

# Leveraging the Asymmetry between CPU Design Lifetime and Depreciation Lifetime to Reduce the Datacenter Total Cost of Ownership

## Abstract

To mitigate the degradation of processor reliability due to transistor aging and extend the lifetime of the processor, modern server processors typically increase the timing and/or voltage margins. This mitigation increases the power consumption in proportion to the processor lifetime (e.g., 7–10 years). However, despite modern processors having typical lifetimes of 7–10 years, the IT equipment-renewal cycle in many datacenters is 3–5 years.

In this work, we evaluate how modern processors with lifetimes longer than the renewal cycle of datacenters affects the datacenter total cost of ownership (TCO). Toward this end, we (1) estimate a modern processor’s power consumption and performance with timing and voltage margins corresponding to lifetimes of 3 and 10 years, and (2) input the power and performance values into a TCO calculator.

Our analysis shows that the TCO of a 30 000-server datacenter decreases by 4.3% if we reduce the server processor lifetime from 10 years to 3 years. This technique would reduce TCO even more if applied to more IT equipment, such as datacenter switches.

## 1. Introduction

In the past decades, the semiconductor industry has continuously propelled VLSI technologies toward extraordinary frontiers. This progression has been accompanied by several significant trends; for example, new process nodes have miniaturized transistors to nanometric dimensions. Novel three-dimensional integration techniques now stack multiple layers of integrated circuits (ICs), allowing for increased functionality and improved performance with a smaller footprint. New devices and materials have pushed performance and offer increased energy efficiency. Unfortunately, such developments have made semiconductors more vulnerable to reliability issues, specifically those triggered by transistor aging. The gradual decline in a transistor’s performance over time is known as “transistor aging” and is caused by hot carrier injection and bias temperature instability (BTI) [39, 10, 8, 35, 38]. This study focuses on BTI, which is the primary aging mechanism in modern ICs [13, 35, 24, 9].

Transistor aging increases the threshold voltage, which increases the switching delay and degrades transistor performance. The degree of transistor aging depends strongly on the mission profile, which represents the specific operational conditions and workload requirements that datacenter ICs are designed to handle. It includes factors such as the type and

intensity of computational tasks, the duration of each task, and the operating temperature [5, 9]. Common approaches for mitigating transistor aging include imposing extra timing margins on the clock cycle time, increasing the operating voltage, or a combination of both [14, 26, 44]. However, adding extra timing margins on the clock cycle would require faster logic elements, such as low-threshold voltage (low- $V_t$ ) cells, leading to an increase in leakage current and energy consumption. Similarly, increasing the operating voltage increases not only the leakage current but also the switching current, further increasing the energy and power consumption [26, 37]. Apart from the power impact, transistor aging also increases the complexity of IC design, which significantly affects the time to market [36]. Therefore, transistor aging significantly affects the total cost of ownership (TCO) of large-scale computing systems, especially those of datacenters.

With the growing reliance on digital technology and the need for supporting critical business functions and enabling the seamless flow of information, the scale of datacenters has increased dramatically over the years. Today, datacenters house large-scale computer servers and offer a wide range of services, including cloud storage, data analytics, and web hosting. The cost of building, operating, and maintaining a datacenter is substantial, with factors such as energy consumption, cooling, and reliability contributing to the overall cost.

Previous studies [18, 31, 11, 30, 42] suggest using a blend of cost models to evaluate datacenter expenses, encompassing infrastructure, servers, networking equipment, operating and maintenance costs, and staff expenses. In this paper, we argue that the cost of transistor aging, which has yet to be considered by hyperscaler operators, significantly affects the following datacenter cost models:

1. Today, datacenter IC vendors typically assume conservative lifetime requirements and operating conditions, which require excessive aging margins in clock cycle time or operating voltage to compensate for the speed degradation over time, resulting in a significant increase in power consumption. For instance, CPU vendors often assume lifetimes of 5–10 years for datacenters and hyperscalers, with an operating temperature range of 90–125 °C, but, in practice, datacenter servers operate at lower temperatures and depreciate within 3–5 years [28]. Similarly, network IC vendors assume lifetimes of 10 years and an operating temperature range of 105–125 °C. However, in datacenters, these components typically depreciate within 4 to 5 years

and operate in relatively relaxed thermal conditions [28]. As a result, transistor aging increases infrastructure costs related to datacenter power distribution and cooling equipment. Additionally, it increases electricity costs for servers, networking equipment, and cooling.

- Overdesign assumptions for transistor aging can increase the design complexity of datacenter ICs and thereby impact not only server and network costs but also the time to market.
- Reliability considerations may also affect maintenance and staff expenses. The reduced lifetime of datacenter ICs increases the need for repairs. However, ICs with extended lifetimes and higher power consumption increase the number of personnel needed to support power and cooling infrastructure.

This study proposes a model that accurately and adaptively determines the aging margins needed to comply with the required IC lifetime and mission profile in datacenters. Our approach reduces the TCO of datacenters by providing hyper-scale operators with different trade-offs between IC lifetime, design complexity, performance, energy consumption, and operating conditions. This is shown in Figure 1, which illustrates possible ranges for IC operating condition for voltage, frequency, and temperature in datacenters. Examples of two types of ICs are shown: Chip 1 and Chip 2. Chip 1 includes different stock keeping units (SKUs) shown as IC1–IC5, where IC1 is designed for the nominal target parameters of 10 year lifetime, maximum operating temperature  $T$ , and a nominal voltage  $V$ . Figure 1 presents a qualitative example of four different SKU flavors for IC1 that makes the following trade-offs to reduce lifetime to 3 years:

- IC2 increases the clock frequency, which improves performance compared with IC1 but at the expense of a reduced lifetime of 3 years.
- IC3 trades off reduced lifetime against operating at lower voltage, which reduces both dynamic and static power consumption.
- IC4 combines moderate voltage reduction and frequency increase, thereby achieving a moderate performance increase and reducing both static and dynamic power consumption.
- IC5 operates at a higher clock frequency than IC1, reducing its lifetime and lowering its operating temperature.

Chip 2 is designed and manufactured with the same nominal conditions as Chip 1 except for the lifetime, which is 3 years. In this case, Chip 2 is designed with a smaller clock cycle time aging margin and therefore uses slower logic elements, resulting in less leakage current than Chip 1.

Our experimental analysis uses the estimation and exploration TCO tool from Ref. [18] to provide qualitative and quantitative guidelines of datacenter design decisions with respect to TCO while taking into account aging considerations. Our experimental analysis examines several SKU flavors, such as those in Figure 1 for Chip 1. In addition, we offer a sensitivity analysis of transistor aging design margins, which includes

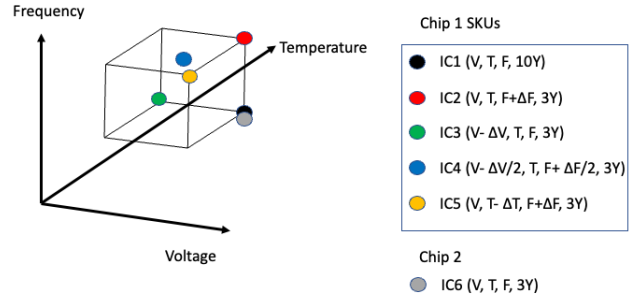


Figure 1: IC lifetime, frequency, operating voltage, and temperature trade-offs.

both clock cycle time and voltage margins, and their impact on power consumption at the design stage (e.g., Chip 2). The data acquired for this analysis is used by the TCO tool to analyze the suggested IC options with reduced lifetime. As a case study, we choose a RISC-V core to demonstrate the suggested analysis. Our TCO analysis considers a single server, a small-scale datacenter, and a large-scale datacenter with a variable number of servers. The experimental results show that reducing the lifetime of datacenter ICs from 10 to 3 years reduces the normalized TCO per queries per second (QPS) for a single server, a small-scale datacenter, and a large-scale datacenter by up to 4%, 5.2%, and 5.4%, respectively.

