UCLouvain

epl

École polytechnique de Louvain

Promoting integrated circuits reusability thanks to thermal annealing: an environmental and financial profitability study

Author: **Antoine PERCY**
Supervisors: **Jean-Pierre RASKIN, Sébastien TOUSSAINT**
Readers: **Anne FALIER, David BOL, Jean-Brieuc FERON**
Academic year 2021-2022
Master [120] in Electrical Engineering

I am proud and happy to finish my five beautiful study years at the Ecole Polytechnique de Louvain-la-Neuve with this master thesis subject.

Writing this master thesis was an extremely enriching experience, but it was not always easy. Indeed, writing a master thesis involves learning and passion, but also many twists and turns.

It is important for me to thank all people and organisms that supported me and enabled me to complete this thesis.

First of all, I would like to thank my supervisor, Jean-Pierre Raskin, for having inspired me and for having trusted me to realize this master thesis, and for making me meet the right people.

Then, I want to thank Sébastien Toussaint, who accompanied me on this journey. I could always count on him for guidance for his exceptional presence. I could also count on him for changing my mind with his long stories and explanations about just any subject in engineering.

I would also like to show my gratitude to the EPL grand master of life cycle assessment Thibault Pirson who helped me to realize the life cycle part of this master thesis.

Of course, I thank Marnick Vanlee, who gave me his time, taught me so much about the world of e-waste, and allowed me to visit the facility of CTG-circular in Mechelen.

I also think to Justine Lebrun and Mieszko Di Gregorio Botek, my co-memorants with whom we met every day to advance our research at the Open Hub, which I thank for the welcome. I think that we bring each other to the top during this master thesis. I was lucky to share this adventure with them.

Finally, I could not finish those thanks without mentioning my relatives. I thank my family, my parents as well as my brother and my sisters, for having always supported me and for having instilled in me great values, including those of hard work and surpassing oneself. I also thank my friends for allowing me to relax me when I needed it. The last but not least, I would like also to heartily thank my girlfriend who supported and helped in this difficult journey.

Abstract

The exponential growth associated with semiconducting knowledges translates inevitably, in a linear economy, into ever-increasing carbon emissions. The so-called silicon 'roadmap' was built with the objective of delivering, at a high rate, new products to fulfill bigger consumer needs. In such framework, no efforts needed to be oriented toward the reusability of building blocks such as integrated circuits (ICs). With the recent evolution of the semiconductor supply chain, Moore's law shortness of breath motivates the elaboration of sustainable semiconductor businesses. In this thesis, we chose to investigate the relevance of a circular model for ICs'life cycle based on remanufacturing ICs thanks to a thermal annealing process. Particularly, we focused on a prominent ICs' family: the flash memories. The recovering of such memory pristine electrical properties is investigated thanks to 'Erase/program' time evolution curves before and after the annealing treatment. From there, we estimated the economic relevance of reconditioning ships thanks to the life cycle assessment of a hypothetical remanufacturing plant. We showed that the reuse of flash memories could reduce the European Greenhouse gas emissions by 0.6 to 1.4 MegaTons CO2 equivalent. To inferring benefits of such approach, the structural and functional costs were estimated as well as market trends. It was shown that the profitability range around -0.9 and 13.9M€ depending highly on the market demand for remanufactured ICs. Nevertheless, we promote here a value chain for reconditioned ICs with ICs reliability and testability at its core.

Contents

List of abbreviations

BGA	Ball Grid Array
CT	Charge Trap
CTG	Composite Transfer Grant
DDR3 SDRAM	Double Data Rate 3rd generation Synchronous Dynamic Random Access Memory
DDR4 SDRAM	Double Data Rate Fourth Generation Synchronous Dynamic Random Access Memory
DIP	Dual Inline Package
DRAM	Dynamic Random Access Memory
EEE	Electric and Electronic Equipment
EEPROM	Electrically-Erasable Programmable Read-Only Memory
FG	Floating Gate
eMMC	embedded Multi-Media Card
ESD	Electrostatic Sensitive Discharge
FPGA	Field Programmable Gate Array
GHG	Greenhouse Gases
GPIO	General Purpose Input/Output
GWP	Global-warming potential
IC	Integrated Circuit
ICT	Information and Communications Technology
ISPP	Incremental Step-Pulse Programming Procedure
JEDEC	Joint Electron Device Engineering Council
LCA	Life-Cycle-Assessment
LPDDR SDRAM	Low Power Double Data Rate SDRAM
LSB	Least-Significant Bit
MLC	Multi-Level Cell
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
MSB	Most Significant Bit
MSDA	Marketing, Selling, Distribution, and Administration
ONFI	Open NAND Flash Interface
PCB	Printed Circuit Board
SILC	Stress-Induced Leakage Current
SLC	Single-Level Cell
SMD	Surface-Mounted Device
SoC	System on Chip
TAT	Trap-Assisted Tunneling
TLC	Triple-Level Cell
TSMC	Taiwan Semiconductor Manufacturing Company, Ltd
TSOP	Thin Small-Outline Package
UFS	Universal Flash Storage

List of Figures

List of Tables

Chapter 1

Introduction

Since 2020, the semiconductor industry has been facing an unprecedented supply crisis. This crisis impacts all the pants of our economy, and semiconductors are now considered as the new 'oil' of our modern society [21, 22]. Microchip production capacities are an essential resource that influences global geopolitics. The current tensions around Taiwan are partly due to the fact that Taiwanese companies produce 67 percent of semiconductors used worldwide[23]. If China invades Taiwan, the world will be at the mercy of an almost instantaneous shutdown of half of the global production of electronic chips [24].

Consequently, in March 2022, the European Parliament launched a budget of 48 billion euros to increase the European market share from 10% today to 20% in 2030 [25].

Besides that, the climate stakes have never been so high. The global temperature is already at +1.1 °C since the pre-industrial era, and the impacts of global warming have never been so visible [26]. Here again, the ICT (Information and Communications Technology) has a responsibility. The environmental impact of ICT is at its highest level, and it is accounted for 3.5 to 3.8 % of the global GHG (Greenhouse gases) emissions [27, 28]. 54% of the emissions are allocated to the equipment manufacturing [27] and can reach up to 70% in the case of a smartphone [29].
The GHG emissions of terminals use phase are less and less important. Indeed, the exponential evolution of transistor integration has enabled the fabrication of systems with billions of transistors with more than 1,000 times higher energy efficiency [30]. The downside is that the manufacturing process of semiconductors becomes so energy intensive that, for example, integrated circuit manufacture impact is 36-40% of the carbon footprint of smartphones [29, 1]. All sectors must focus on reducing their carbon emissions. For the ICT, this can be done partly by introducing sobriety principles in the use and manufacturing phases of equipment.

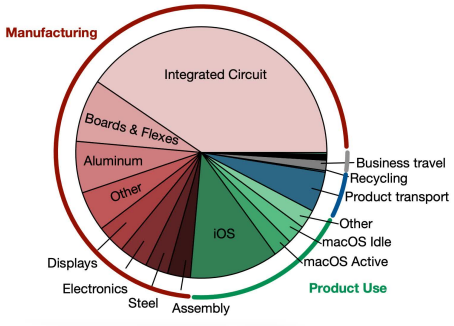


Figure 1.1: Apple's carbon-emission breakdown. Reproduced from [1]

One way to reach sobriety in the semiconductor industry while reducing our dependence on foreign markets is to reuse chips available in end-of-life devices. The idea of reusing semiconductors is not new, and Fraunhofer-Institut, Europe's largest application-oriented

research organization, demonstrated that it is technically feasible [8, 31, 32]. However, industries have not yet taken the plunge. One of the reasons is the lack of accepted quality statements of recovered components and the lack of studies about the economic viability of reusing components. [33]. Moreover, it has never been proven that reused ICs (integrated circuits) are better from an environmental point of view than new ones, and the problem is not so trivial because new ICs consume orders of magnitude less power than older ones. This master thesis will address these questions.

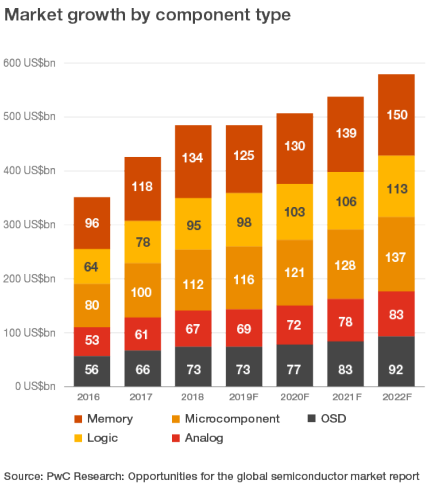


Figure 1.2: Market growth by component type. Reproduced from [2]

The world of semiconductors is vast, and the ICs have much heterogeneity. Hence, impossible to attack these three problems without targeting a specific type of IC. The first intuition was to find industry standards. Indeed standards allow the development of a single remanufacturing procedure applicable to many chips. Standards exist in the semiconductor industry, especially in the memory segment, which is the most significant segment in market share, i.e., 26% of the total semiconductors market. The memory segment is divided into two parts: RAM and flash memories. Flash memory will degrade throughout its life due to the intrinsic physic of the transistors used in it [13, 14, 15]. The reuse of degraded ICs will lead to a loss of market confidence in second-hand chips. From here, a research question arises, how to affirm the reliability during the remanufacturing process of a second-hand chip that degrades over time?

To answer this question, it will be studied how the primary electrical properties of flash memories can be nearly fully recovered thanks to thermal annealing in a reuse of semiconductors context and how it can be used as a guarantee of quality.

On top of that, a remanufacturing process of ICs will be studied, and the first published results on the environmental impact and the costs related to reusing semiconductors will be presented. The financial results obtained are put into perspective with new flash memory market prices, and the optimizations required to reach profitability are presented. Moreover, a complete LCA (Life-Cycle-Assessment), including the use phase of new and second-life flash memories, is also present in the manuscript.

Master thesis roadmap

First of all, Chapter 2 will introduce the fundamental concepts needed to understand what is a semiconductor, a flash memory, the thermal annealing process, and the process to reuse ICs.

Then, Chapter 3 will present and detail the research question regarding thermal annealing for flash memories.

Chapter 4 will explain the methodology used for studying the thermal annealing effects on flash memories, and a preliminary experiment will be presented.

Afterward, Chapter 5 will provide the results and discuss the thermal annealing effects on the "program/erase" time of flash memories.

Chapter 6 investigates reused flash memories' carbon footprint. To this end, the remanufacturing process designed for this master thesis is presented. Moreover, due to the infinity of possible use cases, an adaptative approach will be proposed to assess if yes or not the use of reused ICs make sense for a specific use case.

Chapter 7 deals with the financial profitability of a business based around semiconductor remanufacturing. The cost and the optimization that must be done to reach profitability are presented and discussed.

Finally, Chapter 8 will conclude the manuscript by responding to the three research questions and will give perspectives for further research.

Chapter 2

State of the art

This chapter exposes the key concepts necessary to understand the present thesis.

A brief overview of the semiconductor industry and the package terminology used will be presented before the reusing of semiconductors theory, to finish with a personal reflection on the types of ICs that could help to reach financial profitability. The final part will be dedicated to flash memory and thermal annealing theory.

2.1 An overview of the semiconductors' ecosystem

2.1.1 Integrated circuit definition and Moore's law

Integrated circuits, also called semiconductors, are the foundation of our digital world and are defined as "a circuit of transistors, resistors, and capacitors constructed on a single semiconductor wafer or chip, in which the components are interconnected to perform a given function" [34]. Integrated circuits are the building blocks of computers, smartphones, smart TVs, smart cars,... For example, A car embeds more than three thousand microchips [35] or a USB key integrates two main ICs, the NAND flash memory IC and the controller. These two ICs are shown in Figure 2.3.

Transistors are the building blocks of modern electronics. "A transistor is a semiconductor device used to amplify or switch electrical signals and power" [3]. The scheme of one MOSFET (Metal-Oxide-Semiconductor Field-Effect Transistor), the type of transistor most used today, can be found in Figure 2.1. To make it simple, pin G is the control pin. The pin G controls whether if the transistor connection between S and D is passing or blocking in the case of digital usage of the transistor. In the case of an analog usage of the transistor, the current between S and D will be proportional to the voltage applied in G.

The miniaturization of transistors has a direct impact on the computing power per available area or the amount of data that one single chip can store. Indeed, the smaller the transistor, the more it is possible to put on a given surface. The miniaturization law of transistors is known as the Moore's law. Gordon Earle Moore, the co-founder of Intel, postulated in 1965 that the size of the transistor would be reduced by a factor of two every two years [36] and as Figure 2.2 shows. This law has been confirmed empirically. It is also essential to understand that miniaturization is not only a story of transistors per area but directly impacts the exponentially growing computing power available to humanity. The

state-of-the-art volume production technology is the 5nm (for the channel length) one introduced in 2020. The industry roadmap calls for the introduction of 3nm technology in 2023 and the 2nm technology in 2024 [3]. 2nm is the size of tens of silicon atoms. Going further may become difficult to reach. Therefore other ways to integrate more and more transistors into a single chip will have to be found like 3D integration and new materials used,... If this is not done, this could be the end of Moore's law and, therefore, the end of ICs obscolescence caused by Moore's law. This could be very advantageous for the reuse of semiconductors.

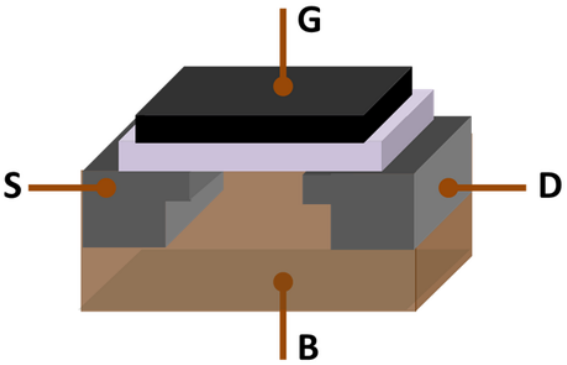


Figure 2.1: Scheme of a MOSFET transistor. Reproduced from [3]

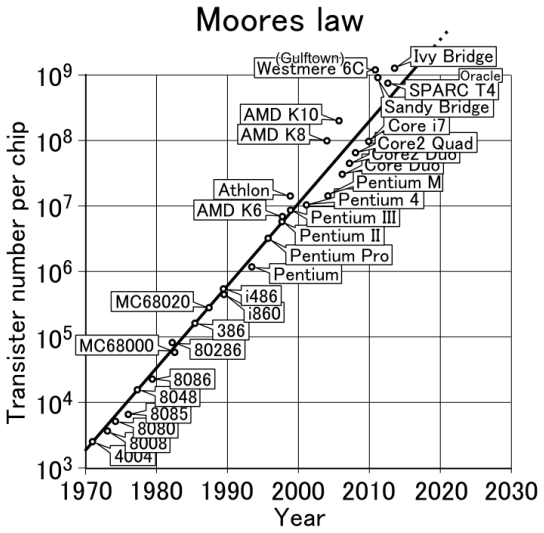


Figure 2.2: Moore's Law. Reproduced from [4]

2.1.2 Some basics for a better understanding

Before diving into the theory of semiconductor reuse, it is important to introduce some basics and to introduce the package terminology used in this master thesis.

First, microchips are pieces of silicon wafer on which the electronic circuits have been engraved. These small pieces of silicon are called dye and are embedded in packages. An IC in its package is embedded into a larger system and mounted on a PCB (Printed Circuit Board). In the example shown in Figure 2.3, a NAND flash memory and a USB controller are embedded on a USB drive PCB. Both are ICs, and we notice that the controller connectors and the NAND flash memory one do not look the same. Actually, there are embedded in different types of packages. Whether and how they can be reused depend on the type of packages used. In order to clearly discern the different types of packages that will be discussed in the manuscript, these are presented in the rest of this section.

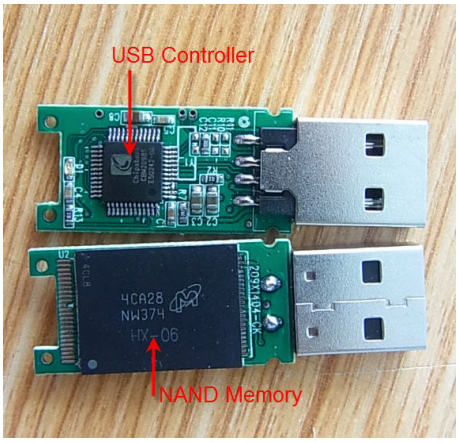


Figure 2.3: Nand flash memory and its controller embedded in a USB drive PCB. Reproduced from [5]

Through-hole Technology The definition given by [37] is "Through-hole components are wired leads (legs) that are pushed through drilled holes on the circuit board.". The most famous through-hole package is the dual hole in array packaging, also known as DIP (Dual Inline Package), represented in Figure 2.4. Everyone who has ever had electronic labs has handled this kind of package.

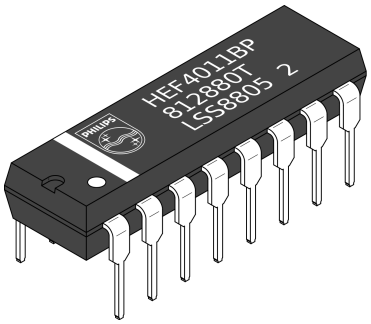


Figure 2.4: example of DIP package. Reproduced from [6]

Surface Mounted Device The second main type of package is the family of surface-mounted devices. Surface-mounted devices are chips mounted directly onto the surface of the PCB. This type of package reduces mounting cost and place used on the PCB [5]. SMD (Surface-Mounted Device) includes a lot of subpackage categories represented in Figure 2.5.

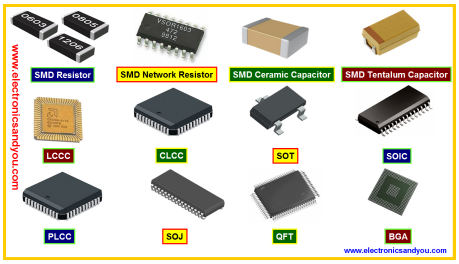


Figure 2.5: SMD packaging types, not representative. Reproduced from [5]

The focus must be placed on ball grid array packaging (BGA packaging). BGA packaging is represented in Figure 2.6. The connection is done with small solder balls placed under the IC. This configuration allows more interconnections than other SMD packaging because all the bottom surfaces of the device can be used, instead of just the perimeter [38]. Focus on this packaging is done because BGA chips are the most studied in the reuse of semiconductors context. This trend is driven by the fact that BGA packages are often used for valuable and dense chips. Moreover, planarity problems observed when reusing other SMD or DIP chips due to pin deformation are not present when dealing with BGA.

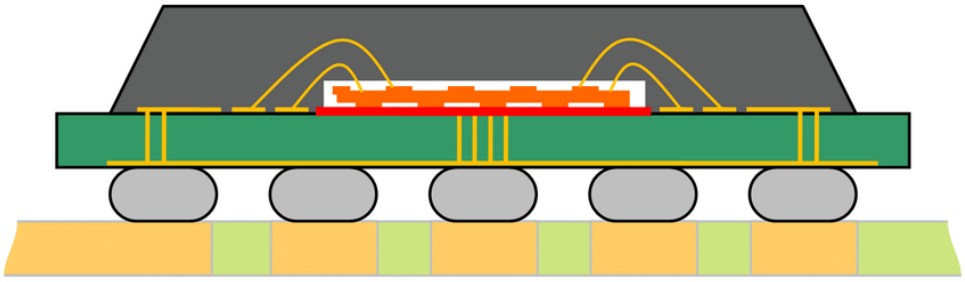


Figure 2.6: Schematic of ball grid array (BGA) IC. Reproduced from [7]

2.2 Reuse of semiconductors

The idea of semiconductors reuse was topical more than 20 years ago and the motivations were quite the same as today. European institution was at that time already aware of the E-Waste problem and made some directives to counteract the problem effects. [39]

At that time technological, managerial and cost aspects were the reasons why ICs reuse was not considered by the industry. These lacks are explained in [33] as follow :

- Technical: Absence of proofed and accepted quality statements for recovered component
- Managerial: Lack of delivery reliability in terms of kind, time, and location of components and missing experiences with costs for handling and reintegrating of reworked parts in the workflow
- Costs: There must be a balance among quality, delivery, reliability, and costs.

Since the most valuable work related to this field is the Sustainability Smart project led by the Fraunhofer and funded by the EU-Program Horizon 2020 [31]. This project allowed scientists to study smartphone sorting automation[40] and semiconductor remanufacturing [32]. They also demonstrate the feasibility of reusing BGA chips on new devices [8]. Indeed, they made 50 new wakers with reused controllers and reused flash memories coming from old video recorders. The process of reusing BGA chips was devided in two parts :

1. Desoldering
2. Remanufacturing

These two steps are discussed in the following subsections for the specific case of BGA chips. BGA chips are, in most, more expensive than chips embedded in other types of packages and they do not suffer from pin deformation that can lead to contact fault. Therefore, they are good candidate for revalorisation.

2.2.1 Desoldering process

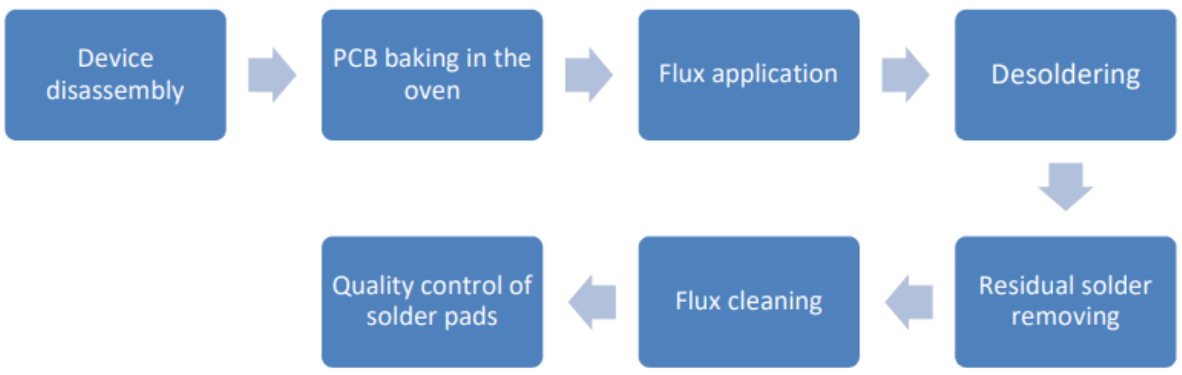


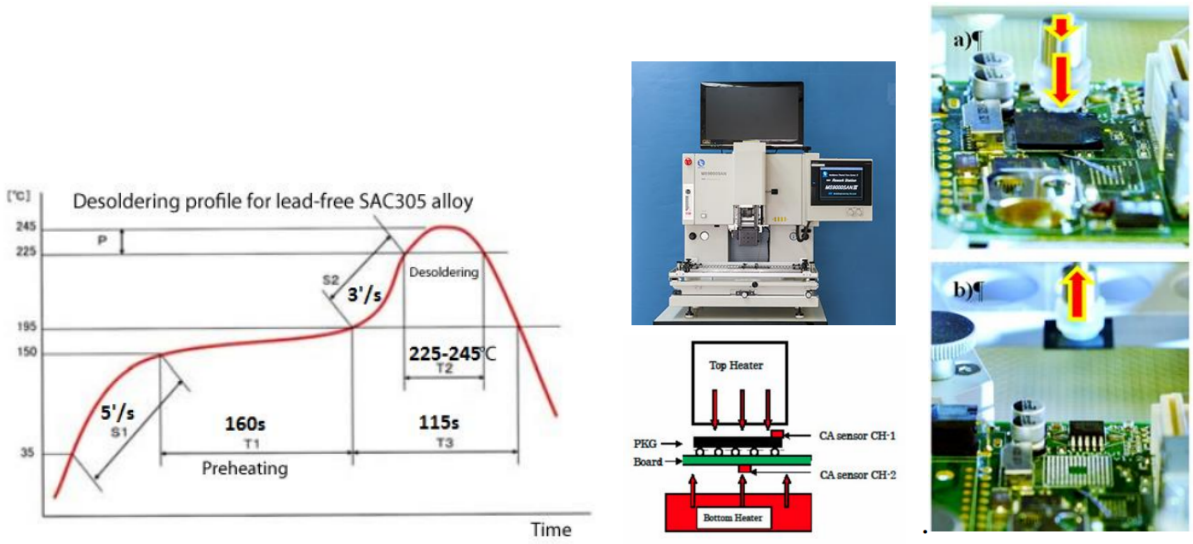
Figure 2.7: Flow chart: desoldering process. Reproduced from [8]

PCB baking in the oven This stage is required for all electronic components before exposing them to high temperatures. The reason is that electronic components embedded in plastic packages absorb moisture that can expend when exposing the component to high temperatures. This can damage the components and this damage might not be visible. Therefore drying the IC before mounting or demounting operation is needed to limit the risk of undetectable damage [32].

Flux application and desoldering The industry uses lead-free solder. The eutectic point of this alloy is located between 217 and 221 $°C$. This solder has a big tension surface. This results in the need of applying a significant force to remove the IC from the PCB. Higher temperature or flux application tends to decrease the tension surface of the alloy. [32, 9]

Exposure of the IC to temperature and use of a drag force is needed the remove components from PCB. Regarding these issues, the position of "SustainabilitySMART" was that the temperature profile needs to be carefully chosen because electronic chips are very temperature sensitive. On the one side, their desoldering strategy was to have a desoldering profile as low as possible to protect electronic chips damage (Fig 2.8a) [32, 9]. This constrained desoldering techniques. On the other side, [41] showed that the temperature resistance of memory BGA components to multi-reflow processes was above expectation (compared with datasheet information) and resist up to eleven thermal reflow processes (against three in datasheets).

There exist different desoldering techniques which have their pros and cons. For example, for the automatic reflow machine used to repair PCBs (Figures 2.8b and 2.8c) the thermal reflow profile is highly controllable but this machine is too slow for industrial application. The solution proposed by the "SustainabilitySMART" project can be found in [32], the desoldering line designed uses hot air for heating the PCB and a pic and place robot to remove components from the PCB. The components encounter a maximum temperature of 240 degrees. This system have a minimal desoldering time of 75 seconds per BGA component.



(a) (b) (c)

Figure 2.8: (a) Example of a lead-free temperature profile for desoldering process. (b) The MS9000SAN-(III) Rework Station – an example of equipment for desoldering. (c) An example of the component separation during desoldering. All reproduced from [9]

.

2.2.2 Remanufacturing process

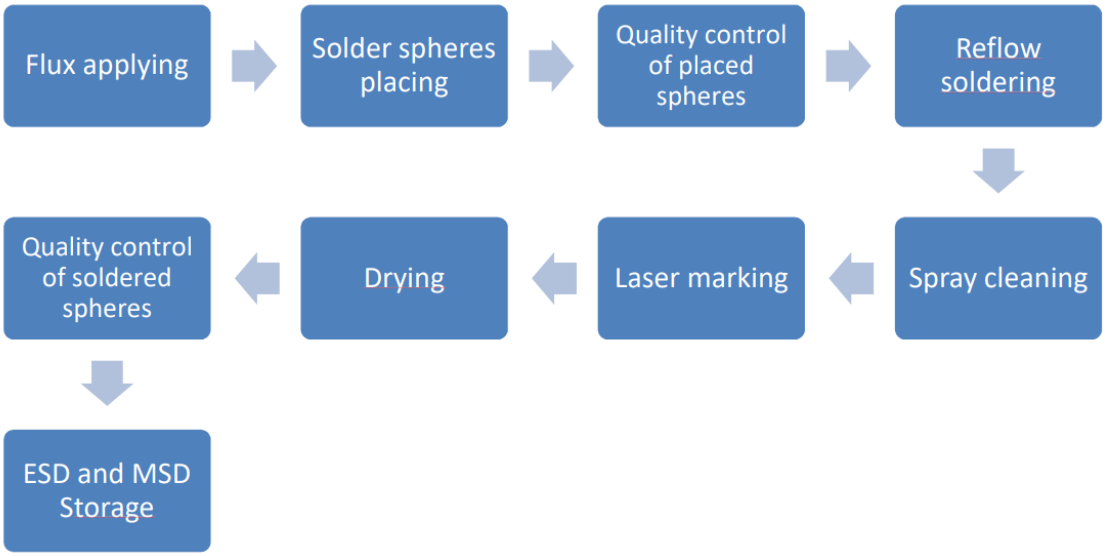


Figure 2.9: Flow chart: remanufacturing process. Reproduced from [8]

The flow chart proposed by [8], can be found at Figure 2.9. The key challenge of the process is the re-balling of the BGA devices. Three different techniques are studied and benchmarked in [42]. All of these three techniques expose the IC to one or more thermal reflow processes. Solutions without thermal reflow process were developed by industrial. For example, the SB2-jet of PacTech is a machine capable of placing solder balls without the need for one thermal process. The solder ball reflow is performed by a laser system

using an infrared wavelength. The ball placement throughput of the machine is between 6 to 8 balls per second, which means that an eMMC (embedded Multi-Media Card) flash memory can be reballed in 23 seconds [43]. A demonstration video can be found by following this link.
This machine allows outsourcing all the key challenges linked with the re-balling of BGA electronic components. We only have to complain about two main technical issues associated with the re-work of de-soldered components defined in [9] :

- Remove the remaining solder
- Prepare surfaces for optimal re-soldering

2.2.3 Refresh and testing phase

Semiconductors' components can be returned to the market provided that refreshing operation and check test verification are processed [9]. These steps are crucial to improve market confidence about second-life ICs. Therefore, a test procedure is required for each type of components. For that we need :

- An electrical interface, this need to be done through a plug and play socket. (If we resolder the IC in a test PCB we need to remake all the manufacturing processes)
- A low-level driver to make the communication with each IC.
- A Program testbench to detect faulty ICs.

