FFHS

Fernfachhochschule Schweiz
Zürich | Basel | Bern | Brig

Mitglied der SUPSI

Disposition Masterthesis

Master-Thesis

im Studiengang MAS Smart Engineering and Process Management (Industrie 4.0)

von

Silvan Zahno

Eingereicht beim Referenten Dr. Sharam Dadashnia

am 12.02.2021

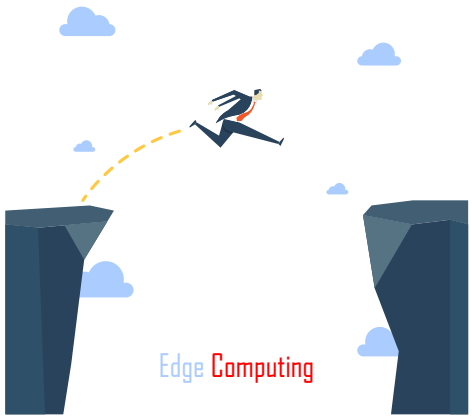
FFHS
Fernfachhochschule Schweiz
Zürich | Basel | Bern | Brig
Mitglied der SUPSI

Masterthesis

Disposition Masterthesis

The way to efficient edge computing for industry 4.0

Master of Advanced Studies in Industry 4.0



Created by: Silvan Zahno

Register Number: 10-964-112

Phone: +41 79 665 41 75

Mail: silvan.zahno@students.ffhs.ch

Professors: Dr. Sharam Dadashnia

Submission Date: 12.02.2021

Version: v1.1

FFHS
Fernfachhochschule Schweiz
Zürich | Basel | Bern | Brig

Mitglied der SUPSI

Departement Wirtschaft & Technik

Antrag Weiterbildungsmaster-Thesis

Dieser Antrag ist vom Studierenden auszufüllen als auch zu unterschreiben und anschliessend an den Referenten zur Genehmigung weiterzuleiten.

Studierende/r (Antragsteller/in):

Name, Vorname: Zahno Silvan

Anschrift: Hannigstrasse 6 3930 Visp

E-Mail: silvan.zahno@students.ffhs.ch

Studiengang/Jahrgang/Ort/Gruppe: MAS Industrie 4.0

Datum: 20.01.2021 Unterschrift:

Titel der Master-Thesis (Disposition siehe Anlage):

☐ Ja, die Master-Thesis soll als **streng vertraulich** eingestuft werden.

(Bitte beachten: Nur gegen eine Sperrgebühr von CHF 500.-- möglich.)

Sprache der Master-Thesis: Deutsch ☐ Englisch ☑

(Bitte beachten: Die mündliche Prüfung wird nur in Deutsch abgenommen.)

Referent:

Name, Vorname: Sharam Dadashnia

Anschrift: Wilhelm-Leuscher-Strasse 1 DE-66352 Grossrosseln

E-Mail: sharam.dadashnia@ffhs.ch

Antrag vom Referenten genehmigt:

Datum: 11.01.2021 Unterschrift:

Der vom Referenten genehmigte Antrag ist zusammen mit der Disposition der Master-Thesis termingerecht per E-Mail einzureichen an den jeweiligen Studiengangsleiter und das jeweilige Prüfungssekretariat:

EMBA SGL claudia.stadelmann@ffhs.ch / Prüfungssekretariat tatjana.schmid@ffhs.ch

MAS Industrie 4.0 SGL david.gemmet@ffhs.ch / Prüfungssekretariat carole.schnydrig@ffhs.ch

MAS Arbeit 4.0 SGL annina.coradi@ffhs.ch / Prüfungssekretariat kim.zenhaeusern@ffhs.ch

MAS Digital Education SGL marc.garbely@ffhs.ch / Prüfungssekretariat marina.blum@ffhs.ch

Der Antrag wird sodann von der Studiengangleitung in der dafür vorgesehenen Zeitspanne überprüft (die Daten entnehmen Sie bitte dem Terminplan für die Master-Thesis). Zu einem späteren Zeitpunkt erhalten Sie die Informationen über den Co-Referenten.