As part of our experimental analysis, we also examine how aging affects the datacenter CO<sub>2</sub> footprint. Our analysis indicates that reducing the IC lifetime from 10 to 3 years not only reduces energy consumption and TCO but also reduces the CO<sub>2</sub> footprint by up to 10%, which mitigates climate change, lowers greenhouse-gas emissions, and promotes a more sustainable and environmentally friendly operation. Additionally, reducing CO<sub>2</sub> emissions aligns with regulatory requirements, corporate social responsibility goals, and the increasing demand for energy-efficient and environmentally conscious practices in industry.

This paper makes the following primary contributions:

- We introduce a new model to adaptively determine the clock cycle time and voltage aging design margins for datacenter ICs based on their lifetime and mission-profile requirements.
- We demonstrate how various SKUs with different design targets allow trade-offs between IC lifetime, operating voltage, temperature, and frequency.
- We use a RISC-V core as a case study to perform a sensitivity analysis and demonstrate how clock cycle time and voltage aging margins affect overall chip power.
- Our experimental analysis using the the estimation and exploration TCO tool indicates that we can decrease by more than 5% the normalized server TCO per QPS and reduce the CO<sub>2</sub> footprint by 10%.

The remainder of this paper is organized as follows: Section 2 presents the background for the study. Section 3 introduces

the proposed adaptive aging model. Section 4 presents the experimental analysis. Section 5 presents prior works. Finally, Section 6 summarizes the study and our conclusions.

## 2. Background

This section overviews the background information regarding transistor aging and its relationship to IC power consumption. In addition, we discuss how common BTI-mitigation approaches affect power and thermal conditions.

### 2.1. Transistor aging

The deterioration of transistors over time, known as transistor aging, is mainly governed by BTI, which occurs when a static voltage (representing a constant logical state) is applied to the gate of a transistor with no current flow for an extended period, typically ranging from 10 s to several weeks [40]. BTI increases the transistor threshold voltage, which results in a longer switching delay. The reaction-diffusion model is commonly used to represent the  $\Delta V_{th}$  shift as a result of BTI [40]. The increase  $\Delta V_{th}$  due to BTI stress is

$$\Delta V_{th} = C_T e^{-E_a/k_B T} t^{1/n}, \quad (1)$$

where  $C_T$  is a technology-dependant constant,  $n$  is the time exponent,  $E_a$  is the activation energy,  $T$  is the operating temperature,  $k_B$  is the Boltzmann constant, and  $t$  is the overall time. Equation (1) indicates that a significant fraction of the  $\Delta V_{th}$  shift occurs early in the IC lifetime. For example, approximately 70% of the  $\Delta V_{th}$  degradation occurs within the first year of a 10-year-lifetime IC. Additionally, Equation (1) demonstrates a strong relationship between the operating temperature and the shift in threshold voltage. An operating temperature of 105°C will result in a  $\Delta V_{th}$  that is approximately 2.5 times larger compared with 90°C.

The relation between the propagation delay  $t_{pd}$  of logical elements and the transistor threshold voltage  $V_{th}$  is provided through the alpha power law [34]:

$$t_{pd} \propto \frac{V_{DD}}{(V_{DD} - V_{th})^\alpha}, \quad (2)$$

where  $V_{DD}$  is the operating voltage and  $\alpha \approx 1.3$  is the velocity saturation index. The combination of Equations (1) and (2) expresses the propagation delay that occurs when considering aging degradation.

### 2.2. Static and dynamic power

The power consumption of CMOS circuits consists of two main categories: *dynamic power* and *static power* [32, 17]. Dynamic power is further divided into switching power and short-circuit power. This section explains the parameters that affect power consumption in these categories.

#### 2.2.1. Dynamic power

Dynamic power is consumed by circuit activity and depends mainly on the circuit activity factor, circuit capacitance, clock

frequency, and supply voltage. The two main sources of dynamic power consumption are *short-circuit power* and *switching power*.

**Short-circuit power** is the power dissipated during the brief transitional period when both the n and p transistors of a CMOS gate are “on.” Short-circuit current typically represents about 10%–15% of the total power consumption.

**Switching power** is the power dissipated due to switching the transistor from 0 to 1 and vice versa.

Dynamic power can be estimated as follows

$$P_{dynamic} = AF \times C \times V^2 \times F \quad (3)$$

where  $AF$  is the activity factor,  $C$  is the load capacitance,  $V$  is the supply voltage, and  $F$  is the operating frequency.

#### 2.2.2. Static power

Static power is consumed due to the transistor leakage currents while they are “off.” The leakage power is the leakage current ( $I_{leakage}$ ) times the supply voltage  $V$  (i.e.,  $P_{leakage} = I_{leakage}V$ ). The leakage current ( $I_{leakage}$ ) has five main current sources: subthreshold leakage, junction reverse bias current, gate-induced drain leakage, punch-through current, and gate tunneling current [32, 29].

### 2.3. Effect of BTI-mitigation approaches on power and thermal conditions

One of the common techniques to mitigate transistor aging is to increase the circuit supply voltage. Doing so reduces the propagation delay due to the BTI’s increase in threshold voltage (as discussed in Section 2.1). However, increasing the supply voltage to alleviate the effect of BTI has drawbacks. The static and dynamic powers are exponential and quadratic, respectively, proportional to the supply voltage. Therefore, increasing the supply voltage significantly increases the power consumption. Additionally, the power increase increases the circuit temperature, which degrades the performance of thermally limited systems, such as datacenter servers [15].

Another approach for mitigating transistor aging is to impose tighter timing constraints on the clock cycle time. For instance, if we assume a clock cycle time of 1 ns and anticipate a 10% degradation over the lifetime of the IC, it would be necessary to tighten the clock cycle time in the physical design implementation to 0.9 ns. Such an approach involves the usage of fast low  $V_t$  logical cells, which may exhibit high leakage power. However, this approach increases the development time of the physical design stage, potentially delaying the time to market.

### 2.4. Datacenter total cost of ownership

Total cost of ownership (TCO) is a key optimization metric for datacenters [23]. The TCO consists of two main costs: (1) capital expenses (CAPEX) and (2) operational expenses (OPEX). CAPEX include the cost of acquiring a building, the power costs, including electricity payments, the cost of

acquiring cooling equipment, and the cost of acquiring servers, including all their components and networking equipment. OPEX include operation power and maintenance costs.