Such a procedure is technically feasible for some IC and was demonstrated for flash memories [41]. Some reverse engineering might be required but this is far from impossible to develop. The Achilles heel of the test phase is that engineering a test procedure for complex IC can take time and will cost tens of thousands of dollars for each type of IC. So economic analysis needs to be done before diving into the development of testing procedures to ensure return on investment.

2.2.4 Standard interfaces to the rescue of component reuse

This section is the result of personal reflections and is the consequence of the holistic approach used during the present thesis.

On the one hand, we saw in Section 2.2.3 that the testing phase was very important because without advanced testing, it would be impossible to confirm that an IC can be returned to the market. Moreover, market demand will depend strongly on the guarantee that the testing phase can provide.

On the other hand, the development of testing procedures can be very costly and must be amortized on many components. This characteristic limits the type of ICs that can be reused. Hopefully, standard interfaces exist in the EEE (Electric and Electronic Equipment) industry. The reuse strength of ICs based on standard is that all the ICs respecting one standard are interchangeable. Hence, a system that is designed to embed one ICs respecting one standard can embed all the ICs that respect this standard, whatever the

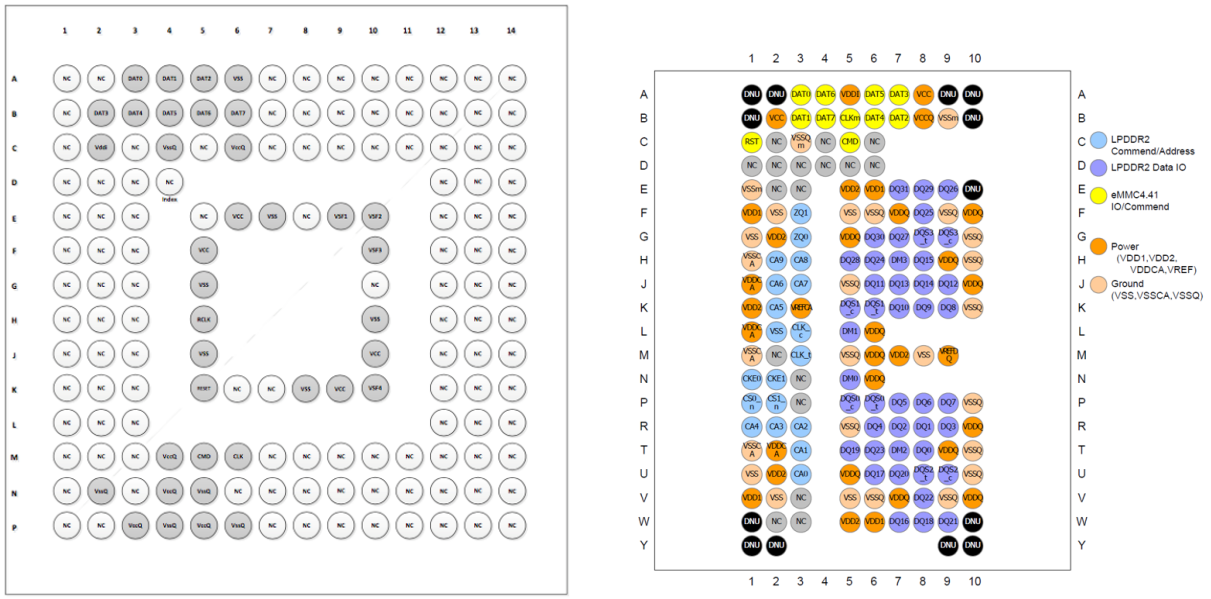
manufacturer.

Table 2.1 contains a non-exhaustive list of standards in the semiconductor industry. All these standards define the electrical interface and the communication protocol to communicate with the IC. There are developed and managed by the engineering trade association Solid State Technology Association, commonly known as JEDEC (Joint Electron Device Engineering Council).

For some of these standards, it is possible to have a finite number of different electrical interfaces defined, but the communication protocol is always the same. For example, BGA 159 pinout scheme used for eMMC is shown on Figure 2.10a. This electrical interface is the most used for eMMC and UFS (Universal Flash Storage). The differences between UFS and eMMC are the pin allocation and communication protocol. Thanks to these standardizations, sockets to connect with the IC can be bought on the market, and drivers can be found on the internet. So the testing problem is reduced to a programming problem.

Moreover, these standards are widely used in the industry. For example, the market of DRAM (Dynamic Random Access Memory) has a size of USD 94.19 Billion in 2021 and the market share of DDR4 SDRAM (Double Data Rate Fourth Generation Synchronous Dynamic Random Access Memory) and DDR3 SDRAM (Double Data Rate 3rd generation Synchronous Dynamic Random Access Memory) are respectively equal to 90%, and 8% [44]. Such trends are the same as the other standards. If you dismantle a smartphone from any producers, you have a high probability of finding an LPDDR SDRAM (Low Power Double Data Rate SDRAM) and an eMMC, eUFS, or an ONFI (Open NAND Flash Interface) in it. Sometimes RAM and storage are embedded in the same SoC (System on Chip) (Figure 2.10b).

Standards can be the gateway to the economic feasibility of second-life ICs because initial investments are amortized on a large amount of ICs.



(a) (b)

Figure 2.10: Pinout of (a) an eMMC SoC (b) LPDDR and eMMC embedded in one SoC.

Table 2.1: Standards in the memory industry

Standards	**Year**	**Introduced by**	**Complete name**	**Usage**
DDR3 SDRAM SDRAM	2007	JEDEC	Double Data Rate Synchronous Dynamic Random-Access Memory	RAM for Severs, PCs, Workstations, High-end laptops, ...
DDR4 SDRAM	2015			
DDR5 SDRAM	2020			
LPDDR3 SDRAM SDRAM	2007	JEDEC	Low Power Dynamic Double Data Rate Synchronous Random-Access Memory	RAM for Smartphones, Tablets, Notebook, Smartwatches, ...
LPDDR4 SDRAM	2014			
LPDDR5 SDRAM	2019			
eMMC 4.5	2012	JEDEC	Embedded Multi Media Card	Storage for Smartphones, Tablets, Notebook, Smartwatches, ...
eMMC 5.0	2013			
eMMC 5.1	2015			
eUFS 2.0	2012	JEDEC	Embedded Universal Flash Storage	Storage for Smartphones, Tablets, Notebook, Smartwatches, ...
eUFS 3.0	2018			
eUFS 4.0	2022			

2.3 Nand flash memories

2.3.1 Introduction

2.3.1.1 Definition

A Nand Flash memory is an EEPROM (Electrically-Erasable Programmable Read-Only Memory). This type of memory is a non-volatile memory that can be electrically erased and reprogrammed. Nand memory is the most used non-volatile memory type. We can find these memories in many devices like USB drives, solid-state drives, smartphones, or smart cars [45]. Flash memories are segmented into different types: Single-level cell, Multi-level cell, Triple-level cell,... defined by the number of bits stored in each cell, more information can be found by looking at Tab 2.2 [20]. Historically, flash cells were composed of floating gate transistors, but modern flash memories, with the improvements of 3D flash memory, incorporate charge trap transistors [46].

Flash memories have a finite number of erase/program cycles before breaking down (We call that wearing out for flash memories). As shown in Table 2.2, the number of cycles before wear out tends to decrease with the number of bits stored in each cell [12, 47].

Table 2.2: NAND flash memory comparison. Reproduced from [20]

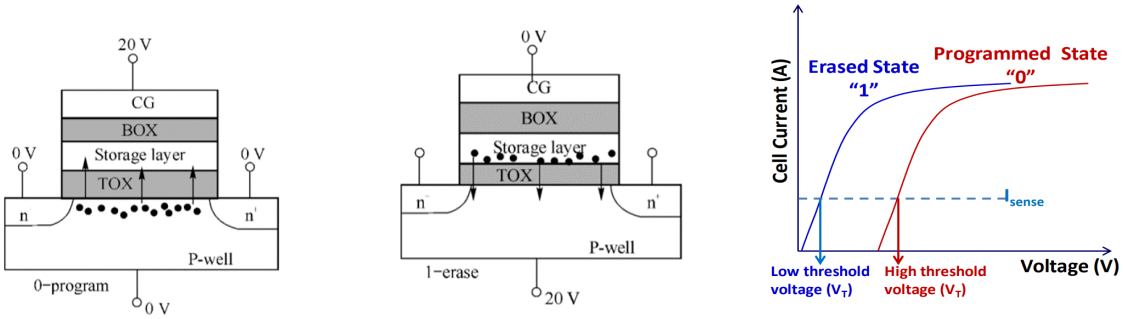
Type	**Description**	**2D Nand endurance (P/E cycles)**	**3D Nand endurance (P/E cycles)**
Single-level cell (SLC)	Stores one bit per cell and two levels of charge	50000 to 100,000	Not manufactured !!
Multi-level cell (MLC)	Stores two bits per cell and four levels of charge	3000	30000 to 35000
Triple-level cell (TLC)	Stores three bits per cell and eighth levels of charge	300 to 1000	1500 to 3000
Qradruple-level cell (QLC)	Stores four bit per cell and sixteen levels of charge	Not manufactured	150 to 1000

2.3.2 Floating gate transistor for storing bits

The most popular flash memories' cells incorporate the use of floating-gate transistors to store bits, Figure 2.11. A floating gate transistor is a MOSFET transistor built with a full isolated gate overlap with one control gate. The isolated gate is surrounded on top by one inter-poly oxide layer and at the bottom by one tunnel oxide layer. This gate is an excellent 'trap' for electrons, it allows charge retention for years [48, 15].

Two operations can be applied to a floating gate transistor: program and erase. The program operation consists of applying a high voltage on the control gate to inject electrons into the floating gate as shown in Figure 2.11 [49, 10]. This operation will increase the threshold voltage (V_{th}) and the transistor will stand in the programmed state or P1 state as shown on the current, voltage characteristics in Figure 2.11c [10]. Conversely, erase operation consists of applying null voltage to the control gate and high voltage to the base.

Electrons in the storage layer will be tunneled through the tunnel oxide layer, V_{th} will decrease, and the transistor will be in the erase state (ER). This binary state concerns only the single level cells [48, 49, 11].



(a) Floating gate transistor, program operation. [49]

(b) Floating gate transistor, erase operation [49]

(c) Output signals of comparator [11]

Figure 2.11: Comparaison of output signals at −45 dBm and −55 dBm

In presence of MLC (Multi-Level Cell) and TLC (Triple-Level Cell), V_{th} can be discriminated respectively in 4 states and 8 states as shown in the Figure2.14 [48, 10]. We understand that for retrieving the bits stored in a cell we need to locate the threshold voltage of the transistor thanks to discriminant values like V_a, V_b, or V_c for MLC for example. These voltage levels are used for the read mechanism which is explained in the section2.3.4.1. [12, 49]

Figure 2.14 helps us to understand why longevity decreases with the number of bits in each cell. The distance between consecutive voltage distributions decreases with the number of bits stored in the cell. A smaller distance between voltage levels means that a smaller voltage shift will lead to reading errors than for SLC. Section 2.3.5.1 present the impact of non-idealities and their impact on the voltage level.

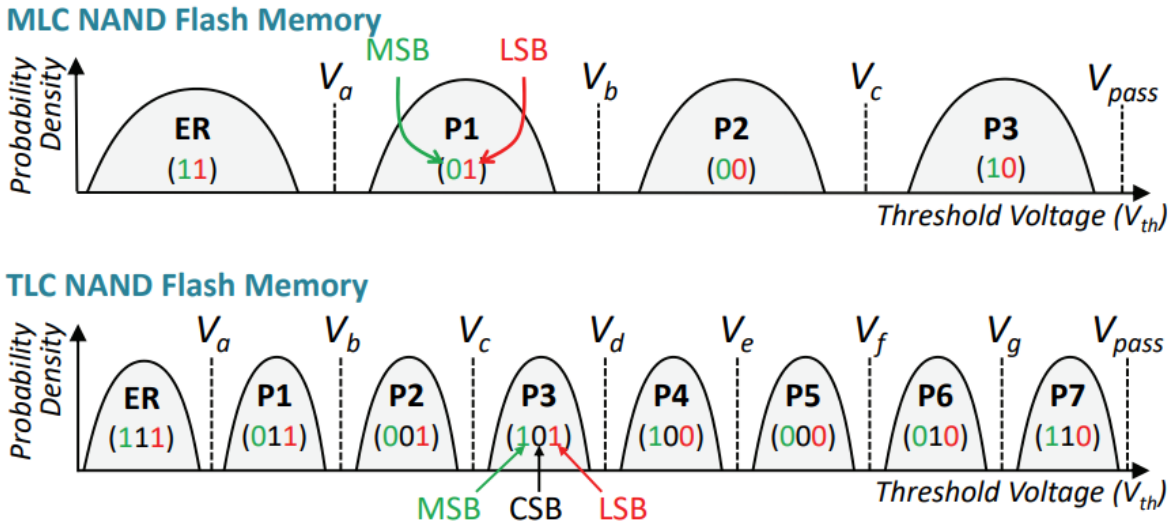


Figure 2.12: Threshold voltage distribution of MLC (top) and TLC (bottom) of NAND flash cell. Reproduced from [10]

2.3.3 Nand memory

2.3.4 Architecture

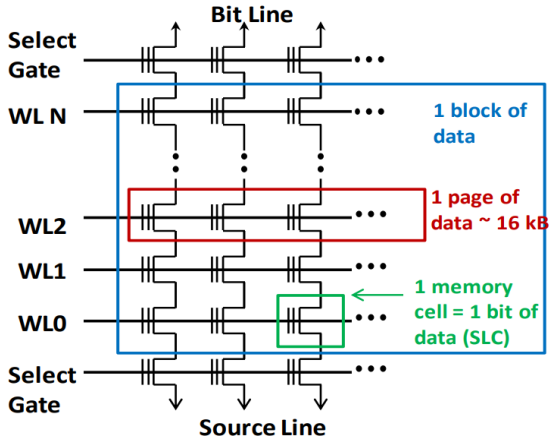


Figure 2.13: Flash structure [11]

The high-level architecture of NAND memory is shown in figure 2.13. It consists of an array of rows and columns of floating gate MOSFET cells. Multiple rows form a block (typically 128–512 rows), a block is the smallest unit that can be erased. Each row is made up of cells connected with a word line (typically 32K–64K cells). All cells connected to one word line form a page, the smallest unit we can read or write. A bit line electrically connects all cells in the same column within a block. [48, 49, 10]

2.3.4.1 Operations

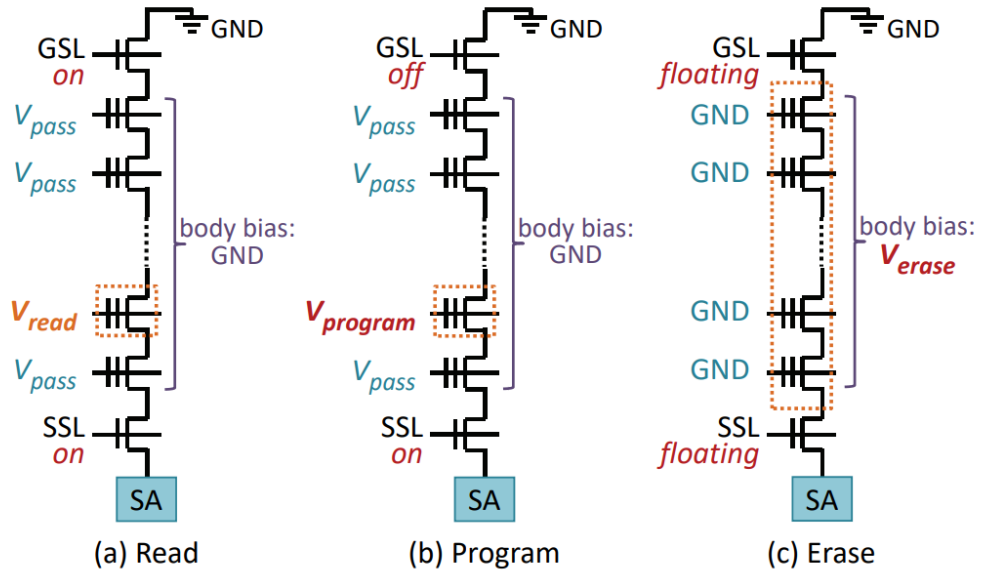


Figure 2.14: Voltages applied to flash cell transistors on a bit line to perform (a) read, (b) program, and (c) erase operations. Reproduced from [10]

Read Page can be read by applying the read voltage on the word line assigned to this page. V_{pass} is applied to the control gate of all other transistors of the block. The pass voltage allows the transistor to be turned on whatever the state of the cell. For a single gate cell, only one read reference voltage is needed because there are only two states (P_0 and ER). The passing or blocking state of the transistor is read thanks to a sense amplifier at the end of the bit line. [48, 12, 10]. For multiple-level cells, first, the voltage V_b is applied to read the LSB (Least-Significant Bit) page and after one other voltage to

read the MSB (Most Significant Bit) page. One step must be added when dealing with a three-level cell [15, 12].

Erase and program When programming a cell, electrons are ejected in the floating gate of the transistor to achieve this the FowlerNordheim (FN) tunneling effect is exploited. [10] The tunneling current density can be modeled as [10] :

$$J_{FN} = \alpha_{FN} E_{ox}^2 e^{-\beta_{FN}/E_{ox}} \tag{2.1}$$

where α_{FN} and β_{FN} are constant, and E_{ox} is the electric field in the tunnel oxide. For injecting or ejecting electrons of the floating, the tunnel oxide is subject to a high electric field generated with high voltage [49].

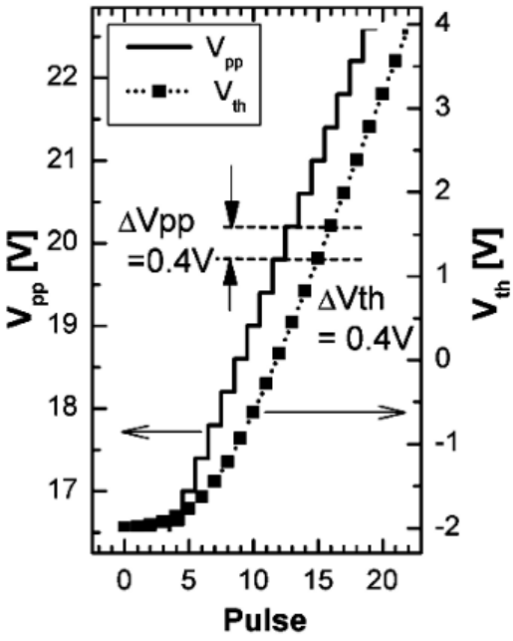


Figure 2.15: Constant V_{th} shift during incremental gate stepping, V_{pp} is the gate voltage. Reproduced from [12]

In reality, for tuning precisely V_{th} the incremental step-pulse programming procedure (ISPP) is used. At each step of the process, the voltage on the gate is incremented by a constant voltage, leading to an increase of V_{th}. After each step, V_{th} is sensed like for a read operation, and the process is stopped if the transistor is in the correct state. Figure 2.15 illustrates the ISSP process. The gate voltage is represented on the graph by V_{pp}. It exists an equivalence between ΔV_{pp} and ΔV_{th}, and this is due to the self-limitation of the tunneling current [12]. Accumulation of electrons in the floating gate leads to a constant electric field between the gate and the channel whatever the voltage applied to the gate. This equilibrium in the electric field is reached independently of any device parameter. More detail about this behavior can be found here [12, 10].
The erase operation is done on all the cells of one block simultaneously. To proceed, a high positive voltage (20V) is applied to the body. This has the effect of decreasing V_{th} since the electrons are tunneled out of the floating gate [12, 50].

2.3.5 Flash memory reliability

2.3.5.1 Charge trap generation

Electrons pass across the tunnel oxide at each P/E operation. This electron tunneling is responsible for a continuous increase in the trap density in the oxide and at interfaces [18, 50].

The presence of traps results in charge trapping/de-trapping into the tunnel oxide or unwanted charges who leave/enter the storage layer. Electrons trapping also reduces the tunneling efficiency, it is for this last reason that an iterative algorithm for a program or erase makes sense [15].

The interface and oxide trap generation have a power dependence on the number of P/E cycles. Δ_{nit} for interface trap and Δ_{not} are proportional to $cycle^m$ where m is 0.62 for interface trap and 0.3 for oxide/bulk trap [13]. This formula has been verified experimentally, the result can be found in Figure 2.16.

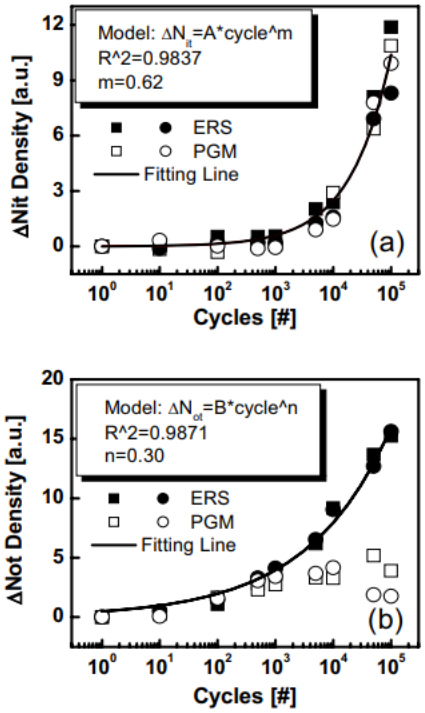


Figure 2.16: The measured and calculated trap generation as a function of program/erase cycle count. (a) Interface trap generation case; (b) bulk trap generation case. Reproduced from [13]

All unwanted charges trapped or additional charges that flow from or into the storage layer lead to unwanted voltage threshold modification. This can result in read operation errors [50, 14].

2.3.5.2 Data retention

A basic requirement for NAND cells is to keep charges in the floating gate over time. Retention loss mechanisms will affect the reliability in time of flash cells. There exist three main mechanisms that induce leakage tunneling current. Each unwanted charge entering or quitting the floating gate transistor will induce a threshold voltage shift. This same threshold voltage is used to read the data present in the cell. Therefore voltage shift can lead to reading errors. Three mechanisms will induce unwanted displacement of charges. They are illustrated in Figure 2.17.

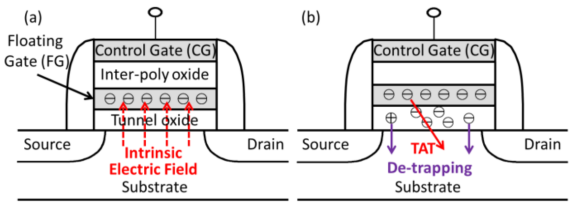


Figure 2.17: (a) Cross-sectional view of a flash cell, (b) retention loss mechanisms. .Reproduced from [14]

Stress-induced leakage current (SILC) An intrinsic electric field will be induced by the charge present in the floating gate. This will result in a stress-induced leakage current of electrons from the floating gate toward the substrate. SILC results in a decrease in the threshold voltage drops [10, 13, 50, 14].

Trap-assisted tunneling (TAT) As explained in the Endurance paragraph, more and more charges will be trapped in the tunnel oxide. These trapped charges amplify the SILC. This lead to a decrease in the threshold voltage because electron leak away from the floating gate [13, 15, 50, 14].

Charge de-trapping Positive or negative charges trapped in the tunnel oxide can also spontaneously be de-trapped. This results in an increase or decrease of the threshold voltage depending on the electrical polarization of the charge [13, 15, 50, 14].

At the end of the day, the sum of all the retention charge leakage mechanisms presented induces a negative shift of the threshold voltage distribution, see Figure 2.18. This can lead V_{th} to cross a threshold voltage state border, and the read can rise to. Indeed, if V_{th} crosses a voltage border between two states, when a read operation will be done, the read operation will output the wrong state.

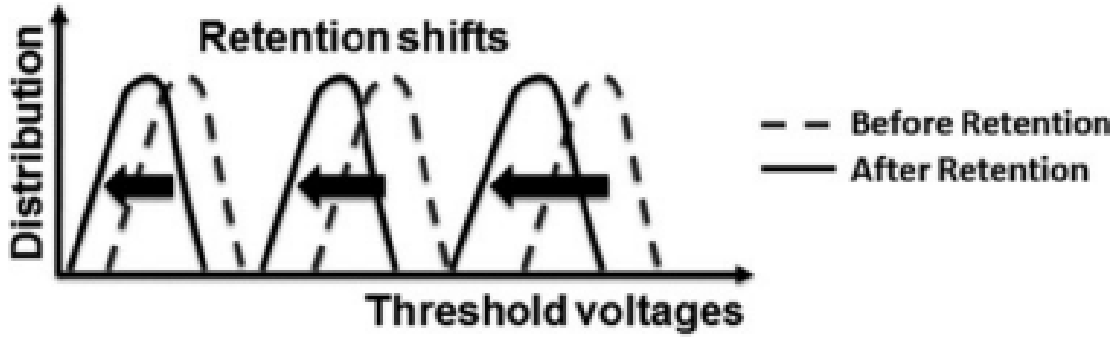


Figure 2.18: Threshold voltage shifts induced by retention. Reproduced from [15]

2.4 State of the art NAND technology

Charge-trap transistor For some years, most NAND flash memories are composed of charge-trap transistors, schematics of a CT (charge-trap) transistor can be seen in Figure 2.19. The difference with a floating gate cell is that the floating gate is replaced by an electrically insulating silicon nitride layer called the current trap. This type of transistor has the advantage to be less impacted by charge leakage. Therefore, it allows us to reduce the size of the cells and to have higher endurance. The nitride layer is surrounded by a tunnel oxide and a blocking layer. The tunnel oxide suffers from the same issue explained in section 2.3.5.2, and the wear-out mechanisms are the same as the ones for floating gate transistors [51]

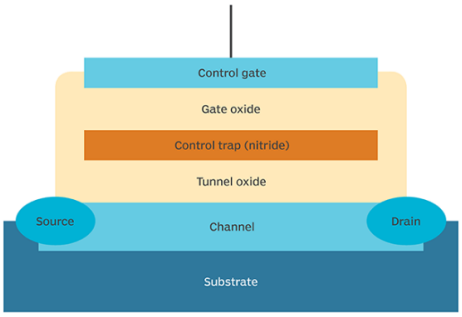


Figure 2.19: Charge trap memory cell [16]

3D integration The major reason why the industry moved from floating technology to charge trap technology is that the NAND architecture moved from 2D architecture to 3D architecture [17]. Figure 2.20, shows the evolution of the 3D architectures, it is a good example that when the size of the transistor cannot be divided by two every two years that other ideas, like 3D integration, allow continuing Moore's law.

