Review of Novel Computing Architectures for Neural Applications

ADVANCED SEMINAR submitted by

Alexander Kuhn, B.Sc.

NEUROSCIENTIFIC SYSTEM THEORY
Technische Universitüt München

Supervisor: Dr. Christoph Richter
Final Submission: 27.06.2014

TECHNISCHE UNIVERSITÄT MÜNCHEN
LEHRSTUHL FÜR STEUERUNGS- UND REGELUNGSTECHNIK
ORDINARIUS: UNIV.-PROF. DR.-ING./UNIV. TOKIO MARTIN BUSS

2014-03-31

A D V A N C E D S E M I N A R

Review of novel computing architectures for neural applications

Problem description:

The traditional approaches to simulate neural networks are to use either analog VLSI or general purpose computing architectures like (clusters of) CPUs or GPUs. A promising new approach that can be situated in between is the SpiNNaker platform [1]. SpiNNaker endows commercial RISC processor cores with a very fast, special purpose interlink. Its many-core architecture enables massively parallel and asynchronous computing, its communication bus enables the efficient exchange of small data packets. Each processor can simulate a number of spiking neurons, the packets then resemble spikes sent through synapses in between the neurons. Although SpiNNaker is tailor made for this very application, it is just as well able to run other types of neural networks or even physical simulations. SpiNNaker is still a purely academic product. There are, however, a number of commercial products emerging that offer massively parallel computing. Two prominent and very different examples are Intel's Many Integrated Core (MIC) [2] and Adapteva's Epiphany [3] architecture.

The objective of this advanced seminar is to evaluate the hardware capabilities of SpiNNaker with respect to contemporary commercial products built on the MIC or Epiphany architecture, namely Xeon Phi and Parallella. The evaluation is to focus on applications relevant to neuromorphic engineering: e.g. robotics, artificial- and spiking neural networks.

Bibliography:

[1] `http://apt.cs.man.ac.uk/projects/SpiNNaker/`
[2] `http://www.intel.com/content/www/us/en/processors/xeon/xeon-phi-detail.html`
[3] `http://www.adapteva.com/introduction/`

Supervisor: Christoph Richter

(J. Conradt)
Professor

Abstract

The human brain has always been an unsolved riddle with its complex structure and its immense computational power. Containing about 10^{11} neurons and 10^{15} synapses firing at 10Hz it is a vast network with 10^{18} actions per second [12]. Every neuron sends information out to several thousand neurons with every neuron receiving information from thousands of synapses [24]. Over the past decades there have been numerous attempts to better understand the human brain, which has proven rather difficult. On the one hand it is not easy to operate on a live subject for tests and on the other hand there are only insufficient ways to model the brain appropriately. A neural network seems a good way to simulate the way the brain works, but suitable hardware is hard to find. With all the neurons working in parallel, a simulation using a neural network would be extremely complex and probably not possible for a conventional computing architecture. Recently, massively parallel computing architectures have become better and more available leading to big advances. On the one hand, there is the SpiNNaker board suited for these kinds of simulations and aiming for very large spiking neural networks while also offering small boards for mobile robots. On the other hand, there is the Parallella board, started through Kickstarter and costing only 99$ while being efficient and small enough for mobile applications. Both competitors seem to provide a lot of computational power while being both relatively small and power saving.
This review aims to investigate the hardware and software aspects of these milestones, to outline all the advantages and strengths as well as the drawbacks and to evaluate their capabilities in terms of simulating a brain and in mobile robotics.

Contents

Chapter 1

Introduction

In recent years, the human brain has been getting a lot of attention from engineers. But even though computers have become more powerful, the large scale neural networks needed to even come close to simulating the human brain are still very hard to process. And those architectures calculating neural networks need too much power to be feasible let alone be suitable for any kind of mobile application [16].
Massively parallel computing is addressing exactly these issues: Using less powerful cores but large amounts of them is a step closer to neurons while keeping the power consumption down. After a short insight into the biological inspiration and neural networks, two of these massively parallel architectures are examined and evaluated in detail. On the one hand is the university based SpiNNaker board while on the other hand there is the Parallella board funded through Kickstarter. This paper investigates which of the two performs better when simulating a (part of a) brain and how they fare in mobile robotics.