Existing TCO models [18, 31, 11, 30, 42, 23, 19, 22] estimate the TCO by summing the datacenter infrastructure cost ( $C_{\text{infrastructure}}$ ), the server acquisition cost ( $C_{\text{server}}$ ), the networking equipment cost ( $C_{\text{network}}$ ), the power cost ( $C_{\text{power}}$ ), and the maintenance cost ( $C_{\text{maintenance}}$ ):

$$\begin{aligned} TCO &= CAPEX + OPEX, \\ CAPEX &= C_{\text{infrastructure}} + C_{\text{server}} + C_{\text{network}}, \\ OPEX &= C_{\text{power}} + C_{\text{maintenance}}. \end{aligned} \quad (4)$$

Although these tools take into account several datacenter parameters, such as server performance, power, cost, age, and mean time to failure, they *do not* explore how processor lifetime due to transistors aging affects the datacenter TCO.

### 3. An Adaptive Model for Transistor-Aging Margins in Datacenters

This section presents an integrated model that adaptively determines the necessary aging design margins for datacenter ICs based on their specific lifetime and mission profile requirements. Unlike common existing design approaches that assume worst-case scenarios, the proposed approach avoids over-designing ICs for datacenters, thereby reducing datacenter TCO and minimizing their carbon footprint.

By using Equation (2), the ratio  $R_{\text{tpd}}$  of the propagation delay of aged logical elements to fresh elements can be expressed by

$$R_{\text{tpd}} = \frac{t_{\text{pd aged}}}{t_{\text{pd fresh}}} = \frac{(V_{\text{DD}} - V_{\text{th}})^\alpha}{(V_{\text{DD}} - V_{\text{th}} - \Delta V_{\text{th}})^\alpha} = \left(1 - \frac{\Delta V_{\text{th}}}{V_{\text{DD}} - V_{\text{th}}}\right)^{-\alpha}. \quad (5)$$

Additionally, by combining Equations (5) and (1), we can express  $R_{\text{tpd}}$  as

$$R_{\text{tpd}} = \left(1 - \frac{C_T e^{-E_a/k_B T} t^{1/n}}{V_{\text{DD}} - V_{\text{th}}}\right)^{-\alpha}. \quad (6)$$

Figure 2 presents a set of curves illustrating the shift in logical cell delay over a lifetime range of up to 10 years. The curves assume constant operating junction temperatures of 60, 70, 80, 90, 100, and 105 °C. Note that the speed degradation is proportional to the lifetime raised to the power  $1/n$  and to an exponential dependent on the operating temperature. For example, when operating at 105 °C, a three-year lifetime incurs an 8% speed degradation, whereas a 10 year lifetime degrades by 10%. However, when operating at 90 °C, speed degrades by 2.2%–2.6% after a two to three years, whereas the degradation is 4% for a 10 year lifetime.

As opposed to IC vendors, who typically specify chips under worst-case conditions (e.g., 10 year lifetime and junction temperatures of 105 °C), datacenter ICs operate under different

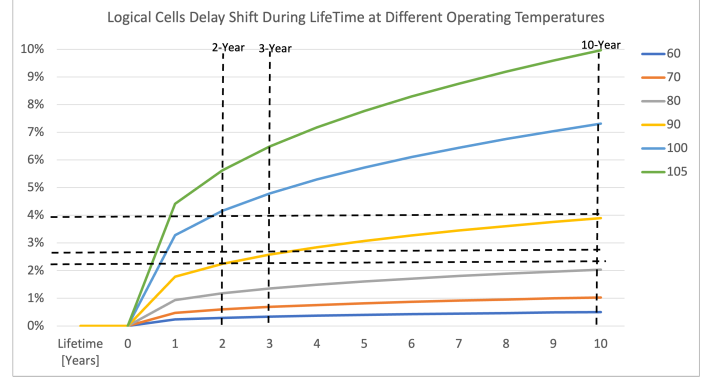


Figure 2: Logical cells delay shift during lifetime at different operating temperatures.

workloads and temperatures. The mission profile is commonly used as a simplified representation of all relevant passive and dynamic load conditions to which a population of computational elements is exposed during its entire life cycle [7]. Two possible examples of a datacenter mission profile [41, 20, 33] are given in Table 1.

T <sub>j</sub> [°C]	Mission Profile 1 [%]		Mission Profile 2 [%]	
	Active	Passive	Active	Passive
25	0%	0%	0%	50%
30	0%	0%	0%	0%
40	0%	48%	3%	0%
50	0%	0%	3%	0%
60	35%	0%	8%	0%
70	16%	0%	5%	0%
80	1%	0%	4%	0%
90	0%	0%	28%	0%
100–105	0%	0%	0%	0%

Table 1: Possible mission profiles of datacenters.

Given that datacenters operate on different workloads and temperatures, we suggest a model based on the Arrhenius equation [25] [see Equation (7)] to determine the aging degradation under changing conditions. The original Arrhenius equation serves to determine thermal acceleration factors for time-to-failure distributions of semiconductor devices:

$$AF = \exp\left[\frac{E_a}{k_B} \left(\frac{1}{T_s} - \frac{1}{T_i}\right)\right], \quad (7)$$

where  $AF$  is the acceleration factor due to changes in temperatures,  $T_i$  is the absolute temperature of the tested system, and  $T_s$  is the absolute temperature of the system.

The proposed model searches for an effective constant temperature  $T_{\text{eff}}$ , which produces an aging degradation equivalent to the degradation induced by the mission profile. Let  $\{T_i\}$ ,  $i = 1, 2, \dots, N$  denote the set of operating temperatures in the mission profile. For example,  $T_1 = 25^\circ\text{C}$  and  $T_2 = 30^\circ\text{C}$  for the mission profile presented in Table 1. Let  $AF_i$  be the



acceleration factor for temperature  $T_i$ :

$$AF_i = \exp \left[ \frac{E_a}{k_B} \left( \frac{1}{T_i} - \frac{1}{T_t} \right) \right]. \quad (8)$$

Let  $P_i$  denote the percentages of time that an IC is at a temperature  $T_i$ , and let  $AF_{\text{eff}}$  be the effective acceleration profile that corresponds to the IC mission profile [7], as described in the following equation:

$$AF_{\text{eff}} = \sum_{i=1}^n AF_i \times P_i = \sum_{i=1}^n P_i \exp \left[ \frac{E_a}{k_B} \left( \frac{1}{T_i} - \frac{1}{T_t} \right) \right]. \quad (9)$$

We can also express  $AF_{\text{eff}}$  by assuming  $T_s = T_{\text{eff}}$  as shown by the following equation:

$$AF_{\text{eff}} = \exp \left[ \frac{E_a}{k_B} \left( \frac{1}{T_{\text{eff}}} - \frac{1}{T_t} \right) \right]. \quad (10)$$

We can extract  $T_{\text{eff}}$  by combining Equations (9) and (10):

$$T_{\text{eff}} = \left\{ \frac{1}{T_t} + \frac{k_B}{E_a} \ln \sum_{i=1}^n P_i \exp \left[ \frac{E_a}{k_B} \left( \frac{1}{T_i} - \frac{1}{T_t} \right) \right] \right\}^{-1}. \quad (11)$$

For example, for mission profile 1 and mission profile 2 from Table 1 we obtain  $T_{\text{eff}} = 57.3^\circ\text{C}$  and  $T_{\text{eff}} = 73^\circ\text{C}$ , respectively.  $T_{\text{eff}}$  can then be assigned to Equation (6) to calculate the degradation ratio:

$$R_{\text{tpd}} = \left( 1 - \frac{C_T e^{-E_a/k_B T_{\text{eff}} t^{1/n}}}{V_{\text{DD}} - V_{\text{th}}} \right)^{-\alpha}. \quad (12)$$

As illustrated in Figure 2, the aging degradation for  $T_{\text{eff}} = 57.3^\circ\text{C}$  is nearly 0.3% while  $T_{\text{eff}} = 73^\circ\text{C}$  will be degraded by 0.7% assuming a 3 year lifetime.