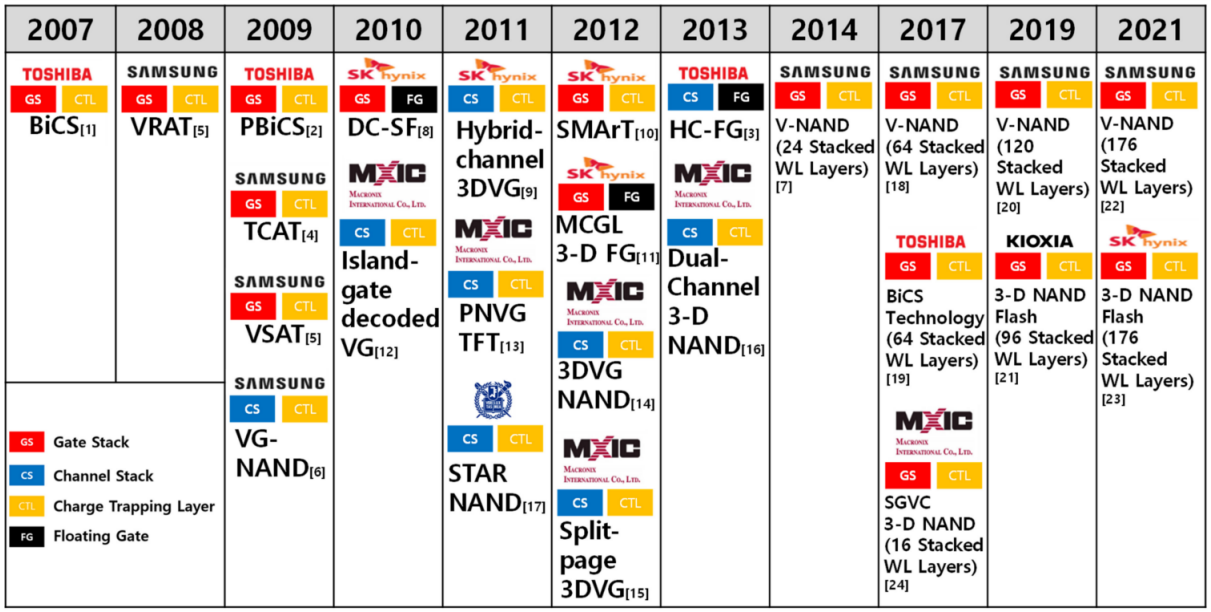


Figure 2.20: Chronological development of 3D NAND flash technologies. Reproduced from [17]

2.5 Thermal Annealing for flash memories

In Section 2.3.5.1 is discussed the different mechanisms which lead to the fact that cycling endurance is limited in flash memories. This is highly due to tunnel oxide degradation induced by the slow increase of trap density present at the interfaces or in the oxide [13, 15, 50, 14]. Interface traps recover with time and this mechanism can be accelerated by increasing the temperature [18, 19].
Temperature allows more mobility to charges trapped in the oxide or at the interface. The two main parameters for the thermal annealing process are the temperature and time needed to allow the major part of charges to be detrapped. The time of the operation

is dependent on the temperature and the relation between both is an Arrhenius law [19, 18, 13]. Calculation of the time and temperature required for one specific IC can be done and are discussed in detail in Appendix A.

2.5.1 Thermal annealing approach

When talking about thermal annealing for semiconductors, two main heater categories exist.
The in-situ heaters which are heaters directly embedded within the device or which are device functionalizing elements used as a heater, allow reaching bigger temperatures locally without damaging the device [52]. Varied techniques are proposed in the literature, [53] demonstrates that it was possible to generate heat by the current flowing between two split source contacts. [19] modifying the word line of flash memories array to generate Joule heating. In-situ heaters are not applicable in a reuse environment. Indeed in-situ heater need to be added at the design phase of the device or to use device characteristic but this requires advanced knowledge of the device.

The out-situ heaters are conventional ovens [19]. Out-situ heater thermal annealing is limited by the package thermal limit and mechanical stress induced by heat can damage the device because the device is not heated locally [52]. However, [19] show BE-SONOS (for band gap engineering silicon–oxide–nitride–oxide–silicon) charge-trapping NAND Flash recovery after 100K P/E (100,000 program/erase) 250 °C/2hr baking cycles for 9 times.

Chapter 3

Research question regarding thermal annealing for flash memories

Previous studies have demonstrated the feasibility of reusing semiconductors [8, 31, 32]. They also stated the importance of the testing phase to detect whether an IC is functional or not. The testing phase needs to be very accurate to minimize extra costs due to malfunctioning ICs used to build larger systems.

Flash memories degrade over time, and the flash memories' lifespan is limited by the number of erase/program cycles. These degradations are due to traps present in the tunnel oxide of charge trap or floating gate transistors [13, 15, 50, 14]. The presence of traps in the tunnel oxide leads to retention leakage, and a flash memory that is no more able to retain data is inoperable. Consequently, when we pick up a flash memory from an E-Waste, it is possible that the flash memory is close to his life's end. Therefore, we need, therefore, to find a way to detect or process the degradations present in the flash memory. Hopefully, it has been shown in previous research that the traps could be recovered thanks to thermal annealing [18, 19]. In this section, we aim to study how to recover the initial properties of flash memories thanks to thermal annealing in the hope to demonstrate that the process could be useful in a reuse of flash memories context.

First, the research aims to assess if thermal annealing on flash memories has the desired effect in a reuse context. The reuse context implies that the internal structure of the flash memories remains unknown. Consequently, the optimal bake temperature and bake time can not be calculated precisely. Moreover, this applies also that we have to use out-situ thermal annealing techniques. Therefore a standard laboratory oven is used.

Second, four thermal annealing cases will be studied. The four cases are the following ones ; 2h at 125 °C; 12h at 125 °C; 2h at 250 °C; 12h at 250 The metrics used for the experiments are the program and erase time of data in flash memories. Therefore, the erase and program time of 100k P/E (Program/Erase) cycles will be measured before and after each thermal annealing process.

Chapter 4

Methodology for thermal annealing

4.1 Study of the thermal annealing effects on NAND flash memories

This section presents the objectives, the experimental methodology as well as the experimental setup used to demonstrate the effectiveness of thermal annealing in a reuse context.

4.1.1 Objectives

The objective is to demonstrate the utility of using thermal annealing for extending flash memories' lifespan in a reuse context. The reuse context implies that in-situ heaters can not be used. Therefore, a conventional laboratory oven will be used to bake the ICs. So, the manipulation can easily be repeated in a semiconductors' reuse factory because the only thing needed is an oven. Moreover, data about the ICs' architecture are assumed to be unavailable. This leads to difficulties in estimating the right temperature and bake time to make a good thermal annealing process. Therefore five cases will be studied : without thermal annealing; 125°C - 2h; 125°C - 12h; 250°C - 2h and 250°C - 12h.

4.1.2 Methodology

Our measurement tools will be program and erase times over the life of flash memories. These two variables are correlated with the number of Erase/Program cycles applied on blocks and, therefore, the wear out level. The relation can basically be assumed as: Erase/Program times correlated with the number of Erase/Program cycles and, therefore, with the number of degradation present in the device.
The flash memories' used are listed in Table F.3. Macronix manufactures these flash memories with 2 and 4 GB capacity. The communication protocol is ONFI compatible, and the packaging is a 48-TSOP (Thin Small-Outline Package). They are single-level cell memories. Therefore, cells theoretically wear out after 100k Erase/Program cycles. Each block is composed of 64 pages. The page size for these flash memories equals 2048 bytes, and each page has an additional 64 bytes for error correction code. A flash memory that has been used for this master thesis is shown in Figure 4.1.

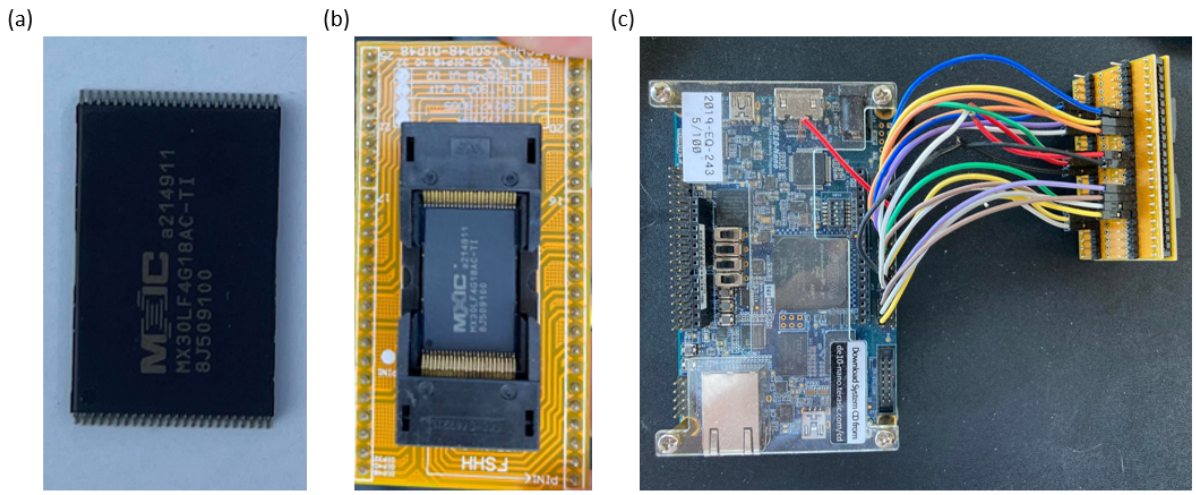


Figure 4.1: (a) A flash memory studied for this master thesis, (b) 48 TSOP to 48 DIP socket to allow to connect the flash memories with GPIO (c) DEO10 NANO board connected with the socket. This set-up has been used to wear out the flash memories and to measure program and erase times

Table 4.1: List of flash memories used

IC name	**Size [GB]**	**Page size**	**Page per block**	**Technologies**
MX30LF2G18AC	2	2048+64	64	SLC
MX30LF4G18AC	4	2048+64	64	SLC

4.1.2.1 Methodology for wearing out the flash memory

The block erase and page program times will be our measurement tools to assess the thermal annealing effects on flash memories. These two variables will be measured while Erase/Program cycles are done to reach the end of life of flash memories. Thereby, we will have measured the impact of degradations on erase and program times and we will be able to compare measurements made before and after the thermal annealing process.

The experimentation setup can be found in Figure 4.1. One DEO10-NANO FPGA (Field Programmable Gate Array) board is used to communicate with the flash memory. The communication protocol is the ONFI (Open Nand Flash Interface) one, and the four operations programmed are the read, erase, program, and reset operations presented in Appendix B.1.

The FPGA board is programmed so that a finite number of Erase/Program cycles are done to wear out the flash memories, typically 100k to simulate an entire SLC lifespan.

Program and erase times are measured at each iteration and saved into a file. The measure time scale resolution is 40ns. This limitation is due to the fact that the time measure is based on a clock with a frequency of 25MHz. For timing reasons, erase operations are made on specific blocks, and only the first page of these blocks is programmed. Hence,

flash memories are worn out locally and not on the full memory, which would take a few days to achieve.

In parallel to these operations, the content written on the page is verified every thousand cycles to detect failed blocks.

4.1.2.2 Data Post-processing

Due to the fact that erase and program times behave like random variables, the mean over 16 blocks is computed to mitigate block-to-block variations.

After that, for more readable data, the average mean is calculated with a window of size 100.

4.1.2.3 Baking Methodology

Flash memories are baked at a specific temperature and for a specific duration defined by the experiment case. Four cases are taken into account : 2h at 125 ;; 12h at 125 °C; 2h at 250 °C; 12h at 250 Wear out operations on the ICs are done before and after each passage in the oven to measure the effect of temperature and bake time on flash memories.

4.1.2.4 Effects of wear out on the program and erase operation speed

In this section, an initial experiment is presented to analyze wear out effects on the program and erase operations' speed.

Figure 4.2a shows the page program time and the block erase time in function of the number of Erase/Program cycles. Operations' delays are stochastic variables that take value around discrete levels. This is explained by the fact that the internal controller of flash memories completes both operations with an iterative algorithm that repeats Erase/Read cycles for completing an erase operation or Program/Read cycles for completing a read operation.

The oxide degrades as Erase/Progam cycles are repeated. This results in a modification in the number of iterations needed to complete Program and Erase operations. Program operations become easier with oxide degradation, and erase operations become harder. This leads to our observations that operations' times tend to decrease for program and increase for erase.

Figure 4.2b shows page program times, and the block erase times in function of the number of Erase/Program cycles when taking into account the mean over 16 blocks. Curves on Figure 4.2c are the moving average of curves present in Figure 4.2b with a window width of 100. We remark that the data post-processing allows for having more readable curves.

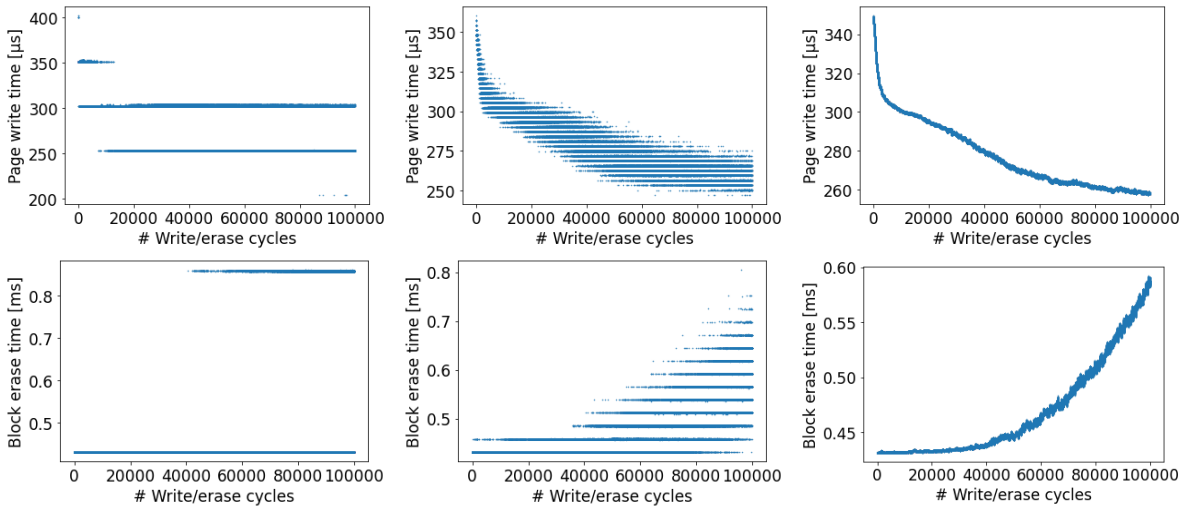


Figure 4.2: Page program and block erase times as a function of the number of Erase/Program cycles for (a) A unique block. (b) A mean over 16 unique blocks. (c) A moving average of the mean over 16 unique blocks with a convolution window width of 100

Chapter 5

Thermal annealing results

5.1 Thermal annealing

In this section, we study the thermal annealing effects on the page program and block erase times. The goal is to study the wear out effects in the hope of using a thermal annealing process when reusing flash memories.

To do this, measures on the erase time and program time will be taken while repeating 100K program and erase cycles (100k P/E cycles) before and after baking the ICs. A primary experiment will show the behavior of both parameters when two 100k P/E cycles are repeated without thermal annealing between them.

5.1.1 100,000 program/erase cycles repetition without thermal annealing

In this section we present the results of two repetitions of 100k P/E (Program/Erase) cycles without thermal annealing between them. These results will serve as comparison for the following section where thermal annealing effects on flash memories are studied.

5.1.1.1 Description

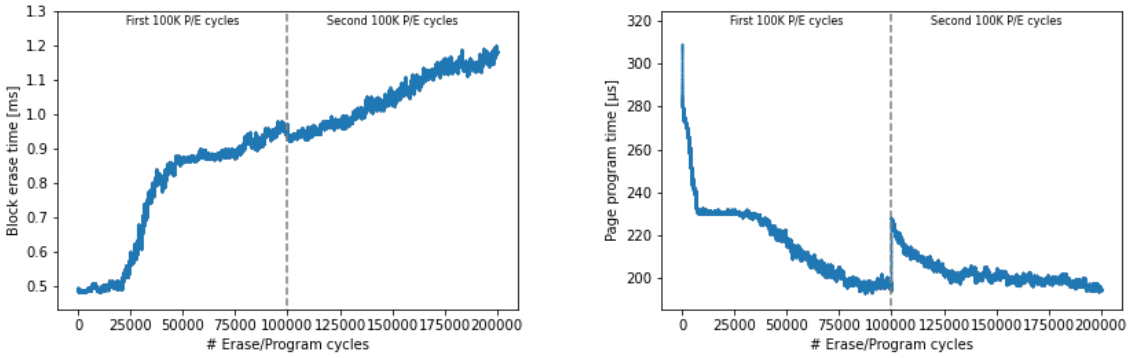
Figure 5.1 shows the block erase and page program times in function of the number of P/E cycles. First 100k P/E cycles were done on a flash memory. After the flash memory has been put to rest for 24 hours and finally 100k P/E cycles were done again. Plot of the measures can be found on Figure 5.1.

5.1.1.2 Observations

Figure 5.1 shows that without thermal annealing between two 100k P/E cycles, the behavior of block erase time is more continuous at the interface of both cycles than the page program time. Indeed we remark in Figure 5.1b that the page program time at cycle 0 for the second 100k P/E cycle is approximately equals to the cycle 40k of the first 100k P/E cycle.

5.1.1.3 Discussions

Erase time seems more correlated than program time to the wear out level. Indeed, the unexpected improvements in the program time can not be explained only by an oxide or interface traps density reduction. Therefore we can state that another mechanism interferes with the program time. Unfortunately, this mechanism has not been identified and may require further investigations to be identified. Without certainty, this can be, for example, explained by the fact that close-in-time P/E repetitions can modify the program time.



(a) (b)

Figure 5.1: (a) Erase time and (b) Program time for two 100k P/E consecutive cycles without bake process between them

5.1.2 100,000 program/erase cycles repetition with thermal annealing

In this section, we study the effects of thermal annealing on block erase time and page program time as a function of the temperature and the bake time.

5.1.2.1 Description

Figures 5.2 and 5.3 respectively show the thermal annealing impact on the block erase time and the page program time for four thermal annealing cases.

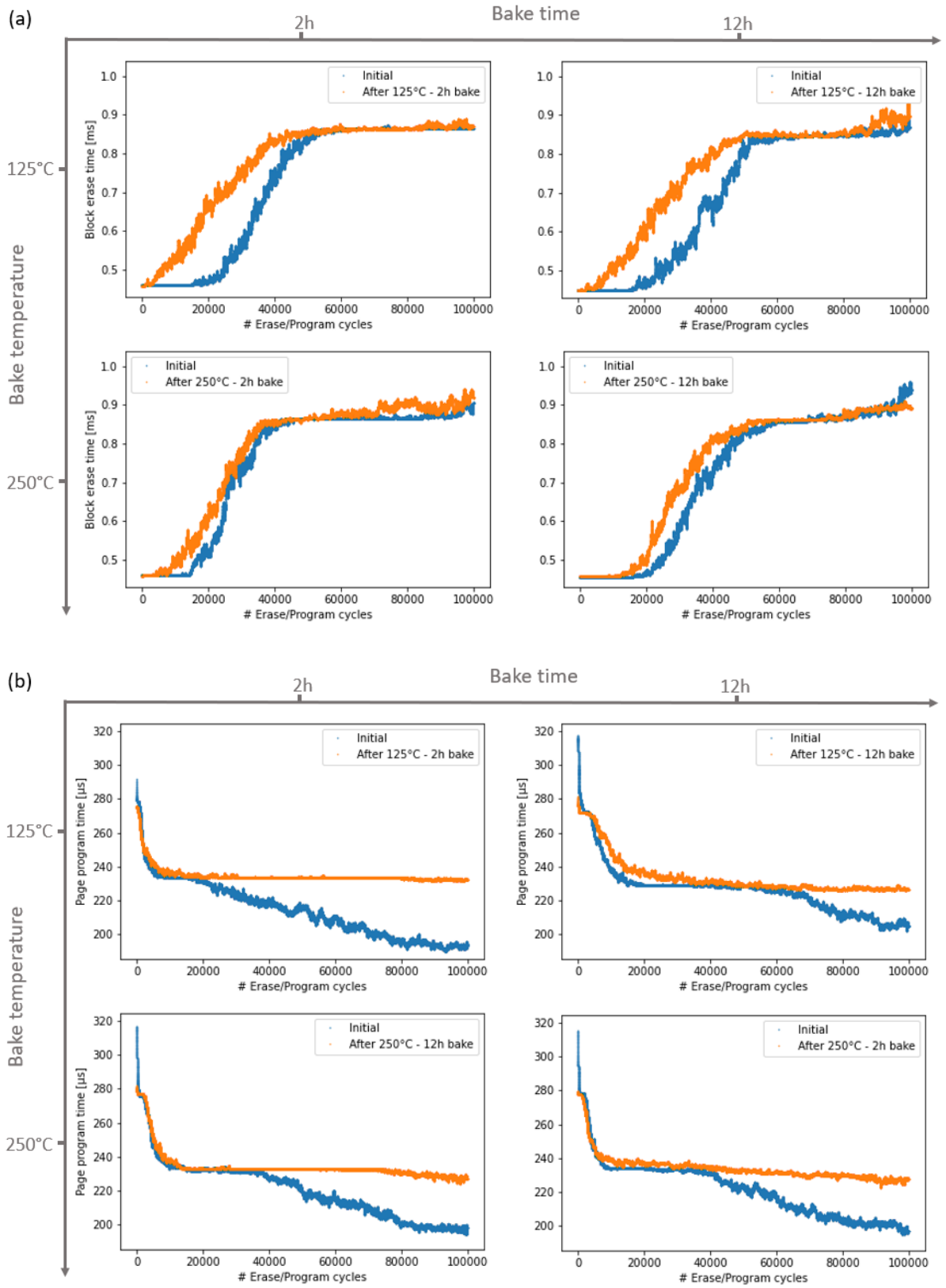


Figure 5.2: Thermal annealing results for (a) Block erase time, (b) Page program time. How to read the graph : the blue curve is the measure of the first 100k P/E process. After that, the flash memory is baked and a second 100k P/E process is done and measured. This is the orange curve.

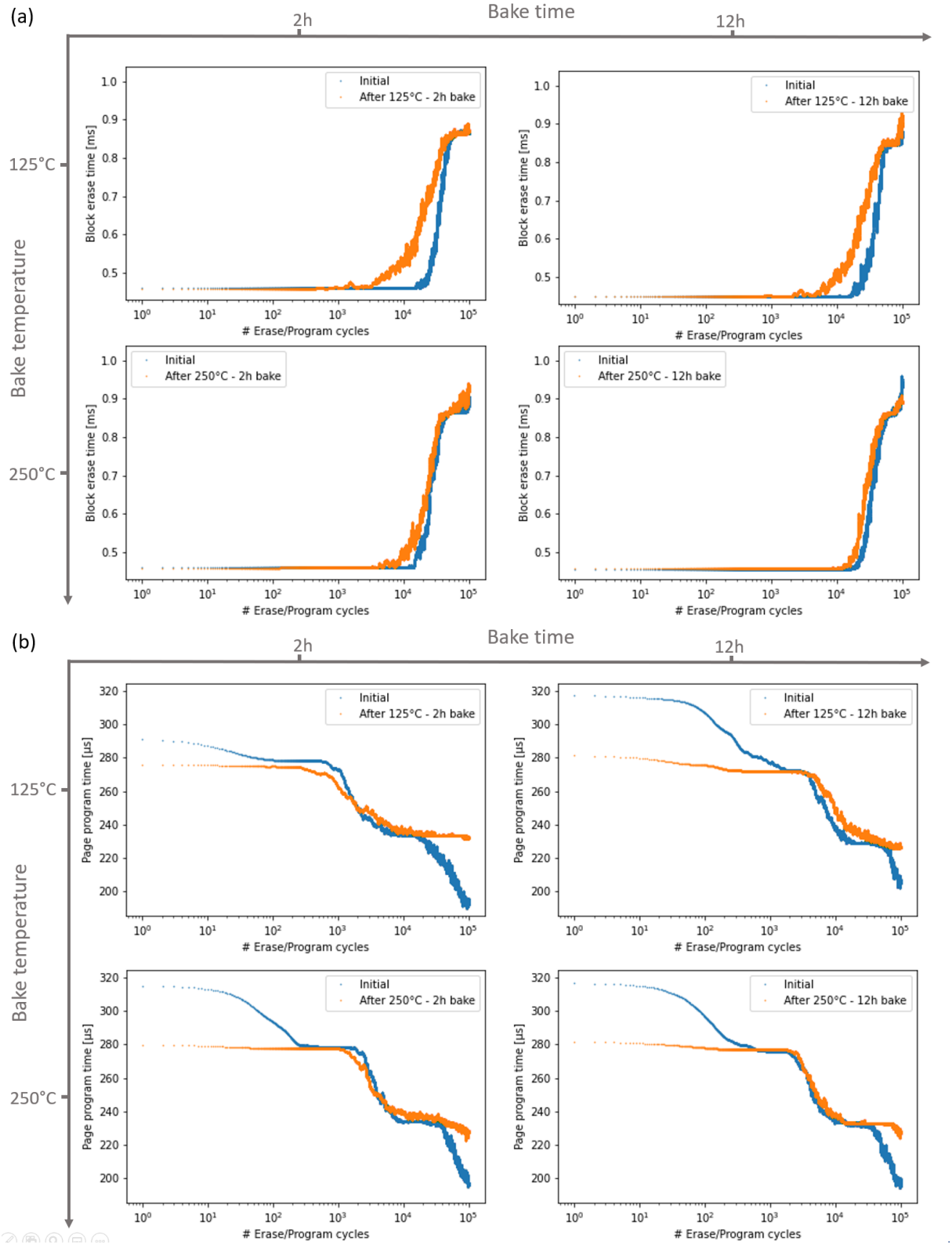


Figure 5.3: Thermal annealing results for (a) Block erase time, (b) Page program time. How to read the graph : the blue curve is the measure of the first 100K P/E process. After that, the flash memory is baked and a second 100k P/E process is done and measured. This is the orange curve. (Logarithmic scale for x-axis)

5.1.2.2 Observations

Erase First of all, compared to Figure 5.1 where no thermal annealing was applied to flash memories, we can remark that the thermal annealing has an effect on the second 100k P/E cycles regarding block erase time. This effect occurs even for the worst case i.e. 125 °C - 2h. Moreover, the curves after baking are closer (more similar) to the initial ones when the temperature and bake time are higher. We also remark that after 55k cycles, no big differentiation can be done between cases. Indeed, the erase time is between 0.8 and 1ms for all the second 100K P/E cycle cases. This was not the case for the example without thermal annealing, where the value at the same moment was between 1 to 1.2ms. Here again, ameliorations are present, whatever the thermal annealing temperature and bake time. Appendix C contains other complementary Figures that confirm our observations.

Program For page program measures, it is easier to make observations regarding Figure 5.3 where a logarithmic scale for the x-axis is used. As for erase operations, we remark a big difference in the second cycle curves between with and without thermal annealing. As for erase, page program time for the second 100k P/E cycles is more similar to the initial one with a bigger bake temperature and a longer bake time. We notice that for each second cycle below the 1000 iteration, the page program time is close to 280 µs. However, the correlation between program time, temperature, and bake time is less pronounced than for erase measures. Indeed, the curves are more chaotic in terms of distance with the initial curves than for erase cases, and no relation with the temperature or bake time can be extracted. Appendix D contains other complementary Figures that confirm our observations.

5.1.2.3 Discussions

A relation exists between the thermal annealing process temperature and the block erase time. This is less relevant for the page program time. It can be due to the fact that program time is correlated with other mechanisms as explained in Section 5.1.1.

If erase time is confirmed to be a good representation of the traps' density in the oxide, we could conclude that the properties of flash memories are recovered at a high rate and this rate increases as a function of the temperature and bake time. This has to be confirmed by further studies.

However, data are not statistically representative, given that only one sample was used per case. Results are therefore not immune to device-to-device variations.

To conclude, further studies need to be done to tune the optimum bake time and bake temperature. Also, the retention rate after thermal annealing was not studied. Indeed, it is essential to study the impact of thermal annealing on data retention of flash memories. As a reminder, the primary goal of flash memories is to conserve data for as long as possible. Therefore, retention time is crucial. Moreover, the market needs certitude regarding the lifespan of re-used flash memories, and this will be done by developing a thermal annealing process whose effects will be known and quantified in a representative way.

Chapter 6

Life cycle assessment of new and reused flash memories

The previous chapter showed us that further investigations need to be done regarding the use of thermal annealing applied to flash memories. Indeed, a thermal annealing process coupled with an elaborate testing phase can guarantee the quality of second-life flash memories. Whereas the reuse of ICs is possible, it is time to verify how it can reduce the semiconductor industry's carbon footprint.

Hence, the second goal of this master thesis is to assess the carbon footprint of second-life and new flash memories.

First, this chapter will present the scope of the LCA for new and second-hand flash memories. Then, the result of new flash memories manufacturing GWP (Global Warming Potential) will be presented.

Second, we will present the GWP of second-hand flash memories remanufacturing.

Third, a comparison between both flash memory types global warming potential will be done by taking into account all the life cycles of new and reused flash memories.

Finally, the macroscopic carbon footprint reduction that second-hand flash memories could achieve is presented.