1 Introduction

1.1 Title

Edge Processing Unit (EPU) - Hardware accelerated edge computing for Industry 4.0.

1.2 Initial situation

Cloud computing was and still is a big topic for the Industry 4.0, during these last years it was generally thought that processing of data into the cloud was beneficial from a development, infrastructure as well as a cost standpoint. But there are some tasks, especially when high data-rate or realtime capabilities are needed, where cloud computing is non optimal. The development of communication technologies like 5G and others, helped extending the use case of cloud computing further, but in the last years a trend back to on-site computing with a looser connection to a cloud system could be seen (Panetta, 2017). Such processing are called edge or fog computing. This field is still quite new and a lot of possibilities lies in a the combination of cloud and edge processing (Katz, 2020). The term Edge Processing Unit (EPU) has not been coined yet, therefore I would like to coin this term for specialized on-site hardware which can include FPGA or ASIC chips harboring coprocessors henceforth called Hardware Accelerators.

1.3 Problem

Most companies which started using Industry 4.0 will own or rent cloud infrastructure such as Amazon AWS, Microsoft Azure, Google Cloud, etc.. Cloud processing is limiting for some types of processing. Especially for image analysis with or without machine learning, the high data rate of video feeds is limiting and the transfer can be unstable to the cloud. Additionally for real-time and security application on-site processing is still required. Typically such systems require costly on-site processing platforms normally with a combination of CPU and GPU chips. Hardware Accelerators such as FPGAs can highly improve efficiency and bring costs down (*Introduction to FPGA Acceleration*, 2020). But the complexity of development tools and lack of development knowhow are hindering its use outside of cloud platforms (Baker, 2015) (Higginbotham, 2020).

1.4 Research Question

What edge computing technology can industrial companies use for flexible on-premise processing.

This research question is linked to various thematic areas such as:

- Cloud Computing: this is an Information Technology (IT) infrastructure that is made available through the internet. It usually includes storage space, computing power or application software as a service.
- Edge Computing: in contrast to cloud computing, refers to decentralized data processing at the border of the network, the so-called edge.
- Fog Computing: also known as fogging refers to an architecture which uses edge devices in combination with cloud computing. Part of the computation especially which uses a high bandwidth is done at the edge. High-level computing is send to a cloud infrastructure via the data plane and analysed further. Fog computing is especially used in big data use cases.
- Machine to Machine Communication (M2M): communication between machines and system is essential for Industry 4.0
- Big Data: Big data is a field that systematically analyzes and extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional data-processing application software.
- Artificial Intelligence (AI): The term ”artificial intelligence” is often used to describe machines (or computers) that mimic ”cognitive” functions that humans associate with the human mind, such as ”learning” and ”problem solving”.
- Machine Learning (ML): Machine learning is a subfield of AI, used since the term ”artificial intelligence” is often used to describe machines (or computers) that mimic ”cognitive” functions that humans associate with the human mind, such as ”learning” and ”problem solving”.

2 Form

This thesis will be a project work in collaboration with an industrial use case. A limited system of one practical use case will be identified and implemented using an edge computing device including a hardware accelerator. The developed solution will be tested in an industrial environment on a remote controlled machine.

The company called Syrto is based in Steg, Valais and has agreed to showcase the solution on one of their many industrial applications currently running at customers. The specific use case will be identified during the thesis. In the subsection 2.1 the company will be showcased.

During the project an agile work methodology called Scrum is used. The prototype development will pass through 7 different stages: Conception, initiation, analysis, design, construction, testing, deployment. The described workpackages are split into smaller tasks labeled: Theoretical Research, writing, development and implementation. These cards are placed in a virtual Kanban board which allows for an agile development cycle. According to the project phases research & synthesis, development and integration & testing the tasks are placed within the Kanban sections backlog, sprint, doing, verify and done.

