to be capable of measuring it and reducing it [33]–[36].

Thus, we describe in this article a two-CPU renewal model and attempt to determine the ideal frequency and proportions to replace an old-generation CPU by a new one, along several criteria and with different scenarios in order to tend to a carbon-neutral system. We then extend this model to three generations of CPUs, allowing to target more realistic scenarios. Our models use a precise quantitative assessment of the manufacturing costs.

2 OUTLINE

The section “Methodology” describes the model, built on a comparison of the performances, the manufacturing

energy, and the energy use between the cores of two different CPUs. This model offers three degrees of liberty: the proportion of CPUs to replace, the replacement frequency and the date of the first replacement. As well, three renewal policies are proposed: a CPU-to-CPU match, a core-to-core match, and a workload capacity-to-workload capacity match. Then, we extend this model to take into account different levels of hardware failure (MTBF) for three successive generations of CPUs.

The section “Case studies” discusses the results given by this model. It is separated in two parts. We first validate the two-CPU renewal model and determine the best CPU-replacement scenarios in a Data Center, according to the CPU respective specifications. We also focus on the sensitivity of the results to MTBF. The second part is a forecast on the three-CPU renewal model, that compares the costs of upgrading a CPU twice in a row, or only once. In both parts, the model takes into account the market availability constraint of the newer-generation CPUs. The impact of these renewal scenarios is assessed on economical, environmental and energetic criteria, and we compare the optimal points of each of them.

The section “Conclusion and future works” closes the article by discussing potential improvements and uses of our model.

3 METHODOLOGY

We assume that the environmental impact of using a processor is proportional to its electric energy consumption. In other words, we consider that the power consumption of a processor and of the infrastructure (cooling and global Data Centers power supply) is directly related to the environmental impact of processors. Similarly, we assess the manufacturing environmental costs in terms of energy.

3.1 Comparison between two different versions

Let P_1 and P_2 be two processors that start being manufactured at date d_1 and d_2 , respectively. We consider in this study that P_1 and P_2 belong to the same family and therefore have characteristics that are meaningful to compare, as it can be the case for the Intel® Xeon® processor family for instance. It is nevertheless unfair to compare global characteristics and performances, since P_1 and P_2 may enjoy different amount of cores, so we have to normalize.

For each processor P_i , let us define its manufacturing energy per core (Wh/core):

$$E_i^m = \frac{\mathbb{E}_i^m}{n_i} \quad (1)$$

where $i \in \{1; 2\}$. \mathbb{E}_i^m is the total manufacturing energy cost of the processor P_i . It has been studied in-depth by the working group Boavizta [37], and it depends on the lithography technology, the amount of cores, and the die size. n_i is the quantity of cores of P_i .

The annual usage energy per core (Wh/core) reads,

$$E_i^u = \frac{TDP_i}{n_i} \cdot y \quad (2)$$

where TDP_i is the Thermal Design Power (TDP) of P_i and y corresponds to the yearly use time – in this article, we

let $y = 8760\text{h}$, corresponding to an uninterrupted round-year usage, as it is standard in data centers. We consider the time t as being discrete with one-year steps.

We consider that the Data Center is running only P_1 processors at $t_0 \geq d_1$ and that, at this date, it starts to operate. At the end of the t_{\max} time step, the Data Center stops to operate. In order to evaluate the impact of different renewal policies, we introduce two more time parameters : t_S , the starting date of the renewal scheme, and t_I , a fixed time interval that separates two renewals. Each renewal will replaces part of P_1 processors by P_2 ones. To make calculations, we distinguish four stages in which the Data Center can be over its lifespan :

- the Data Center contains only P_1 's (initial stage);
- a certain amount of P_1 's is replaced by some P_2 's (replacement stages);
- the mix of P_1 's and P_2 's is stable, but a replacement will occur later (interval stages) ;
- no further replacement is scheduled (final stage).

In order to track the evolution of the quantity of cores of processors P_1 and P_2 , and their impact, we introduce the following sequences: $(N_{1_k}^r)_{k \in \llbracket 0, k_{\max} \rrbracket}$ the amount of P_1 cores that are removed and $(N_{2_k}^r)_{k \in \llbracket 0, k_{\max} \rrbracket}$ the amount of P_2 cores that replace them at the k^{th} replacement date. Correspondingly, $(N_{1_k})_{k \in \llbracket 0, k_{\max} \rrbracket}$ and $(N_{2_k})_{k \in \llbracket 0, k_{\max} \rrbracket}$ represent the respective amount of P_1 and P_2 cores after the k^{th} replacement date.

3.1.1 Initial stage : only P_1 processors

The Data Center contains no P_2 until t_S , the initial renewal date (that respects the constraint $t_S \geq d_2$). This spans over $\Delta t_S = t_S - t_0$ steps. The total amount of cores in the Data Center is labelled $N_{1_{\text{tot}}}$. The total energy that is spent during Δt_S is :

$$E_{\text{init.}}(P_1) = N_{\text{tot},1}(E_1^m + E_{u,1} \cdot \Delta t_S) \quad (3)$$

The value of $E_{\text{init.}}$ depends only on the characteristics of P_1 (E_1^u and E_1^m) and is indeed a function of P_1 .

3.1.2 Replacement stage

The renewal dates t_k are evenly scheduled. It happens $k_{\max} \in \mathbb{N}^* + 1$ times, and for all $k \in \llbracket 0, k_{\max} \rrbracket$,

$$t_k = t_S + k \cdot t_I \quad (4)$$

$t_{k_{\max}}$ is the last renewal date. Therefore, we have to ensure that the following condition is respected

$$t_{k_{\max}} \leq t_{\max} \quad (5)$$

or, equivalently,

$$k_{\max} \leq \frac{t_{\max} - t_S}{t_I} \quad (6)$$

Therefore k_{\max} can be computed as the integral part of the right-hand side of eq. 6, since the parameters of our model are t_S , t_I and t_{\max} .

When t reaches such a renewal date t_k , we consider that the replacement happens at the first day of t_k (a calendar year, here). An amount of P_2 cores replaces

some P_1 cores and we count the annual usage energy consumption for each new P_2 core, symmetrically, we discount for the replaced P_1 cores.

At t_k , the total cost aggregates two factors : the usage cost and the manufacturing cost.

$$E_k(P_1, P_2) = N_{1_k} E_1^u + N_{2_k} E_2^u + N_{2_k}^r E_2^m \quad (7)$$

It is important to notice that the total energy consumption is now a function of P_1 and P_2 since it depends on the characteristics of both processors.

We investigate degressive renewal strategies. Therefore, the sequence $(N_{1_k}^r)_{k \in \llbracket 0, k_{\max} \rrbracket}$, that determines the number of removed P_1 cores, is defined by

$$(N_{1_k}^r)_{k \in \llbracket 0, k_{\max} \rrbracket} : N_{1_k}^r = N_{1_k} \cdot \rho \quad (8)$$

with ρ a percentage, the replacement fraction, that is a parameter of the model. Equation 8 is an approximation, as it is impossible to replace a fraction of a core or of a processor. The exact equation is

$$N_{1_k}^r = n_1 \cdot \lceil \frac{N_{1_k}}{n_1} \cdot \rho \rceil \quad (9)$$

Discretization is negligible when the processor number is large. Section 4.1.1 addresses the question quantitatively. Here and below, we only presents the continuous model.

The quantity of remaining P_1 cores, $(N_{1_k})_{k \in \llbracket 0, k_{\max} \rrbracket}$, is the result of accumulating the successive removals (that are themselves computed on the basis of the remaining cores, in a mutually recursive fashion). It reads as follows,

$$(N_{1_k})_{k \in \llbracket 0, k_{\max} \rrbracket} : \begin{cases} N_{1_0} & = N_{1_{\text{tot}}} - N_{1_0}^r \\ N_{1_{k+1}} & = N_{1_k} - N_{1_{k+1}}^r \end{cases} \quad (10)$$

The number of added P_2 cores, $(N_{2_k}^r)_{k \in \llbracket 0, k_{\max} \rrbracket}$, depends on the replacement policy, defined by a renewal ratio R :

$$(N_{2_k}^r)_{k \in \llbracket 0, k_{\max} \rrbracket} : N_{2_k}^r = R \cdot N_{1_k}^r \quad (11)$$

The quantity of P_2 cores in the Data Center, $(N_{2_k})_{k \in \llbracket 0, k_{\max} \rrbracket}$, reads as follows,

$$(N_{2_k})_{k \in \llbracket 0, k_{\max} \rrbracket} : \begin{cases} N_{2_0} & = N_{2_0}^r \\ N_{2_{k+1}} & = N_{2_k} + N_{2_{k+1}}^r \end{cases} \quad (12)$$

We now examine three scenarios that induce different renewal ratios R , that we call the **fixed-CPU scenario**, the **fixed-core scenario** and the **fixed-workload scenario**.