6.1 Scope

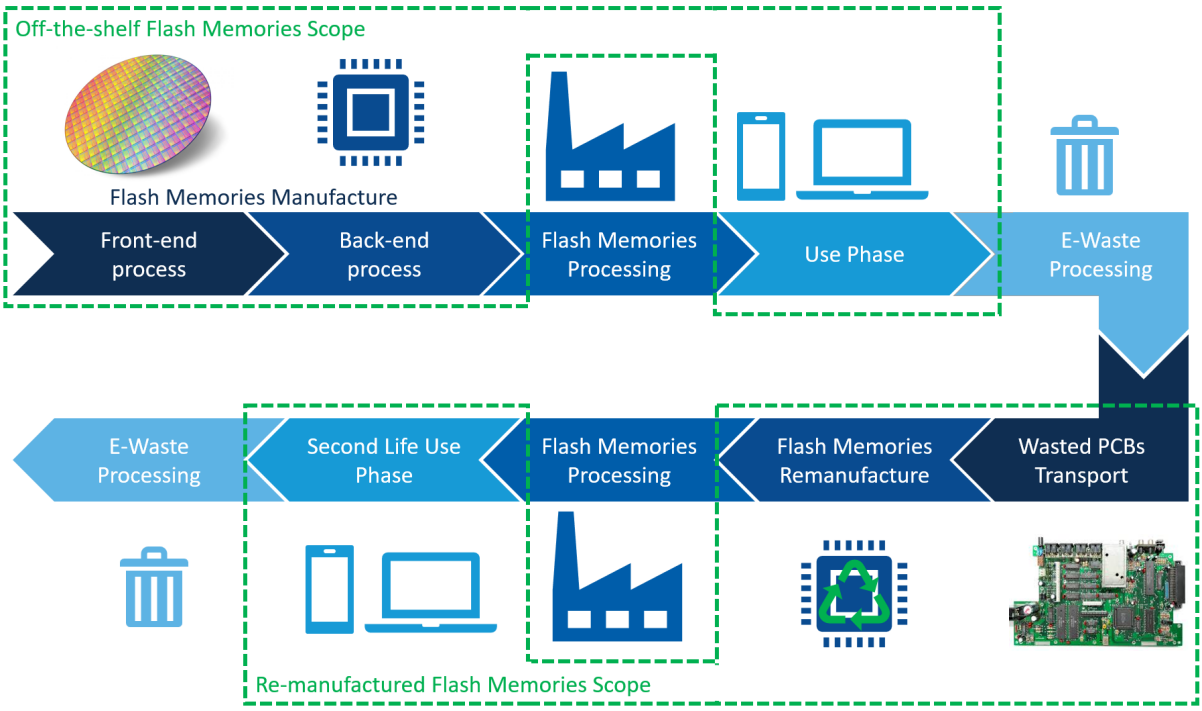


Figure 6.1: Scope of the life cycle assessment of new flash memories and remanufactured flash memories

The scope of the life cycle assessment is defined in Figure 6.1 and was designed to compare new and old flash memories' global warming potential. Hence, the LCA is a mono-parameter one where only the GWP (global warming potential) is considered.

6.1.1 New flash memory scope

The scope of off-the-shelf flash memory LCA includes the flash memories manufacturing (Front-end and back-end processes) and the use phase of the new flash memories. Note that for the flash memories manufacturing, three methods based on three different sources were used and are explained in detail in Appendix F.1.

6.1.2 Remanufactured Flash Memory Scope

The scope of remanufactured flash memories' LCA includes transport from the wasted PCB supplier, the flash memorie' remanufacturing process, and the use phase of second-life flash memories. Hypothesis and calculation are discussed in detail in Appendix F.0.1.

The remanufacturing process and the machines used for each process step can be found in Figure 6.2. In the aim to compute the remanufacturing process GWP, the raw materials input, the energy input, and the manufacture of the machines were taken into account. A complete explanation of the process and the machines' parameters used are presented in Appendix E.

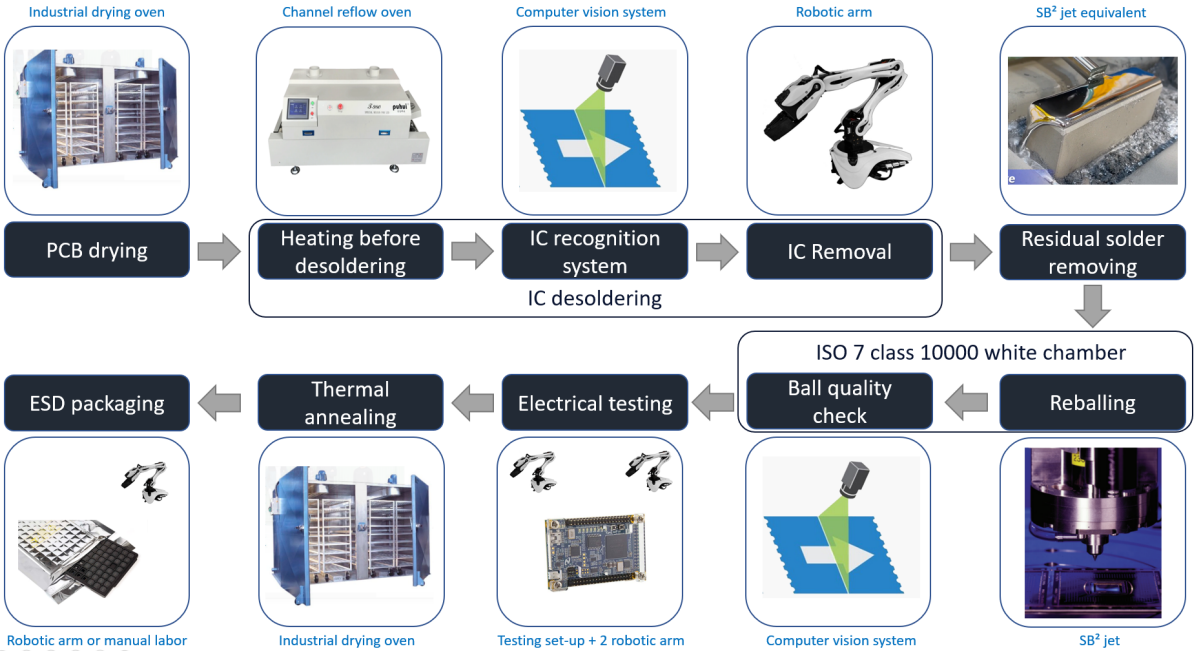


Figure 6.2: ICs' remanufacturing process with the selected machines for each step

6.1.3 Use phase scenarios

The carbon footprint of reused flash memories must be assessed by taking into account the use phase. New flash memories consume orders of magnitude less power, therefore the emissions gap present between manufacturing and remanufacturing processes could be annihilated during the use phase.

We know that the GWP of the use phase is linearly correlated with the energy consumption. We need therefore to compute the energy consumption of flash memories on their entire use phase. Two calculation methods are presented and discussed.

First, the energy consumed is deduced from the amount of GB written in the use phase. It is relevant to deduce the energy from the amount of GB written but the number of GB read can not be neglected because this operation also consumes energy. Therefore, they are accounted for as a number of GB read per GB written. For example, if 5 GB are read for each GB written, the total energy consumed per GB written is one times the energy of one GB written and five times the energy of one GB read. This gives for eMMC4.5 a total energy per GB written of 48 J (see Table 6.1).

This approach assumes that the ICs are embedded in a system designed to take advantage of the best flash memory parameters and optimized to use as little energy as possible. This is utopic, and the flash memories will never be used in such conditions. Therefore a second method is required.

The second approach is the same as used in [54]. It means using flash memories for 6 hours per day. The active power taken into account is the biggest one between read and write active power for second life flash memories and the smallest one for new flash memories.

6.1.4 Flash memories' parameters

Flash memories used are listed on Table 6.1. Flash memories references for eMMC4.5 and eMMC5.1 were found while dismantling the Samsung Galaxy S3 4G edition and the Samsung Galaxy A10. Note that the eMMC5.1 will be considered as a second-hand IC and a new IC

The parameters for the UFS4.0 flash memory were found from values found on the Samsung semiconductor website. This last flash memory was chosen because it is the last generation of flash memories from Samsung and, therefore, the one with the best features. The power consumption is not known, but Samsung declares reaching 6.0MBps per 1 of sequential read speed [55]. The voltage level of the IC remains unknown. Therefore, one hypothesis must be posed to complete the $P = V \cdot I$ relation. Flash memories have two voltage levels, one for the controller and one for the flash cells array. For UFS 3.0 of Samsung, these voltages were respectivelly 1.2 V, and 2.6 V [56]. To take the worst case into account, it is assumed that the voltage level for UFS4.0 is equal to 1.2 V.

Table 6.1: Parameters of flash memories used for the life cycle assessment

Flash memory	**Manufacturer**	**Used in**	**Type**	**Capacity [GB]**	**Write/ read Speed [MB/s]**	**Active power read/write [mW]**	**Write/Read energy per GB [J/GB]**
KLMxGxGEAC	Samsung	Samsung Galaxy S3 4G edition, 2013	eMMC4.5	64	50/150	1029/1029	18/6
KLMCG2UCTA	Samsung	Samsung galaxy A10, 2019	eMMC5.1	64	200/300	984/984	4.3/2.6
\\	Samsung	State-of-the-art, 2022	UFS4.0	64	2800/4200	560/840	0.17/0.17

6.1.5 What is not included in the LCA

The transport of flash memories after the manufacturing or remanufacturing process is not included because it is assumed to be equal for both cases.

The processing of the flash memories is not accounted for because it is assumed to be the same process for both types of flash memories (for example, creating USB drives will not have a bigger GWP for second-life flash memories than for new ones).

The e-waste processing is also assumed to be equal for both types of flash memories. Therefore, this part is not included in the LCA.

6.2 Assessments of the carbon footprint of the manufacturing and remanufacturing process of flash memories

This section is dedicated to the study of the GHG emissions from the flash memories manufacturing and remanufacturing process. The aim is to use these results to calculate the carbon footprint of new and reused flash memories when accounting for the complete life cycle.

First, the GHG emissions of flash memories fabrication will be given and discussed.

Then, the GWP of remanufacturing process of flash memories will be presented.

6.2.1 Production of new flash memories

In this subsection, the GHG emissions of the flash memories manufacturing are discussed. Three sources were used to output the numbers and are presented in Appendix F.1.

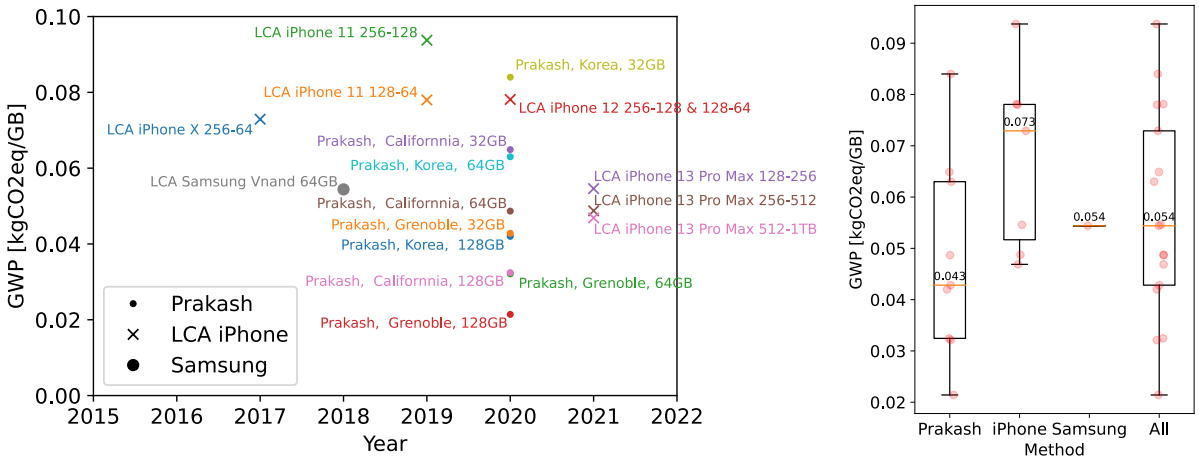
6.2.1.1 Description

In Figure 6.3 are plotted the GWP per GB of flash memories manufacturing. These GWP emissions come from the three different methods defined in F.1 and are plotted on the graph as a function of the publication year of the data. Note that the first method, inspired by the database made by Prakash at al. in 2013 and dye area information published by Fraunhofer in 2021, was arbitrarily plotted in 2020.

For method 1, the annotations mention the region in which the energy mix was used and the flash memory size that is taken into account. For example, 'Prakash, Korea, 64GB' mean that the energy mix used is the one of Korea and that the dye area used is the one for 64GB, i.e., 300 mm^2. All these parameters can be retrieved at Table F.2.

For method 2, the annotations mention the iPhone model and the two capacities studied. For example, 'LCA iPhone 11 256-128' means that the model is the iPhone 11 and that the difference in the GWP emission was made with the 256GB and 128GB models.

For method 3, the data is a raw value, so no supplementary information is given.



(a) (b)

Figure 6.3: (a) GWP of the flash memories manufacturing process per GB. Data were calculated thanks to 3 methods listed in Appendix F.1. Note that for the Prakash method data were arbitrarily placed in the year 2020. (b) Boxplot of the data classed per method

6.2.1.2 Observations

The graphs show the fact that the flash memories manufacturing GWP per GB is concentrated between 0.1 and 0.02 kgCO2eq/GB. The Prakash calculation method has the lowest median GWP, i.e., 0.043 kgCO2eq/GB while the iPhone method has the biggest median emission, i.e., 0.073 kgCO2eq/GB. The median values of Samsung LCA and when all the data are reunited are equal to 0.054 kgCO2eq/GB. The LCA of the iPhone is 1.5 to 3 times higher than the Prakash California ones when the capacities are similar, i.e., when the difference of iPhone capacities equals the one used for Prakash California.

6.2.1.3 Discussions

The low values of the Prakash method can be explained by the fact that the calculations of the manufacturing emissions include only the energy use and PFC emissions emitted during back-end and front-end processes. This choice tends to have results lower than the true values. Moreover, the chosen electricity mix emissions are significant and will influence significantly flash memories manufacturing emissions. Indeed, the lowest value is induced by the fact that the region of Grenoble has a GWP per kWh three to five-time lower than the two other locations. The case of Grenoble was taken as an example because a new state-of-the-art semiconductor factory will be built there in the following years. However, at the moment, the production capacities have more chance to be located in California or Korea than in France. The Korean and Californian cases lead to emissions per GB which are closer than other method results.

By comparison, another study done in 2019 that used the Prakash database had a GWP per GB equals to 0.1 kgCO2eq/GB [57]. This tends to confirm our observation that our application of the Prakash database underestimates the GWP of the manufacturing process.

The high value of the iPhone method is difficult to explain due to the lack of transparency

in the LCA made by Apple.

Also, Samsung did not state the methodology to obtain the results in the Samsung case. Therefore these results are difficult to analyze.

For the rest of the LCA part, the value from the Prakash, California method will be used for different capacities and locations. This choice allows us to include the lowest emissions factor without using the one of Grenoble, which seems unrealistic regarding other studies. Indeed, our research aims to prove that reused flash memories have a lower environmental impact than new flash memories. Therefore taking into account the lowest realistic values for GWP of flash memory manufacturing will lead to better confidence in the result. Moreover, the Prakash method allows us to be more case-specific while dealing with different capacities.

In summary, the GWP that will be used for the manufacturing of flash memories in the rest of this chapter will be 2.048 kgCO2eq for 32GB, 3.072 kgCO2eq for 64GB and 3.84 kgCO2eq for 128GB.

6.2.2 Remanufactured flash memories

This subsection discusses the GHG emissions emitted by the remanufacturing process of flash memories. The remanufacturing process and the hypothesis made for the GWP calculation can be found in Appendices E and F.0.1.

The results will be discussed as a function of the yield[1] and the travel distance of wasted PCBs. As a reminder, the PCBs are provided thanks to wasted PCB suppliers, and the ICs can only be tested at the end of the process. Therefore, the GWP of a remanufactured IC is proportional to $1/yield$ and depends on the distance between the supplier facility and the remanufacturing IC factory.

6.2.2.1 Description

Figure 6.4 shows the GWP of the remanufacturing process per entering IC for the different process steps. The total GWP is highlighted for the following cases, without PCB transport and when the PCB provider is located in Mechelen, Ankara, and Shanghai. As a reminder, it is assumed that each entering IC goes through all the process steps and is tested at the end.

Figure 6.5 shows the GWP of the remanufacturing process per IC as a function of the yield for the three different cases with the transportation included. These results are given thanks to the following formula :

$$GWP_{IC} = \frac{GWP_{entering}}{\alpha} \tag{6.1}$$

Where GWP_{IC}, $GWP_{entering}$ and α are the GWP per IC, the GWP per IC entering the process and the yield.

[1]Yield is the percentage of ICs that pass the testing phase and therefore can be reused

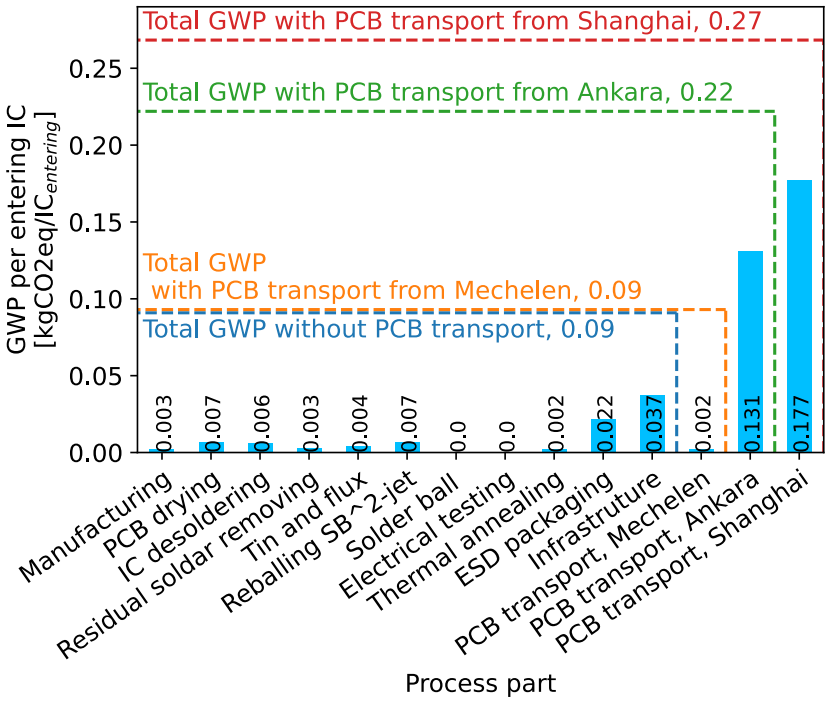


Figure 6.4: Flash memories remanufacturing process GWP for the different process steps and PCB transport GWP as a function of three different geographical locations of the wasted PCBs' suppliers. The GWP emissions are expressed in KgCO2eq/IC$_{entering}$, which means that no yield is considered. Total GWP emissions are highlighted for four cases.

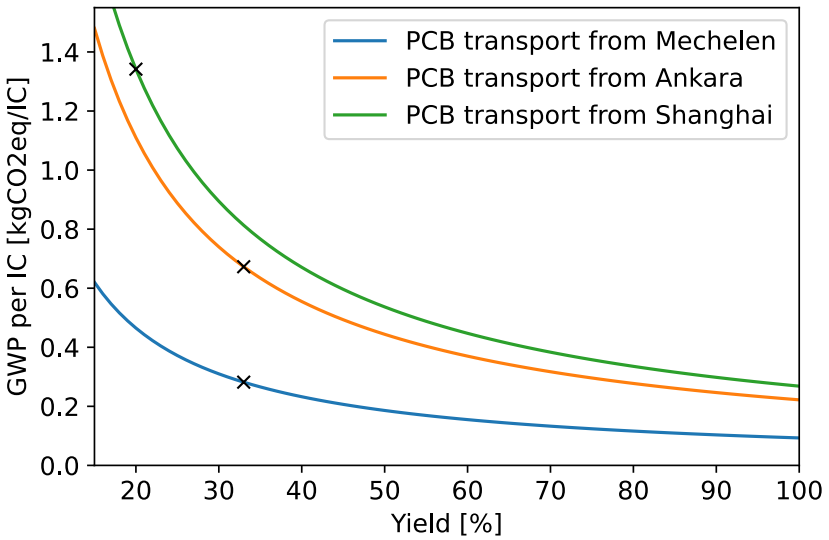


Figure 6.5: Remanufacturing process GWP as a function of the yield and the 3 wasted PCB suppliers' locations. As a reminder, the yield is the percentage of ICs that pass the testing phase and can be reused. The testing phase takes place only at the end of the process, therefore the GWP is proportional to $1/yield$

6.2.2.2 Observations

Figure 6.4 shows that the remanufacturing process emissions for each IC entering the process are equal to 0.09 kgCO2eq when the transport is not accounted for. When the transport is accounted for the remanufacturing emissions for each IC entering the process are equal to 0.09 kgCO2eq for the Mechelen case, 0.22 kgCO2eq for the Ankara case, and 0.27 kgCO2eq.

The energy used in-house per entering IC equals 0.14 kWh. This result is based on machines' energy consumption and is therefore accurate. Details of machines' energy consumption can be found in Appendix E.

6.2.2.3 Discussions

First, **the share of wasted PCBs' transport in the GWP can be insignificant but also very important depending on the cases.**. Indeed, for the worst case presented, the PCBs' transport was responsible for more than 70% of the GWP. Hence, it is very important that the flash memories are remanufactured within a radius of 500-1000km.

The GWP share of the infrastructure is not negligible and can be reduced by 68% by using a more energy-efficient factory, i.e., to shift from a D-class building to an A-class building (see hypothesis).

The low power consumption per IC (0.14 kWh), combined with the fact that raw materials are scarce, lead to the fact that the energy and material inputs are responsible for only 0.031kgCO2eq (0.0073kgCO2eq for raw materials and 0.0237kgCO2eq for electricity). This low impact is also influenced by the energy mix used, i.e., the one of Belgium with 0.169 kgCO2eq/kWh is relatively 'green'. In South Africa[2], for example, the energy use will be responsible for 0.102 kgCO2eq of GHG emissions per entering IC. Hence, **the energy mix used is significant**.

The ESD (Electrostatic Sensitive Discharge) packaging GWP is difficult to reduce because the packaging needs to be ESD and is limited to special plastic packaging.

Figure 6.5 shows that the GWP of remanufactured IC evolves in 1/yield. Having a too-bad yield will lead to emitting way more greenhouse gases. **The three crosses on the graph are the three cases that will be used for the next section, where the GWP of second life and new IC will be compared.** Their denomination will be :

- Mechelen case : Yield of 33%, GWP = 0.28 kgCO2eq
- Ankara case : Yield of 33%, GWP = 0.67 kgCO2eq
- Shanghai case : Yield of 20%, GWP = 1.34 kgCO2eq

[2]GHG emissions of South Africa per kWh of electricity = 0.732 kgCO2eq

6.2.2.4 Conclusion and how to use the results

The results presented are the only ones available in the literature. An adaptive method is therefore proposed to be able to derive results according to the main assumptions.

The most significant part of the work was to evaluate the electrical energy used per IC entering the process, i.e., 0.14 kWh, and the emissions of raw materials input, i.e., 0.0073 kgCO2eq. All other parameters are highly adaptable and depend on the assumptions posed. The data can be adapted to your need by using the following equation :

$$GWP_{IC} = \frac{W_{elec} \cdot GWP_{elec} + GWP_{raw} + GWP_{ESD} + GWP_{infra} + GWP_{transportation}}{\alpha} \tag{6.2}$$

Where the parameters are defined in Table 6.2. Note that the energy was calculated for a BGA 159 IC. Using other BGA type will not influence the result much because only the energy used for the re-balling part will change proportionally with the number of balls, but this will remain insignificant regarding the other source of emissions.

Table 6.2: Parameters used in equation 6.2.

Parameter	**Definition**	**Value**	**Unit**
GWP_{IC}	GWP of the flash memory remanufacturing process	Adaptable	kgCO2eq
W_{elec}	Quantity of electrical energy used for the remanufacturing process	0.14	kWh/IC
GWP_{elec}	GWP of the electricy mix used per kWh	Adaptable	kgCO2eq/(kWh IC)
GWP_{raw}	GWP of the flux and tin used	0.0073	kgCO2eq/IC
GWP_{ESD}	GWP of the ESD packaging	Adaptable	kgCO2eq
GWP_{infra}	GWP of infrastructure	Adaptable	kgCO2eq
$GWP_{transportation}$	GWP of the transport of wasted PCB from the supplier to the remanufacture factory	Adaptable	kgCO2eq
α	Yield of the process	Adaptable	%

6.3 LCA comparison between new and second life flash memories

In this section, we study the total GWP of new and reused flash memories as a function of different use phase scenarios. Four flash memories will be benchmarked, and their parameters can be found in Table 6.1.

First, an LCA will be done depending on the number of GB written and read.

Then, an LCA will be presented depending on the number of use days of flash memories. For this, the IC is assumed to be in the active state for 6 hours each day.

Finally, an adaptable method will be presented because both methods have not been able to show the limit cases where the reuse of semiconductors would be less attractive. This method allows every engineer to calculate for a specific case whether the IC reuse is beneficial or not regarding our mono-parameter LCA. The most valuable results of this chapter will be presented at the end of the section where the adaptive method will be used.

6.3.1 Life cycle assessment as a function of the number of gigabytes read and written

This section presents results related to the total life cycle GWP as a function of the number of GB written and read in the use phase for new and reused flash memories.

6.3.1.1 Description

Figures 6.6c show the total life cycle GWP as a function of the number of GB written in the use phase for new and reused flash memories. However, the GB read cannot be neglected. Therefore, the energy of read operations is accounted for by assuming a certain number of GB read per GB written.

The total life cycle GWP is studied for two remanufactured ICs and two new flash memories. IC parameters can found in Table F.3.

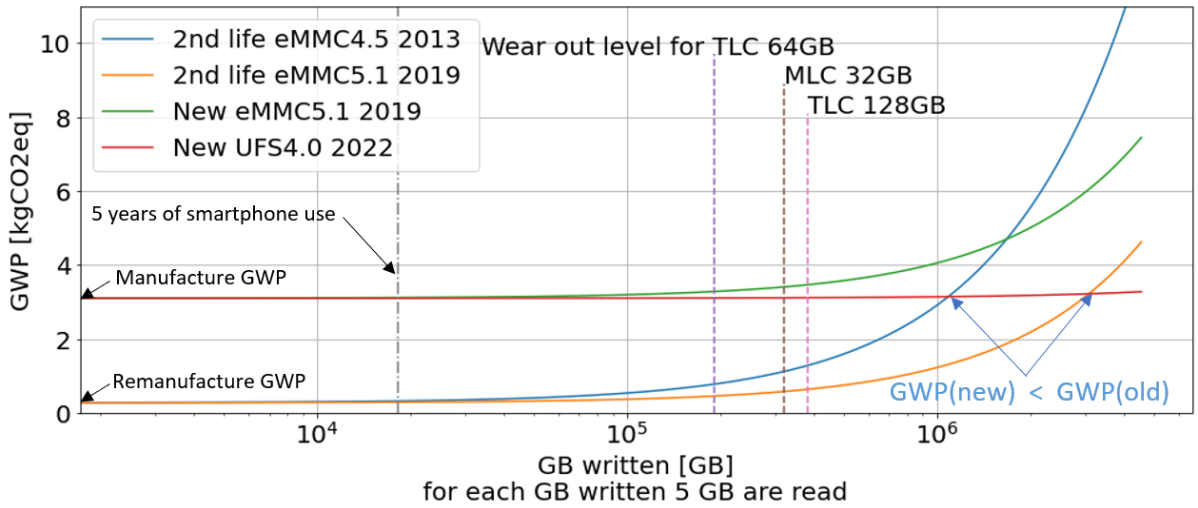
To ease the understanding, markers have been placed on the graphs. The first one (the dashed-point line) is the number of GB written after five years of smartphone usage[3]. The other ones (the dashed lines) are the wear-out level for TLC 64GB, MLC 32GB, and TLC 128 GB.

Two extreme cases are presented, the first where the parameters are favourable to second-hand ICs and the second where the parameters are favourable to new ICs.

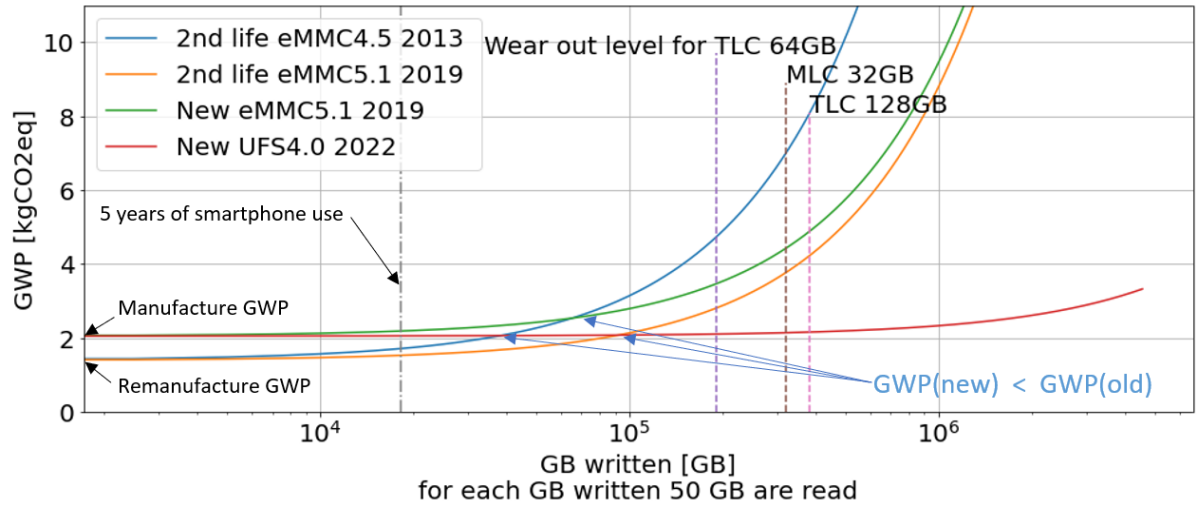
[3]The daily number of data written in a flash memory embedded in a smartphone is equal to 10 GB [58].