1.1 Biological Inspiration

As [14] found, the human brain works well, even under noisy circumstances. It thus is interesting to see how information is processed inside of it.
On the most basic level, information is carried through neurons via action potentials. The action potential works as a signaling device for connecting the many operations of the human brain such as contacting the muscles or transporting information from the sensory inputs to the central nervous system [21]. An impulse will trigger a neuron to raise its potential above a certain threshold which results in an action potential - also called a spike. This signal then runs along the neuron where it is transferred to thousands of other neurons through the synaptic gaps. Looking at the function and the firing rates, it performs asynchronously, event based and massively parallel [19].

1.2 Neural Networks

The most similar implementation is a neural network in terms of structure and behavior of the human brain. There are different mathematical neuron models like the leaky integrate and fire (LIF) and the Izhikevich model. The latter is more complex than the former, but also closer to the biological reality [17].
The way a neural networks functions is by using one of these models and connecting several layers consisting of numerous neurons. The connections are weighted to achieve the desired behavior and can be updated to improve the results. Since in a neural network many neurons work in parallel, they are computationally expensive and best performed on parallel architectures.
There are two types of neural simulation strategies: clock driven and event driven. The former method periodically updates all the neurons simultaneously while the latter works more like the brain with the different neurons waiting for an event. This holds the advantage of only needing to perform computations when it really is necessary which reduces power consumption. A more detailed overview can be found in [24].
Thus, copying the brain promises a combination of low power consumption, computational power and adaptability to various situations along with the added ability to compensate for noise. With conventional computing architectures being unsuited, the question arises whether novel computing approaches can deliver these benefits.

1.3 Criteria

As the products at hand are novel computing architectures, they both are not publicly available at the time of this review. The evaluation thus relies on the papers investigating the SpiNNaker board and information from the Adapteva website [2], the data sheets and manuals.
Based on the hardware, assumptions concerning the applications and possibilities can be made. To introduce a metric for a comparison, the boards are evaluated regarding two applications: simulating a brain and use in mobile robotics. As [20] found that a factual evaluation of neural simulators still is rather difficult, these two aspects are used to help the evaluation.

Chapter 2

SpiNNaker

The SpiNNaker project was designed at Manchester University and has a clear focus: Simulating a brain by connecting a large number of efficient processor cores with overall low power consumption, high message throughput and advanced routing facilities [15]. It is supposed to provide a system that can be varied in its size and thus in its processing power and power consumption. Depending on the size of the board, they contain a different number of SpiNNaker chips, which are the heart of the project.
Just like the neurons in a brain are not computationally powerful, the SpiNNaker Chip runs with a frequency of 200 MHz. However, the board also features a highly sophisticated router that allows the chips to be connected to each other on larger boards, and these boards again can be connected to one another while still using a seamless connection. Thanks to a high band width of up to 900MB/s and fast connections, the packets between the chips can be sent in vast numbers and received quickly. So just like neurons the SpiNNaker chips alone are limited in their power, but capable of complex simulations when put together in larger arrays. Combining the architecture and the communications of the current version, the result is a massively parallel manycore system. It can contain different numbers of components ranging from only one SpiNNaker chip on the smallest board to the (not yet reached) ultimate goal of 1,036,800 ARM9 cores with 7TB of memory distributed over the system. Apart from it being massively parallel, SpiNNaker works asynchronously and event driven [10] to properly mimic the biological inspiration. It sends neural signals just like spikes that only contain the address of the neuron it originated in as atomic events [19]. This event based programming allows for a low power consumption in idle while being able to provide high computational power when needed. These features combined make for a highly versatile product, being able to be put to use in both mobile and high power applications while being an important step towards understanding the brain better and to ultimately building an artificial one.

2.1 The SpiNNaker Chip

The most important part of the SpiNNaker board is the newly developed chip. Each chip contains 18 identical processing subsystems, where one of them is selected to be the monitor processor while 16 others simulate groups of neurons. One of the 18 cores is left out, because chips with only 17 functioning cores are accepted during production [10]. To improve fault resistance, the choice of the monitor processor is not fixed but performed dynamically [6]. To further avoid faulty behavior, defect detection and isolation mechanisms are included [6].
While Complex Instruction Set Computers (CISC) are state of the art for computationally heavy tasks, SpiNNaker uses Reduced Instruction Set Computer processors (RISC) which consume significantly less power and generally are cheaper. They do not provide the same performance as the complex ones, but putting them in a parallel order and using them in large numbers eliminates this problem [15].
Every chip contains 64KBytes of data memory and 32 KBytes of instruction memory [6]. The SpiNNaker system utilizes the ARM968 which is a small and power saving processor of the ARM9 family. The cores run at 200MHz and use fixed point operations. Every core is capable of modeling several hundred neurons with about 1000 synapses connected to each neuron.