The **fixed-CPU scenario** preserves the number of CPUs: there is a one-to-one match between a removed P_1 processor and an added P_2 processor. The fixed-CPU renewal ratio reads therefore as follows,

$$R = \frac{n_2}{n_1} \quad (13)$$

The **fixed-core scenario** preserves the number of cores. The fixed-core renewal ratio is such that:

$$R = 1 \quad (14)$$

This readily entails

$$N_{2_k}^r = N_{1_k}^r \quad (15)$$

and as well $N_{2_k} + N_{1_k} = N_{\text{tot},1}$.

The **fixed-workload scenario** preserves the workload. The renewal ratio is then the ratio of the CPUs' performances per core, $\frac{B_i}{n_i}$, where B_i are the CPU PassMark® Software Pty Ltd performance benchmark values [38]. The ratio then reads as follows

$$R = \frac{B_1 \cdot n_2}{B_2 \cdot n_1} \quad (16)$$

3.1.3 Interval stage : between two replacement stages

This stage happens when $k \in \llbracket 0, k_{\max} - 1 \rrbracket$ and t satisfies

$$t_k = t_S + k \cdot t_I < t < t_S + (k + 1) \cdot t_I = t_{k+1} \quad (17)$$

In this case, one only has to take into account the energy consumption of the usage of the processors, which are the blend between P_1 and P_2 that has been set during the replacement stage t_k above. We multiply this energy consumption with the length of the interval, that is $t_I - 1$ time steps. Indeed the missing time step has been handled by the replacement stage above and interval stages occur only if $t_I \geq 2$. During those stages, for all $k \in \llbracket 0, k_{\max} - 1 \rrbracket$, the total energy consumption then reads,

$$E_{\Delta k}(P_1, P_2) = (N_{1k} \cdot E_1^u + N_{2k} \cdot E_2^u) \cdot (t_I - 1) \quad (18)$$

3.1.4 Final stage

Here, the replacement scheme has come to its end. This means that $t \in \llbracket t_{k_{\max}} + 1, t_{\max} \rrbracket$. In this case, P_1 might be fully replaced by P_2 processors or one still might be in presence of a blend, which typically happens with a degressive replacement strategy.

As this stage makes sense only if $t_{k_{\max}} < t_{\max}$, the condition on k_{\max} of eq. 6 becomes strict :

$$k_{\max} < \frac{t_{\max} - t_S}{t_I} \quad (19)$$

The total energy consumption of this stage then reads as follows, where $\Delta t_{k_{\max}} = t_{\max} - t_{k_{\max}}$

$$E_{\text{end}}(P_1, P_2) = (N_{1k_{\max}} E_1^u + N_{2k_{\max}} E_2^u) \cdot \Delta t_{k_{\max}} \quad (20)$$

3.1.5 Additional manufacturing energy consumption

The mean time between failure (MTBF) is the life expectancy between two failures. From the failure probability point of view, it means that, every year, a certain average proportion $x\%$ of the processors P_1 and P_2 break down. As a small simplification, we suppose that no such event occurs during the initial stage.

If a processor failure occurs on a P_1 processor at a replacement stage t_k , we consider that it is part of the N_{1k} replaced processors of this stage. This only requires, that x is lower than the replacement fraction ρ , which is, by an order of magnitude, the case in practice and there is no extra cost associated with it. Otherwise, we consider that P_i processors are replaced with new P_i processors, in which case we must add the manufacturing costs.

We neglect replacement during the initial stage (P_1 CPUs are new) and, for all $k \in \llbracket 0, k_{\max} - 1 \rrbracket$, the additional annual energy consumption functions for the k -th replacement and interval stages read:

$$E_{\text{add}k}(P_1) = \frac{x \cdot N_{1k}}{100} \cdot E_1^m \cdot (t_I - 1) \quad (21)$$

$$E_{\text{add}k}(P_2) = \frac{x \cdot N_{2k}}{100} \cdot E_2^m \cdot t_I \quad (22)$$

For the last k_{\max} -th replacement and final stages, the annual additional energy consumption is:

$$E_{\text{add}end} = \frac{x \cdot N_{1k_{\max}}}{100} \cdot E_1^m \cdot \Delta t_{k_{\max}} + \frac{x \cdot N_{2k_{\max}}}{100} \cdot E_2^m \cdot (\Delta t_{k_{\max}} + 1) \quad (23)$$

The global energy consumption. The global energy function is defined as being a function of P_1 and P_2 and corresponds to the energy consumption that is emitted during $t_{\max} - t_0 + 1$. It reads,

$$G(P_1, P_2) = E_{\text{init.}}(P_1) + \sum_{k=0}^{k_{\max}-1} \left[E_k(P_1, P_2) + E_{\Delta k}(P_1, P_2) + E_{\text{add}k}(P_1) + E_{\text{add}k}(P_2) \right] + E_{k_{\max}}(P_1, P_2) + E_{\text{end}}(P_1, P_2) + E_{\text{add}end} \quad (24)$$

3.2 Comparison between three different versions

It is possible to extend our definition of the global energy function to more than two processors. Let P_1 , P_2 and P_3 be three different processors of the same family that are manufactured at t_{P_1} , t_{P_2} and t_{P_3} . This yields a combinatorial explosion of the strategies, and in order to ease the comparison between them during the case studies, we restricted the replacement fraction to $\rho = 100\%$ and the replacement stages to occur only at t_{P_2} or at t_{P_3} (k_{\max} in eq. 24 below is then at most 1). The following global energy consumption function expressed in eq. 24 depends now on three variables since we take into account three processors. In its most general form, it reads,

$$G(P_1, P_2, P_3) = E_{\text{init.}}(P_1) + \sum_{k=0}^{k_{\max}-1} \left[E_k(P_1, P_2, P_3) + E_{\Delta k}(P_1, P_2, P_3) + \sum_{i=1}^3 E_{\text{add}k}(P_i) \right] + E_{k_{\max}}(P_1, P_2, P_3) + E_{\text{end}}(P_1, P_2, P_3) + E_{\text{add}end} \quad (25)$$

4 CASE STUDIES

4.1 Comparison between two different versions

In order to evaluate the economic, energetic and environmental impacts of CPU renewal in a Data Center, we estimate the model presented in section 3.1 with two processors: Xeon® Silver 4114 and Xeon® Gold 5220. Both of them enjoy the Skylake® micro-architecture. For simplicity, the two processors are abbreviated Silver and Gold respectively. Their specifications are shown in table 1. The three impacts are defined as follows.

- **Economical impact:** the manufacturing cost is defined to be the launch price listed on Intel's website [39]–[41] and the use cost is defined to be equal to the price of the electric energy consumed by the CPUs. We assume in this article that the fictional Data Center studied with our model is localized in France. In this case, TotalEnergies® [47] suggests a price of €0.1740 per 1 kWh.

Table 1: Specification of three processors Intel® Xeon® of micro-architecture Skylake®.

	Unit	Xeon® Silver 4114	Xeon® Gold 5220	Xeon® Platinum 8380
Release date		2017 [39]	2019 [40]	2021 [41]
PDT	W	85 [39]	125 [40]	270 [41]
Core amount		10 [39]	18 [40]	40 [41]
Lithography	nm	14 [39]	14 [40]	10 [41]
Die surface [42]	mm ²	325	485	640
Performance		13 125 [43]	25 740 [44]	62 318 [45]
Manuf. primary energy [37]	kWh	67.26	79.03	126.62
Manuf. carbon weight [46]	kgCO _{2e}	27.77	32.62	52.27
Manuf. price	€	704 [39]	1664 [40]	8666 [41]
Run on primary energy	kWh	744.60	1 095.00	2 365.20
Run on carbon weight	kgCO _{2e}	42.37	62.31	134.58
Run on price	€	129.56	190.53	411.54
Performance per core		1 312	1 430	1 558
	kWh/core	6.73	4.39	3.17
E^m	kgCO _{2e} /core	2.78	1.81	1.31
	€/core	70.40	92.44	216.65
	kWh/core	74.46	60.83	59.13
E^u	kgCO _{2e} /core	4.24	3.46	3.36
	€/core	12.96	10.59	10.29

• **Energetic impact:** we consider the primary energy (PE) – renewable or not – required to manufacture the processors and to run them.