For the first case, in Figure 6.6a it was assumed that the remanufacturing process GWP was equal to 0.27 kgCO2eq (Mechelen case with a yield of 33%, see Figure 6.5) and that for each GB written, 5GB are read. Moreover, the manufacturing GWP is assumed to be equal to 3.84 kgCO2eq (Prakash, California, 128 GB case)

For the second case, in Figure 6.6b it was assumed that the remanufacturing process GWP was equal to 1.32 kgCO2eq (Shanghai case with a yield of 20%, see Figure 6.5) and that for each GB written, 50GB is read. Moreover, the manufacturing GWP is assumed to be equal to 2.048 kgCO2eq (Prakash, California, 32 GB case)



(a) 5 GB read per GB written, the remanufacturing process GWP = 0.27 kgCO2eq/IC and manufacturing process GWP = 3.84 kgCO2eq/IC



(b) 5 GB read per GB written, the remanufacturing process GWP = 1.34 kgCO2eq/IC and manufacturing process GWP = 2.048 kgCO2eq/IC

(c) Total life cycle GWP of four flash memories (listed in Table 6.1) as a function of the number of data written in the use phase. Two use phase scenarios were studied. Scenarios differ from each other with the amount of GB read per GB written, and the manufacturing and remanufacturing process GWP (defined respectively in Section 6.2.1.3 and Figure 6.5). Markers have been placed to highlight the lifespan of different flash memories and the number of GB written in a smartphone use case (5 years of use). How to read the figures: The two lowest values at the origin are the second-hand ICs, and the biggest ones are the new ICs. We can state that before curves' intersections, the second-ICs are more eco-friendly than the new ICs. After curves' intersections, this is the opposite.

6.3.1.2 Discussions

Looking at the GB written method in Figure 6.6c, it comes out that the GWP for remanufactured flash memories, even though the oldest ones, will never exceed emissions of new ones for the case where the flash memories are embedded in smartphones. As a matter of fact, second-life flash memories' curves stay almost flat before the 18250 written GB threshold (5 years of smartphone usage). Before this threshold, the GHG emissions are

induced mainly by the remanufacturing or manufacturing phase.

Figure 6.6a shows us that in some cases, the use of second-life flash memories can reduce the GWP up to 75-80% on the complete life cycle of flash memories.

Figure 6.6b shows that use cases exist where the use of second-life flash memories is more polluting than the use of new ones. Indeed, the remanufactured curves indeed cross the curves of new ICs before reaching the end of life. It means that intensive use cases exist where the use of second-life flash memories is not beneficial regarding our mono-parameter LCA.

The GB written approach has imperfections. By using this method, it is assumed that the engineers using flash memories in their products have designed their system with the aim to optimize the flash memory power consumption. For example, to write 18250 GB and to read 91250 GB respectively take, respectively 2h18 and 6h01 to complete for UFS4.0. This seems to be a really low usage time in the total life cycle. **Therefore, only accounting for the energy per GB written and read declared in datasheets is an unrealistic hypothesis.**

6.3.2 Life cycle assessment as a function of the number of use days (Adaptative method)

We saw in the previous section that the method of GB written was based on inaccurate assumptions. Therefore, another method needs to be presented.

This section presents results related to the total life cycle GWP as a function of the number of days used for new and reused flash memories. The method used is adaptative so that everyone can use the data to assess the environmental impact of one specific use case.

6.3.2.1 Explanations

Figure 6.7 shows the GWP of new and second-life flash memories as a function of the number of use days. We assume that each chip is in the active state for 6h per day and that the flash memories manufacturing process GWP is equal to 3.072 kgCO2eq (Praskash, California for 64 GB see Figure 6.3). Moreover, we assume the flash memories remanufacturing process GWP to be equal to 0.27 kgCO2eq/IC (Mechelen case with a yield of 33% , see Figure 6.4).

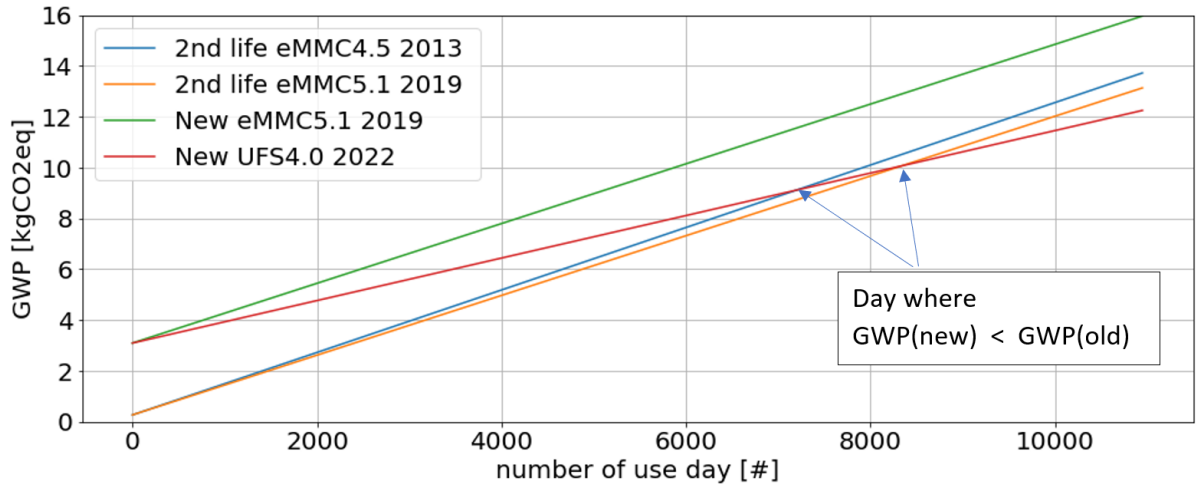


Figure 6.7: Total life cycle GWP of four 64GB flash memories as a function of the number of use days. Two of them are new ones and two of them are second-life flash memories. Their power consumptions are listed in F.3. We assume that the flash memories manufacturing GWP is equal to 3.072 kgCO2eq (Praskash, California for 64 GB see Figure 6.3). We also assume that the flash memories remanufacturing GWP is equal to 0.27 kgCO2eq/IC (Mechelen case with a yield of 33% , see Figure 6.4). The ICs are assumed to be used for 6h per day in an active state.

6.3.2.2 Generalization

By observing Figure 6.7, we note that new and second-hand flash memories' curves intersect around the 7000-9000 days. This intersection is the moment where the GWP of new flash memories becomes smaller than the GWP of second-life flash memories. The results obtained have been largely influenced by hypotheses on the embedded carbon footprints (the carbon footprint at day 0), and the electricity mix used (the one from Belgium). Therefore, the results are not representative of all the possible cases.

Consequently, we propose to generalize the approach and to present a formula that will output as a function of the parameters inputted the day D where the curves will cross. Engineers will thus be able to know that if their product is designed to last longer than D days then it might not be worth to use a reused IC from the point of view of our mono-parameter LCA. The formula that gives D is the following one :

$$D = \frac{(GWP_{new} - GWP_{old}) \cdot 1000}{(P_{old} - P_{new}) \cdot GWP_{elec,use} \cdot h} \tag{6.3}$$

Where parameter definitions can be found in Table 6.4.

6.3.2.3 One example for the adaptative method

This subsection presents one example where the Formula 6.3 is used.

The example is shown in Table 6.3. The day D where GWP for new flash memories becomes smaller than GWP on second-hand flash is calculated as a function of :

- Three use phase locations (consequently, three different electricity mix: France 0.119 kgCO2eq, Germany 0.46 kgCO2eq and South Africa 0.732 kgCO2eq)
- Three remanufacturing GWP. The three cases of Figure 6.5 (Wasted PCB suppliers in Shanghai - yield of 20 % : 1.32 kgCO2eq/IC, Wasted PCB suppliers in Ankara - yield of 33 % : 0.67 kgCO2eq/IC, Wasted PCB suppliers in Mechelen - yield of 33 % : 0.28 kgCO2eq/IC)
- Three flash memories size: 32 GB, 64GB, 128GB with manufacturing GWP emissions respectively equal to 2.04, 3.072, 3.84 kgCO2eq/IC (Prakash, California method, see Figure 6.3).
- Three hours per day. Hours per day vary between 4h, 8h, 12h. It is worth to mention that, on average, people spend 4h48 a day on their phones. [59]

New and old IC consumptions are the UFS 4.0 and eMMC4.5 one, i.e., 0.57 W and 1.049 W.

6.3.2.4 Discussions related to the adaptative approach

Results of Table 6.3, show situations where remanufactured ICs have a significant advantage over new ones from an environmental point of view. However, for others, the trend is inverted. In the worst case, the carbon emissions reduction during the production phase is wiped out in less than one year. Therefore, the calculation of day D needs to be done before designing a product with second-hand ICs in it.

Table 6.3: Results of equation 6.3 as a function of three use phase locations (France, Germany, and South Africa), of three capacities (32GB, 64GB, 128GB), of three remanufacturing GWP (Shanghai case, Ankara case, and Mechelen case, see Figure 6.4) and of three number of hours use per day (4h, 8h, 12h). It is assumed that new and old flash memories' power consumptions are respectively 0.56 W (UFS4.0) and 1.049 W (eMMC4.5). As a reminder, the equation 6.3 outputs the day from which the total life cycle GWP of new flash memories is smaller than the one of second-life ICs as a function of the five parameters. It means that if the result is 1000 for your set of parameters and if you plan to embed flash memories in products designed to last more than 1000 days. Then, it is worth to use new flash memories regarding our mono-parameter LCA.

Use phase location : France

Use hour per day :	**4h**			**8h**			**12h**		
Flash memory size :	**32GB**	**64GB**	**128GB**	**32GB**	**64GB**	**128GB**	**32GB**	**64GB**	**128GB**
WPCBs from Shangai	3042	7441	10740	1521	3721	5370	1014	2480	3580
WPCBs from Ankara	5920	1039	13619	2960	5160	6809	1973	3440	4540
WPCBs from Mechelen	7596	11995	15294	3798	6809	7647	2532	3998	5098

Use phase location : Germany

Use hour per day :	**4h**			**8h**			**12h**		
Flash memory size :	**32GB**	**64GB**	**128GB**	**32GB**	**64GB**	**128GB**	**32GB**	**64GB**	**128GB**
WPCBs from Shangai	787	1925	2779	393	962	1389	262	642	926
WPCBs from Ankara	1532	2670	3523	766	1335	1762	511	890	1174
WPCBs from Mechelen	1965	3103	3957	982	1552	1978	655	1034	1319

Use phase location : South Africa

Use hour per day :	**4h**			**8h**			**12h**		
Flash memory size :	32GB	64GB	128GB	32GB	64GB	128GB	32GB	64GB	128GB
WPCBs from Shangai	494	1210	1746	247	605	873	165	403	582
WPCBs from Ankara	962	1678	2214	481	839	1107	321	559	738
WPCBs from Mechelen	1235	1950	2486	617	975	1243	412	650	829

Number of year :	<1	1-2	2-3	3-4	4-5	5<

Table 6.4: Parameters used in equation 6.3

Parameter	**Definition**	**Value**	**Unit**
D	Number of use day after which $GWP_{LCA,new} < GWP_{LCA,old}$	Result	#
GWP_{new}	GWP emission of new IC manufacture	Adaptable see Figure 6.3	KG CO2-eq/IC
GWP_{old}	GWP emission of second life IC remanufacture	Adaptable see Figure 6.4	KG CO2-eq/IC
$GWP_{elec,use}$	GWP emission of the electricity mix used during the use phase	Adaptable see Electricity Map	KG CO2-eq/Kwh
P_{old}	Power of old second life IC	Adaptable See Table F.3	W
P_{new}	Power of new IC	Adaptable see Table F.3	W
h	Number of active state hour per day	Adaptable	#

6.3.3 Macroscopic carbon footprint reduction due to second-hand flash memories use

This section presents two macroscopic study cases to assess the carbon footprint reduction that might be expected at the European level. The first study is about flash memories embedded in smartphones, and the second is about flash memories embedded in SSDs.

The results were generated thanks to other results presented in the LCA section of this master thesis.

6.3.3.1 European smartphones case

This subsection presents the GHG emissions reduction that we might expect at the European level by reusing, in new smartphones, flash memories embedded in EoL (end-of-life) smartphones.

In 2020, 185.9M smartphones were sold in Europe [60]. Each of these smartphones embeds a flash memory, and these smartphones will reach their end-of-life around 2023. If at this

moment, 75% of the flash memories embedded in them can be reused in new smartphones. The European GHG emissions could be reduced by 185-435 KTCO2eq[4]. It is equivalent to the yearly emissions of 30k-71k Europeans.

6.3.3.2 European SSDs case

This Subsection presents the GHG emissions reduction that we might expect at the European level by reusing, in new SSDs, flash memories embedded in EoL (end-of-life) SSDs.

In 2020, 1195 Exabytes of SSDs were sold at the European level[5]. Each of these SSDs embeds one or more flash memories. The end-of-life of these SSDs will reach approximately in 2025. If at this moment, 75% of the flash memories embedded in them can be reused in new SSDs. The European GHG emissions could be reduced by 405-943 KTCO2eq[6]. It is equivalent to the yearly emissions of 66k-154k Europeans.

[4]185.9M smartphones were sold in 2020 in Europe [60]. GWP per kWh in Europe = 0.45kgCO2eq/kWh. Bad case, 32GB flash memories, good case 128GB flash memories. Hour use per day 4h48 in average [61]. Smartphone average lifespan = 2.5 year [62], yield = 75%

[5]20.739 Exabytes of SSDs were sold in 2020 worldwide. Europeans represent 5% of the worldwide population. Therefore European sales of SSDs were assumed to be five percent of the worldwide sales

[6]GWP per kWh in Europe = 0.45 kgCO2eq/kWh. Flash memories are assumed to be 64GB or 128GB. Use hour per day is assumed to be 10 hours, and yield is assumed to be equal to 75%. Lifespan assumed to be five years

Chapter 7

Financial study of the reuse of flash memories

The two previous chapters showed that it was possible to use thermal annealing in a reuse context to recover the initial properties of flash memories and that the flash memories reuse could reduce the European emissions of 0.6-1.4 MTCO2eq. This is equivalent to the yearly emissions of 97k-226k Europeans.

The ecological benefits are significant but will never take place if the reuse of ICs is an economic disaster.

Hence, the third goal of this master thesis is to study the financial feasibility of IC remanufacturing.

First, the price of ICs available on the market will be assessed. This stage is crucial because it is utopian to believe that the price of second-hand ICs will be higher than the price of new ones. The price of new flash memories would rather be a high limit for the price of remanufactured flash memories.

Second, the cost of the remanufacturing process will be investigated, and its most significant parameters will be discussed.

Finally, we will study the profit that can be generated as a function of the sale price and process optimization. Again, it is essential to remember that initial investments, engineering, and MSDA costs are not accounted for.

The scope, the hypotheses, and the parameters used for the cost calculation of the remanufacturing process are gathered in Appendix F.3.

7.1 Results: new flash memory prices

In this section, we will present the market prices of new flash memories. Knowing their prices is crucial because it will highly influence the prices of reused flash memories. Indeed, we might expect that the selling prices of second-hand flash memories will be below those of new ones.

Hence, the prices of eMMC5.1 flash memories were studied. This standard was created in 2015 and is always used today. Therefore, eMMC5.1 can be recovered from old PCBs and is always used in new ones. The prices were found on market analysis platforms (platforms that monitor the price of something) and on marketplaces (Websites where we can buy ICs). A complete explanation of the methodology used is detailed in Appendix F.2.2.

7.1.1 Description

Figure 7.1 shows the prices for eMMC5.1 coming from market analysis platforms and marketplaces as a function of the memory type and the memory capacity. Three sources are compared, two market analysis platforms ELINFOR and DRAMeXchange, and marketplaces such as RS, Digikey, Arrow, Octopart, and Mousser. The smallest prices found on these marketplaces were retained.

Figure 7.1: eMMC5.1 minimum prices from three different sources, eMMC5.1 spot prices from DRAMeXchange, eMMC5.1 spot prices from ELINFOR, and eMMC5.1 minimum prices on marketplaces (RS, Digikey, Arrow, Mousser, ...). Data for DRAMeXchange were not available for 64GB MLC and 16GB TLC.

7.1.2 Discussions

The prices on marketplaces tend to be higher than the ones on exchange platforms, especially for TLC flash memories. Two facts might explain this. First, marketplace quantities were under 3000 ICs. This quantity is far from what we might expect when smartphone manufacturers deal directly with flash memories manufacturers. There can be a wholesale discount in such deals. Second, marketplaces take a commission for each IC sold, which can also be one of the reasons why prices are higher.

For the rest of the financial study, the prices retained are 2.5€for 32GB TLC, 5.2€ for 64GB TLC and 15€ for 128GB TLC.

7.2 Cost of the flash memories remanufacturing process

In this section, the cost of reusing flash memories is studied. The main cost sources will be identified and discussed to reduce the process costs.

The costs were calculated thanks to the flash memories remanufacturing process detailed in Appendix E. The methodology and hypotheses are discussed in Appendix F.4.

The results of this part will be used for the economic profitability part. Results will highlight the main cost sources related to the reuse of ICs.

As a reminder, it is important to differentiate the two cost units. The first cost unit is the price per entering IC. As a reminder, each entering IC is assumed to be remanufactured and tested at the end of the process. Therefore, even if the IC does not pass the testing phase, it will cost one remanufacturing process for one IC. The second cost unit is the cost per IC. This is the cost per salable IC[1]. The relation between both costs is :

$$C_{IC} = \frac{C_{entering}}{\alpha} \tag{7.1}$$

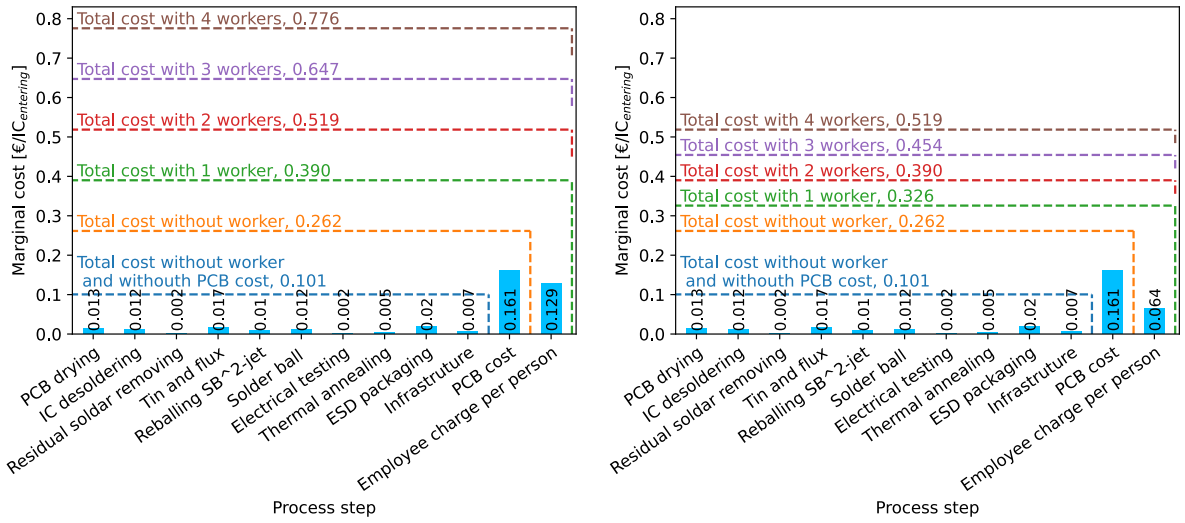
where C_{IC} is the cost per IC, $C_{entering}$ is the cost per entering IC and α is the yield, i.e., the percentage of ICs that pass the test.

7.2.1 Description

Figure 7.2 shows the process step, infrastructure, and human resources costs per entering IC. Two cases are studied: First, it is assumed in (a) a daily output of 1440 ICs. This occurs when the warehouse operates only 8 hours a day when the workers are present. Second, in (b), a daily output of 2840 ICs is assumed. This occurs when we add to a typical day the fact that the slowest machine, the SB^2, Jet, works autonomously for 8 hours at night. Note that initial investment, engineering, and MSDA costs are not considered.

Figure 7.3 shows remanufacturing process costs per salable IC as a function of the yield and the autonomous work that the slowest machine (the SB^2 jet) might do outside the work hours.

[1]Salable ICs are ICs that pass the testing phase and are functional.



(a) (b)

Figure 7.2: Remanufacturing costs per entering IC with a yield equal to 1 for the different process steps, infrastructure, and human resources costs. It is assumed in (a) a daily output of 1440 ICs (when the warehouse operates for 8 hours a day, i.e., only when the workers are present) and in (b) a daily output of 2840 ICs (when we add the fact that the slowest machine, the SB^2, Jet works autonomously for 8 hours at night).

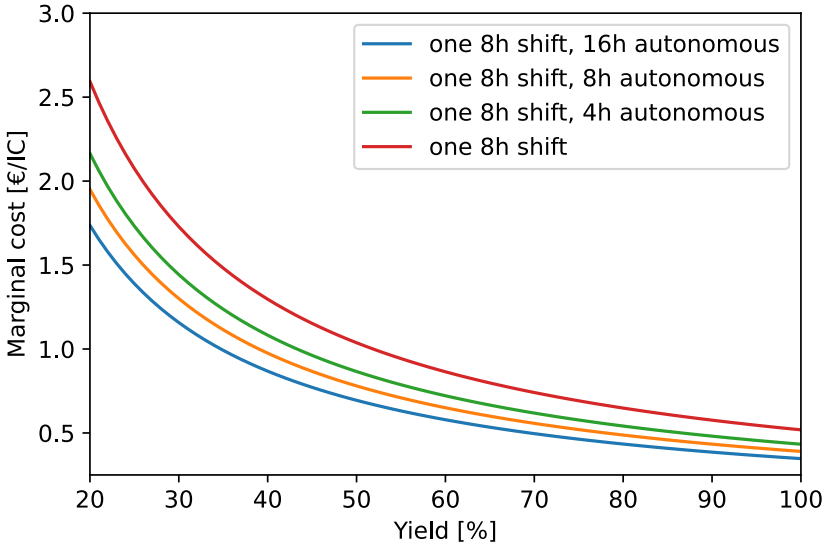


Figure 7.3: Remanufacturing process costs per salable IC as a function of the yield and the autonomous work that the slowest machine (the SB^2 jet) might do outside the work hours. 'one xh shift, yh autonomous' means that the work day is composed of one shift of x hours with three workers in the factory and y hours where the SB^2 jet work in autonomy.

7.2.2 Observations

The cost without accounting for the PCB cost and employees charges equal 0.09€. This amount is divided between energy cost, i.e., 0.046€ and raw materials purchase of 0.048€. The PCB cost is equal to 0.161€ and highly depends on the hypothesis made for the price

per kg of PCBs and the average weight of one PCB. It was assumed that the average PCB weight is 350g (discussed in Appendix F.2.2). Employees charges equal 0.129€ per worker without process optimization and 0.064 € with process optimization. This could be even more optimized if the autonomous operation of slow machines is allowed for 16 hours.

The employees charges become the primary cost source (with PCB cost) once the number of employees grows to 3 or 4. Process optimization help to reduce costs significantly.

7.2.3 Discussions

To realize the scenario (b), where the SB^2-jet operates autonomously, the design of an autonomous robot to feed the SB^2-jet is required. The design of such robots is possible but requires initial investments in R&D.

Then, it is essential to mention that the results include only the operational costs. The costs of engineering and the MSDA (marketing, sales, distribution, and administration) were not considered. Therefore, the actual cost is higher than this one.

The purchase cost of PCBs can vary as a function of the type of PCBs received. For example, the weight of a phone PCB is only 17g while, for a computer, it can grow to 600 or 800g. Therefore, the price of PCBs can vary positively or negatively, given that the PCBs are bought per kg. For this research, we assume that the mean PCB weight was 350g.

We remark in Figure 7.3 that the more the machine is autonomous, the fewer the marginal cost is, and the higher the yield is, the fewer the marginal cost is. Therefore, developing and optimizing the process to operate autonomously is important. The yield can be adapted and will depend on the quality of the wasted PCBs. Indeed, we will have a different yield if the PCBs come from a landfill or from a telephone dismantling plant.

7.3 Results : economic profitability

In this section, the study of the economic viability of a business based on the reuse of ICs will be studied. It will be shown that adaptation and economy of scale are required to reach profitability.

As a reminder, despite all the environmental benefits of reusing ICs, semiconductors' reuse will be effective at large scales only once financial profitability if proven.

7.3.1 Description

Figure 7.4 shows the cumulative marginal benefits over one year as a function of the yield and the price discount to which second-hand ICs will be sold compared to market prices. Note that the remanufactured flash memories distribution is assumed to be 60% TLC 32GB, 25% TLC 64GB, 15% TLC 128GB. Initial investments, engineering costs and MSDA costs are not considered, therefore an interpretation of the graph should be drawn knowingly. For example, considering initial investments of 1M€ amortized in 5

years, two sales and three engineers (for total employees charges per year equal to 500K€), and 100K€ of other costs, the cumulative benefit per year must be higher than 800K€ to meet profitability.

For information, the prices as a function of the discount rate are presented in Table 7.1. As a reminder, **the remanufacturing factory is able to process 691,200 ICs per year**. Therefore for a yield of 50%, the factory output 345,600 salable[2] ICs per year.

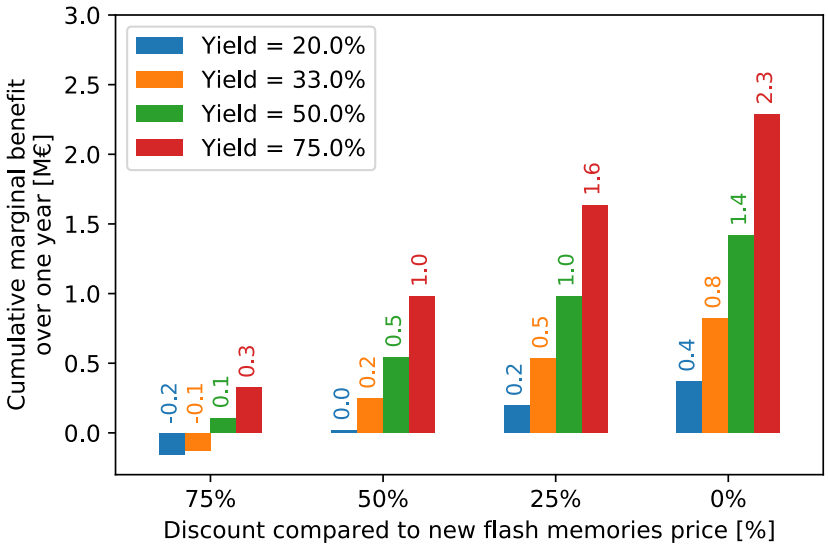


Figure 7.4: The cumulative benefit over one year as a function of the yield and the price discount compared with the market prices (prices as a function of the discount rate can be found in Table 7.1. The remanufactured flash memory distribution is assumed as follow 60 % TLC 32GB, 25% TLC 64GB, 15% TLC 128GB. The throughput of the factory is assumed to be 691,200 ICs per year (2440 ICs per day). As a reminder, the initial investments, engineering costs, and MSDA costs are not considered in the calculation of the costs. The graph needs to be read while considering, for example, the following: if initial investments were 1M€ amortized in 5 years, that two sales and three engineers are needed for total employees charges per year equal to 500K€, and that 100K€ of other costs are accounted for. To meet profitability, the cumulative benefit per year must be higher than 800K€.

Table 7.1: Price as a function of the discount rate

Discount compared to market price [%]	**TLC 32 GB Price [€]**	**TLC 64 GB Price [€]**	**TLC 128 GB Price [€]**
0 (Defined in Figure 7.1)	2.5	5.2	15
25	1.9	3.9	11.2
50	1.2	2.6	7.5
75	0.6	1.3	3.7

[2]Salable ICs are ICs that pass the testing phase and are functional.