The proposed model shows that, given an mission profile, the degradation of the aging speed can be significantly less than the worst-case assumption of IC vendors assuming an operating temperature of  $105^\circ\text{C}$ . The effective temperature, which is calculated by the proposed model for a given mission profile, can potentially relax aging margins, reduce datacenter TCO, and reduce the  $\text{CO}_2$  footprint.

## 4. Experimental Explorations

Our experimental analysis explores the trade-offs between IC lifetime, design complexity, performance, energy consumption, and operating conditions for datacenter TCO. Our analysis consists of two stages.

The first stage is a sensitivity analysis of transistor aging design margins for power consumption and includes the following:

1. The impact of transistor aging clock cycle time margins on power consumption. This analysis focuses on the lifetime of new ICs at the design stage.
2. The impact of operating voltage aging margins on existing ICs.

In the case study, we use a RISC-V CPU core and perform full synthesis, place-and-route, and power analyses.

In the second stage, we explore the trade-offs between lifetime, voltage, and frequency for datacenter TCO per QPS and  $\text{CO}_2$  footprint. We examine various possible IC SKU options in conjunction with a redesign of new ICs with reduced lifetime assumption, as described in Table 2.

The IC1 option involves a baseline 10-year-lifetime server operating at 1 V with a 2 GHz clock frequency. IC2 is a similar alternative for a 3-year-lifetime server but with a 2.2 GHz clock frequency. IC3 considers the same server as IC1 but operating at 0.93 V and with a reduced lifetime of 3 years. IC4 combines IC2 and IC3 by considering a slightly lower voltage in conjunction with a slightly higher clock frequency. IC6 considers a new server design with a reduced lifetime of 3 years in conjunction with reduced clock cycle time aging margins but operating under the same nominal conditions as the baseline option (IC1). Additionally, we added to the analysis two more configurations (IC6-5% and IC6-10%, corresponding to a reduced processor cost of 5% and 10%, respectively, relative to the IC1 processor cost). These configurations show the cost savings achieved by relaxing lifetime requirements.

These scenarios are discussed in detail in section 4.2, which more thoroughly analyzes the trade-offs. Our analysis covers a single server, a small datacenter, and a large datacenter with a variable number of servers.

### 4.1. Impact of transistor aging on power

The sensitivity analysis uses the CV32E40P RISC-V CPU core<sup>1</sup> [1] to examine how transistor aging design margins affect the power consumed. We conduct full synthesis, place and route, and timing analyses and simulate the power consumed by the RISC-V core in a 28 nm process node. The synthesis, place-and-route, and power analyses were done using Cadence® *Genus*<sup>TM</sup>, *Innovus*<sup>TM</sup>, and *Joules*<sup>TM</sup>, respectively. Table 3 summarizes the physical design parameters of our experiments. The physical design flow incorporates a multiple threshold voltage design flow (multi- $V_t$ ), which is widely used by common EDA tools to optimize both power consumption and performance. In this approach, transistors with different threshold voltages are selectively used in different parts of the design based on their specific requirements. For example, standard  $V_t$  (SVT) transistors are used in regions where power efficiency is the primary concern, whereas low- $V_t$  (LVT) transistors are used in areas that require high performance.

Figure 3 illustrates a sensitivity analysis showing how the clock cycle time aging margin depends on the leakage power and LVT cell count. The aging margin represents an additional guardband applied to the CPU clock cycle time, which corresponds to different lifetime targets. A higher aging margin and longer lifetime target increase the leakage power, primarily because of a significant rise in the number of LVT cells. For

<sup>1</sup><https://github.com/AI-Vector-Accelerator/cv32e40p>

IC Product	Nominal voltage	F [GHz]	$f_s$	$f_d$	Lifetime (years)
IC1	1	2	$P_s$	$P_d$	10
IC2	1	2.2	$P_s$	$1.1 P_d$	3
IC3	0.93	2	$0.7 P_s$	$0.8649 P_d$	3
IC4	0.96	2.10	$0.9 P_s$	$0.9676 P_d$	3
IC6	1	2	$0.8 P_s$	$1 P_d$	3

Table 2: Possible trade-offs between lifetime, operating voltage, and temperature in a datacenter for different IC products.  $P_s$  is the static CPU power consumption and  $P_d$  is the dynamic power consumption.

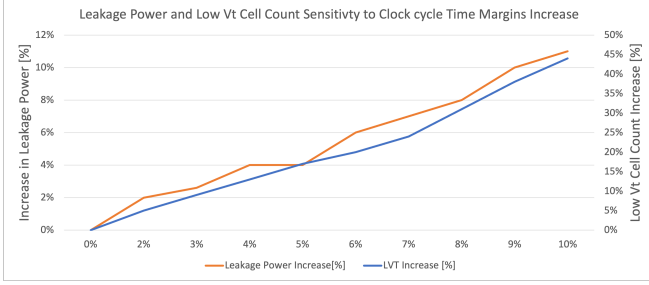


Figure 3: Leakage power and LVT cell count sensitivity analysis with respect to clock cycle time margin.

instance, increasing the clock cycle time margin by 10% (corresponding to a 10 year lifetime) produces an 11% increase in leakage power and a nearly 45% increase in the number of LVT cells. However, a 2% increase in the clock cycle time margin (corresponding to a 3 year lifetime) produces only a 2.6% increase in leakage power and only a 5% increase in the number of LVT cells.

Our experimental analysis also includes a sensitivity analysis of dynamic and static power consumption for a range of operating voltages (see Figure 4). Leakage power grows exponentially with operating voltage, whereas dynamic power consumption is proportional to the voltage squared. For example, the IC3 IC product from Table 2, which operates at 0.93 V, decreases static power by 30% and dynamic power by more than 13% compared with the baseline IC product (IC1). In addition, IC4 operating at 0.96 V decreases static power by 10% and dynamic power by more than 3% relative to IC1. The dynamic power of IC4 also takes into account that it runs 5% faster than IC1.

In the second stage, we use the data acquired from this section and evaluate how it affects the TCO of datacenters.

Physical design parameter	
Process node	28 nm
Nominal voltage	1 V
Junction temperature	105 °C
Clock frequency	366 MHz
Standard cell library	SVT and LVT
Clock cycle time margins	0%–10%
Voltage range	$V_{nom} - 10\%$ , $V_{nom}$ , $V_{nom} + 10\%$

Table 3: Physical design parameters.