• **Environmental impact:** it is estimated by the global warming emissions (GWP) which is expressed in carbon dioxide equivalent (CO_{2e}). The processor manufacturing weight of CO_{2e} is determined by using the weighted average of the electrical mix from the factories location [48]. The electrical mix of each country for 2021 has been got from the website Our World In Data [46]. The carbon intensity of electricity for the processor manufacturing is then 0.413 kgCO_{2e} for 1 kWh. The carbon intensity of electricity for the processor usage in France is 0.05693 kgCO_{2e} for 1 kWh [49].

The validation of our model and the determination of the best-case scenario of the Silver CPU replacement can be summarized in three questions:

- **Question 1:** Why does equation 24 approximates the actual global energy and by which margin?
- **Question 2:** Does the Mean Time Between Failure (MTBF) have a significant effect on the impacts of CPU replacement (economical, PE, and GWP)?
- **Question 3:** Given the Silver and Gold CPUs, how and when should we renew them in the Data Center in order to minimize the impacts ?

Unless otherwise specified, we consider that the Data Center is initially composed of 100 Silver in 2017. It stops at the end of 2026. Given the specifications of the Silver and Gold CPUs, listed in table 1, the fixed-CPU scenario yields a replacement ratio of 1.8, the fixed-core scenario yields a replacement ratio of 1, and the fixed-workload scenario yields a replacement ratio of 0.92. When the replacement is total, i.e. when the replacement fraction ρ is 100%, this respectively leads to 100 Gold CPUs (1800 cores), 55.56 Gold (1000 cores), and 50.99 Gold CPUs (917.83 cores).

4.1.1 Question 1: Why does equation 24 approximates the actual global energy and by which margin?

In order to confirm the replacement model that we suggest, one has to know for certain that the global energy we

compute is in good agreement with the reality. Equation 24 is an approximation since potentially fractional ratios (eq. 13,16) can generate fractional amount of Silver and Gold cores (eq. 10,11,12,15). The same holds on CPUs since the amount of cores is not necessarily a multiple of the amount of cores of a single CPU.

To assess the difference, we have investigated two cases: a high amount of Silver (100 CPUs – 1 000 cores) and a low amount of Silver (5 CPUs – 50 cores). For both initial amounts of Silver, we have determined two amounts of Gold (labelled “rational” and “integer”) for each of the three scenarios. A rational value corresponds to the computed values of cores and CPUs from equations 11,12,15. An integer value stand for upper rounded values of cores and CPUs. They are gathered in table 2. For each scenario, we computed the standard deviation of the global energy. They have been determined for the three impacts – PE, GWP, and Price. The standard deviations are gathered in table 2 as well.

We see that for a fixed-workload replacement, the calculated amounts of Gold are 50.99 (51 if rounded upper) and 2.55 (3 if rounded upper). The standard deviations of the three impacts are 0.01 % in the case of 100 Silver. However, the standard deviations in the case of low amount of Silver are higher, between 9.99 % and 11.61 %. Such an increase of the standard deviation is expected intuitively since rounding errors are absolute and benefit from the dampening of larger numbers.

A similar behavior of the global energy standard deviations for a fixed-core replacement is observed as well. The standard deviations are lower for 100 Silver while it is higher for 5 Silver. Since the fixed-CPU replacement implies a conservation of the amount of CPUs, the standard deviations are 0 % in this case.

The conclusion to this question is that the non-discrete replacement model that we suggest is valid in the case of large values of CPUs, 100 being already large enough. Moreover one has to keep in mind that we only take into account the manufacturing and the use of processors. We do not consider the processors end-of-life due to a

Table 2: Integer and rational amounts of Gold that are manufactured to replace Silver for the three scenarios and the related standard deviation of the global energy from the computed amount of CPUs for the three impacts: PE (Primary Energy), GWP (Global Warming Potential), and Price.

	Integer	Rational	$\sigma(\text{PE})$	$\sigma(\text{GWP})$	$\sigma(\text{Price})$
Fixed-workload					
100 Silver	51	50.99	0.01 %	0.01 %	0.01 %
5 Silver	3	2.55	11.61 %	11.10 %	9.99 %
Fixed-core					
100 Silver	56	55.56	0.60 %	0.58 %	0.52 %
5 Silver	3	2.78	5.73 %	5.47 %	4.93 %
Fixed-CPU					
100 Silver	100		0 %	0 %	0 %
5 Silver	5				

significant lack of available information, which means that the global energy is underestimated.

4.1.2 Question 2: Does the Mean Time Between Failure (MTBF) has a significant effect on the impacts of CPU replacement (economical, PE, and GWP)?

Table 3: Global energy for a total renewal in 2019 in the case of the three scenarios without MTBF $G_0\%$ and standard variations of the global energy with various MTBF ($\sigma_{3.00 \cdot 10^{-3}\%}$, $\sigma_{0.35\%}$, $\sigma_{2.90\%}$). ‘‘CPU’’ stands for fixed-CPU scenario, ‘‘Core’’ stands for fixed-core scenario, and ‘‘Workload’’ stands for fixed-workload scenario.

	$G_0\%$	$\sigma_{3.00 \cdot 10^{-3}\%}$	$\sigma_{0.35\%}$	$\sigma_{2.90\%}$
Primary energy (MWh)				
CPU	1 039.55	$2.00 \cdot 10^{-4}\%$	0.02 %	0.17 %
Core	646.70	$2.00 \cdot 10^{-4}\%$	0.02 %	0.17 %
Workload	606.35	$2.00 \cdot 10^{-4}\%$	0.02 %	0.17 %
Global warming potential (tCO _{2e})				
CPU	64.36	$1.00 \cdot 10^{-3}\%$	0.13%	1.15%
Core	40.75	$1.00 \cdot 10^{-3}\%$	0.13%	1.10%
Workload	38.33	$1.00 \cdot 10^{-3}\%$	0.14%	1.09%
Price (K€)				
CPU	415.14	$8.00 \cdot 10^{-3}\%$	0.90%	8.63%
Core	273.44	$8.00 \cdot 10^{-3}\%$	0.92%	7.61%
Workload	258.88	$9.00 \cdot 10^{-3}\%$	1.04%	7.44%

By definition, the higher the MTBF the longer a processor operates without any failure. The MTBF of electronic devices is not always given by manufacturer. However, B. Schroeder and G. A. Gibson [50] have analyzed failure data from a high-performance computing Data Center. 23 000 failures have been recorded between 1996 and 2005, 64 % of them are coming from hardware breakdowns and among them 42.8 % result from CPU breakdowns. Thus 700 failures are annually caused by CPU breakdowns over 24 101 processors, which gives a MTBF value of 301 607 h (around 34 years). In the end, it means that every year 2.90 % CPUs fail due to hardware reasons.

Two theoretical studies on hardware and software failures in Data Centers tackle CPU MTBF [51], [52]. In [51], the MTBF of CPU under normal operation condition is estimated to be 260 000 000 h, equivalently, each year 0.003 % of the CPUs break. In [52], the MTBF value chosen as input parameter for their study is 2 500 000 h. It corresponds to 0.35 % of CPUs failing every year.

In order to evaluate the effect of the MTBF on the costs, we consider that the 100 Silver CPUs are all replaced in 2019. The Gold are then used until the end of 2026.

We consider four MTBF values (0 %, $3.00 \cdot 10^{-3}\%$, 0.35 % and 2.38 %) and we compute their corresponding global energies ($G_0\%$, $G_{3.00 \cdot 10^{-3}\%}$, $G_{0.35\%}$, $G_{2.90\%}$). In table 3, we gather $G_0\%$ and the variation of it with the global energies computed with MTBF values: $\sigma_{3.00 \cdot 10^{-3}\%}$, $\sigma_{0.35\%}$, $\sigma_{2.90\%}$. The global energies are expressed in MWh, tCO_{2e} and K€ in order to evaluate the energetic, environmental and economic impact.