7.3.2 Discussions

As stated in Figure 7.3, we can assume that at least 800K€ of cumulative benefit is required to cover unconsidered costs. This threshold of profitability with all costs considered is reached for a discount rate above 25% and a yield above 50%. A lot of cases are under the threshold of profitability, therefore, some adaptations need to be done to reach profitability in a maximum of cases.

Allowing the slow machines to work for 16h autonomously could increase the ICs throughput by 25%. Indeed, the throughput limit is defined by the slowest machines. If these machines can work in autonomy for 16h it can help to increase the ICs throughput per year. It is possible that **more workers will be required**. We can assume that the factory needs five workers in this case.

Increasing the factory capacity could help to cover managerial and engineering costs. Indeed, if more ICs are threatened, more ICs sales will allow covering other costs. Increasing the capacity is easy and requires only duplicating the process line.

Increasing the capacity is possible, but the bottleneck could be the finite number of wasted PCBs available in the market. To obtain a substantial supply of PCBs, it will be necessary to partner with several companies involved in e-waste dismantling. Our plant requires 691,000 PCBs per year to work at full speed and 1,382,000 if the capacity is doubled. To compare, The belgian company CTG circular (company active in refurbishing computers and smartphones) processes a little more than 100,000 units per year. This is far from the capacity that we need for.

The Belgian market is too small to support profitability because not enough PCBs are available. Indeed, 2.6 million smartphones are sold each year. If ten percent of these smartphone PCBs transit yearly per our ICs remanufacturing warehouse, it will provide only 260,000 wasted PCBs. This number is not sufficient to support profitability.

However, if the business is scaled at the European level, profitability could be assured. Indeed, 185.6M smartphones were sold in Europe in 2020. If less than 2.5% of these smartphone PCBs are treated by our warehouses, there will be enough ICs to feed the equivalent of 4 ICs remanufacturing lines, adapted in consequence with the previous points presented. Figure 7.5 shows the cumulative benefit that can be expected if the points discussed here are applied, i.e., a factory that works 24h/24h with five workers. The total capacity is 4,147,200 ICs per year, divided into four sites.

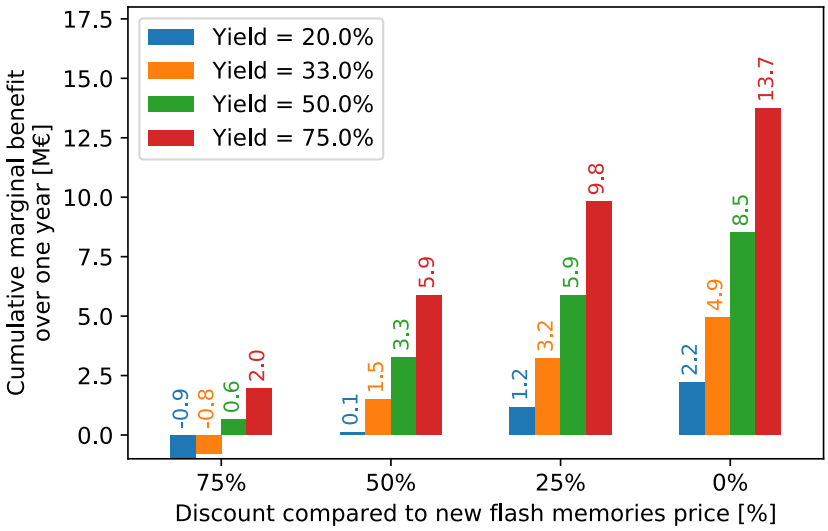


Figure 7.5: The cumulative benefit over one year as a function of the yield and the price discount compared with the market prices. The remanufactured flash memory distribution is assumed as follow 60 % TLC 32GB, 25% TLC 64GB, 15% TLC 128GB. For this case, it is assumed that four factories are deployed and that each factory has five workers and can remanufacture 1,036,800 ICs annually. As a reminder, the initial investment, engineering, and MSDA costs are not considered in the cost calculation. The graph needs to be read while considering, for example, the following: if initial investments were 3M€ amortized in 5 years, that five sales and seven engineers are needed for total employees charge per year equal to 1.2M€ and that 200K€ of other costs are accounted for. To meet profitability, the cumulative benefit per year must be higher than 2M€.

Remanufacturing other types of ICs can reduce the PCB cost per IC. In the current state of affairs, we assume that only one IC where reused per PCB, and therefore, the PCB cost per IC is 0.16€. If other types of ICs are reused, it could help reduce the PCB purchase price per IC.

The profitability depends on the prices, which depend on the market's confidence in reused ICs. Indeed, market confidence will influence the prices of reused ICs. Therefore the development of an accurate testing phase and an accurate thermal annealing process could guarantee the quality of second-life flash memories and therefore enhance market confidence.

Conclusions and perspectives

Thermal annealing: Useful in a reuse context?

Thermal annealing is beneficial to recover from flash memories' aging effects. This affirmation was already topical long before this master thesis and will not revolutionize the world of semiconductors, but this was not the aim. The goal was to make a primary investigation of the use of thermal annealing in a reuse context.

The results are positive and show a recovery of the flash memories' initial electrical properties after a thermal annealing process and this is even more marked while using higher temperatures. Regarding only block erase measures, it can be stated that the initial properties of flash memories can be quasi-totally recovered. These observations open the door to the Development of a thermal annealing process that could be used in a flash memories remanufacturing plant. Thermal annealing coupled with a complete electrical testing process as presented in [41] could guarantee the quality of second-hand flash memories and be a significant sales argument.

To reach market confidence, further research should be done on the retention time after a thermal annealing process on a representative number of flash memories. Indeed, it is essential to show that the retention can be recovered at X% thanks to thermal annealing, and this percentage needs to be known. The primary goal of flash memories is to store data as long as possible. Hence, the possibilities of reuse will depend on the certifications that we can give on the retention time and on the lifespan of second-hand flash memories.

Can the use of second-hand flash memories reduce the carbon footprint of the semiconductor industry ?

According to the numbers presented in this thesis, the answer is yes! Indeed, it was demonstrated that the use of remanufactured flash memories in the European market could help to reduce GHG emissions up to 1.4MT CO2 equivalent. These results are mainly due to the fact that the GWP of the production phase is reduced by 36-96% compared with new flash memories.

However, although minor, there exist cases where the emissions saving made during the production phase will be reduced to nothing during the use phase. Therefore, it is very important to assess, for each use case, if remanufactured ICs are beneficial for the environment thanks to the adaptative method developed in this thesis.

Could a business based around the reuse of semiconductors be profitable ?

This question is the most important and questionable one because semiconductor reuse will never be implemented without financial profitability.

This master thesis demonstrates that the benefits range around -0.9 and 13.7M€ on a yearly basis without accounting for initial investment, engineering, and MSDA costs. The benefits depend mainly on the remanufacturing process IC throughput, the yield, and the market prices of reused flash memories. The yield will depend on the ability of an economy to process the e-waste cleanly. The remanufacturing process IC throughput will depend on the number of wasted PCBs available on the market because replicating the process is easy. The market response will define the prices. Hence, an excellent thermal annealing process and a reliable testing phase are the keys to guaranteeing the quality of second-hand flash memories and therefore selling at high prices.

However, the profitability also depends on externalities like the possibility of the end of Moore's law or the possible arrival of disruptive technology on the market. On the one hand, it has been years since we announced the end of Moore's law due to physical limits, but this could also come from capital limits [63]. In an increasingly uncertain world, capital may no longer be available to reach the next technological node. The end of Moore's law would be an incentive to start a business based on semiconductor remanufacturing. On the other hand, the arrival of disruptive technology on the market would render all integrated circuits present in e-waste obsolete and, therefore, destroy any hope of reusing semiconductors.

However, under the present circumstances, semiconductor reuse should be considered as a cheap way to reduce our dependence on the foreign market for chip supply and to reduce the carbon footprint of ICT.

What are the next steps ?

There is still some work to do before the reuse of semiconductors becomes a reality. Researches need to be done on ICs' identification, faulty integrated circuit detection, demonstrators' design, and the implementation of an industrial line destined for ICs reuse.

Studying flash memories' retention rate before and after thermal annealing is also required. This can be done by using flash memories compatible with the read-retry mechanism. The read-retry mechanism allows us to modify the read reference voltage. This behavior can be used to measure the distribution of V_{th} over time. Therefore, it is a really accurate tool to measure the retention rate over time.

Moreover, it might be interesting to replace flash memories from PCBs in new smartphone with old flash memories, with the aim to monitor their impact on the system's speed,

power consumption, and lifespan. This experiment could be a game-changer if the outputs are positive.

Finally it could be interesting to reuse other valuable ICs like RAM or CPU but before that we need to assess the technological and financial feasibility of the reuse of these ICs.

Bibliography

[1] U. Gupta, Y. G. Kim, S. Lee, J. Tse, H.-H. S. Lee, G.-Y. Wei, D. Brooks, and C.-J. Wu, "Chasing Carbon: The Elusive Environmental Footprint of Computing," Oct. 2020.

[2] PWC, "Opportunities for the global semiconductor market," 2019.

[3] Wikipedia contributors, "Transistor — Wikipedia, the free encyclopedia," 2022. [Online; accessed 12-August-2022].

[4] shigeru23, "English: Moores'law (Major CPU'S 1970-2011 Transister number's trend chart)," Sept. 2011.

[5] Wikipedia contributors, "Surface-mount technology — Wikipedia, the free encyclopedia," 2022. [Online; accessed 1-July-2022].

[6] "Dual Inline Package," *Wikipédia*, Sept. 2021.

[7] W. Wits, "Integrated cooling concepts for printed circuit boards," *Water Resources Management - WATER RESOUR MANAG*, Jan. 2008.

[8] P. Ciszewski, "Device with reused components: Technology demonstration," 2019.

[9] S. Janusz, K. Marek, A. Aneta, S. Wojciech, and J. Kamil, "Challenges of reuse and remanufacturing of modern chips in smart mobile devices," in *2016 Electronics Goes Green 2016+ (EGG)*, (Berlin), pp. 1–7, IEEE, Sept. 2016.

[10] Y. Luo, "Architectural Techniques for Improving NAND Flash Memory Reliability," p. 255.

[11] S. Sakib, P. Kumari, B. Talukder, M. Rahman, and B. Ray, "Non-Invasive Detection Method for Recycled Flash Memory Using Timing Characteristics †," *Cryptography*, p. 17, 2018.

[12] M. R. C. L. M. A., *Inside Nand Flash Memories*. Springer Dordrecht Heidelberg London New York, 2010.

[13] H. Yang, H. Kim, S.-i. Park, J. Kim, S.-h. Lee, J.-k. Choi, D. Hwang, C. Kim, M. Park, K.-h. Lee, Y.-k. Park, J. Shin, and J.-t. Kong, "Reliability Issues and Models of sub-90nm NAND Flash Memory Cells," in *2006 8th International Conference on Solid-State and Integrated Circuit Technology Proceedings*, (Shanghai, China), pp. 760–762, IEEE, 2006.

[14] Y. Cai, Y. Luo, E. F. Haratsch, K. Mai, and O. Mutlu, "Data retention in MLC NAND flash memory: Characterization, optimization, and recovery," in *2015 IEEE 21st International Symposium on High Performance Computer Architecture (HPCA)*, pp. 551–563, Feb. 2015.

[15] R. Micheloni, ed., *3D Flash Memories*. Dordrecht: Springer Netherlands, 2016.

[16] "What is floating gate transistor (FGT)? - Definition from WhatIs.com." https://www.techtarget.com/searchstorage/definition/floating-gate.

[17] G. H. Lee, S. Hwang, J. Yu, and H. Kim, "Architecture and Process Integration Overview of 3D NAND Flash Technologies," *Applied Sciences*, vol. 11, p. 6703, Jan. 2021.

[18] Q. Wu, G. Dong, and T. Zhang, "A First Study on Self-Healing Solid-State Drives," in *2011 3rd IEEE International Memory Workshop (IMW)*, (Monterey, CA, USA), pp. 1–4, IEEE, May 2011.

[19] H.-T. Lue, P.-Y. Du, C.-P. Chen, W.-C. Chen, C.-C. Hsieh, Y.-H. Hsiao, Y.-H. Shih, and C.-Y. Lu, "Radically extending the cycling endurance of Flash memory (to > 100M Cycles) by using built-in thermal annealing to self-heal the stress-induced damage," in *2012 International Electron Devices Meeting*, (San Francisco, CA, USA), pp. 9.1.1–9.1.4, IEEE, Dec. 2012.

[20] "What is write endurance?."

[21] Lepatre, "Semi-conducteurs : la faille chinoise," *Le Monde.fr*, Apr. 2021.

[22] P. Dr Scröder, "Solutions for resilient semiconductor supply chains," *Chatham House – International Affairs Think Tank*, Jan. 2022.

[23] "• Top semiconductor foundries market share 2022 | Statista." https://www.statista.com/statistics/867223/worldwide-semiconductor-foundries-by-market-share/.

[24] K. Leswing, "Apple chipmaker TSMC warns Taiwan-China war would make everybody losers." https://www.cnbc.com/2022/08/02/apple-chipmaker-tsmc-warns-taiwan-china-war-would-make-everybody-losers.html, 2022.

[25] "Paquet législatif sur les semi-conducteurs." https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-chips-act_fr.

[26] R. P. Allan, C. Cassou, D. Chen, A. Cherchi, L. Connors, F. J. Doblas-Reyes, H. Douville, F. Driouech, T. L. Edwards, E. Fischer, G. M. Flato, P. Forster, K. M. AchutaRao, B. Adhikary, E. Aldrian, and K. Armour, "Summary for Policymakers," p. 32.

[27] F. Bordage, "Study - The environmental footprint of the digital world," 2019.

[28] A. Monnin, *Towards Digital Sobriety*. Mar. 2019.

[29] N. Moreau, T. Pirson, G. Le Brun, T. Delhaye, G. Sandu, A. Paris, D. Bol, and J.-P. Raskin, "Could Unsustainable Electronics Support Sustainability?," *Sustainability*, vol. 13, June 2021.

[30] P. J. Denning and T. G. Lewis, "Exponential laws of computing growth," *Communications of the ACM*, vol. 60, pp. 54–65, Dec. 2016.

[31] "About the project sustainablySMART."

[32] S. Janusz, P. Ciszewski, P. Dawidowicz, M. Koscielski, A. Arazna, K. Janeczek, W. Steplewski, G. Podhradsky, and R. Ambrosch, "DESOLDERING AND REMANUFACTURING OF SEMICONDUCTOR COMPONENTS FROM ELECTRONIC MOBILE DEVICES," *2019*.

[33] I. Stobbe, H. Griese, H. Poetter, H. Reichl, and L. Stobbe, "Quality assured disassembly of electronic components for reuse," pp. 299 – 305, 02 2002.

[34] "Definition of integrated circuit | Dictionary.com." https://www.dictionary.com/browse/integrated-circuit.

[35] "What the global semiconductor chip shortage means for UK car buyers | Auto Express." https://www.autoexpress.co.uk/news/355016/global-chip-shortage-hits-car-makers.

[36] Wikipedia contributors, "Moore's law — Wikipedia, the free encyclopedia." `https://en.wikipedia.org/w/index.php?title=Moore%27s_law&oldid=1104715676`, 2022. [Online; accessed 17-August-2022].

[37] A. Coombs, "SMT vs SMD vs Through-hole vs BGA: What works best in PCBA? - MIS," Dec. 2020.

[38] Wikipedia contributors, "Ball grid array — Wikipedia, the free encyclopedia," 2022. [Online; accessed 1-July-2022].

[39] "COM(2000)347-1 - Waste electrical and electronic equipment - EU monitor."

[40] R. Johanna, "Automated Identification and Sorting of Collected Smartphones," 2019.

[41] J. Sitek, M. Koscielski, P. Dawidowicz, P. Ciszewski, M. Khramova, D. N. Quang, and S. Martinez, "Investigations of temperature resistance of memory bga components during multi-reflow processes for circular economy applications," in *2017 21st European Microelectronics and Packaging Conference (EMPC) Exhibition*, pp. 1–7, 2017.

[42] J. Sitek, M. Koscielski, A. Arazna, K. Janeczek, and W. Stęplewski, "Investigations of bga components'balls remanufacturing techniques for circular economy applications," in *2018 7th Electronic System-Integration Technology Conference (ESTC)*, pp. 1–6, 2018.

[43] PADTECH, "SB^2 Laser Solder Jetting Systems."

[44] "Dynamic Random Access Memory Market Share, Analysis | 2022-27." https://www.mordorintelligence.com/industry-reports/dynamic-random-access-memory-market.

[45] Wikipédia, "Mémoire flash — wikipédia, l'encyclopédie libre," 2022. [En ligne; Page disponible le 2-mai-2022].

[46] R. Sheldon, "Charge trap technology advantages for 3D NAND flash drives," 2018.

[47] D. Price, "NAND and eMMC: All You Need to Know About Flash Memory," Sept. 2016. Section: Technology Explained.

[48] R. Micheloni, L. Crippa, and A. Marelli, *Inside NAND Flash Memories*. Dordrecht: Springer Netherlands, 2010.

[49] L. Shijun and Z. Xuecheng, "Analysis of 3D NAND technologies and comparison between charge-trap-based and floating-gate-based flash devices," *The Journal of China Universities of Posts and Telecommunications*, pp. 75–96, 2017.

[50] H. Watanabe, *Tutorial: Trap-related Reliability Issues in NAND Flash Memory*. Apr. 2012.

[51] Wikipedia contributors, "Flash memory — Wikipedia, the free encyclopedia," 2022. [Online; accessed 19-August-2022].

[52] S. Amor, V. Kilchytska, D. Flandre, and P. Galy, "Trap recovery by in-situ annealing in fully-depleted mosfet with active silicide resistor," *IEEE Electron Device Letters*, vol. 42, no. 7, pp. 1085–1088, 2021.

[53] P. Galy, R. Lethiecq, and M. Bawedin, "In-situ heater for thermal assist recovery of MOS devices in 28 nm UTBB FD-SOI CMOS technology," *Solid-State Electronics*, vol. 168, p. 107737, June 2020.

[54] S. Boyd, A. Horvath, and D. Dornfeld, "Life-cycle assessment of nand flash memory," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 24, pp. 117 – 124, 03 2011.

[55] "UFS 4.0 | Universal Flash Storage." https://semiconductor.samsung.com/content/semiconductor/g
4-0.html, 2022.

[56] "UFS 3.0 | Universal Flash Storage." https://semiconductor.samsung.com/content/semiconductor/g
3-0.html, 2022.

[57] A. Busa, M. Hegeman, J. Vickers, N. Duque-Ciceri, and C. Herrmann, *Life Cylce Assessment of DELL R740*. 2019.

[58] "Gestion de l'usure flash dans Android Automotive." https://source.android.com/docs/core/perf/flash-wear?hl=fr.

[59] A. Annie, "State of Mobile 2022," 2022.

[60] P. N. L. e. S. Fafard and M. à. j. l. m. . j. . à 15:00, "Part de marché smartphone : des ventes en baisse, mais Samsung reste leader." https://www.cnetfrance.fr/news/part-de-marche-smartphone-39884221.htm.

[61] A. Lucina, "L'homme passe près de cinq heures par jour sur son mobile." https://www.20minutes.fr/high-tech/3218291-20220117-un-humain-passe-en-moyenne-4-8-heures-par-jour-sur-son-smartphone.

[62] "French Repairability Index: Fairphone takes the lead." https://www.fairphone.com/fr/2021/02/04/french-repairability-index/, Feb. 2021.

[63] "Moore's second law," *Wikipedia*, Dec. 2021.

[64] R. Europe, "Obsolescence."

[65] S. ELECTRONICS, "Precision BGA Reball."

[66] "Rebillage Laser – Retronix Europe."

[67] A. C. Ciabai, "Motherboard Size Comparison Chart | Techosaurus Rex," May 2021.

[68] E. A. limited, "Drying oven - Essa Australia - Page - PDF Catalogs | Technical Documentation | Brochure." https://pdf.directindustry.com/pdf/essa-australia/drying-oven/33731-510121-_4.html, 2022.

[69] "Stainless Steel Electric Heating Drying Machine Fruit Vegetable Hot Air Dry Oven Dryer - Buy Chinese Traditional Medicine Dehydrator Chemical Food Hot Air Dry Oven Dryer,Stainless Steel Drying Oven Mango Slice Pineapple Slice Drying Oven Hot Air Dryer,Tray Dryer Oven Hot Air Circulating Drying Oven Industrial For Fruit Vegetable Product on Alibaba.com." https://www.alibaba.com/product-detail/Electric-Drying-Oven-Electric-Drying-Oven_1600450710490.html?spm=a2700.7724857.normal_offer.d_title.390d204eauTCYP&s=p.

[70] Puhui, "Channel reflow oven T-980." https://~(?<subdomain>.+)\\.te168\\.com$:443/channel-reflow-oven-t-980-products/.

[71] Mecademic, "The world's smallest, most precise and compact six-axis robot arm." https://www.mecademic.com/en/meca500-robot-arm.

[72] "Bras Robotique à 4DOF (Assemblé) - Lynxmotion (LSS)." https://www.robotshop.com/be/fr/bras-robotique-a-4dof-assemble—lynxmotion-lss.html.

[73] "Power Options for IP Surveillance Solution - Mistral Solutions." https://www.mistralsolutions.com/articles/power-options-ip-surveillance-solution/.

[74] J. Draper, "How Much Electricity Does A Gaming PC Use? (Explained)," June 2022.

[75] "Zx3000 Ce Rohs Hot Air Heat Heater For Paper Cup Machine - Buy Blower Heater,Plastics Air Blower,Air Heat Blower Product on Alibaba.com." https://www.alibaba.com/product-detail/ZX3000-CE-ROHS-Hot-air-heat_62455018161.html.

[76] PacTech, "SB^2 Laser Solder Jetting Systems."

[77] T. Xu and W. Tschudi, "Cleanroom Energy Benchmarking Results," 2001.

[78] ABN, "What type of cleanroom are you in need of?." https://www.abn-cleanroomtechnology.com/solutions/rent-a-cleanroom.

[79] n. Sypla, "Clean-Room Partitions | Sypla." https://www.sypla.com/en/product/clean-room-partitions/.

[80] TBH, "CR-GL 265." https://www.tbh.eu/en/products/cr-series.html.

[81] Terasic, "DE10-Nano : User manual," 2019.

[82] ADEME, "Documentation Base Carbone." https://bilans-ges.ademe.fr/documentation/UPLOAD_DOC_FR/index.htm?electromenager.htm.

[83] Apple, "13-inch MacBook Air Product Environmental Report," p. 9, 2019.

[84] "Amazon.com: Laptop Motherboard for HP, Professional I5 6200 Main Board for HP 440 G3/430 G3 Laptop Notebook, ABS+Chip, CR2023 Battery (not Included) : Electronics." https://www.amazon.com/Motherboard-Professional-I5-6200-Notebook-Included/dp/B081YP97NZ.

[85] "CM MSI *A320M-A Pro Max* AM4/DDR4/M-ATX *7665/P2+A : Amazon.fr: Informatique." https://www.amazon.fr/MSI-A320M-A-Pro-Max-A320/dp/B07W838R4F/ref=sr_1_9?crid=VEJFIK10ELE9&keywords=macbook+motherboard&c
9.

[86] ELECTRICITYMAPS, "Live 24/7 CO_2 emissions of electricity consumption." http://electricitymap.tmrow.co.

[87] S. Prakash, L. Ran, K. Schischke, S. Dr. Lutz, and G. Gensch, "Schaffung einer Datenbasis zur Ermittlung ökologischer Wirkungen der Produkte der Informations- und Kommunikationstechnik (IKT). Teilvorhaben C," 2013.

[88] K. Schischke, C. Clemm, A. Berwald, M. Proske, G. Dimitrova, J. Reinhold, C. Prewitz, and C. Neef, *Ecodesign Preparatory Study on Mobile Phones, Smartphones and Tablets - Final Task 4 Report Technologies*. 2021.

[89] Semicondutor industry assocaition, "STATE OF THE U.S. SEMICONDUCTOR INDUSTRY 2020," 2020.

[90] The World with AFP, "GlobalFoundries and STMicroelectronics announce the construction of a factory in Grenoble." https://newsinfrance.com/globalfoundries-and-stmicroelectronics-announce-the-construction-of-a-factory-in-grenoble/, July 2022.

[91] A. Riva, S. D'Angelosante, and C. Trebeschi, "Natural gas and the environmental results of life cycle assessment," *Energy*, vol. 31, pp. 138–148, Jan. 2006.

[92] Apple, "iPhone X Environmental Report," 2017.

[93] Apple, "iPhone 11 Environmental Report," 2019.

[94] Apple, "iPhone 12 Product Environmental Report," 2020.

[95] Apple, "iPhone 13 Pro Product Environmental Report," 2021.

[96] "Combien coûte un manœuvre dans la construction ? | Je veux engager." https://www.jeveuxengager.be/personnage/combien-coute-un-manoeuvre-dans-la-construction-/construction.

[97] “Le prix de l’électricité en Belgique en 2022 | Comparateur-Energie.be.” https://www.comparateur-energie.be/blog/prix-electricite-belgique/.

[98] RURAWEB, “» Flux sans résidus 1 litre.”

[99] “étain: Achat à 20.00€ le kilo - Prix de l’étain en gros.” https://www.metalcash.be/achat/metaux/etain.

[100] “13.94€ 4% de réduction|PMTC 250K Étain Boule De Soudure BGA 0.2/ 0.25/ 0.3/ 0.35/ 0.4/ 0.45/ 0.5/ 0.55/ 0.6/ 0.65/ 0.76mm Pour BGA Reballing Kit | AliExpress.” //fr.aliexpress.com/item/32841002030.html?src=ibdm_d03p0558e02r02&sk=&aff_platform=&aff_

Appendix A

Temperature and time estimation

Different approaches co-exist in the literature to estimate both parameters' requirements.

Trap interface recovery evolution First, [13, 18] have based their estimation by modeling the trap interface recovery evolution in function of the time, N_{it} :

$$\frac{dN_{it}}{N_{it}} = -k \cdot dt \tag{A.1}$$

Where t is the thermal annealing time, and k is a function of temperature based on the Arrhenius approximation :

$$k = k_0 \cdot \exp -\frac{E_a}{RT} \tag{A.2}$$

Where k_0 is the intrinsic rate constant, E_a is the activation energy that depends on the device and R is the gas constant. k_0 and E_a take value dependent on the device and technology studied. For the case study in the paper, it was the activation energy of hydrogen on Si-H bond which was 0.52eV. The result can be found in Figure A.1.

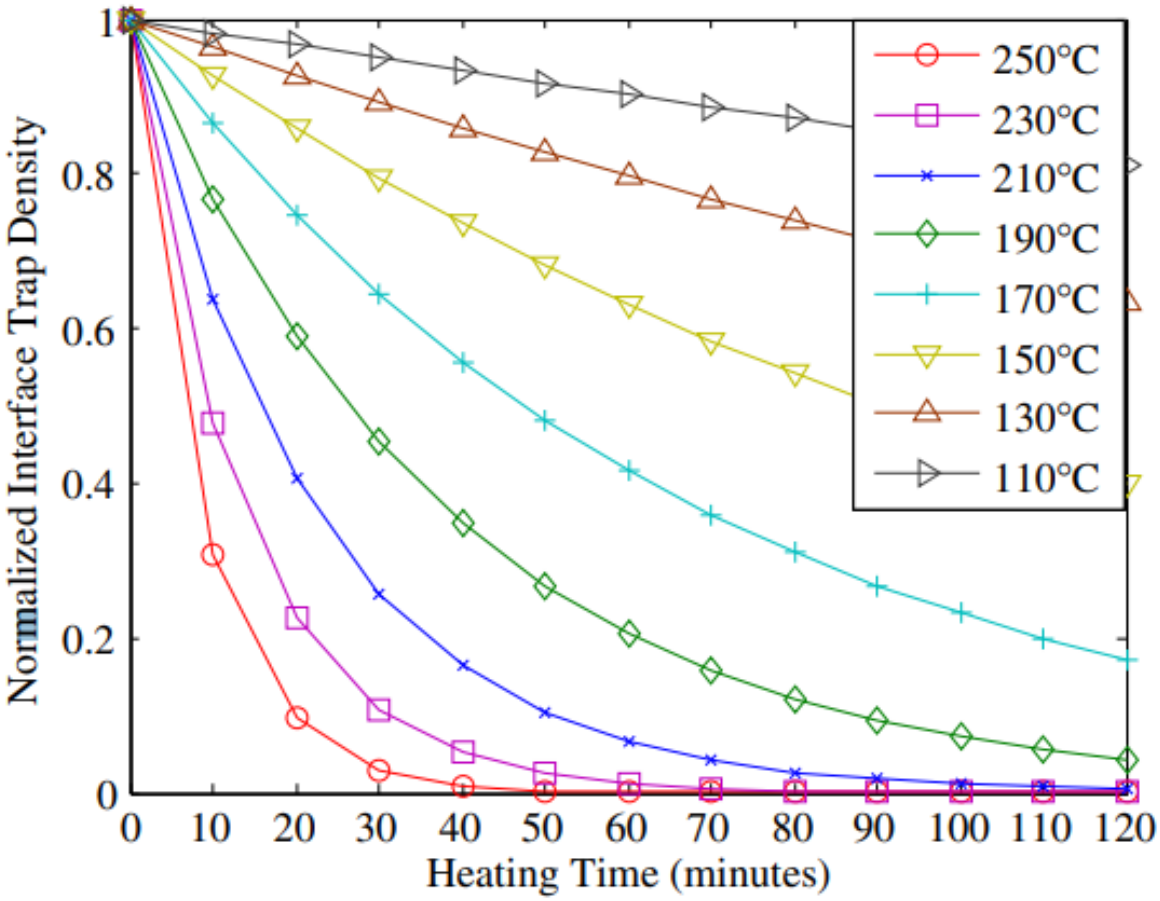


Figure A.1: Interface trap recovery as a function of heating time under different temperature. Reproduced from [18]

Experimental approach k_0 and E_a might remain unknown, at this moment an experimental approach can be used. For example in [19], primary experimentation proves that the studied flash memory characteristics remain the same after doing 10 times, 100k P/E cycles, and 250 °C - 2h bake. The 250 °C - 2h was a point reference to estimate k_0 the function of different activation energy, the result can be found in Figure A.2.