2.2 Overview and Architecture

Figure 2.1 shows how the SpiNNaker nodes are designed. It can be seen that each chip features 128MB of SDRAM shared among the chip's cores. All the cores have access to the router, which then allows connecting a SpiNNaker chip to others. On top of the figure the 6 bidirectional connections to neighbors are visible connecting the chips in a hexagonal grid[6].

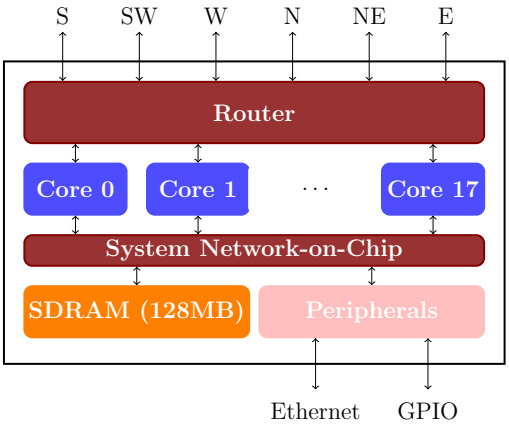

Figure 2.1: Architecture of a SpiNNaker node from [10]

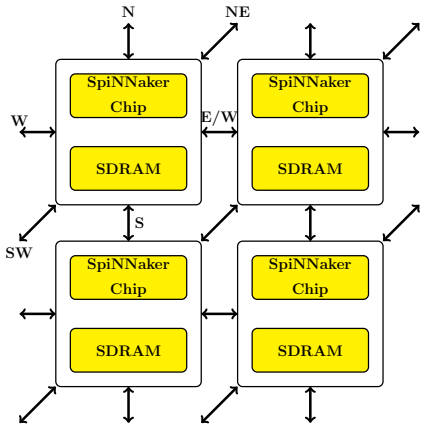

Figure 2.2: Hexagon layout of the chips from [6] and [18]

The way these layouts are connected can be seen in Figure 2.2. Dedicated FPGA based transceivers enable this connection to be extended seamlessly beyond the border of a physical board. With these connections and the different types of routing explained later on, up to 255x255 SpiNNaker chips can be connected, thus enabling huge networks.

2.3 Router and Memory

Managing and distributing the packages used for communication is the router's responsibility. It provides the monitor processor with all the information required to perform the necessary management [6]. Neural spikes are sent as atomic events that are broadcast as discrete packages. With the data processing handled by the router there still remains the memory management, because the SpiNNaker board has both tightly coupled memory (TCM) and chip-wide distributed SDRAM. The brain does not use a central memory. Likewise, the SpiNNaker processors also can modify any memory without synchronizing with others first. [19]
As in any network, congestion can occur between nodes, but there is a low-level solution, an emergency routing that redirects packages along an alternate route in case the preferred one is blocked for a programmable amount of time [9].
The cores can access the SDRAM with a bandwidth of 8Gb/s [15]. It takes the SpiNNaker network about 0.2-μs per node hop allowing it to work at a similar pace as biological neurons do [11].
Information in the brain is sent using small packets to many recipients, while conventional high performance computers are best suited for point-to-point messaging using large data packets. This is unsuited to simulate the brain and thus the routers are not only capable of default and point-to-point routing but also of multicast routing, allowing them to send packages to a large number of target nodes. If no match for the incoming message's routing key can be found in the lookup table of the router for multicast, the message is forwarded to the diagonally opposite connection (default routing). This helps keeping the lookup table reasonably small. [6]

2.4 Power Consumption

With the human brain working at an estimated 20W the bar is set very high for any kind of computing device. So when evaluating any novel architecture aiming to work similar to a brain, the power consumption has to be taken into account. Since SpiNNaker can not perform floating point operations, the computational power per Watt is measured by using instructions per second per Watt - the SpiNNaker delivers 2,200MIPs per Watt [11]. Even though that value does not suit the evaluation of a neural network well and neurons or synapses per Watt are a better metric, it allows the comparison to other architectures. Using a 48 Chip SpiNNaker board, [22] measured the power needed to simulate neural networks and found that a network

of 250,000 neurons with over a billion synaptic events per second only needs about 30W. Even though this is far from the brain's 20W for 10^{11} neurons, it is efficient with less than 1W per Chip.
Thus every SpiNNaker chip is able to simulate about 19,000 neurons each with a thousand synapses per Watt.