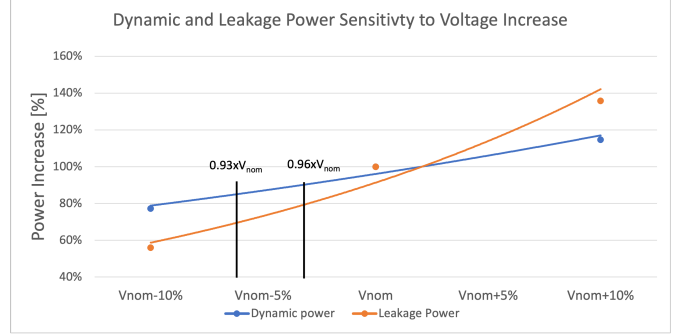


Figure 4: Leakage and dynamic power sensitivity with respect to  $V_{DD}$  margin.

## 4.2. Total cost of ownership analysis

TCO is a holistically optimized metric that can be used either at the design time of a system or at run time to compare the cost efficiency of solutions. This paper uses TCO to explore the benefits of customizing transistor aging design margins based on specific lifetime requirements. TCO is determined by the sum of CAPEX, which includes all costs such as building and cooling acquisition and server costs, and OPEX, which includes operational power costs (both static and dynamic), performance costs, and maintenance costs (based on the system’s reliability requirements). In this analysis we mainly explore OPEX that can be used to adjust run-time parameters such as voltage and frequency based on a given aging requirement. For comparison purposes we assume that the initial CAPEX are the same for all configurations explored. Operational expenses are determined by the sum of the power cost ( $C_{power}$ ) and maintenance cost ( $C_{maintenance}$ ) [18, 31].  $C_{maintenance}$  encapsulates the redundant components needed to satisfy the availability requirements and is mainly affected by the mean time to failure ( $MTTF$ ), mean time to repair ( $MTTR$ ), and temperature based on the Arrhenius function [Equation (7)]. Conversely,  $C_{power}$  is strongly affected by the total server power, the power usage effectiveness of the data center and the electricity cost per kW h. For this analysis we consider both peak and idle power. Thus, to estimate TCO based on the average power, we use the following equation for the CPU component (and for all the other components):

$$P_{avg CPU} = (uP_{peak CPU}) + [(1 - u)P_{idle CPU}], \quad (13)$$

where  $P_{peak CPU}$  is the peak CPU power consumption,  $P_{idle CPU}$  is the idle CPU power consumption, and  $u$  is the average

IC Product	$P_i$ (W)	$P_a$ (W)	#QPS	Processor cost (\$)
IC1	85	99	379 747	600
IC2	87	102	387 097	600
IC3	63	75	379 747	600
IC4	78	91	382 166	600
IC6	72	86	379 747	600
IC6-%5	72	86	379 747	570
IC6-%10	72	86	379 747	540

Table 4: CPU information for each IC product.  $P_i$  is the CPU idle power consumption,  $P_a$  is the CPU active power consumption, and QPS is the number of queries per second with the 400  $\mu$ s constraint.

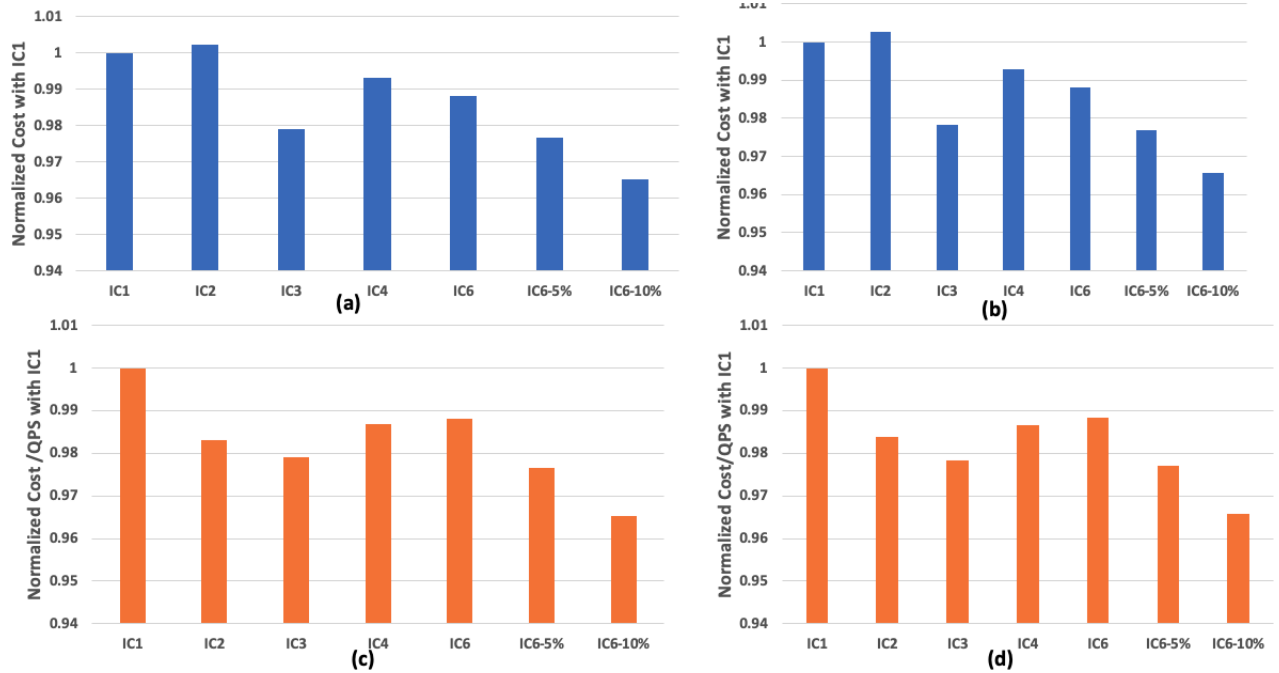


Figure 5: Single server normalized with IC1 (a) cost including average power, (b) cost including peak power, (c) cost per QPS including average power, (d) cost per QPS including peak power.

utilization. For TCO estimations based on the peak power consumption the utilization can be set to unity. In this paper we analyze how both peak and average power affect the TCO to capture small- and large-scale facilities.

To evaluate all configurations presented in Table 2, we use Skylake servers running Memcached [12], which is a lightweight distributed in-memory object-caching system (key-value store) used to accelerate user-facing applications with stringent latency requirements by alleviating database load. We ran Memcached in real machines using a cluster of six nodes, one node for the server process and five nodes for client processes running the mutilate load generator [27] that recreates the ETC workload from facebook [6]. In these experiments we disabled turbo mode and core and uncore frequency dynamic scaling and set the frequency at 2 GHz for both core and uncore components. We also disabled c-states (C0) and p-states to provide comparable power and performance values for all IC products. To monitor idle and active CPU power we

use turbostat. We ran each experiment five times for each IC product (IC1–IC6) and each time we collected the 99th tail latency and peak CPU power. The final power, temperature, and tail latency results are calculated by removing the minimum and maximum values and averaging the three remaining values. For this analysis we assume a 400  $\mu$ s constraint on the quality of service end-to-end, 99th tail latency application.