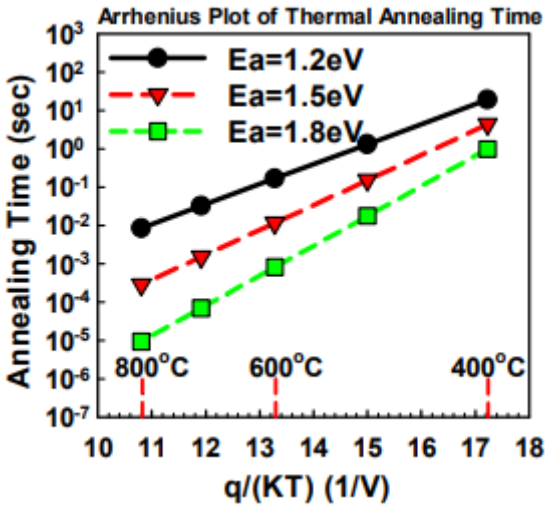


Figure A.2: Arrhenius plots for thermal annealing. The reference annealing time is 2hr at 250 °C for this estimation. Reproduced from [19]

The degradation in the tunnel oxide can be repaired thanks to thermal annealing. Thermal annealing means that we use heat to treat a material or here a semiconductor device.

Appendix B

Thermal annealing

B.1 Implementation of the communication with the flash memories

One DEO10-NANO FPGA board is used to communicate with the flash memory. A 48TSOP to 48DIP socket is used to link the flash memories with the GPIO of the FPGA board. The Deo10 Fpga Board is composed of one cyclone 5 FPGA chips and one HPS CPU with two cores. A finite-state machine is programmed onto the FPGA and is controlled by the HPS CPU.
Four ONFI operations are programmed on the FPGA :

- Reset operation: Resets the read/program/erase operation and clears the status register to ready.
- Read operation: Allow external read of the content of a page
- Erase operation: Erase all the content of a block, ie set one in the memory.
- Direct write operation: Write data inputted on a page without using the cache in the flash memory.

Correctness of operations For giving the reader confidence about the process of communicating with flash memories, erase and write operation will be demystified. The important are listed here under :

- CE# is the chip enable signal. Chip is enabled when CE# is low
- RE# is the read enable signal. At each rising edge of RE#, IO is set by the flash memory to the value that needs to be outputted.
- WE# is the write enable signal. At each rising edge of WE#, the flash memory read the value of the IO.
- CLE is command latch enable signal, command inputs are enabled when CLE is high.
- ALE is address latch enable signal, address inputs are enabled when ALE is high.
- IO[7:0] is the bidirectional connexion in which bytes are inputted or outputted

- R/B# is the ready busy mode of the IC. If R/B# is high the IC is ready, if not the IC is busy.

Different phases are represented on the figures and are explained by :

1. Command input phase is when you input the command code. Command code for read, erase and write operations are respectively 00h, 60h and 70h.

2. Address input phase is when you input the address. This is done by consecutively inputting bytes representing the address. Address for an erase operation is a block address that has a length of 3 bytes and needs to be terminated by the D0h command. Address for a write or a read operation is a block and page address combined and has a length of 5 bytes and the address input phase terminates after the fifth bytes for a write operation and after the input of 30h command for a read operation.

3. Data input phase is when the data that have to be programmed during a write operation are sent to the flash memory. 2112 bytes need to be sent, and each byte is read at a rising edge of W/E#. The data input cycle is finished by inputting 10h. (Only for Write)

4. Data read phase is when the content of a page is read. Bytes are outputted at the rising edge of RE# starting with the first byte and finishing after 2112 bytes read.

5. Processing phase is the phase when the flash memory is processing the command. This phase begins and finishes respectively with a falling edge and a rising edge on R/B#.

6. Status Read operation is the operation for reading the output status. For it, 70h command is sent to the device and after that the status is outputted by the IC at the rising edge of RE#. If this status equals E0h it means that the operation succeeds and that the IC is ready for the new operation.

Figure B.1 shows a read status operation

Figure B.2 shows a page program operation

Figure B.3 shows that a read operation on a page after a Program operation on the block. It output 00h and this was expected.

Figure B.4 shows a block erase operation

Figure B.5 shows that a read operation on a page after a Program operation on the block. It output FFh and this was expected.

Figure B.5 shows that a read operation on a page after a Program operation on the block. It output FFh and this was expected.

It is exactly the behavior expected, therefore It can be concluded that the operations are correctly implemented.

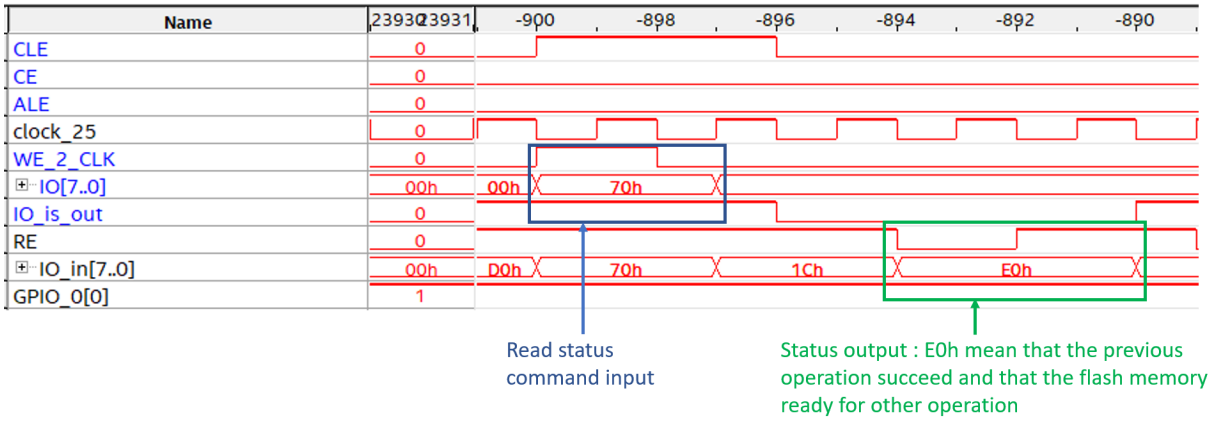


Figure B.1: Read status operation

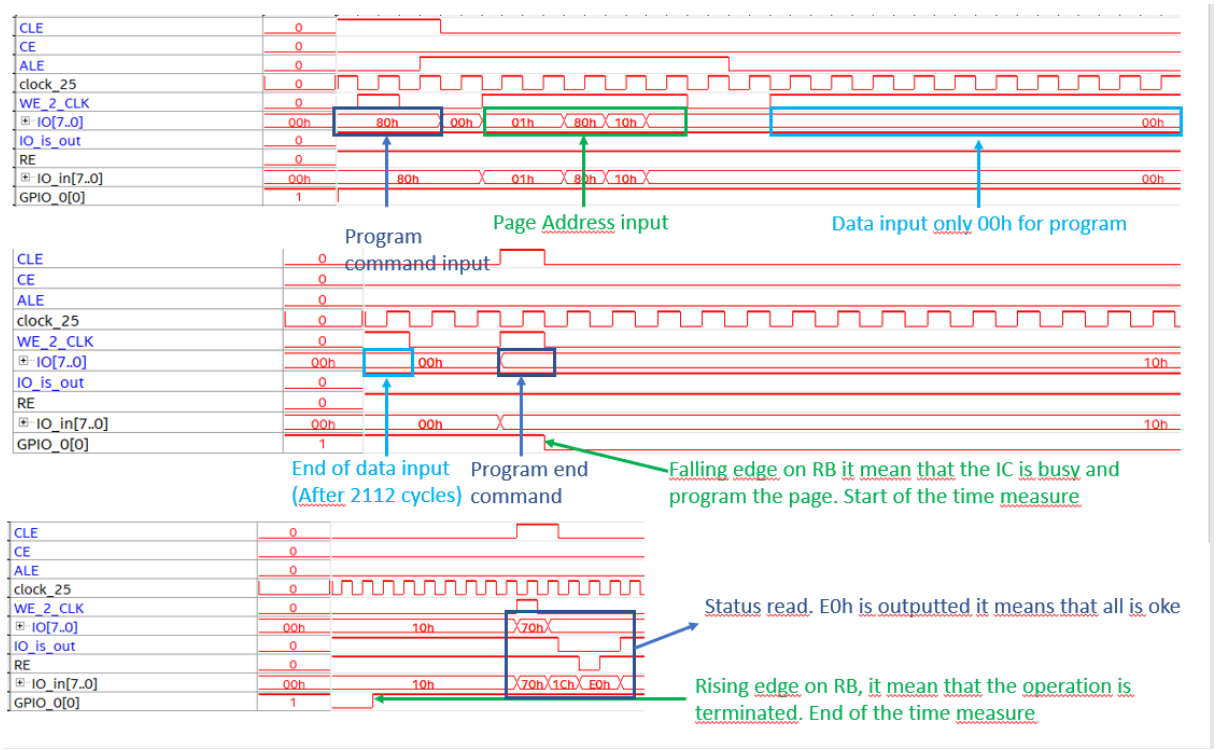


Figure B.2: Program operation

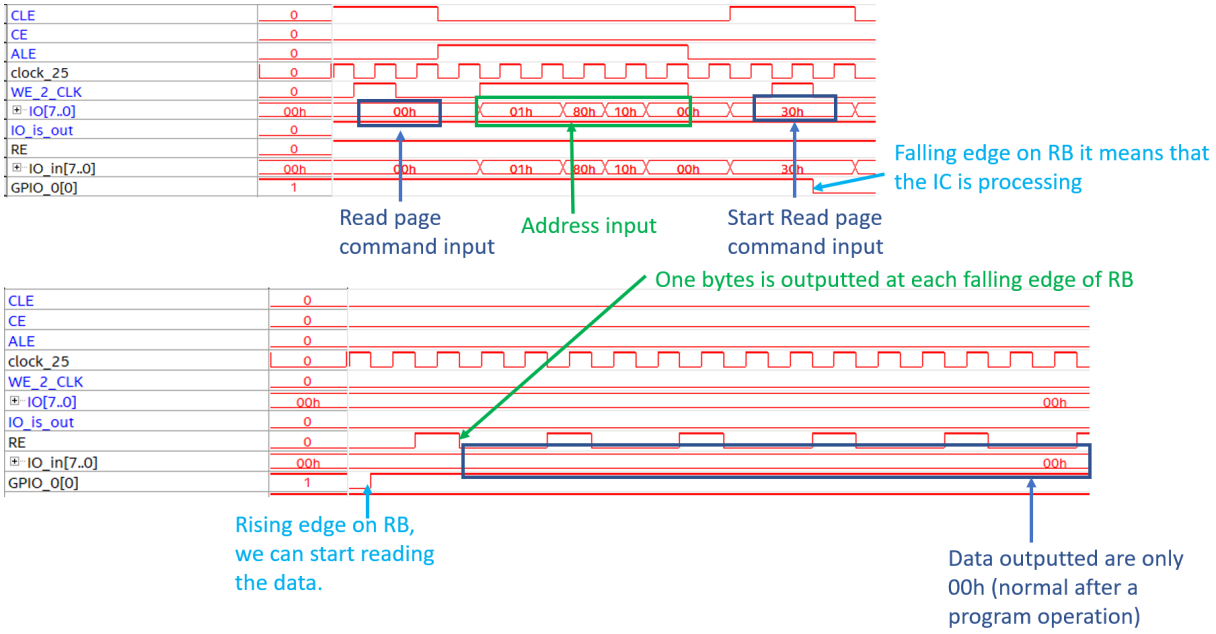


Figure B.3: Read operation on a page after a Program operation on a page. This outputs 00h as expected

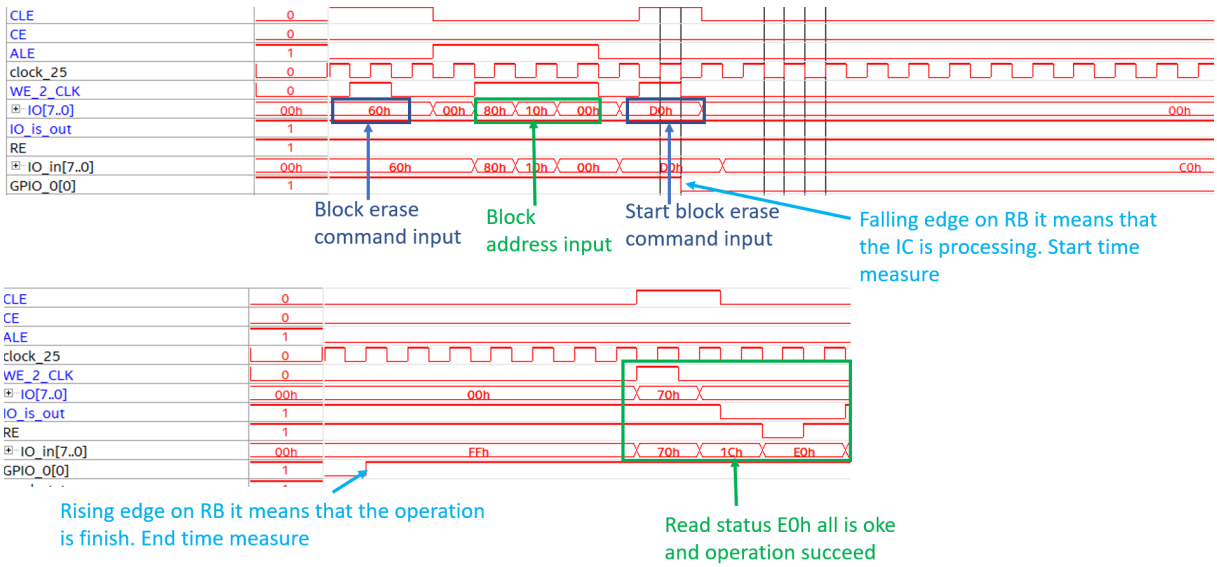


Figure B.4: Block erase operation

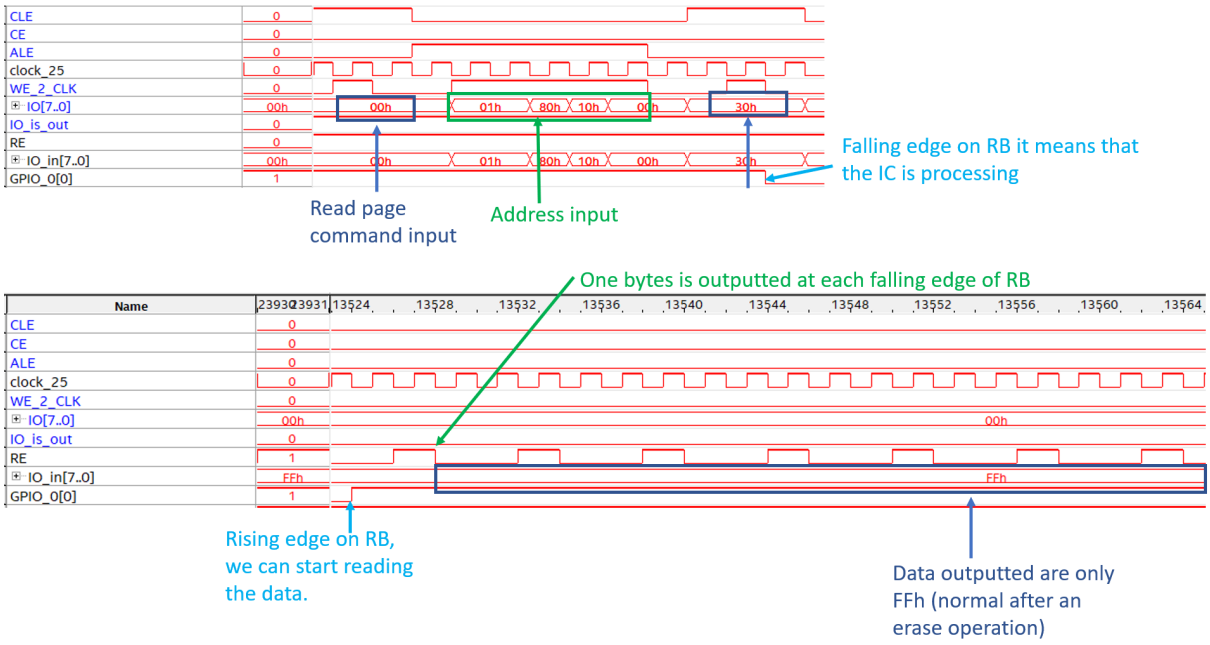


Figure B.5: Read operation on a page after a Program operation on the block. This outputs FFh

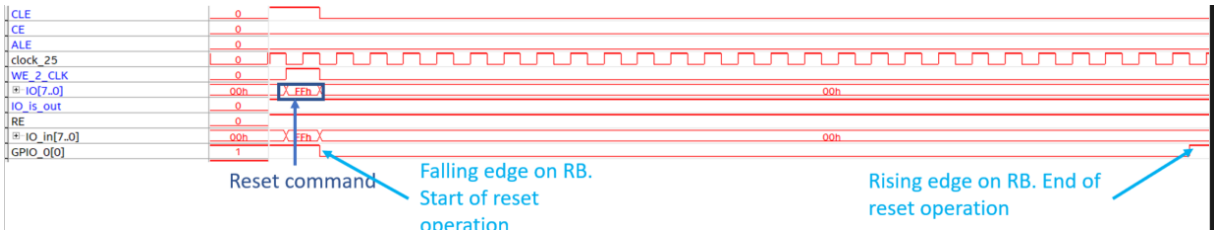


Figure B.6: Reset operation

Appendix C

Thermal annealing: curves features extraction for erase

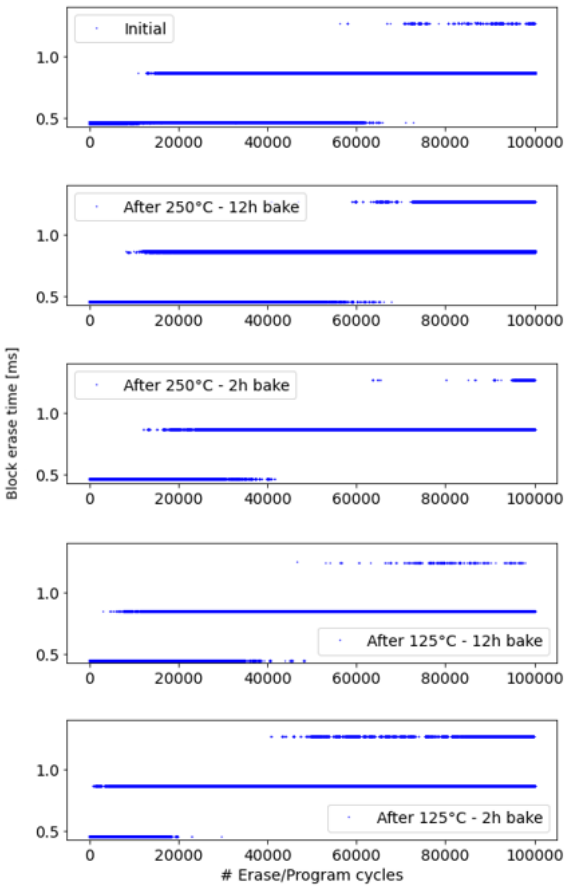


Figure C.1: Raw measure of block erase time

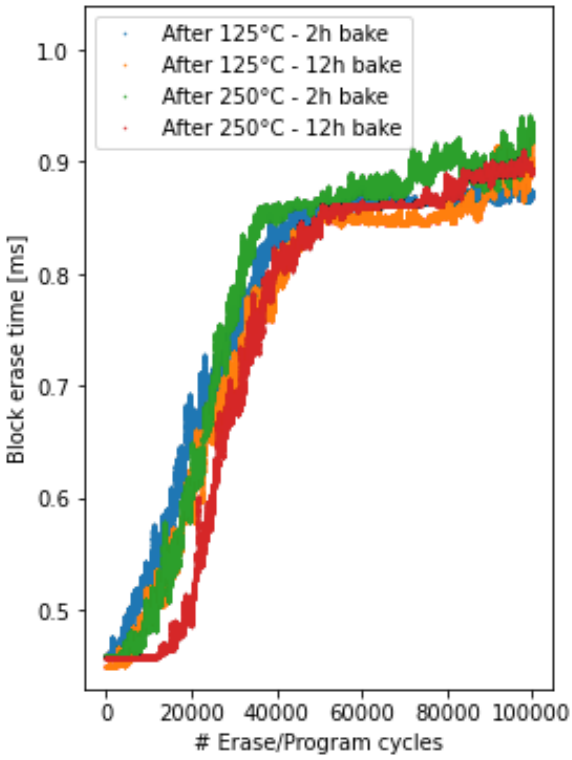


Figure C.2: Boxplot of the mean over the last 1000th measure of block erase time.

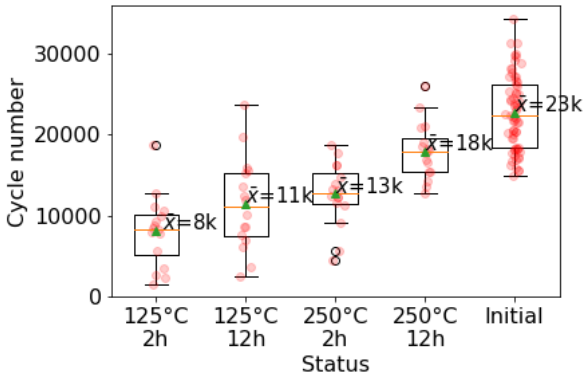


Figure C.3: Boxplot of the distribution of the 100th cycles respecting this condition: erase time > 0.7ms

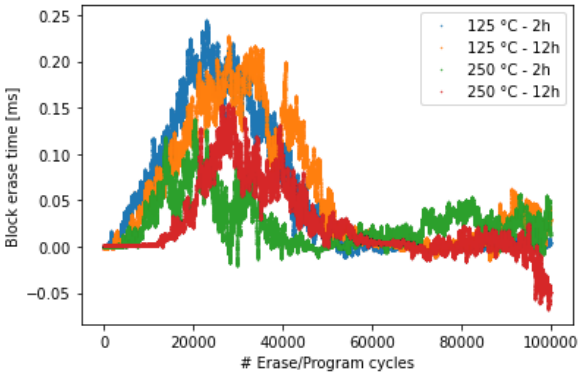


Figure C.4: Result of $t_{after} - t_{initial}$ for block erase time.

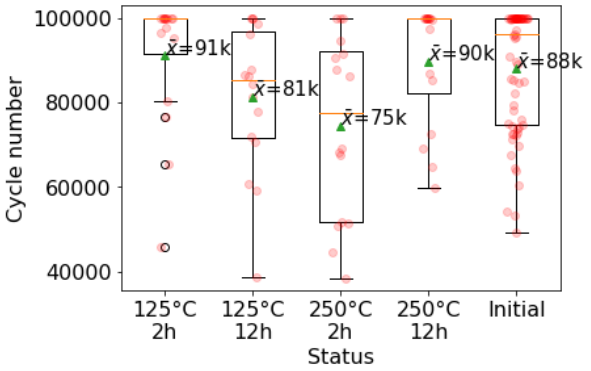


Figure C.5: Boxplot of the distribution of the 100th cycles respecting this condition: erase time > 1ms

Appendix D

Thermal annealing: curves features extraction for program

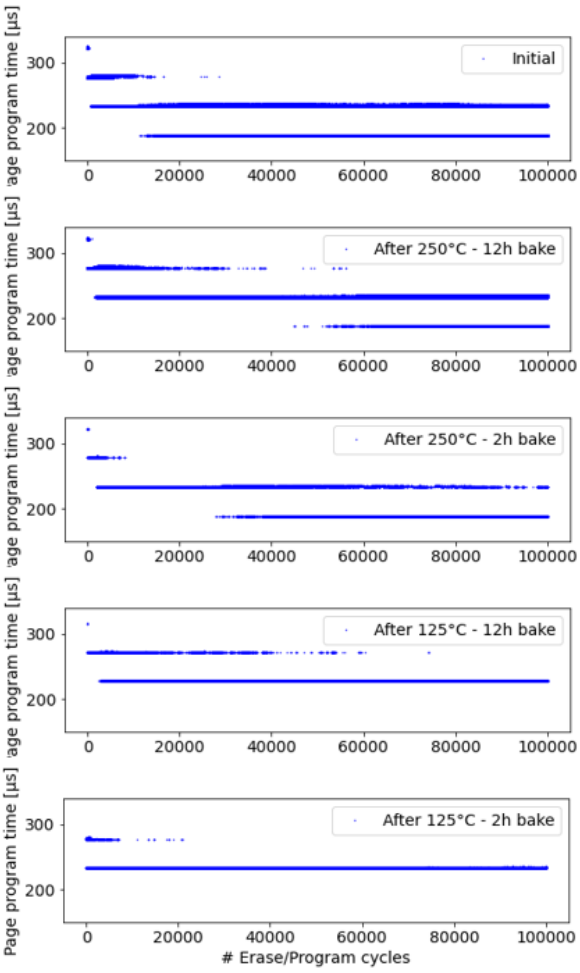


Figure D.1: Result of $t_{after} - t_{initial}$ for page program time

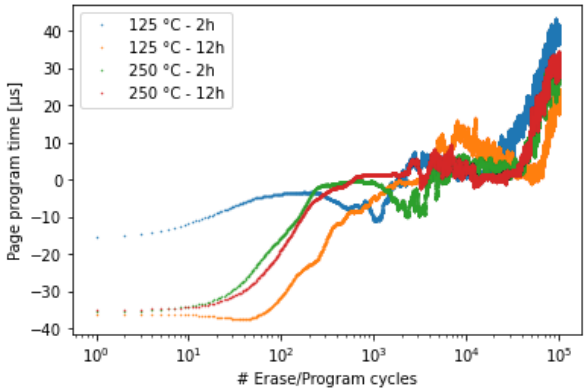


Figure D.2: Result of $t_{after} - t_{initial}$ for page program time

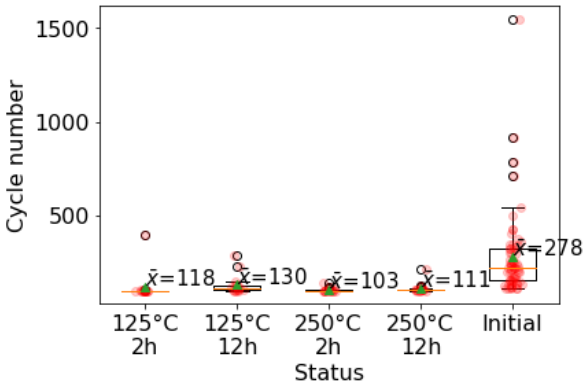


Figure D.3: Boxplot of the distribution of the 100th cycles respecting this condition: $300\mu s >$ program time $> 250\mu s$

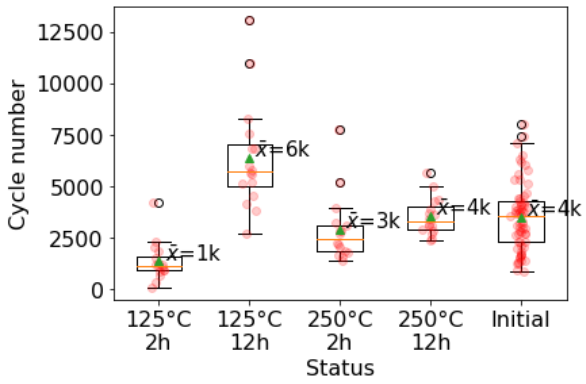


Figure D.4: Boxplot of the distribution of the 100th cycles respecting this condition: 250μs > program time > 200μs

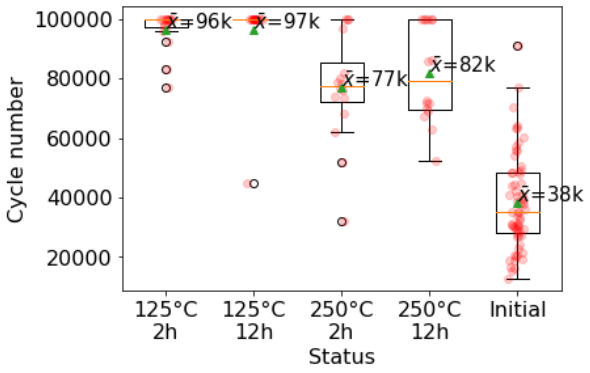


Figure D.5: Boxplot of the distribution of the 100th cycles respecting this condition: 200μs > program time

Appendix E

Design of a reuse IC factory used as a tool for the LCA and the back-of-the-envelope economic benchmarking

E.1 Objectives

The objective is to design a reuse IC factory to benchmark the ecological impact and economic profitability of IC's remanufacturing process, specifically the ones flash memories.