The quick answer to **question 2** is that yes, MTBF has a non-negligible impact on the global energy consumption considering the energetic, carbon and economic impacts.

The lower MTBF value ($3.00 \cdot 10^{-3}\%$) triggers very few variations on the global energy. The percentage variations are respectively $2.00 \cdot 10^{-4}\%$, $1.00 \cdot 10^{-3}\%$ and $8.00 \cdot 10^{-3}\%$ for PE, GWP, and the price for each of the three scenarios.

The intermediate MTBF value (0.35 %) implies less than one percent variations for 8 out of 9 combinations impact/scenario. Only the percentage deviation of the price in the fixed-workload replacement case shows an increase of 1.04 %. However it remains still low in regards to the initial values of the global cost $G_0\%$.

The higher values of the MTBF (2.90 %) have a non-negligible impact on the variation of the global energy. Even though it is lower or equal to 1.15 % for the three scenarios for primary energy and global warming potential, one cannot bypass the effect of the MTBF on the price. Indeed, considering a MTBF of 2.90 % leads to an increase of around 8 % for the three scenarios. In the worse case of the fixed-CPU scenario, the increase corresponds to a global cost of $G_{2.90\%} = \text{€}450.96 \text{ K}$. It is $\text{€}35.82 \text{ K}$ higher than the global cost for the fixed-core scenario with no MTBF ($G_0\% = \text{€}415.14 \text{ K}$).

From now on, we consider an MTBF value of 2.90 % for Silver and Gold CPUs.

4.1.3 Question 3 : how and when should we replace the Silver by the Gold CPUs to minimize the impacts ?

The short answer is that, in the case of the fixed-workload scenario, replacing all Silver by Gold in 2019 has the lowest economical, energetic, and environmental impacts. Otherwise said, if we plan to eventually upgrade all our processors, then it is best to do it as soon as the new generation of CPUs becomes available. However, it is important to notice that this answer holds only in the two-CPU case, and that CPUs with different characteristics, or run/produced in different countries,

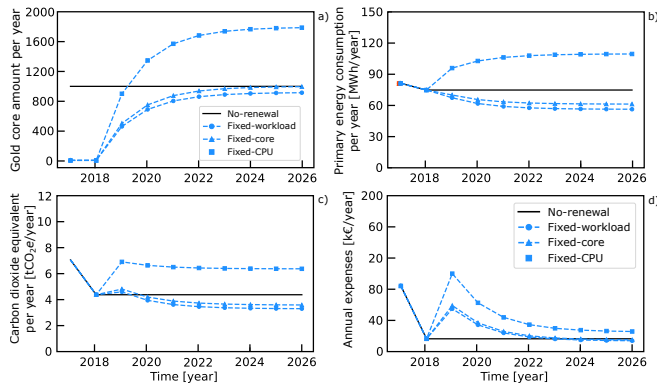


Figure 1: Multi-criteria evolution – core (a), PE [MWh] (b), GWP [tCO₂e] (c), and spendings [K€] – for a 50% replacement every year, considering four scenarios: no-renewal at all (plain black line), fixed-workload replacement (blue dashed line with squares), fixed-core replacement (blue dashed line with triangles), and fixed-CPU replacement (blue dashed line with circles).

could have given different answers. We analyse the results on three parameters of our model: the CPU replacement fraction ρ , the frequency of the replacements (in the case of a partial replacement policy) and the starting date of the replacement scheme.

The fraction approach.

We assume that for the three scenarios, the first replacement date occurs in 2019, at the Gold release date.

In figure 2, the dotted orange lines represent the evolution of the core amount (a), primary energy (b), global warming potential (c), and price (d) for a total replacement in 2019. In figure 1, the dashed blue lines represent the same quantities (a), (b), (c), and (d), for an annual 50% replacement rate of the remaining Silver CPUs. In both figures, we have computed the remaining Silver CPUs. The lines with filled squares correspond to the fixed-CPU scenario, the lines with filled triangles correspond to the fixed-core scenario, the lines with filled circles correspond to the fixed-workload scenario, finally, the black plain lines represent a no-renewal scenario, where the Data Center remains composed of 100 Silver CPUs.

All scenarios of the 100% fraction of fig. 2 quickly reach constant annual primary energy consumption, annual carbon dioxide equivalent emissions and annual expenses. Moreover, the fixed-CPU scenario increases the consumption (primary energy – 109.73 MWh/year, carbon dioxide equivalent – 6.33 tCO₂e/year, and price – €23.81 K/year), compared to the no-renewal scenario. However, the two other scenarios yield lower annual consumption than for the no-renewal scenario, for each of the three impacts. The only excessive consumption – in regard to the no-renewal scenario – is observed in 2019 and is due to the manufacturing costs.

The annual 50% replacement means that every year after 2019 included, half of the remaining Silver CPUs are replaced with Gold CPUs. We can observe in fig. 1a), in the case of the fixed-core renewal: in 2018 the Data Center is composed of 1000 Silver cores, in 2019 it is

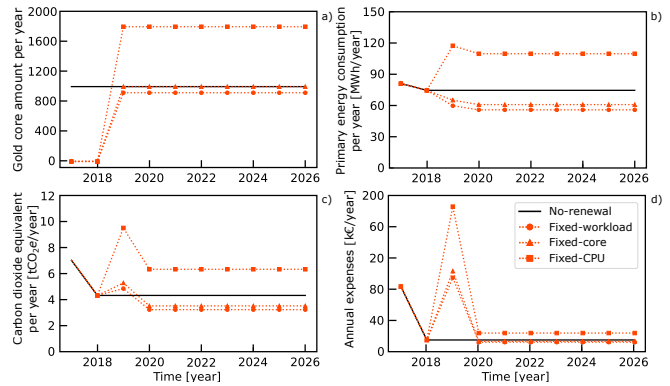


Figure 2: Multi-criteria evolution – core (a), PE [MWh] (b), GWP [tCO₂e] (c), and spendings [K€] – for a total replacement in 2019, considering four scenarios: no-renewal at all (plain black line), fixed-workload replacement (orange dotted line with squares), fixed-core replacement (orange dotted line with triangles), and fixed-CPU replacement (orange dotted line with circles).

composed of 500 Silver cores and 500 Gold cores, in 2020 it is composed of 250 Silver cores and 750 Gold cores and so on. Such fractionated replacement allows to soften the replacement in order to not have less important impacts on 2019 in regard to a total replacement. In the end, the limit is the same as the total replacement. The primary energy consumption in 2019 is equal to 95.93 MWh (for an annual 50% replacement) while it is equal to 117.40 MWh (for a total replacement). Similarly, the carbon dioxide equivalent emission is equal to 6.87 tCO₂e while it is equal to 9.49 tCO₂e for a total replacement in 2019. The annual expenses in 2019 is almost divided by two for the fractionated replacement (€99.20 K in comparison to €185.45 K for the total replacement). In addition, we see that the limit of the evolution of the core amount per year, primary energy consumption, carbon dioxide equivalent and annual expenses are reached very quickly. The core amount of Gold CPUs in 2023 is 1743.75 (fixed-CPU scenario), 968.75 (fixed-core scenario) and 889.15 (fixed-workload scenario), which corresponds to 97% of the total amount of Gold CPUs. The replacements can then be considered as being over in 2023.

Figures 3 and 4 present the cumulative primary energy (a), the cumulative carbon dioxide equivalent (b), and the cumulative expenses (c) over the years for the three scenarios in the case of a total replacement and of an annual 50% replacement. We can see that in the case of the fixed-CPU replacement – for both total and fractional replacement – the three cumulative impacts are always higher than the ones in the case of no replacement. It rules out the possibility that the fixed-CPU replacement may allow to reduce at least one of the three impacts.