For the datacenter and server costs (DRAM, CPU, SSD, and other components) we use available industrial data [3, 2, 4]. For IC products 1–4 and 6, Table 4 shows the CPU idle power consumption ( $P_i$ ), CPU active power consumption ( $P_a$ ), the number of queries per second (QPS) with the 400  $\mu$ s constraint, and the processor’s cost in USD. To monitor  $P_a$  for IC1 we ran Memcached, as described above, at 2 GHz core frequency, and we monitored the CPU power on the server node. For idle power we monitored CPU power when the server was idle at 2 GHz core frequency on the server node. To estimate  $P_a$  for all the other IC products we use the power

factors given in Table 2 in the following equations:

$$P_{\text{alC}x} = f_s P_s + f_d x P_d + f_d k P_d, \quad (14)$$

$$P_{\text{ilC}x} = f_s P_s + f_d k P_d, \quad (15)$$

where  $P_{\text{alC}x}$  and  $P_{\text{ilC}x}$  are the estimated active and idle power for each IC product, respectively.  $P_s$  is the static CPU power consumption,  $P_d$  is the dynamic power consumption, and  $x$  and  $k$  are the percentage of active power when the machine is idle. To estimate  $P_s$  we monitor the CPU power when the server is idle with the C1 c-state enabled to stop the CPU clocks from capturing the leakage power.

The number of QPS in Table 2 is the total number of QPS served by the 400  $\mu\text{s}$  99th tail latency threshold. Since the different IC products in Table 4 keep the power within the TDP limit, we conservatively<sup>2</sup> assume that the dynamic temperature  $T_d$  is unchanged compared to the baseline product. Finally, the processor cost is taken from publicly available data for an Intel core i7 processor [3]. For the last two configurations (IC6-5%, IC6-10%), we applied a percentage cost reduction at the respective percentages on the initial value of \$600.

To analyze the TCO we leverage the TCO tool from Ref. [18] providing as input to the tool all the parameters from Table 4 and many others.

#### 4.2.1. Single server cost analysis

We first analyze the cost and cost per QPS for all IC products assuming a single server. Figures 5(a) and 5(c) present the estimated normalized cost and cost per QPS, respectively, taking into account the average power, as shown in Eq. (13). Additionally, Figures 5(b) and 5(d) show the estimated normalized cost and cost per QPS, respectively, taking into account the peak power consumption. For the single server analysis, we exclude infrastructure, network, and maintenance because these components are applicable only to larger-scale datacenters. Figure 5(a) shows that the IC2 configuration is not a good option because it increases the cost more than IC1 because of the higher power consumption. However, Figure 5(a) shows that, including the QPS, IC2 provides approximately a 1.8% savings compared with the baseline (IC1). Figures 5(b) and 5(d) express cost numbers very similar to the graphs that take into account average power consumption [Figures 5(a) and 5(c)] for estimating the TCO.

For the remaining analysis, we include the peak power consumption.

#### 4.2.2. Analysis of datacenter total cost of ownership

In this case study, we investigate how the TCO per QPS of each configuration is affected in a large datacenter with 30 000 servers. Figure 6(a) shows the TCO per QPS normalized by IC1 for all the IC products for a datacenter with 30 000 servers. Figure 6(a) is similar to Figure 5(d) in that it shows that IC3

is the preferred product because it provides a lower TCO per QPS for a large-scale datacenter. Moreover, the savings of a larger-scale datacenter increase with datacenter size. For example, IC6-10% provides the best TCO per QPS with a 4.3% improvement over the baseline. The second-best configuration is IC3%, which delivers an improvement of almost 3.8%. IC3 provides the highest OPEX per QPS, as shown in Figure 6(b), because it has a lower active power (see Table 4).

#### 4.2.3. Impact of different IC products on CO<sub>2</sub> footprint

This analysis explores how IC products affect the CO<sub>2</sub> emissions. Figure 7 presents the normalized CO<sub>2</sub> per QPS emission per year. The results show that IC1 emits more CO<sub>2</sub> per QPS than all other configurations, whereas IC3 and IC4 reduce the CO<sub>2</sub> emissions by up to 14% and 5%, respectively. Additionally, the CO<sub>2</sub> emissions for IC6, IC6-5%, and IC6-10% decrease compared with the baseline (IC1) by almost 8%. This result is explained by the strict correlation of power consumption with CO<sub>2</sub> emissions: the higher the power consumption, the higher the CO<sub>2</sub> emissions.

Overall, the results of this analysis show that most ICx configurations that operate at lower voltages outperform the baseline configuration (IC1) and significantly reduce the TCO per QPS (up to 4.3%). Such reductions represent a cost savings of about \$103 323 per month for a datacenter hosting 30 000 servers.

## 5. Related Work

This paper shows how over-provisioning modern processors with higher reliability margins than the actual lifetime affects a datacenter’s TCO. Several prior works discuss how processor design parameters affect datacenter TCO (see, e.g., Refs. [31, 16, 21, 43]).

Panagiota *et al.* [31] discuss the implications of DRAM failures and protection techniques for datacenter TCO. They present a modeling framework and simulator to analyze the TCO implications of various protection techniques and report that DRAM failure can significantly affect TCO and that the choice of protection technique is crucial. ECC techniques reduce failures but increase power consumption and reduce memory capacity, leading to higher TCO. Grot *et al.* [16] explore the specialized scale-out processor architecture to maximize on-chip computing density, thereby maximizing the performance for a given TCO. They optimize the TCO in data centers using scale-out processors and analyze how various design parameters, such as core count, frequency scaling, and power management, affect performance and power consumption. Through experimental evaluation and modeling, they demonstrate that scale-out processors, which use many low-power cores, can significantly improve energy efficiency and reduce the TCO of datacenters. Kleantous *et al.* [21] emphasize the need to consider multiple layers of the system, including architecture, microarchitecture, and circuit levels, to comprehensively evaluate TCO. They argue that

<sup>2</sup>A product that operates at lower power than the baseline dissipates less heat so it operates at a lower temperature than the baseline.