2.1 Syrto AG

Syrto AG connects the old world with the new digital one and thus creates opportunities for its customers to carry out work more cost-effectively and efficiently. It changes the way of working of tomorrow and sets with its technologies and concepts the conditions for the future growth of its customers and the satisfaction of their employees. With the concept of Syrto AG and the offered technologies the clients can tap the enormous potential of productivity, quality and speed. Syrto AG analyzes the processes, workflows and procedures for the achievement of the customers goals and additionally offers various solutions in teleoperation and telerobotics, image evaluation and automation for cost reduction. They are working with highly motivated mechanical, electrical, process and industrial engineers. They work pragmatically, goal-oriented and coordinate the activities closely with their clients. The customers of Syrto AG are companies that want to reduce costs in their operational processes. The customers of Syrto AG see the benefit in digital networking and automation and take calculable risks to implement projects.

3 Objectives

The thesis has multiple SMART objectives:

- Overview about the edge computing and its use cases during workpackage 3
- Overview about hardware accelerators and its use cases during workpackage 3
- Comparison of computing power, cost as well as power efficiency with and without hardware accelerators.
- Comparison and implications of datatransfer between edge to a cloud system with and without hardware accelerators.
- One practical showcase of edge computing realized during the workpackages 4-7.
- Using a combination of traditional hardware as well as a custom Hardware Accelerator implemented in a FPGA
- The processing hardware should be cheap (<300CHF)
- 2x energy reduction compared to a non hardware accelerated on-side processing platform
- 4x data-rate reduction compared to a cloud processing platform

4 Limitation

The thesis will showcase one example of image analysis, it can not be used as implementation guide. The end product is a proof of concept and is not intended to be used in an industrial environment.

The research question can be limited to edge devices to lower bandwidth, energy consumption and maintenance costs while performing image analysis and maintaining a cloud connectivity.

5 Research Methods

The source information will be gathered with help of the course documents and many internet researches as well as systematic literature research. The data will be categorized into three groups:

- Primary resources - high level of detail

  - Reports
  - Thesis
  - Company reports
  - Conference proceedings
- Secondary resources - medium level of detail
  - Journals
  - Books
  - Newspapers
  - Publications
- Tertiary resources - low level of detail
  - Indexes
  - Databases
  - Encyclopedias
  - Bibliographies

The sources for the literature review will be from IEEE Xplore, Science Direct, Springer Link, Google Books, Gartner, ProQuest, JSTOR. In order to approach the topic from a high level and to gain a better understanding of the various branches in the area of edge computing, machine learning and hardware accelerators, a potential analysis for the individual techniques will be made. Each subject will be looked at in the context of an industrial application. At last, the evaluation of methods will be applied to an example of a Swiss SME company. The literature search strategy will be defined in the beginning with selecting different Keyword combinations. The word combinations will be described the the thesis document.

6 Structure

This is only a preliminary structure of the thesis and is subject to change.

1. Preface
2. Management Summary Keywords
3. Introduction
   (a) Syrto

(a) Acronyms Glossary

(b) Glossary

(c) Declaration of Authorship

7 Current References

Following the literatures researched used for this disposition. (Putam, 2017) (Ertel, 2013) (Maier, 2019) (Hägele, Nilsson, Pires, & Bischoff, 2016) (Lea, 2018) (Cooper & Hedges, 1993) (Biggio & Roli, 2018) (Fettke, 2006) (Venable, 2010) (Katz, 2020)

8 Workplan

The projectplan is split in three main section: research, development and integration. In the workpackage 1 and 2 an in-depth research will be conducted about existing research, technologies and the market. The methodology for this review will be based on (Fettke, 2006). The milestone ”Research Done” indicates the end of the first phase. Afterwards, approximately 2 months are reserved for the search of a use case within the company Syrto and the development of the prototype. This work will be done in the workpackages 3-6. The last two weeks are planned for finishing the thesis and performing on-site testing within a real industrial use case.