E.2 Inspiration

The design of a reuse IC factory was inspired by the two flow charts from the SustainabilitySMART project, see Figures 2.7, 2.9 and also from contact with Retronix Europe and two videos available on the internet, one from Retronix Europe and one from Spirit Electronics [64, 65]. Both companies are active in the reballing of BGA components. Spirit Electronics can process 40K balls per hour [66] and the one of Retronix Europe does not submit the IC to thermal reflow process [66].

E.3 Hypothesis

Important hypotheses and limits of the study are listed and discussed here under :

- Remunfactured IC are UFS or eMMC flash memories with a BGA159 electrical interface. This interface is the most used for both standards.
- Inputs of the process are Wasted PCB provided by an external provider. It means that the dismantling of E-waste is not done in-house.
- The PCB are sorted so that each input PCB contains an eMMC or UFS chips. Therefore, each PCB entering the process contains one chip to be remanufactured. Note that this hypothesis can be relaxed by adding a computer vision system that sorts PCB with or without eMMC and UFS. As a result, the error induced is low regarding LCA and close to zero regarding cost calculation. Indeed, the resale price

of unused PCBs can be considered equal to the purchase price, and regarding the result of the cost study, the cost of the sorting process will be small.

- PCB size is assumed to be 25 by 30 centimeters. This hypothesis was chosen without an in-depth study of the computers or smartphone PCB dimensions but based on some observations. Smartphone PCBs have a side length between 3 and 10 cm. The ATX computer gaming PCB has a dimension of 22.96x30.48cm. The ATX computer gaming PCB can be considered a big PCBs for computer [67]. Of course, it exists bigger PCB, but there are exceptions. Many PCBs are smaller than 25x30 cm, for example, laptop PCBs, smartphone PCBs, and SSD PCBs. Therefore this dimension seems to be a good estimation for taking one of the worst cases for the application without basing the estimation on outliers.

E.4 Process flow

The remanufacturing process flow chart used for this research can be found in Figure E.1. For each part of the process, a match between the needs of the process and a machine found on the internet is done. Machine used and their parameters are listed in Table E.1. Their parameters are considered for the LCA and the back-of-the-envelope cost calculation. Detail of each part of the process, as well as the calculations and assumptions made for completing the Table E.1 are presented in this section.

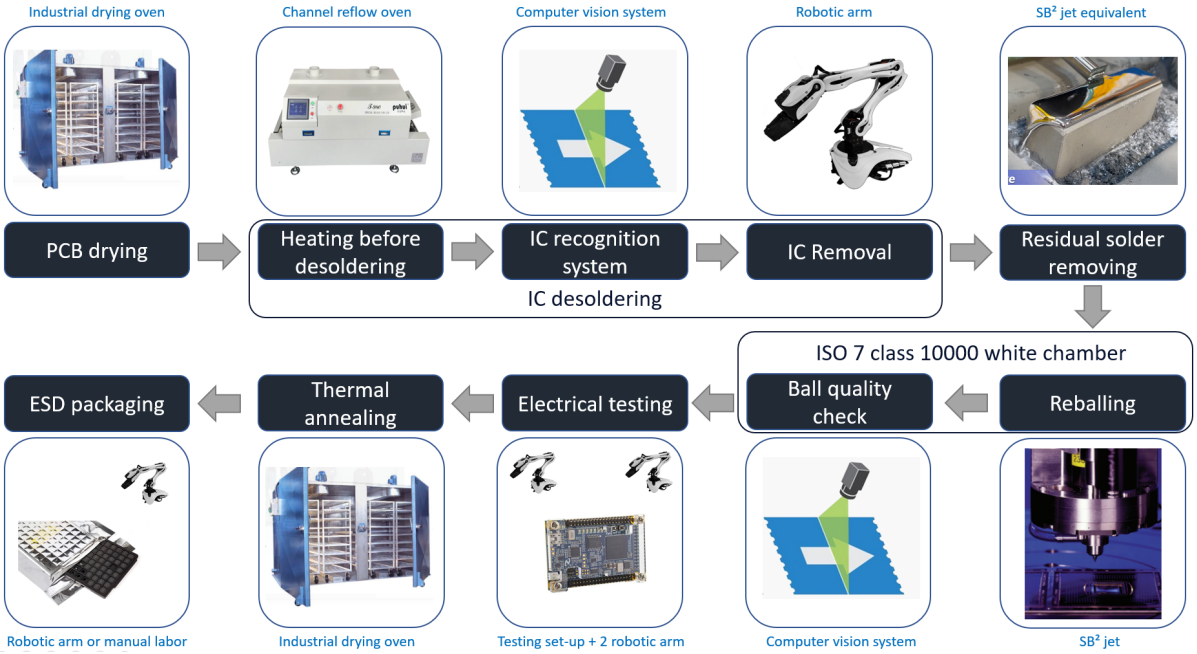


Figure E.1: IC remanufacturing process with selected machines for each step

PCBs drying PCBs need to be baked to remove the moisture present in packages to avoid damage in the rest of the process. A bake time for one hour at 150 degrees is sufficient [42]. For this operation, ESSA model DO06 oven was chosen [68]. The oven has an internal dimension of 3.2 m wide, 1.7 m deep, and 1.9 m high, which gives a volume of 10.1 m^3. The oven is delivered with two nine-stage trolleys for a total surface on the trolleys of 30.1 m^2. PCBs are dried vertically with the 25 cm side at its base (reminder:

the dimensions of the PCBs are set as equal to 25 by 30 cm). The PCBs height is 30 cm which is bigger than the trolley's floor-to-floor distance. Therefore, one stage of two is kept. It remains only four stages with a remaining area of 13.37 m^2. Assuming that 6 PCBs rows of vertical PCB can be placed in the deep and that we space each PCB 8 cm from each other, the oven's capacity is 589 PCBs (See Annex for schematic). The weight of the oven is assumed to be equal to 120 kg, the accurate weight was not available, but one of a similar oven is used [69]. The power consumption of the oven is equal to 24 kW.

IC desoldering For the desoldering process, ICs need to be heated with a precise reflow temperature pattern to melt the solder tin without damaging the ICs, and after a robotic arm comes pic, the ICs [42]. To avoid manual ICs recognition, a computer vision system must be placed before the robotic arm to recognize the valuable ICs automatically. To optimize the process flow, it was chosen to make that in series thanks to a conveyor oven. The oven selected is the channel reflow oven T980 manufactured by Puhui [70]. This oven is initially destinated for the ICs soldering process on PCBs. It contains eight temperature zones to control the temperature of the PCBs with a precision of +- 2 degrees. The total heating zone has a length of 116 cm and is 40 cm deep. Regarding our assumption, the oven is too narrow to put two PCBs next to each other. The different heating zone allows to program the reflow temperature pattern needed for the desoldering and to reach to temperature for desoldering at the output of the oven. Once at the output, the IC recognition system would recognize UFS and eMMC ICs, and one robotic arm would pick up the ICs. This system is similar to the SustainabilitySMART project. The significant difference is that our system is serialized through a conveyor belt oven. It allows us to go from a desoldering time of 75 seconds per IC for SustainabilitySMART to 16 seconds for our configuration [42].

The process is questionable. Indeed, the recognition system and the pic up of ICs need to be done while the PCBs are in the heating zone. The system is close to reality, and errors induced by this simplification are negligible.

The robotic arm chosen is the Meca500 Six-Axis Industrial Robot Arm from Mecademic [71]. An arbitrary hypothesis was made that the maximum repetitive robotic movement duration was 15 seconds, meaning that every robot can manage 240 ICs per hour. Its weight is equal to 4.5 kg, and Mecademic did not communicate its power consumption. Therefore the power consumption of a more miniature robot was taken, i.e., 60 W for a robot that weight 0.65 kg [72]. This power was reduced to the weight of the robot used. Therefore, the robot's power consumption is assumed to be equal to 415 W.

The computer vision system comprises 4 cameras for a total power consumption equal to 0.048 kW (12 W each [73]).

It is assumed that a computer with a power consumption of 0.07 kW is required to analyze the video stream for the IC desoldering phase and the ball quality check phase. This power consumption is the one of computer gaming [74].

Residual solder removal is the process of cleaning solder residue on desoldered ICs. An industrial residual solder removal process can be seen on the video of Spirit electronics

[65]. It can be resumed by one robotic arm which takes the ICs, baths the ICs in flux, heats the ICs with hot air, and finally, the connexion face of the ICs is rubbed against molten tin excited at an ultrasonic frequency to remove residual solder (like an ultrasonic bath but with molten tin). The process is modeled, with one robotic arm, one hot air heater, one sort of vibrant heating stone (Assumed to have the same impact and electricity consumption as the hot air heater), and flux consumption and tin consumption.

The same robotic arm than for desoldering and the ZX3000 CE ROHS for the hot air heater [75] are used. It is assumed that tin and flux consumption are respectivelly equal to 0.5 g and 0.14 mL per IC. These consumptions represent the mass of the tin or flux for a volume equal to 50% of an eMMC chips volume, i.e., 11.5x13x1mm.

Reballing For the BGA reball phase, SB^2-Jet solder ball attach system produced by PadTech is used [76]. The system can place up to 10 balls per second. Solder ball reflow is performed by a laser system using an infrared wavelength which allows for avoiding one thermal reflow process.

As a reminder, UFS and eMMC are assumed to have a BGA159 electrical interface. Therefore 159 solder balls need to be placed per IC. Assuming that for each IC, the reball time is 16 seconds and the displacement time between the ICs storage rack and process zone is 4 seconds. The throughput of the BGA reball machine is 180 IC per hour.

ISO 7 clean room According to the Padtech datasheet, the SB^2-Jet needs to be in an ISO 7 class 10000 clean room [76]. It is assumed that the volume of the clean room is equal to 40 m^3 (4x4x2.5m). The air change rate for an ISO 7 clean room is 100 times per hour, and the efficiency for the recirculation air unit is 2718.41 m^3/kW [77, 78]. So the power consumption of a 40 m^3 ISO 7 clean room is given by :

$$P = \frac{40 \cdot 100}{2718.41} = 1.47\,\text{kW} \tag{E.1}$$

The weight of the wall of a clean room is equal to 2.2 kg/m^2 [79], and the weight of clean room ventilation and filter system is assumed to be equal to 280 kg [80] (Note that the weight given by the source were arbitrarily multiplied by four to take security margin). The total weight of the clean room is assumed to be equal to 259.2 kg.

Ball quality check Ball quality check is done through an automated computer vision system. This system is assumed to be composed of two cameras with a power consumption of 0.012 kW each and a computer shared with the IC recognition system included in the ICs desoldering part with a power consumption equal to 0.7 kW [74].

Testing phase Electrical testing phase is there to assess if ICs are in a state of working. It is assumed that the electrical test lasts two minutes per IC with the same set-up used for this master thesis. Therefore, eleven tests set up are needed to assume the ICs flow. 4 Robotic arms move the ICs from the storage rack to the testing socket and vice versa.

The power consumption of one test set-up composed of one DE10-Nano equals 10W [81].

Thermal annealing Thermal annealing is done according to the results of this thesis, i.e., bake at 250 degrees for 12 hours. The oven used is a second unit of the ESSA model DO06 oven [68].
The surface of the trolleys is equal to $30.1m^2$. The surface of one eMMC or UFS is $155.25\,\mathrm{mm}^2$ (dimension of eMMC or UFS 11.5x13mm). It is theoretically possible to fill the oven with 194525 eMMC or UFS if they are placed next to each other. It is assumed that the thermal annealing process simultaneously takes place with 20000 ICs.

ESD packaging ESD packaging means electrostatic discharge packaging, a standard in the electronics industry. This is assumed to be done manually or with the help of a robotic arm.

Table E.1: Machines and there parameters present in the remanufacturing process. When "+++" is in the IC throughput it means that the IC throughput is limited by another machine.

Machine	**Weight [kg]**	**Power [kW]**	**IC throughput [#/h]**	**GWP Method Label**
Industrial oven: drying	120	24	589	Ademe
Industrial oven: thermal ann.	120	24	20000	Ademe
Robotic arm	4.5	0.415	240	Ademe
Vision system: Camera	NA	0.12	+++	Mac book air
Vision system: Computer	NA	0.7	+++	Mac book air
Channel reflow oven	130	7	223	Ademe
Hot air heater	1.1	3.3	+++	Mac book air
Heating and vibrating stone	1.1	3.3	+++	Ademe
SB^2 jet	800	3.52	188	Ademe
Testing set-up	NA	0.01	30	Mac book air
Clean room	359.2	5.51	+++	Ademe

Appendix F

Methodology for the lifecycle assessment comparison between old and reused flash memories

Different IC will be taken into account for the lifecycle assessment comparison. The IC from the two groups will be compared. First, the group of last generation off-the-shelf flash memories, and second reused old flash memories recovered from old PCBs.

F.0.1 Methodology for the GWP calculation for the remanufacturing process and the PCB transport

F.0.2 Hypothesis

For the remanufacturing process GWP calculation, the hypothesis issued at Section E are applied, and the following hypothesis are also taken into account :

- Impact per entering IC will be calculated from the process flow on Figure E.1. It is assumed that each entering IC goes through all the process steps. Indeed, the ICs are tested quasi at the end of the process.
- The impact of ICs, which are not discarded after the test phase, is given thanks to this equation.

$$GWP_{IC} = \frac{100 \cdot GWP_{entering\_IC}}{\alpha} \tag{F.1}$$

  Where α is the yield, i.e., the fraction of ICs that are not discarded during the remanufacturing process expressed in percentage. GWP_{IC} is the final GWP emission of the ICs that will be sold as second-hand ICs.
- Remanufacturing impact between chips is assumed to be the same for all the entering IC. This hypothesis is based on the fact that all eMMC and UFS have the same mechanical dimension and electrical interface regardless of the manufacturer, the year of production, or the storage capacity.
- Manufacturing impact of machine used is accounted for. The methodology used depends on the label of the column "GWP Method Label" in Table E.1. If the label is "Ademe" the methodology is the one proposed by Ademe (Agence De l'Environnement et de la Maîtrise de l'Énergie): "For machines (and production lines), a default value

is proposed: it corresponds to the emission factor for vehicle manufacturing, i.e. 5500 kgCO2eq of the machine." [82]. If the label is "Mac book air," the methodology is the following: The GWP emissions are considered equal to the production and transport emissions of an Apple 13-inch MacBook Air, i.e., 135 kgCO2eq [83].
This method is not accurate but allows giving an order of magnitude.

- The impact of the manufacturing process is amortized for five years of usage. For example, if the production impact of machine A is 5000 kgCO2eq and his output throughput is 1000 IC per year. The impact of the manufacturing of the machine per IC is 1 kgCO2eq.

- For the LCA, it is assumed that the factory is working at a constant speed of 180 IC per hour and 16 hours per day for five days per week, so 260 days per year. It results in the production of 3744000 remanufactured flash memories in 5 years.

- Infrastructure impact is accounted for. It is assumed that the remanufacturing process takes place in an industrial hangar with a surface of 400 m^2 and is of energy class D (GWP emissions from 21 to 35 kgCO2eq/m^2 yearly, the biggest value is retained). It is assumed that the life span of the hangar is 50 years of which 5 years are for the remanufacturing of IC and his construction impact is equal to 825 kgCO2eq/m^2[82].

- PCB transport GWP emissions between the PCB supplier and the remanufacturing factory assumed to be located in Louvain-la-Neuve, Belgium, are accounted for. Three PCB suppliers' locations will be studied, Mechelen, Ankara, and Shangai. Parameter are listed in Table F.1.
  The PCB weight is a trade-off between different PCB weights. Smartphone PCBs weights 17 g (Measure done), one example of a Notebook from HP PCB weights 368.5 g [84] and the "CM MSI *A320M-A Pro Max*" (Computer gaming PCB) weights 700 g [85]. It is important to notice that flash memories can also be embedded on SSD PCBs with a weight closer to the smartphone PCB weight. With all this observation, it was decided to assume that the weight of one PCB equals 350 g.

Table F.1: PCB delivery impact parameter

Parameter	**Value**	**Unit**
Truck distance, Mechelen	50	km
Truck distance, Ankara	3121	km
Truck distance, Shangai	72	km
Cargo boat distance, Shangai	22000	km
GWP emission factor truck	0.12 [82]	kgCO2eq/(T km)
GWP emission factor cargo boat	0.0018 [82]	kgCO2eq/(T km)
Mean weight of one PCB	0.35	kg

- ESD packaging impact is taken into account, the weight of the packaging of the IC received for RS-component for this master thesis was equal to 8 IC^{-1}. The GWP used for the ESD packaging is the one of primary polystyrene, i.e., 2.824 kgCO2eq kg^{-1}.

- Electricity consumption is considered, and each step is supposed to work at a constant speed of 180 IC per hour. Exceptions are made for PCB drying and thermal annealing steps. The capacity for each operation is 550 IC for the PCB drying step and 10000 IC for the thermal annealing step.
- Electricity mix used is the one of Belgium, i.e. 0.168 kgCO2eq/kWh [86].

F.1 Methodology for the GWP calculation for the manufacturing process

The manufacturing process of new flash memories includes the back-end and front-end processes. The GWP of the manufacturing process, expressed in kgCO2eq GB^{-1}, was extracted or calculated thanks to three methods listed here.

F.1.1 The Prakash method

The first method used is based on the database done by Prakash. et al. in 2013 [87]. Thanks to the data available, the GWP of flash memories production per GB can be calculated thanks to the following equation:

$$GWP(loc, cap) = \frac{PFC_{gwp} + \sum_{Pr\in\{front,back\}} \sum_{E\in\{elec,gaz\}} E(E, Pr) \cdot GWP(E, loc)}{0.01 \cdot A(cap) \cdot cap} \tag{F.2}$$

Where loc is the location where the fabrication takes place. Where cap is the capacity of the flash memory in GB. Where PFC_{gwp} is the PFC emissions of the front-end process in kWh/cm^2 per error free chips. Where $E(E, Pr)$ is the energy type E input for the process Pr, where $GWP(E, loc)$ is the energy type E emission factor in the location *loc*. Where A is the area of the dye in mm^2 in function of the storage capacity *cap*.

Thanks to the parameters listed in Table F.2, nine data points were generated.

The parameters are composed of :

- Electricity and natural gas input in front-end and back-end processes expressed in kWh/cm^2 of error-free chips and PFC emission of the front-end processes expressed in kgCO2eq/cm^2 of error-free chips.
- Dye area for 32, 64, and 128 GB of flash memories published by the Fraunhofer IZM [88] (see Table F.2).
- The GWP emission of electricity in function of the location. Three production places are studied Grenoble, Korea, and California. Korea and California were chosen because the market shares of memory production in Korea and the US are respectivelly 63% and 23% in 2020 [89]. Grenoble is the third production place considered. This place was chosen because a state-off-the-art semiconductor factory will be built in this region. Moreover, it allows having an example where the GWP of electricity is small [90].

Table F.2: Parameters used for the Prakash method

Criteria	**Quantity**	**Unit**	**Source**
Electricity input front-end processes	1.27	kWh/cm^2 per error free chip	Prakash, 2013 [87]
Natural gaz input front-end processes	0.16	kWh/cm^2 per error free chip	Prakash, 2013 [87]
PFC emissions of the front-end processes	0.252[1]	kWh/cm^2 per error free chip	Prakash, 2013 [87]
Yield of flash memories production	75	%	Boyd, 2011 [54]
Electricity input back-end processes	0.55	kWh/cm^2 per error free chip	Prakash, 2013 [87]
Natural gaz input back-end processes	0.07	kWh/cm^2 per error free chip	Prakash, 2013 [87]
Dye area for 32GB flash memories	200	mm^2	Fraunhofer IZM, 2021 [88]
Dye area for 64GB flash memories	300	mm^2	Fraunhofer IZM, 2021 [88]
Dye area for 128GB flash memories	400	mm^2	Fraunhofer IZM, 2021 [88]
GWP of electricity mix for Grenoble	0.119	kgCO2eq/kWh	Electricitymaps, 2022 [**?**]
GWP of electricity mix for California	0.313	kgCO2eq/kWh	Electricitymaps, 2022 [**?**]
GWP of electricity mix for Korea	0.481	kgCO2eq/kWh	Electricitymaps, 2022 [**?**]
GWP of natural gaz	0.427	kgCO2eq/kWh	Riva et. al., 2006 [91]

F.1.2 The iPhone method

the second method uses data available in four LCA done by Apple. Life cycle assessments are respectivelly done for iPhone X (2017), iPhone 11 (2019), iPhone 12 (2020), iPhone 13 Pro Max (2021) [92, 93, 94, 95].

Data available in each LCA give the GWP of an iPhone model in function of the storage capacity. The GWP variation between two iPhones from the same model with different storage capacities is mainly induced by the difference in the manufacturing of embedded flash memories. Therefore by making some simplification, the GWP of flash memories manufacturing per GB can be deduced from it using this equation :

$$GWP(Gen, Cap_1, Cap_2) = \frac{GWP_i(Gen, Cap_2) - GWP_i(Gen, Cap_1)}{Cap_2 - Cap_1} \tag{F.3}$$

Where Gen ,Cap1, Cap2 are respectively the generation of iPhone, the capacity of the first iPhone, and the capacity of the second iPhone, where $GWP_i(Gen, Cap)$ is the emission in kgCO2eq of one iPhone of generation Gen and with a storage capacity of Cap.

However, this first approximation has its limits. Firstly, assumptions made by Apple are not available. Therefore it is complicated to be more precise. Second, some simplifications are done, including the fact that the change in power consumption between two flash memories is neglected. However, neglecting the power consumption change is not a significant error. The use phase is less than 20 percent of the GWP impact of flash memories, and power consumption increases between 48 and 50 percent when doubling the memory capacity for the Samsung KLMCG2UCTA eMMC5.1 (see Table F.3). It seems to be an upper margin compared with other flash memories where this increase is limited to 16 %. Using this formula, it is assumed that the difference in the GWP impact between two iPhones of the same generation is only induced by the difference in the manufacturing impact of flash memory with different storage capacities.

F.1.3 The Samsung VNAND method

is an LCA from Samsung that gives the GWP per GB of 64 GB VNAND chips done in 2018. This number can be used but keep in mind that the methodology and scope are unknown.

F.2 Methodology for the use phase GWP emission

The GWP of the use phase of flash memories depends on the power consumption and the read and write speed of flash memories.

F.2.1 Use phase scenarios

GWP of flash memories is studied in function of two use phase scenarios.

First, the energy consumed is deduced from the number of GB written in the use phase. GB read must also be accounted for. They are accounted for as a number of GB read per GB written. For example, if 5 GB are read for each GB written, the total energy consumed per GB written is one times the energy of one GB written and five times the energy of one GB read. This gives eMMC4.5 a total energy per GB written of 48 J.
This method assumes that the ICs are embedded in a system designed to take advantage of the best flash memory parameters and optimized to use the less energy possible. This is utopic, and the flash memories will never be used in such conditions. Therefore a second method is required.

The second method is the same as used in [54]. It means using flash memory for 6 hours per day in the active state and deducing the energy used from it. The active power taken into account is the biggest one between read and write active power for second life flash memories and the smallest one for new flash memories.

F.2.2 Flash memories parameters

Flash memories used are listed on Table F.3. Flash memories references for eMMC4.5 and eMMC5.1 were found while dismantling the Samsung Galaxy S3 4G edition and the Samsung Galaxy A10. Note that the eMMC5.1 will be counted as both a second-hand IC and a new IC

The parameters for the UFS4.0 flash memory were found from values found on the Samsung semiconductor website. This last flash memory was chosen because it is the last generation of flash memories from Samsung and, therefore, the one with the best features. The power consumption is not known, but Samsung declares reaching 6.0MBps per 1 of sequential read speed [55]. The voltage level of the IC remains unknown. Therefore, one hypothesis must be posed to complete the $P = V \cdot I$ relation. Flash memories have two voltage levels, the one for the controller and the one for the flash cells array. For UFS 3.0 of Samsung, these voltages were respectivelly 1.2 V, and 2.6 V [56]. To take the worst case into account

regarding the research field, it is assumed that the voltage level for UFS4.0 is equal to 1.2 V.

Table F.3: List of KLMxGxGEAC

Flash memory	**Manufacturer**	**Used in**	**Type**	**Capacity [GB]**	**Write/ read Speed [MB/s]**	**Active power read/write [mW]**	**Write/Read energy per GB [J/GB]**
KLMxGxGEAC	Samsung	Samsung Galaxy S3 4G edition, 2013	eMMC4.5	64	50/150	1029/1029	18/6
KLMCG2UCTA	Samsung	Samsung galaxy A10, 2019	eMMC5.1	64	200/300	984/984	4.3/2.6
\\	Samsung	State-of-the-art, 2022	UFS4.0	64	2800/4200	560/840	0.17/0.17

chapterMethodology for economic viability study The economic study aims to make a back-of-the-envelope cost calculation for the ICs remanufacturing process. This cost will be put in perspective with the selling prices of eMMC5.1 flash memories in the market. More information about the methodology for each sort of flash memories are presented in Sections F.4 and F.3.

F.3 Methodology for off-the-shelf eMMC5.1 flash memories prices

Price of eMMC5.1 flash memory from different sources where analyzed and classified into different capacities and bit per cell criteria. The different types of eMMC5.1 flash memories retained for the study are : 8GB MLC, 16GB MLC, 32GB MLC, 16GB TLC, 32GB TLC, 64GB TLC, and 128GB TLC.

Market analysis platforms First, the market analysis platforms are platforms that monitor the price of goods in one specific field. The most famous one for flash memories is the DRAMeXchange website which monitors market trends and prices each day. This platform is paid. Non-subscriber users have only access to a reduced amount of data. Therefore some of them are missing. To compensate for the lack of accessible data on DRAMeXchange, the data from elinfor.com at the date 11 May 2022 were used. Three prices are given on such a platform: the smallest of the day, the first price of the day, and the last price of the day. The lowest prices were retained. This choice was made because reused ICs will compete with low-cost electronics components.

Marketplace The second type of data is prices from different marketplaces. The marketplaces were analyzed for each type of eMMC5.1 flash memory, and the lowest prices were retained. The marketplaces analyzed were RS-component, Arrow, Farnell, Digi-Key, and Mouser Electronics.

F.4 Methodology for remanufactured flash memories

To study the viability of IC reuse, the cost of the IC remanufacturing process will be studied according to the process flow discussed in section E. This cost will

Hypotheses The objective is to make a back-of-the-envelope calculation of the cost of the process of IC remanufacturing and to analyze the most influential parameter. The process flow and hypothesis discussed in section E are used. The following hypotheses are also taken into account :

- Cost per entering IC will be calculated from the process flow on Figure E.1. It is assumed that each entering IC goes through all the process steps.
- The yield is taken into account, i.e., the fraction of ICs that are not discarded during the remanufacturing process expressed in percentage. Therefore the remanufacturing

cost for IC which are sent back on the market can be computed with the following equation

$$C_{salable} = \frac{100 \cdot C_{entering}}{\alpha} \tag{F.4}$$

where α is the yield in percentage, $C_{salable}$ and $C_{entering}$ are respectivelly the cost for salable IC and cost for IC entering the process.

- The lease of an 500 m^2 industrial building is considered and is assumed to be equal to 5000€/month. The rent of a 305 m^2 industrial building in Louvain-la-Neuve is 3650€(see Annexe).
- The employee charge is estimated at 44422.08€ yearly [96].
- The electricity price is assumed to be equal to 0.33€/Kwh [97].
- Flux, solder tin, and Soldar ball price are assumed to be equal to 14.3€/l, 20.9€/kg and 15.46€/250000 balls [98, 99, 100].