Chapter 3

Parallella

Funded through Kickstarter and for sale at 99$, this board bears the slogan *A Supercomputer for Everyone*. Just like the SpiNNaker board it claims to be suited for massively multi-parallel operations and it is also based on low power consumption components [5]. Starting with the size, the Parallella is similar to a credit card. Despite its small exterior, the Parallella board comes with either a 16-core or a 64-core Epiphany chip. The Parallella board contains a dual core Zynq FPGA and an Epiphany coprocessor which are connected through the eLink interface and AXI bus [3].
The following figure shows the Parallella board with all the features labeled in blue and the 16-core version of the Epiphany chip. As can be seen it can be connected using a Gigabit Ethernet port. It also contains a μHDMI port as well as 2 USB ports and a microSD slot.

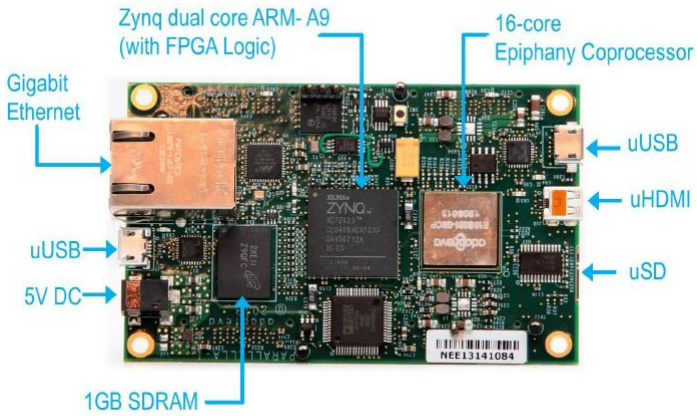

Figure 3.1: The Parallella Board taken from the Reference Manual [3]

While the Epiphany chip is the novel computing part of the Parallella board and

thus investigated in detail, the Zynq chip contains the main processor: It features two 1GHz ARM Cortex-A9 processors along with the fastest 28nm programmable logic. In a Coremark benchmark result provided by Xilinx on [4] the used Zynq 7020 scored 4737.47 - a better value than the processor inside the Samsung Galaxy S4, which scored 4540.00. One of the Parallella's possible applications is in mobile devices according to [2], which this result proves to be reasonable. The Zynq features 256KBytes of tightly coupled memory. Together with the good Coremark results of the different Epiphany chips the Parallella board contains two powerful computing resources.

3.1 The Epiphany Chip

The Parallella uses the Epiphany chip only as a coprocessor but it still is a central part of the Parallella board. The Epiphany chip is Parallella's equivalent to the SpiNNaker chip and as such the focus of this review. At the end of 2013, the operating frequency is stated at 700MHz, but the goal for Q3 of 2014 is set at 1GHz [5]. It is able to perform floating point operations and similar to SpiNNaker it was designed with low power systems in mind while still providing computational power. While the main focus of the Parallella project is not the creation of a million core architecture, the Epiphany chip does have the capability to connect to other Epiphany chips and FPGAs [5]. However, only 4096 cores can be connected to one another, thus limiting the possible networks in size.
These communications are handled by the eLink interface which sends the data as atomic transactions [5]. The layout of the 16 cores with the routers and the four eLinks can be seen in Fig. 3.2.