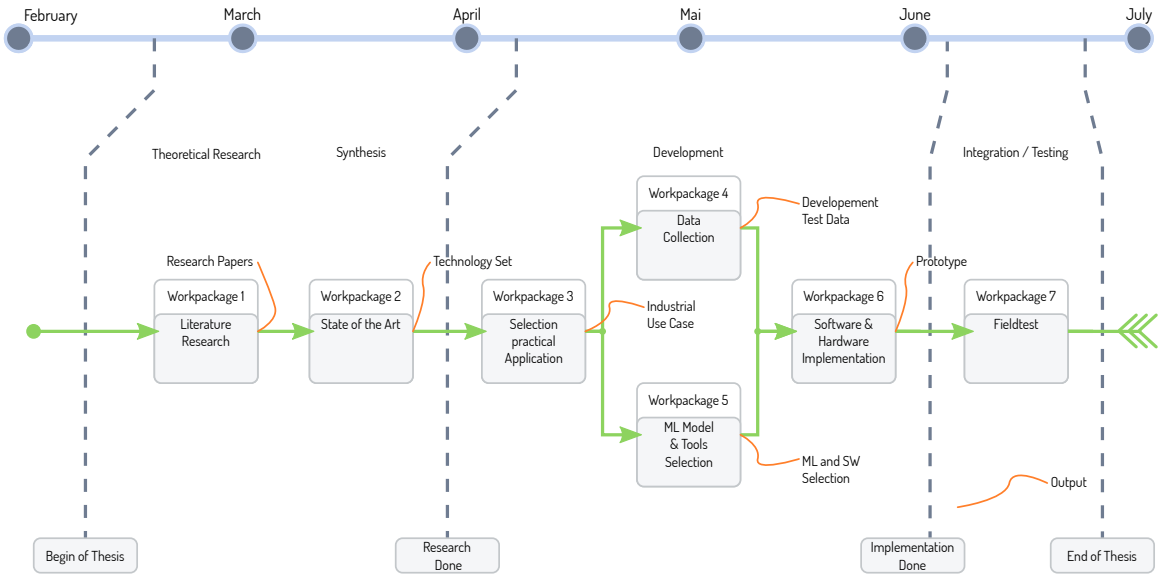


Figure 1: Project Plan

From the project plan, a gantt diagram was created to show the detail and timing of each work-package as well as milestone. The masterthesis will start on February 19^{th} 2021 and will end on June 25^{th} 2021.

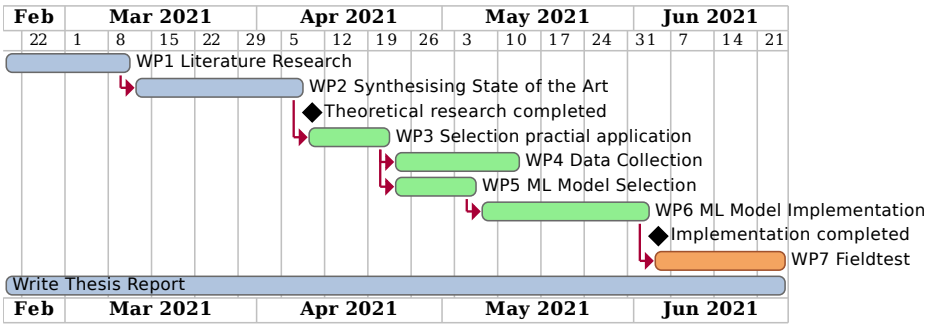


Figure 2: Gantt Chart

Silvan Zahno

References

Baker, L. M., Liana B. (2015, June). Intel to buy Altera for $16.7 billion in its biggest deal ever. *Reuters*.

Biggio, B., & Roli, F. (2018, December). Wild patterns: Ten years after the rise of adversarial machine learning. *Pattern Recognition*, *84*, 317–331. doi: 10.1016/j.patcog.2018.07.023

Cooper, H., & Hedges, L. V. (1993). *The Handbook of Research Synthesis*. Russell Sage Foundation.

Ertel, W. (2013). *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*. Springer-Verlag.