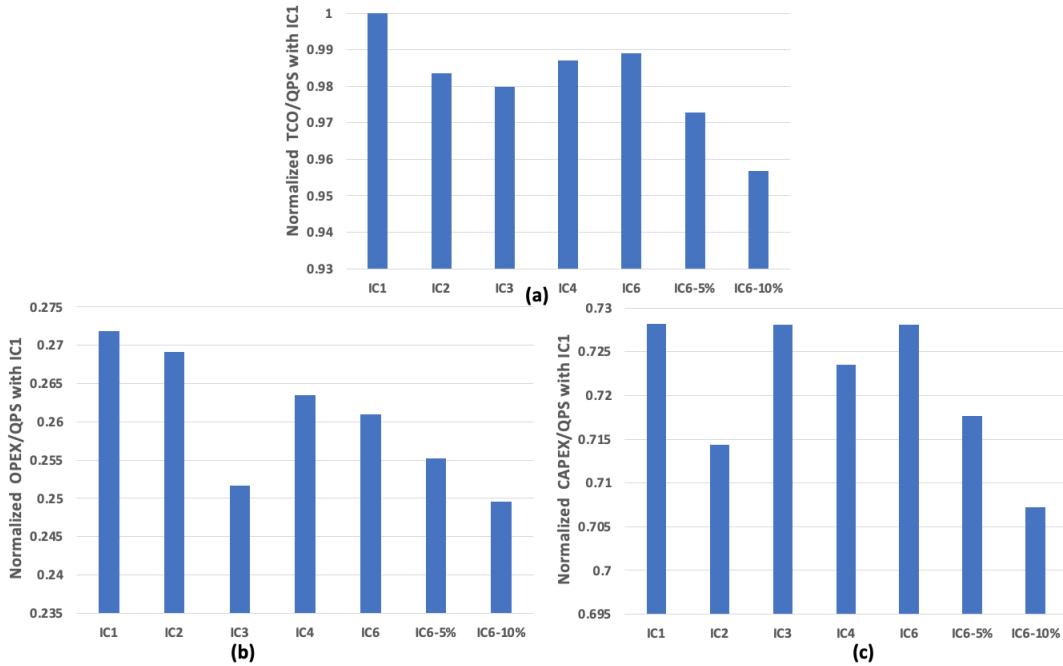


Figure 6: Datacenter with 30 K servers normalized by IC1: (a) TCO per QPS, (b) OPEX per QPS and (c) CAPEX per QPS.

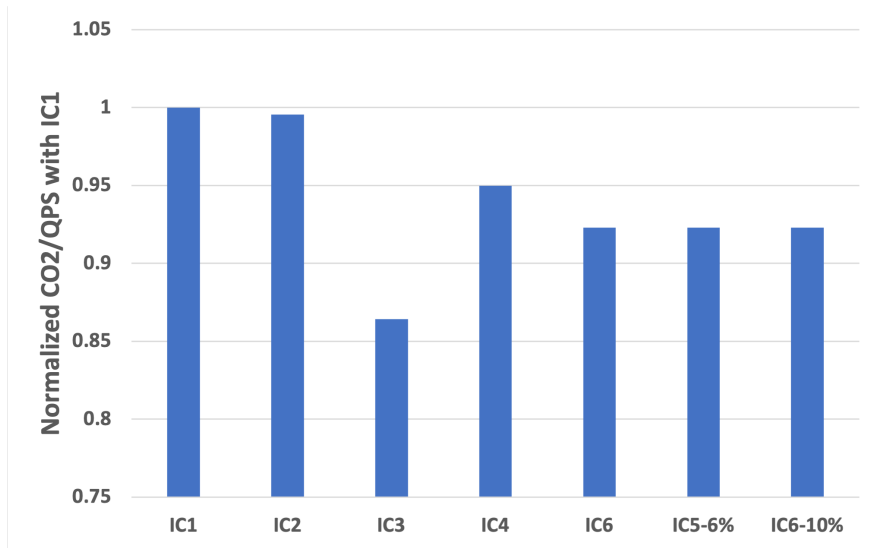


Figure 7: CO<sub>2</sub> emission. The results are normalized by IC1.

optimizing individual layers independently may not improve overall system performance. The paper presents case studies demonstrating the effectiveness of their approach, highlighting the importance of holistic evaluations that consider the interdependencies between different layers for designing efficient and high-performance systems. In a case study, they compare two- versus three-dimensional (3D) processor integration (e.g., DRAM 3D integration) to analyze datacenter TCO. Zhengyu *et al.* [43] explore the use of an optimized flash resource management to improve datacenter TCO. They address the limitations of existing TCO models by considering factors

specific to flash storage, such as endurance, performance, and energy efficiency. They propose a new TCO model that incorporates these factors and experimentally validate its accuracy.

Existing TCO models and calculators [18, 31, 11, 30, 42, 23, 19, 22] and some recent studies [31, 16, 21, 43] take into account several data center parameters, such as server performance, power, cost, age, and mean time to failure. However, these works do not consider how processor lifetime (due to transistor aging) affect datacenter TCO, which is the focus of the present work.

## 6. Conclusions

This paper examines and analyzes how over-provisioning modern processors with higher reliability margins than the actual lifetime affects a datacenter's total cost of ownership (TCO). To evaluate this impact, we estimate a modern processor's power consumption and performance with voltage margins corresponding to lifetimes of 3 and 10 years and feed the power and performance data into a state-of-the-art TCO modeling tool. A rigorous analysis shows that the TCO of a datacenter with 200 servers can be reduced by 5.2% by reducing the server processor lifetime from 10 years to 3 years. We conclude that over-designing modern datacenter processors to extended their lifetime (e.g., 10 years) beyond their actual replacement period (e.g., 3 years) significantly affects datacenter TCO. In addition, this effect increases when considering more IT equipment, such as datacenter switches.