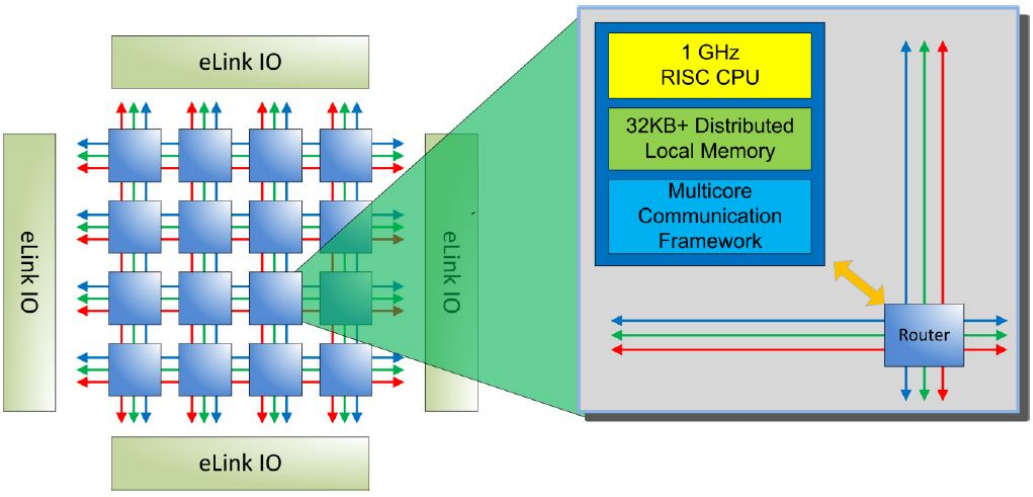

Figure 3.2: Layout of the Epiphany chip from [5]

The Network-On-Chip (NOC) eMesh of the Epiphany has a flat grid like layout dealing with any communication task. It contains three separated grids each responsible for a different task: One for writing on-chip, one for writing off-chip and one for reading [5]. According to the data sheet, the eMesh network does not need any special programming for its transactions which makes it easier to use [5]. Since in neural networks, a lot of actions are performed simultaneously, the bandwidth has to be able to cope with that amount of data. The Epiphany chip has 512 GB/s local memory bandwidths and 8 GB/s off-chip bandwidth [5].
The peak performance is given at 32 GFLOPs for the 16-core and 102 GFLOPs for the 64-core version. To put this into perspective, a high end GPU provides around 2,000 GFLOPs, but requires significantly more power ranging well above 100W.

3.2 Overview and Architecture

The main processor of the Parallella board is the Zynq-7000 AP SoC, which contains a dual-core Cortex-ARM9, which it combines with Xilinx programmable logic (FPGA) as can be seen in Fig. 3.4. How the Zynq and the Epiphany are included in the Parallella board along with the other components can be seen in Fig. 3.3.

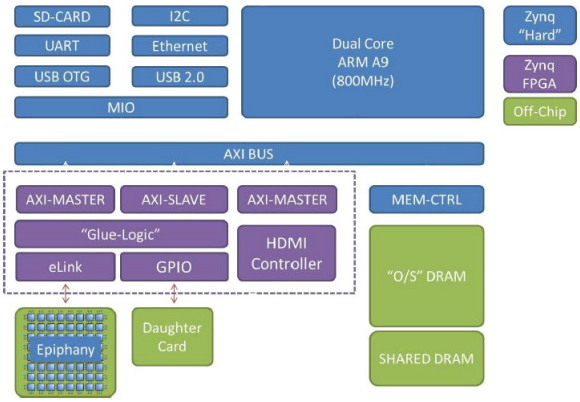

Figure 3.3: The Parallella Board Architecture from the Reference Manual [3]

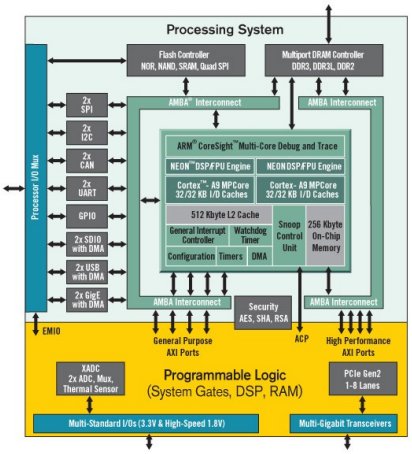

Figure 3.4: Layout of the used Zynq

In terms of memory, the Parallella board's CPUs rely on 512 KB of shared level 2 cache and can use 1GB of SDRAM.[3]
According to the manufacturer it ships with Ubuntu OS, has an HDMI port, 2 USB ports and even a Micro SD Card Slot. The first USB port can be used to connect it to a host PC, which then also supplies the power if the computational load is within limits. The second one can be used for peripherals like e.g. a keyboard. The microSD card is used for the primary boot source and for storage. The μHDMI port can be used to connect the board to a monitor. [3]