However, for a fixed-core and fixed-workload replacement, the cumulative primary energy are lower from 2019 than the one in the no-replacement case. Thus, the total and fractional replacement allow to reduce the primary energy consumption from the very first year of renewal dates. In addition to that, in the case of the fixed-core replacement, the carbon dioxide emission are higher in

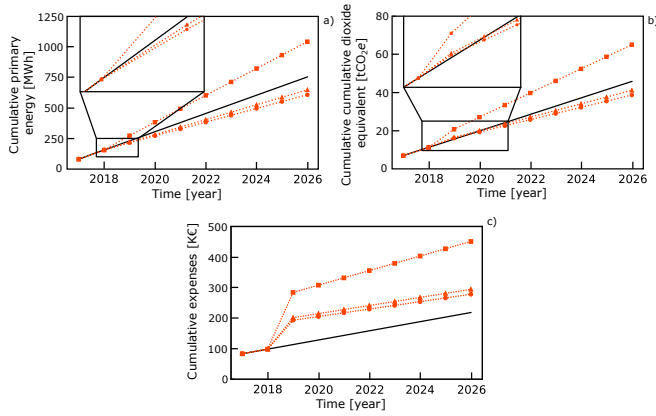


Figure 3: Cumulative multi-criteria evolution – PE [MWh] (a), GWP [tCO₂e] (b), and spendings [K€] (c) – for a total replacement in 2019, considering four scenarios: no-renewal at all (plain black line), fixed-workload replacement (orange dashed line with squares), fixed-core replacement (orange dashed line with triangles), and fixed-CPU replacement (orange dashed line with circles).

2019 and 2020 in regards to the no-replacement case. It is higher only for 2019 in the fixed-workload replacement case. Then, one needs to wait one (fixed-workload replacement) or two years (fixed-core replacement) to reduce the carbon dioxide emissions.

Unlike primary energy and carbon dioxide equivalent, the cumulative expenses, for both of total and fractional replacements (and for both scenarios), are higher than the case of no replacement. If one expects to make savings in these two cases, one has to wait until 2048 – that is to say, when the cumulative expenses for the replacements are lower than the cumulative expenses in the case of no replacement. Such behavior of the cumulative expenses is due to the fact that the manufacture price per core for Gold is 1.31 higher than the one of Silver and the run on price per core for Gold is 1.22 lower than the one of Silver. One has to wait longer to soften the replacement expenses.

The frequency approach.

Figure 5 shows the evolution of the primary energy in MWh (fig. (a), (b) and (c)), the carbon dioxide equivalent in tCO₂e (fig. (d), (e) and (f)), and the expenses in K€ (fig. (g), (h) and (i)), depending on the renewal frequency. For each impact, the three scenarios are considered. With a 10-years 2017–2026 lifespan and a 2019-market availability for the Gold, the frequency varies from 1 to 8 years. We consider five cases: 0% (in black), 25% (in green), 50% (in blue), 75% (in yellow), and 100% (in red) of the remaining Silver CPUs are replaced by Gold CPUs at each stage. The first and last strategies correspond to the no replacement and total replacement. The evolution of the three impacts depends on the frequency of the renewal stages, that start in 2019. For example, a frequency of 1 year⁻¹ means yearly replacements, while a frequency of 3 year⁻¹ yields replacements in 2019, 2022, and 2025.

We see in the nine graphics that the costs for a total replacement and for no replacement do not depend on the

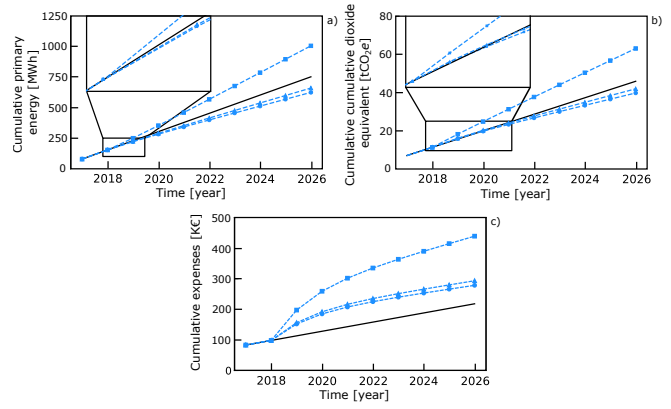


Figure 4: Cumulative multi-criteria evolution – PE [MWh] (a), GWP [tCO₂e] (b), and spendings [K€] (c) – for a 50% replacement every year, considering four scenarios: no-renewal at all (plain black line), fixed-workload replacement (blue dashed line with squares), fixed-core replacement (blue dashed line with triangles), and fixed-CPU replacement (blue dashed line with circles).

frequency. Indeed, the total replacement policy leaves no Silver left after 2019 and the no-replacement policy runs the initial Silver from 2017 to 2026.

Total fixed-workload and fixed-core replacements lead to 917.83 and 1 000 cores respectively. Both manufacture and use costs per core are lower for Gold than for Silver (see table 1), and better performance per core entails a smaller amount of Gold cores compared to Silver. Those factors interfere constructively, so the total replacement is a lower bound for the primary energy and carbon dioxide equivalent, while the no-replacement case is the upper bound. Between these bounds, higher frequencies and fractions of replacements entail lower global primary energy and carbon dioxide equivalent.

The situation is more complex for the economical impact, since the manufacturing cost of the Gold core is higher than for the Silver, but the use cost is lower. As it can be computed, in the case of a total and fixed-workload replacement, the replacement expenses are fully amortized only after 2048. This equilibrium point is out of reach of our 10-years 2017–2026 window. Consequently, the total replacement is an upper bound and the no replacement is a lower bound for global expenses. Because of the high manufacturing price, the lower the renewal frequency and the replacement fraction, the lower the global expenses are. We also observe a strong sensitivity to the frequency for fractional replacements. Indeed, due to the 10-year window effect, a 3 year⁻¹ frequency yields 3 replacements, while 4-7 year⁻¹ yield only two replacements. One has also to notice that for a frequency of 1, 2 and 3 year⁻¹ for a fraction of 75% and for a frequency of 1 year⁻¹ for a fraction of 50%, global expenses slightly exceed the total-replacement upper bound. We suppose that such a behavior is caused by the fact that due to the fractional amount of replaced CPUs, the amount of Silver only tends to zero, but never totally reaches it.

Finally, figures 5 c), f) and i) show that the fixed-CPU scenario, that preserve the amount of CPUs, is not

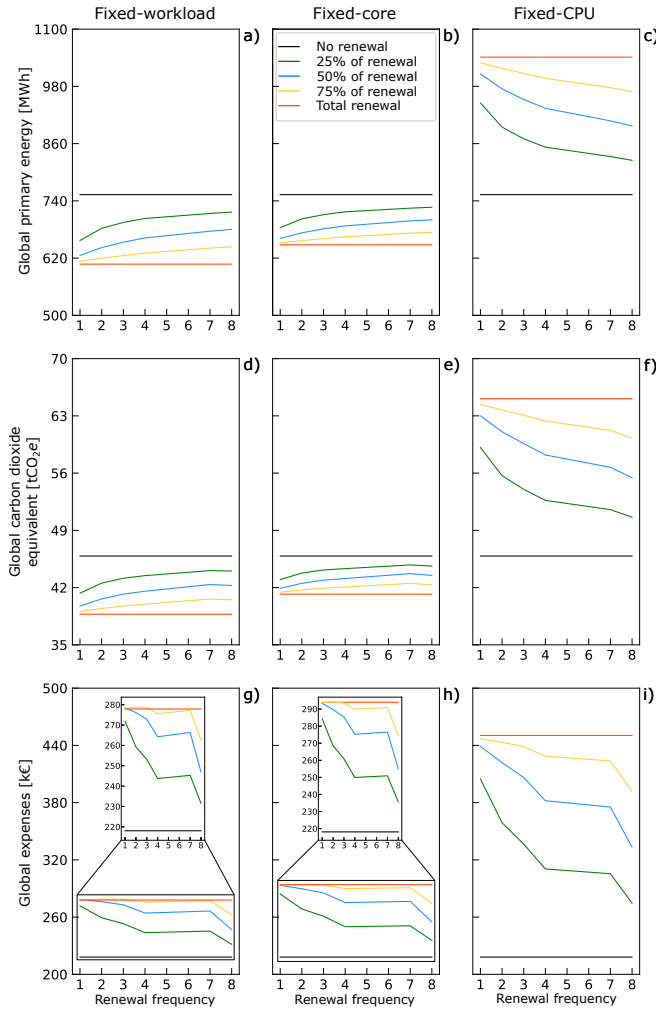


Figure 5: Evolution of (a, b, c) global PE [MWh], (d, e, f) global GWP [tCO_2e], and (g, h, i) global price [K€] depending on the renewal frequencies for the three scenarios (fixed-workload, fixed-core and fixed-CPU). The amount of Silver CPUs that are replaced at each renewal date are: 0% (black plain line), 25% (green plain line), 50% (blue plain line), 75% (yellow plain line), and 100% (red plain line).

efficient. It wastes extra computing power, primary energy, carbon dioxide equivalent, and money. Indeed, no matter the replacement fraction or frequency, a fixed-CPU strategy has always higher costs than no replacement.