## References

- [1] OpenHW Group CV32E40P User Manual. "https://docs.openhwgroup.org/projects/cv32e40p-user-manual/en/latest/index.html", 2003.
- [2] Dram specifications and cost, <https://memory.net/store/>.
- [3] Intel core i7 cost, <https://www.amazon.de/-/en/Intel-i7-11700K-Generation-Desktop-Processor/dp/B08TX5KL5T/>.
- [4] Solid state drive specifications and cost, <https://www.amazon.de/-/en/Samsung-Internal-Solid-State-MZ-V8P2T0BW/>.
- [5] Thomas Aichinger, Gerald Rescher, and Gregor Pobegen. Threshold Voltage Peculiarities and Bias Temperature Instabilities of SiC MOSFETs. *Microelectronics Reliability*, 2018.
- [6] Berk Atikoglu, Yuehai Xu, Eitan Frachtenberg, Song Jiang, and Mike Paleczny. Workload analysis of a large-scale key-value store. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*, pages 53–64, 2012.
- [7] Component Technical Committee Automotive Electronics Council. Failure mechanism based stress test qualification for integrated circuit. *AEC-Q100-REV-G standard*.
- [8] DS Boudreaux, F Williams, and AJ Nozik. Hot Carrier Injection at Semiconductor Electrolyte Junctions. *Journal of Applied Physics*, 1980.
- [9] Sridhar Chakravarthi, Anand Krishnan, Vijay Reddy, CF Machala, and Srikanth Krishnan. A Comprehensive Framework for Predictive Modeling of Negative Bias Temperature Instability. In *IRPS*. IEEE, 2004.
- [10] Kueing-Long Chen, Stephen A Saller, Imelda A Groves, and David B Scott. Reliability Effects on MOS Transistors due to Hot-Carrier Injection. *Transactions on Electron Devices*, page 3, 1985.
- [11] Yan Cui, Charles Ingalz, Tianyi Gao, and Ali Heydari. Total Cost of Ownership Model for Datacenter Technology Evaluation. In *2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, pages 936–942. IEEE, 2017.
- [12] Brad Fitzpatrick and Anatoly Vorobey. Memcached: a distributed memory object caching system, 2011.
- [13] Freddy Gabbay and Avi Mendelson. Asymmetric aging effect on modern microprocessors. *Microelectronics Reliability*, 119:71–81, 2021.
- [14] Freddy Gabbay, Avi Mendelson, Basel Salameh, and Majd Ganaiem. A design flow and tool for avoiding asymmetric aging. *IEEE Design & Test*, 39(6):111–118, 2022.
- [15] Corey Gough, Ian Steiner, and Winston Saunders. CPU Power Management. In *Energy Efficient Servers: Blueprints for Data Center Optimization*. Springer, 2015.
- [16] Boris Grot, Damien Hardy, Pejman Lotfi-Kamran, Babak Falsafi, Chrysostomos Nicopoulos, and Yiannakis Sazeides. Optimizing datacenter tco with scale-out processors. *IEEE Micro*, 32(5):52–63, 2012.
- [17] Jawad Haj-Yahya, Avi Mendelson, Yosi Ben Asher, Anupam Chattopadhyay, Jawad Haj-Yahya, Avi Mendelson, Yosi Ben Asher, and Anupam Chattopadhyay. Power management of modern processors. *Energy Efficient High Performance Processors: Recent Approaches for Designing Green High Performance Computing*, pages 1–55, 2018.
- [18] Damien Hardy, Marios Kleanthous, Isidoros Sideris, Ali G Saidi, Emre Ozer, and Yiannakis Sazeides. An analytical framework for estimating TCO and exploring data center design space. In *2013 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, pages 54–63. IEEE, 2013.
- [19] Damien Hardy, Isidoros Sideris, Ali Saidi, and Yiannakis Sazeides. Eetco: A tool to estimate and explore the implications of datacenter design choices on the tco and the environmental impact. In *Workshop on Energy-efficient Computing for a Sustainable World*, 2011.
- [20] Cinar Kilcioglu, Justin M Rao, Aadharsh Kannan, and R Preston McAfee. Usage patterns and the economics of the public cloud. In *Proceedings of the 26th International Conference on World Wide Web*, pages 83–91, 2017.
- [21] Marios Kleanthous, Yiannakis Sazeides, Emre Özer, Chrysostomos Nicopoulos, Panagiota Nikolaou, and Zacharias Hadjilambrou. Toward multi-layer holistic evaluation of system designs. *IEEE Computer Architecture Letters*, 15(1):58–61, 2015.
- [22] Vasileios Kontorinis, Liuyi Eric Zhang, Baris Aksanli, Jack Sampson, Houman Homayoun, Eddie Pettis, Dean M Tullsen, and Tajana Simunic Rosing. Managing distributed ups energy for effective power capping in data centers. *ISCA*, 2012.
- [23] Jonathan Koomey, Kenneth Brill, Pitt Turner, John Stanley, and Bruce Taylor. A simple model for determining true total cost of ownership for data centers. *Uptime Institute White Paper, Version, 2:2007*, 2007.
- [24] Sanjay V Kumar, Chris H Kim, and Sachin S Sapatnekar. An Analytical Model for Negative Bias Temperature Instability. In *ICCAD*, 2006.
- [25] Keith J. Laidler. The development of the arrhenius equation. *Journal of Chemical Education*, 61(6):494, 1984.
- [26] Yongho Lee and Taewhan Kim. A Fine-grained Technique of NBTI-aware Voltage Dcaling and Body Biasing for Standard Cell-based Designs. In *ASP-DAC 2011*. IEEE, 2011.
- [27] Jacob Leverich. Mutilate: high-performance memcached load generator.(2014), 2014.
- [28] Jialun Lyu, Jaylen Wang, Kali Frost, Chaojie Zhang, Celine Irvine, Esha Choukse, Rodrigo Fonseca, Ricardo Bianchini, Fiodar Kazhamiaka, and Daniel S. Berger. Myths and misconceptions around reducing carbon embedded in cloud platforms. In *ACM 2nd Workshop on Sustainable Computer Systems (HotCarbon'23)*, 2023.
- [29] Saibal Mukhopadhyay, Arijit Raychowdhury, and Kaushik Roy. Accurate estimation of total leakage current in scaled cmos logic circuits based on compact current modeling. In *Proceedings of the 40th annual Design Automation Conference*, pages 169–174, 2003.
- [30] Panagiota Nikolaou, Yiannakis Sazeides, Alejandro Lampropoulos, Denis Guilhot, Andrea Bartoli, George Papadimitriou, Athanasios Chatzidimitriou, Dimitris Gizopoulos, Konstantinos Tovletoglou, Lev Mukhanov, et al. On the evaluation of the total-cost-of-ownership trade-offs in edge vs cloud deployments: A wireless-denial-of-service case study. *IEEE Transactions on Sustainable Computing*, 7(2):334–345, 2019.
- [31] Panagiota Nikolaou, Yiannakis Sazeides, Lorena Ndreu, and Marios Kleanthous. Modeling the implications of DRAM failures and protection techniques on datacenter TCO. In *Proceedings of the 48th International Symposium on Microarchitecture*, pages 572–584, 2015.
- [32] Kaushik Roy, Saibal Mukhopadhyay, and Hamid Mahmoodi-Meimand. Leakage current mechanisms and leakage reduction techniques in deep-submicrometer cmos circuits. *Proceedings of the IEEE*, 91(2):305–327, 2003.
- [33] Vishnu Sai Pilla. Thermal and energy prediction for energy-efficient data centers using machine learning. *Diss. California State University, Northridge*, 2023.
- [34] Takayasu Sakurai and A Richard Newton. Alpha-power law mosfet model and its applications to cmos inverter delay and other formulas. *IEEE Journal of solid-state circuits*, 25(2):584–594, 1990.
- [35] Dieter K Schroder. Negative Bias Temperature Instability: What do we understand? *Microelectronics Reliability*, 2007.
- [36] Dieter K Schroder and Jeff A Babcock. Negative Bias Temperature Instability: Road to Cross in Deep Submicron Silicon Semiconductor Manufacturing. *Journal of applied Physics*, 2003.
- [37] Jayanth Srinivasan, Sarita V Adve, Pradip Bose, and Jude A Rivers. Lifetime reliability: Toward an Architectural Solution. *IEEE Micro*, 2005.

- [38] James H Stathis and Sufi Zafar. The Negative Bias Temperature Instability in MOS Devices: A Review. *Microelectronics Reliability*, 2006.
- [39] E Takeda and N Suzuki. An Empirical Model for Device Degradation due to Hot-Carrier Injection. *Electron Device Letters*, 1983.
- [40] Jyothi Bhaskarr Velamala. *Compact Modeling and Simulation for Digital Circuit Aging*. PhD dissertation, Arizona State University, 2012.
- [41] Yewan Wang, David Nörtershäuser, Stéphane Le Masson, and Jean-Marc Menaud. Potential effects on server power metering and modeling. *Wireless Netw*, 29:1077–1084, 2023.
- [42] Wenrui Yan, Jie Yao, Qiang Cao, and Yifan Zhang. LT-TCO: A TCO Calculation Model of Data Centers for Long-Term Data Preservation. In *2019 IEEE International Conference on Networking, Architecture and Storage (NAS)*, pages 1–8. IEEE, 2019.
- [43] Zhengyu Yang, Manu Awasthi, Mrinmoy Ghosh, and Ningfang Mi. A fresh perspective on total cost of ownership models for flash storage in datacenters. In *2016 IEEE International conference on cloud computing technology and science (CloudCom)*, pages 245–252. IEEE, 2016.
- [44] Lide Zhang and Robert P Dick. Scheduled Voltage Scaling for Increasing Lifetime in the Presence of NBTI. In *ASP-DAC*. IEEE, 2009.