3.3 Router and Memory

A very interesting feature are the USB ports that allow connecting the Parallella board to both a host PC and input peripherals. [3]. It also features a Gigabit Ethernet port for fast connections. The board's performance is achieved in parts by providing a high bandwidth of 512 GB/s in the Epiphany chip and 8 GB/s off-chip bandwidth, as well as 1.3 GB/s between the Zynq and the Epiphany chip [5]. The used eMesh of the Epiphany chip is designed with power saving in mind by making any signal travel from one neighbor to the next, which keeps the traveling distance as short as possible and the operating frequency high. The disadvantage of that strategy is that the routing takes longer in case of jams and that signaling takes a bit longer than when taking a direct path seeing as the signal has to go through several tiles. Reading information takes even longer, because the signal has to travel back and forth [13]. The three parts eMesh network extends beyond the chip, which enables a seamless connection of multiple boards and chips. This is supported by the Epiphany's four eLinks, allowing the connection of multiple Epiphany chips on a board. The north and south eLinks can be used to connect multiple Parallella boards to a maximum of 4,096 connected cores.

3.4 Power Consumption

The article in [13] described Adapteva's goal as maximizing the GFLOPs per Watt and that a lot of power saving capability comes from not using cache but SRAM . The 16-core version of the Epiphany chip typically requires only 270mW , leading to less than 1W for the Parallella board equipped with it. The 64-core Parallella board requires 5W under a typical load with the 64-core Epiphany chip running with less than 2W. Furthermore, the power consumption of the Parallella board was investigated in [23] with a special scheduler. Apart from an overall low power consumption, it found that the Parallella board needs even less power when not using floating point calculations.

Chapter 4

Comparison

The two architectures are evaluated in comparison, regarding two aspects: capability to simulate a brain and applicability in mobile robotics.
According to [13], *Adapteva's goal is to maximize GFLOPs per watt* which is a very different goal than its competitor's since the SpiNNaker project aims to create a million core architecture and is not capable of performing floating point operations. Since the Parallella Board is not available at the time of this review, it is not possible to implement a neural network on it. Thus, the evaluation is centered around the hardware specifications, based on which assumptions can be made regarding the performance. Furthermore, the physical size and the power consumption allow a direct comparison in terms of mobility.

4.1 Brain Scale Simulation

The first part of the evaluation is the simulation of a brain, ideally the human one. Simulating a brain requires the simulation of large numbers of parallel neurons.
The SpiNNaker project is designed to one day simulate the human brain. It is best suited for simple neurons in complex networks with its many low frequency processors. According to [15], it is possible to simulate about 19,000 neurons per SpiNNaker chip [15]. In order to simulate at least a sizable part of the brain, larger numbers are required, leading to the 48-chip board with 48 chips and 19000x48=912000 neurons. Rounding that to a million neurons, the 48-chip SpiNNaker board is able to model roughly 0.0001% of the human brain's neurons in biological real-time. The SpiNNaker architecture allows 255x255 chips to be linked together leading to the million core machine. In that configuration it is able to simulate about 1.2 billion neurons, which is circa 1% of the human brain's neurons.
The Parallella board was not specifically designed for such a simulation, but based on the similar hardware, assumptions can be made. For one, the Epiphany cores can use 8 times more shared memory than the SpiNNaker chip, which would help the simulation. Relying on [11] one of the SpiNNaker's ARM cores with 200MHz is able to simulate about 1000 neurons with 1000 inputs, requiring about 4 bytes per

synapse. Thus 16 cores need 64 MB for the simulation. Applying the same formula to the Parallella board, the cores could simulate more neurons with their higher frequency and the larger memory. Although it is not clear, whether the increased frequency will help for real-time simulations. Furthermore, the Parallella board has a significantly higher bandwidth than the SpiNNaker chip with 1.3GB/s between the Zynq and the Epiphany compared to SpiNNaker's 900MB/s. However, the Parallella board is limited at 4,096 cores, which despite better hardware in certain aspects, allows the SpiNNaker project to simulate larger networks. So in a direct comparison of two similar sized boards below that limitation the Parallella looks like it can perform better at simulating neural networks. But only the SpiNNaker is able to simulate a sizable part of the human brain.
To evaluate the simulation performance with respect to the required power, the number of neurons per Watt is investigated. With less than one Watt per SpiNNaker chip, and about 19,000 neurons per chip this results in more than 19,000 neurons per Watt. There is no number for the Parallella board or the Epiphany chip, but the above mentioned hardware and a similar power consumption suggest a higher number. The Parallella board with the 64-core Epiphany chip requires about 5W, which is five times more than on 18-core SpiNNaker chip needs, but less than five times the number of cores. So based on this comparison and the above estimations, the Parallella theoretically is able to simulate more neurons and synapses per core and thus per board. However, the Parallella probably can simulate less neurons per Watt than its competitor. Furthermore, it is limited to 4096 cores and the SpiNNaker's limit is around a million cores, thus the latter can simulate larger neural networks.