Fettke, P. (2006, August). State-of-the-Art des State-of-the-Art. *Wirtschaftsinformatik*, *48*, 257–266. doi: 10.1007/s11576-006-0057-3

Hägele, M., Nilsson, K., Pires, J. N., & Bischoff, R. (2016). Industrial Robotics. In B. Siciliano & O. Khatib (Eds.), *Springer Handbook of Robotics* (pp. 1385–1422). Cham: Springer International Publishing. doi: 10.1007/978-3-319-32552-1_54

Higginbotham, S. (2020, February). *Why Intel will spend $16.7 billion on Altera.* https://fortune.com/2015/08/27/why-intel-altera/.

Introduction to FPGA acceleration. (2020, February). https://www.stemmer-imaging.com/en/technical-tips/introduction-to-fpga-acceleration/.

Katz, B. (2020). *The Collection Edge: Harnessing Emerging Technologies for Intelligence Collection* (Tech. Rep.). Center for Strategic and International Studies (CSIS).

Lea, P. (2018). *Internet of Things for Architects: Architecting IoT solutions by implementing sensors, communication infrastructure, edge computing, analytics, and security*. Packt Publishing Ltd.

Maier, H. (2019). *Grundlagen der Robotik*. VDE VERLAG GMBH.

Panetta, K. (2017). *Cloud Computing Enters its Second Decade.* //www.gartner.com/smarterwithgartner/cloud-computing-enters-its-second-decade/.

Putam, A. (2017, May). *FPGAs in the Datacenter | Proceedings of the on Great Lakes Symposium on VLSI 2017.* https://dl.acm.org/doi/abs/10.1145/3060403.3066860.

Venable, J. R. (2010). *(PDF) Design Science Research Post Hevner et al.: Criteria, Standards, Guidelines, and Expectations.* https://www.researchgate.net/publication/221581331_Design_Science_Research_Post_Hevner_et_al_Criteria_St doi: 10.1007/978-3-642-13335-0_8

Acronyms

AI Artificial Intelligence. 3

ASIC Application-Specific Integrated Circuit. 2

CPU Central Processing Unit. 2

EPU Edge Processing Unit. 2

FPGA Field Programmable Gate Array. 2, 5

GPU Graphical Processing Unit. 2

IT Information Technology. 3

M2M Machine to Machine Communication. 3

ML Machine Learning. 3

SMART Specific, Measurable, Achievable, Relevant, Time-bound. 5

SME Small and Medium Enterprises. 6

Glossary

Big Data Big data is a field that systematically analyzes and extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional data-processing application software. 3

Cloud Computing cloud computing is an IT infrastructure delivering computing as a service. Thsi can include networking, servers, memory, computing functionalities such as analytics. Typically only services used have to be payed.. 3

Edge Computing In edge computing, computer applications, data and services are moved away from central nodes (data centers) to the outer edges of a network. In other words, the aim is to process data streams in a resource-saving manner, at least partially on site (e.g. directly at the end device or within a factory), while still benefiting from the advantages of the cloud.. 3

Fog Computing also known as fogging refers to an architecture which uses edge devices in combination with cloud computing. Part of the computation which uses a high bandwidth is done

at the edge. High-level computing is send to a cloud infrastructure via the data plane and analysed further. Fog computing is especially used in big data use cases.. 3

Hardware Accelerator Hardware acceleration describes the relief of the main processor by delegation of special computationally intensive tasks to hardware specialized for these tasks. This technique is used in particular for graphic display in computers. Hardware accelerators can be either speciallized asic chips or reconfigurable FPGA chips. 2, 5

Industry 4.0 Industry 4.0 describes the trend towards automation and data exchange in manufacturing technologies and processes which include cyber-physical systems (CPS), the internet of things (IoT), industrial internet of things (IIOT), cloud computing, cognitive computing and artificial intelligence.. 2, 3

Kanban A Kanban board is a tool for implementing the Agile development method. The board is usually divided into four sections: ”to do”, ”in progress”, ”verify” and ”done”. The objects are moved by staff to the current sections on the board.. 4

Scrum Scrum is an agile process framework for managing complex knowledge work, with an initial emphasis on software development, although it has been used in other fields and is slowly starting to be explored for other complex work, research and advanced technologies.. 4