The starting date approach.

The evolution of the global primary energy, carbon dioxide equivalent, and expenses according to the first replacement date for a total replacement are shown in figure 6 (a), (b) and (c), respectively. We see in fig 6 (a), (b) and (c) that the fixed-CPU scenario leads to an over consumption of primary energy, carbon dioxide equivalent and expenses in regards to the no-renewal scenario.

In fig 6 (a), the fixed-workload and fixed-core scenarios allow to reduce the global primary energy. The lower global primary energy is reached when the first replacement date is as soon as possible. In addition to that,

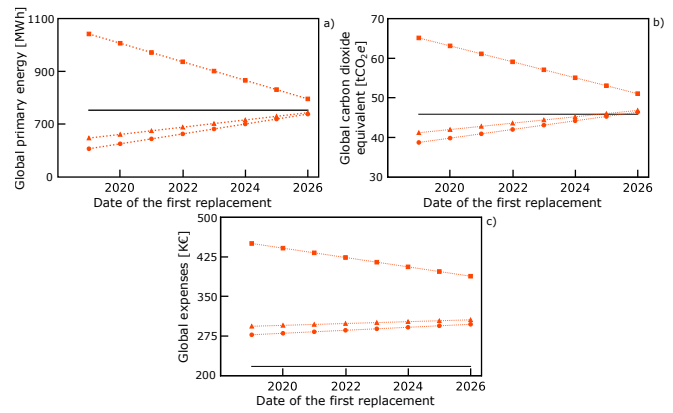


Figure 6: Global multi-criteria evolution – PE [MWh] (a), GWP [tCO_2e] (b), and spendings [K€] (c) – for a total replacement in 2019, considering four scenarios: no-renewal at all (plain black line), fixed-workload replacement (blue dashed line with squares), fixed-core replacement (blue dashed line with triangles), and fixed-CPU replacement (blue dashed line with circles).

the global primary energy of fixed-workload, fixed-core and fixed-CPU scenarios converge to the global primary energy of the no-renewal scenario when the first replacement date increases.

One can see in figure 6 (b) that there is a reduction of the carbon dioxide emissions only if the total replacement is performed in 2025 for the fixed-workload scenario and in 2024 at the latest for the fixed-core scenario. Indeed, we have seen previously that one have to wait one or two years – depending on the scenario – to soften the carbon dioxide emissions due to the CPU manufacturing. Thus, an fixed-workload total replacement in 2026 or a fixed-core replacement in 2025 are over consuming.

Apart from not performing any kind of replacement, which is the more economical scenario, it is better to replace (fixed-workload and fixed-core scenarios) on the very first year when Gold are coming to the market.

We can reach the conclusion to **question 3** that the better replacement is a fixed-workload total replacement in 2019. It allows to save 145.70 MWh, 7.12 tCO_2e and costs 59 801 € in comparison to not performing any replacement at all for a Data Center composed of 100 Silver. However, this specific conclusion is only valid in the case of replacing Silver with Gold, which specifications are shown in table 1. Although the fixed-workload scenario is the most relevant one, we have to keep in mind that it is based on the conservation of the performance per core. Such criterion could be in high-performance computing systems best interest, since calculus depends a lot on the CPU's performances. In the case of Data Centers that host servers, cloud, and virtual machines, it may better to not perform a fixed-workload total replacement but one should perform a fixed-core total replacement, which is based on the conservation of the core amount. Such replacement would save 105.28 MWh and 4.66 tCO_2e and would cost €75 876.

4.2 Extension to a three-CPU case

The values presented in the previous subsection have been computationally obtained by using a program that is based on the two CPUs model (see section 3). However, the program is not able for the moment to compare three CPUs and suggest the better replacement possibilities. Nonetheless, we have used spreadsheet to compare three CPUs and suggest scenarios. The three CPUs that are compared in this subsection are Xeon®Silver 4114, Xeon®Gold 5220 and Xeon®Platinum 8380 (see table 1). They are labelled Silver, Gold, and Platinum, respectively. We investigate here the possibilities to replace Silver by Silver and/or Platinum. We compare the three impacts, as previously.

We assume that the Data Center is filled with 100 Silver in 2017 and the Data Center is running on from 2017 to the end of 2026. Since Gold and Platinum do not appear on the market at the same time, one can enumerate thirteen cases to evaluate the replacement of Silver:

- Case 1: None of the processors P_1 are replaced;
- Case 2: 50% of Silver are replaced in 2019 by Gold;
- Case 3: All Silver are replaced in 2019 by Gold;
- Case 4: 50% of Silver are replaced in 2019 by Gold and the remaining Silver are replaced in 2021 by Platinum;
- Case 5: 50% of Silver are replaced in 2019 by Gold and the remaining Silver are replaced in 2021 by Gold;
- Case 6: 50% of Silver are replaced in 2019 by Gold and then the Gold are replaced in 2021 by Platinum;
- Case 7: 50% of Silver are replaced in 2019 by Gold and in 2021, all of the Silver and Gold are replaced by Platinum;
- Case 8: All Silver are replaced in 2019 by Gold and in 2021, 50% of the Gold are replaced by Platinum;
- Case 9: All Silver are replaced by Gold in 2019 then the Gold are replaced by Platinum in 2021;
- Case 10: 50% of Silver are replaced by Gold in 2021;
- Case 11: All Silver are replaced in 2021 by Gold;
- Case 12: 50% of Silver are replaced in 2021 by Platinum;
- Case 13: All Silver are replaced by Platinum in 2021.

Even though we have shown in the previous subsection that the fixed-workload scenario is the more suitable – the fixed-core scenario as well, to a certain extent – we evaluate the replacement of Silver by two different CPUs according to the three scenarios: fixed-workload, fixed-core, and fixed-CPU. The reader would have noticed that fractional replacement is only considered with a fraction of 50% of Silver and no other fractional values are considered. It would have been difficult to visualize a more global impact of fractional replacement because it would have led to a non-negligible amount of various cases. Since we compute the cases by hand with a spreadsheet, we cannot evaluate every fractional replacement. However, the 50% replacement is giving us a good viewing of the fractional replacement in regards to a total replacement. The table 4 shows the global primary energy

[MWh], the global carbon dioxide [tCO_2e], and the global expenses [K€] of the thirteen cases for a time window of 10 years (from 2017 to 2026). For the fixed-workload scenario, the case 9 – all of the Silver are replaced by Gold in 2019, then the Gold are replaced by Platinum in 2021 – is the one that allows the lowest global primary energy consumption (573.60 MWh) and the lowest global carbon dioxide equivalent (37.66 tCO_2e). However, it is the most expensive one (466.08 K€).

It is quite different for the fixed-core scenario. Case 9 remains the one that allows the lowest global primary energy consumption for this scenario (640.43). Unlike the fixed-workload scenario, case 3 – all of the Silver are replaced in 2019 by Gold – is the case that allows the lowest global carbon dioxide equivalent (41.20 tCO_2e). It is due to the fact that replacing the Gold by Platinum does not allow to soften the first replacement of Silver by Gold (which takes more or less two years, see fig 3). It is then less environmentally impacting to keep Gold and not to replace them with Platinum.

Once again, we can see that the fixed-CPU scenario is not feasible since case 1 – no P_1 processors are replaced – is the one that gives the lowest global primary energy (753.08 MWh), global carbon dioxide equivalent (45.87 tCO_2e), and global expenses (218.33 K€).