4.2 Mobile Robotics

In mobile robotics, five aspects are of great importance: the power consumption, the computational power, the size, the connectivity and the price. In terms of power consumption, the 48-chip SpiNNaker needs approximately 30W, resulting in less than 1W per SpiNNaker chip. This value changes depending on the size of the simulated network and thus the load as shown in [22]. The SpiNNaker uses mains AC and a local supply that produces 12 V DC output for the boards. The Parallella board can be powered either by a 5V DC connector or through one of the USB ports, although due to the lower current, this option limits the computational powers [3]. One Parallella board with the 64 core Epiphany chip needs about 5W. Scaling that performance to the 18x48 cores of the SpiNNaker's 30W, the Parallella setup would require about (5W/(64+2) cores)x(18x48 cores) = 65.5W. This is more than twice the power consumption of the corresponding SpiNNaker board.
Regarding computational power on similar sized boards, the SpiNNaker only contains 16 cores, while the Parallella can offer a (64+2)-core version. The SpiNNaker is listed as delivering 2.2 GIPS/W. Based on the power the most powerful Parallella

board needs with (64+2) cores and on the SpiNNaker's power consumption, this means a 66-core SpiNNaker board delivers about 91 GIPS. The 64-core Epiphany peaks at 85 GIPS. So the 48-chip SpiNNaker can process more instructions per second. However, it is not able to perform floating point operations, which limits its capabilities. Even though the GIPS are not a very strong criterion, they do provide a direct comparison. Additionally, the previous section shows that the Parallella board can simulate more neurons per core than its competitor. Regarding the size, the 48-chip version is significantly larger than the 64-core Parallella board. This makes the Parallella superior in mobile applications with limited space. The Adapteva website [2] even lists mobile phones or tablets as possible applications. Inspecting the fourth aspect, the connectivity, both competitors have GPIO ports. They both feature an Ethernet connection for fast data transferring; on the SpiNNaker it is 100Mbit/s Ethernet, while the Parallella features Gigabit Ethernet. The Parallella board contains two high-speed UARTs, two full-duplex SPI ports, two full CAN bus interfaces and is able to sense the voltage and temperature on the board. The SpiNNaker and some of the Parallella versions feature GPIO ports. The SpiNNaker can be connected using a UART interface or through SPI or TWI. These connectivity options, the small size and lower power consumption make them both suitable for mobile robotics. Last but not least, in mobile robotics, the price tag plays an important role and the Parallella board is extremely cheap at just 99$. The SpiNNaker, since it is not for sale, does not specify its cost. But with a budget of about $8million until December 2013, it is most certainly much more expensive than its competitor at an estimated $5000. Parallella reached almost $900,000 through Kickstarter [1] and according to [13], a total of two million dollars in funding was used to create the board and the Epiphany chip.
Based on this comparison, the Parallella is highly interesting for mobile applications with good performance on a very small board and at a very low price. The SpiNNaker's performance and power consumption are low as well, but its larger size and the unavailability with the estimated high price make it less suited for small or budget applications.

4.3 Conclusion

This review presents two novel computing architectures with their features and capabilities. The comparison shows that both competitors perform well in the areas they were designed for, but that both had to agree to compromises. The SpiNNaker is able to simulate very large neural networks and in the final size it will even be capable to simulate one percent of the human brain. It does so while still providing smaller solutions for mobile robotics with low power consumption. Its main drawback is the lack of availability and the price tag it would be sold for. The Parallella on the other hand provides significantly more cores for its size and is able to perform floating point operations, which makes it suitable for mobile applications. Its small

size and the low power consumption further improve its use in mobile robotics. Its drawback is its limitation of connected cores at 4,096, which limits the capabilities in terms of simulating a sizable part of the human brain.