It is important to notice that both of the fixed-workload and fixed-core scenarios for cases 9 and 3 show significant costs. It is better not to replace Silver if one wants to save money. However, in the replacement cases for the fixed-workload scenario, case 2 – 50% of Silver are replaced in 2019 by Gold – is the one that is the least expensive (247.22 K€), but it leads to extreme global primary energy (680.13 MWh) and global carbon dioxide equivalent (43.80 tCO_2e) values. For the fixed-core scenario, case 2 is the least expensive as well (255.67 K€).

Case 3 for the fixed-core scenario seems to be the best replacement to suggest. Indeed, the global primary energy is not the lowest but among them (647.79 MWh), we have seen previously that the global carbon dioxide is the lowest for this case and the global expenses are 294.24 K€, which is the same order of magnitude as for case 1.

5 CONCLUSION AND FUTURE WORK

In this work, we propose a model of hardware renewal between two CPUs. Our model takes into account the economical, environmental, and energetic impacts of the usage and manufacturing of the CPUs. We have also considered a MTBF of 2.90% for each type of CPUs.

Studying a real-world example, our CPU-renewal model shows that it is more suitable (from the environmental and energetic point of views) to replace all of the Silver with Gold when they come in the market. However, one has to wait a few years (at least 3) to soften the environmental cost of this replacement.

We also study a three-CPU renewal approach in the last part of this article. Our study shows that the fixed-CPU scenario remains the worst scenario to consider when one wants to replace CPUs. Moreover, it appears that replacing all of the Silver by Gold in 2019 and not taking into account the Platinum is the most interesting

Table 4: Global PE [MWh], global GWP [tCO₂e] and global spendings [K€] of the thirteen replacement cases for the three scenarios for a length of time of ten years (from 2017 to 2026).

	Global primary energy [MWh]			Global carbon dioxide eq. [tCO ₂ e]			Global expenses [K€]		
	Workload	Core	CPU	Workload	Core	CPU	Workload	Core	CPU
Case 1		753.08			45.87		218.33		
Case 2	680.13	700.34	897.12	42.27	43.49	55.44	247.22	255.67	333.62
Case 3	607.37	647.79	1 041.35	38.75	41.20	65.10	278.15	294.24	450.96
Case 4	607.13	655.57	1 389.96	38.45	41.38	85.86	332.72	365.17	908.22
Case 5	625.97	661.39	1 006.18	39.80	41.95	63.05	279.95	294.96	441.06
Case 6	663.24	696.66	1 284.74	41.73	43.79	79.83	341.19	370.37	881.58
Case 7	590.25	651.90	1 777.58	37.91	41.68	110.25	426.69	480.27	1 456.18
Case 8	590.48	644.11	792.41	38.21	41.50	50.41	372.12	409.34	548.63
Case 9	573.60	640.43	1 816.60	37.66	41.80	113.88	466.08	524.44	1 546.87
Case 10	698.83	714.03	862.04	43.36	44.30	53.44	250.04	257.00	324.74
Case 11	644.77	675.18	971.20	40.93	42.81	61.08	283.79	297.71	433.19
Case 12	679.99	708.22	1 245.83	42.01	43.72	76.24	302.82	327.22	791.92
Case 13	607.09	663.56	1 738.77	38.23	41.65	106.70	389.34	438.15	1 367.53

choice to make for a fixed-core scenario.

The results shown in this article depends on the specifications of the CPUs we have considered. Thus, others CPUs might have led to different conclusion concerning how and when they should be replaced. However, approach provides the possibility to study various replacement schemes in order to reduce carbon emissions, thus helping with the implementation of carbon-neutral resource-management strategies.

In the scope of this paper, we focus on microprocessors mainly, but the model applies to the replacement of any hardware. We suggest the possibility to replace server hardware in Data Centers separately and not only the replacement of entire servers. Doing so, we would allow to be more selective on the hardware to replace, the amount, the replacement frequency, and the hardware specifications. Being more precise on the replacement scheme would allow Data Centers to tend to a carbon-neutral replacement scheme, by just replacing the necessary devices. Following this idea, we are planing to apply our model to more complex infrastructures implying several hardware devices with different optimal replacement periods. The ultimate goal would be to reach the optimal point for replacement for all devices in a Data Center or any complex infrastructure involving several kinds of hardware.

Finally, another application of our model would be not only to determine the right timing for replacing hardware, but also, the right timing for manufacturing and distributing new hardware. For example, giving a set of data centers equipped with various models of servers and CPUs, would the mass manufacturing of a new generation of hardware would globally have a positive or negative impact? Scaling our model to allow a quantifiable answer to this question is part of the perspectives that could be studied from this work as a starting point.

REFERENCES

- [1] greenIT. <http://www.greenit.fr/>. Accessed: 2022-04-11.
- [2] Nicola Jones. The Information Factories. *Nature*, 561:163–166, 2019.
- [3] Qixiang Cheng, Meisam Bahadori, Madeleine Glick, Sébastien Rumley, and Keren Bergman. Recent advances in optical technologies for data centers: a review. *Optica*, 5(11):1354–1370, 2018.
- [4] Kim Hazelwood, Sarah Bird, David Brooks, Soumith Chintala, Utku Diril, Dmytro Dzhulgakov, Mohamed Fawzy, Bill Jia, Yangqing Jia, Aditya Kalro, James Law, Kevin Lee, Jason Lu, Pieter Noordhuis, Misha Smelyanskiy, Liang Xiong, and Xiaodong Wang. Applied machine learning at facebook: A datacenter infrastructure perspective. In *2018 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pages 620–629, 2018.
- [5] Weng Chon Ao, Po-Han Huang, and Konstantinos Psounis. Joint workload distribution and capacity augmentation in hybrid datacenter networks. *IEEE/ACM Transactions on Networking*, 29(1):120–133, 2021.
- [6] Ayaz Ali Khan and Muhammad Zakarya. Energy, performance and cost efficient cloud datacentres: A survey. *Computer Science Review*, 40:100390, 2021.
- [7] Stephen Lindsay, Kate Cavanagh, and Terijo Lovasz. Data centres and the Australian energy sector. *CSIRO National Energy Analytics Research (NEAR) Program*, 2021.
- [8] Nuoa Lei and Eric R Masanet. Global data center energy demand and strategies to conserve energy. *Data Center Handbook: Plan, Design, Build, and Operations of a Smart Data Center*, pages 15–26, 2021.
- [9] Vehbi Comert, Mustafa Altun, Mustafa Nadar, and Ertunc Erturk. Warranty forecasting of electronic boards using short-term field data. In *2015 Annual Reliability and Maintainability Symposium (RAMS)*, pages 1–6, 2015.
- [10] Anisur Rahman and GN Chattopadhyay. Lifetime warranty policies: complexities in modelling and potential for industry application. In *Proceedings of the Fifth APIEMS Conference 2004*, pages 28.1.1–28.1.8, 2004.
- [11] Xuejiao Han, Maoliang Hu, Zhanqing Yu, Kangsheng Cui, Lu Qu, and Xuejian Zhao. Data center new energy system under the “peak carbon dioxide emissions and carbon neutrality” background and comprehensive evaluation analysis thereof. In *2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2)*, pages 1649–1656, 2021.
- [12] Zhiwei Cao, Xin Zhou, Han Hu, Zhi Wang, and Yonggang Wen. Toward a systematic survey for carbon neutral data centers. *IEEE Communications Surveys Tutorials*, 24(2):895–936, 2022.
- [13] Marcos De Melo da Silva, Abdoulaye Gamatié, Gilles Sassatelli, Michael Poss, and Michel Robert. Optimization of data and energy migrations in mini data centers for carbon-neutral computing. *IEEE Transactions on Sustainable Computing*, pages 1–15, 2022.
- [14] John Judge, Jack Pouchet, Anand Ekbote, and Sachin Dixit. Reducing data center energy consumption. *Ashrae Journal*, 50(11):14, 2008.
- [15] Emerson Network Power. Energy logic: Reducing data center energy consumption by creating savings that cascade across systems. *White paper, Emerson Electric Co*, 9, 2009.
- [16] Arman Shehabi, Sarah J Smith, Eric Masanet, and Jonathan Koomey. Data center growth in the United States: decoupling the demand for services from electricity use. *Environmental Research Letters*, 13(12):124030, 2018.

- [17] Eric Masanet, Arman Shehabi, Nuo Lei, Sarah Smith, and Jonathan Koomey. Recalibrating global data center energy-use estimates. *Science*, 367(6481):984–986, 2020.
- [18] Maria Avgerinou, Paolo Bertoldi, and Luca Castellazzi. Trends in data centre energy consumption under the european code of conduct for data centre energy efficiency. *Energies*, 10(10):1470, 2017.
- [19] Scott Huck. Measuring processor power. *Intel Corporation*, 2011.
- [20] Alon Naveh, Efraim Rotem, Avi Mendelson, Simcha Gochman, Rajshree Chabukwar, Karthik Krishnan, and Arun Kumar. Power and thermal management in the intel core duo processor. *Intel Technology Journal*, 10(2), 2006.
- [21] Brett Battles, Cathy Belleville, Susan Grabau, and Judith Maurier. Reducing data center power consumption through efficient storage. *Network Appliance, Inc*, 2007.
- [22] Carolina Koronen, Max Åhman, and Lars J Nilsson. Data centres in future european energy systems—energy efficiency, integration and policy. *Energy Efficiency*, 13(1):129–144, 2020.
- [23] Jiacheng Ni and Xuelian Bai. A review of air conditioning energy performance in data centers. *Renewable and sustainable energy reviews*, 67:625–640, 2017.
- [24] Alex K Jones, Yiran Chen, William O Collinge, Haifeng Xu, Laura A Schaefer, Amy E Landis, and Melissa M Bilec. Considering fabrication in sustainable computing. In *2013 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 206–210, 2013.
- [25] Erik Brunvand, Donald Kline, and Alex K Jones. Dark silicon considered harmful: A case for truly green computing. In *2018 Ninth International Green and Sustainable Computing Conference (IGSC)*, pages 1–8, 2018.
- [26] Alex K Jones, Liang Liao, William O Collinge, Haifeng Xu, Laura A Schaefer, Amy E Landis, and Melissa M Bilec. Green computing: A life cycle perspective. In *2013 International Green Computing Conference Proceedings*, pages 1–6, 2013.
- [27] Rabih Bashroush. A comprehensive reasoning framework for hardware refresh in data centers. *IEEE Transactions on Sustainable Computing*, 3(4):209–220, 2018.
- [28] Rabih Bashroush and Andy Lawrence. Beyond pue: tackling it’s wasted terawatts. *Uptime Institute*, 2020.
- [29] Rabih Bashroush, Nour Rteil, Rich Kenny, and Astrid Wynne. Optimizing server refresh cycles: The case for circular economy with an aging moore’s law. *IEEE Transactions on Sustainable Computing*, 2020.
- [30] Joseph Doyle and Rabih Bashroush. Case studies for achieving a return on investment with a hardware refresh in organizations with small data centers. *IEEE Transactions on Sustainable Computing*, 6(4):599–611, 2020.
- [31] Fraunhofer IZM Bio by Deloitte. Tpreparatory study for implementing measures of the eco-design directive 2009/125/ec dg entr lot 9 - enterprise servers and data equipment task 5: Environment & economics. *European Commission, DG Internal Market, Industry, Entrepreneurship and SMEs*, 2015.
- [32] Chaoqiang Jin, Xuelian Bai, Chao Yang, Wangxin Mao, and Xin Xu. A review of power consumption models of servers in data centers. *applied energy*, 265:114806, 2020.
- [33] Aurélien Bourdon, Adel Noureddine, Romain Rouvoy, and Lionel Seinturier. Powerapi: A software library to monitor the energy consumed at the process-level. *ERCIM News*, 2013(92), 2013.
- [34] Adel Noureddine. Powerjoular and joularjx: Multi-platform software power monitoring tools. In *18th International Conference on Intelligent Environments*, 2022.
- [35] Constanța Zoie Rădulescu and Delia Mihaela Rădulescu. A performance and power consumption analysis based on processor power models. In *2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pages 1–4, 2020.
- [36] Maxime Colmant, Romain Rouvoy, Mascha Kurpicz, Anita Sobe, Pascal Felber, and Lionel Seinturier. The next 700 cpu power models. *Journal of Systems and Software*, 144:382–396, 2018.
- [37] Boavizta. <https://boavizta.org/blog/empreinte-de-la-fabrication-d-un-serveur>. Accessed: 2022-05-03.
- [38] Passmark® software pty ltd. <https://www.passmark.com/>. Accessed: 2022-05-03.
- [39] Intel xeon silver 4114. <https://www.intel.fr/content/www/fr/fr/products/sku/123550/intel-xeon-silver-4114-proces>
- sor-13-75m-cache-2-20-ghz/specifications.html. Accessed: 2022-06-24.
- [40] Intel xeon gold 5220. <https://ark.intel.com/content/www/fr/fr/ark/products/193388/intel-xeon-gold-5220-processor-24-75m-cache-2-20-ghz.html>. Accessed: 2022-06-24.
- [41] Intel xeon platinum 8380. <https://ark.intel.com/content/www/fr/fr/ark/products/212287/intel-xeon-platinum-8380-processor-60m-cache-2-30-ghz.html>. Accessed: 2022-06-24.
- [42] Cpu specifications die size. https://github.com/Boavizta/boaviztapi/blob/main/boaviztapi/data/components/cpu_manufature.csv. Accessed: 2022-07-08.
- [43] Intel xeon silver 4114 benchmark. <https://www.cpubenchmark.net/cpu.php?cpu=Intel+Xeon+Silver+4114+%40+2.20GHz&id=3095>. Accessed: 2022-06-24.
- [44] Intel xeon gold 5220 benchmark. <https://www.cpubenchmark.net/cpu.php?cpu=Intel+Xeon+Gold+5220+%40+2.20GHz&id=3534>. Accessed: 2022-06-24.
- [45] Intel xeon platinum 8380 benchmark. <https://www.cpubenchmark.net/cpu.php?cpu=Intel+Xeon+Platinum+8380+%40+2.30GHz&id=4483>. Accessed: 2022-06-24.
- [46] Carbon intensity of electricity. <https://ourworldindata.org/grapher/carbon-intensity-electricity?tab=table>. Accessed: 2022-07-13.
- [47] Price of 1kwh in france. <https://www.totalenergies.fr/particuliers/parlons-energie/dossiers-energie/facture-d-energie/les-elements-a-prendre-en-compte-pour-calculer-les-kwh-en-euros#:~:text=Pour%20bien%20calculer%20les%20kWh,Tarif%20Bleu%20pratique%C3%A9%20par%20EDF>. Accessed: 2022-07-13.
- [48] Location of intel factories. <https://www.intel.fr/content/www/fr/fr/architecture-and-technology/global-manufacturing.html>. Accessed: 2022-07-13.
- [49] Carbon intensity of electricity in france. <https://bilans-ges.ade.me/fr/fr/basecarbone/donnees-consulter/liste-element/categorie/64>. Accessed: 2022-07-13.
- [50] Bianca Schroeder and Garth A. Gibson. A large-scale study of failures in high-performance computing systems. *IEEE Transactions on Dependable and Secure Computing*, 7(4), 2010.
- [51] Sanghoan Chang and Gwan Choi. Timing failure analysis of commercial cpus under operating stress. In *2006 21st IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, pages 245–253, 2006.
- [52] Dong Seong Kim, Fumio Machida, and Kishor S. Trivedi. Availability modeling and analysis of a virtualized system. In *2009 15th IEEE Pacific Rim International Symposium on Dependable Computing*, pages 365–371, 2009.



Gabriel Breuil Gabriel Breuil is a green IT researcher at Constellation. His main domain of research is focused on simulation and modelling of the energy consumption in computer science.



Olivier Hermant Olivier Hermant is a Full Professor at Mines Paris, PSL University and a member of the scientific board of GDR GPL, the French research network for Programming and Software Engineering. His main domain of research is focused on Formal Methods and ranges from Theoretical Computer Science to proof assistants and automated theorem proving.



Renaud Pawlak Currently R&D director at ID Capture and founder of CINCHÉO, an R&D advising company specialized in software sustainability and low environmental impact modernization. Former Director of the Mantu group R&D lab specialized in AI, Machine Learning and Data Science. PhD and former Computer Science researcher at INRIA, RPI and ISEP. Entrepreneur (Cinchéo, IDCapture, AOPsys). Author and co-author of more than 40 publications and two books on

Software Engineering.