Bibliography

[1] Parallella: A supercomputer for everyone [online]. https://www.kickstarter.com/projects/adapteva/parallella-a-supercomputer-for-everyone/posts/547314. Accessed: 06/22/2014.

[2] The parallella board [online]. http://www.adapteva.com/parallella-board/. Accessed: 06/22/2014.

[3] Parallella Reference Manual. (Rev. 13.11.25).

[4] Unmatched performance and power [online]. http://www.xilinx.com/products/silicon-devices/soc/zynq-7000/performance.html. Accessed: 06/22/2014.

[5] E16G301 Epiphany 16-Core Microprocessor Datasheet, 13 March 2011.

[6] SpiNNaker datasheet, 6 January 2011.

[7] SpiNNaker - Software Specification and Design, 9 December 2011.

[8] P. Fox. Massively parallel neural computation. PhD thesis, University of Cambridge Computer Laboratory, 2013.

[9] S. Furber and A. Brown. Biologically-inspired massively-parallel architectures - computing beyond a million processors. In Proc. 9th International Conference on the Application of Concurrency to System Design (ACSD'09), July 2009. Event Dates: 1-3 July 2009.

[10] S. Furber, F. Galluppi, S. Temple, and L. Plana. The spinnaker project. Proceedings of the IEEE, 102(5):652–665, May 2014.

[11] S. Furber, D. Lester, L. Plana, J. Garside, E. Painkras, S. Temple, and A. Brown. Overview of the spinnaker system architecture. IEEE Trans. Comput., 62(12):2454–2467, December 2013.

[12] S. Furber and S. Temple. Neural systems engineering. Journal of the Royal Society interface, 4(13):193–206, 2007.

[13] L. Gwennap. Adapteva: More flops, less watts. Microprocessor Report, 6(13):11-02, 2011.

[14] D. Horn and I. Opher. Collective excitation phenomena and their applications. Pulsed neural networks, pages 297–320, 1999.

[15] P. Lenander and A. Fosselius. A survey of the spinnaker project: A massively parallel spiking neural network architecture.

[16] C. Mead. Neuromorphic electronic systems. Proceedings of the IEEE, 78(10):1629–1636, 1990.

[17] E. Painkras. A chip multiprocessor for a large-scale neural simulator. PhD thesis, 2013.

[18] L. Plana, S. Furber, S. Temple, M. Khan, Y. Shi, J. Wu, and S. Yang. A gals infrastructure for a massively parallel multiprocessor. Design & Test of Computers, IEEE, 24(5):454–463, 2007.

[19] A. Rast, F. Galluppi, X. Jin, and S. Furber. The leaky integrate-and-fire neuron: A platform for synaptic model exploration on the spinnaker chip. In IJCNN, pages 1–8. IEEE, 2010.

[20] T. Sharp and S. Furber. Correctness and performance of the spinnaker architecture. In Neural Networks (IJCNN), The 2013 International Joint Conference on, pages 1–8. IEEE, 2013.

[21] C. Smith. Elements of Molecular Neurobiology. Wiley, second edition, 1996.

[22] E. Stromatias, F. Galluppi, C. Patterson, and S. Furber. Power analysis of large-scale, real-time neural networks on spinnaker. In Neural Networks (IJCNN), The 2013 International Joint Conference on, pages 1–8. IEEE, 2013.

[23] Amirhossein Taher Kouhestani. Power-aware scheduler for many-core real-time systems. Master's thesis, KTH, Software and Computer systems, SCS, 2013.

[24] J. Vreeken. Spiking neural networks, an introduction. Institute for Information and Computing Sciences, Utrecht University Technical Report UU-CS-2003-008, 2002.

[25] Weikersdorfer D., Hoffmann R., and Conradt J. (2013) Simultaneous Localization and Mapping for event-based Vision Systems, International Conference on Computer Vision Systems (ICVS), p. 133-142, St. Petersburg, Russia.

[26] Weikersdorfer D., Conradt J. (2012), Event-based Particle Filtering for Robot Self-Localization, Proceedings of the IEEE International Conference on Robotics and Biomimetics (IEEE-ROBIO), pages: 866-870, Guangzhou, China.

[27] Axenie, C., Conradt, J. (2014) Cortically inspired sensor fusion network for mobile robot egomotion estimation. Robotics and Autonomous Systems, Special Issue on "Emerging Spatial Competences: From Machine Perception to Sensorimotor Intelligence", pre-print.