BrainScaleS-1

From Clean Room to Machine Room

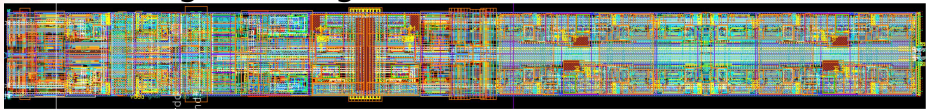
Sebastian Schmitt

On behalf of Electronic Vision(s)

NICE202~~0~~1: 2021-03-16

EBRAINS

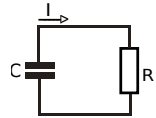
Physical Modelling – Analog Neuron Circuits



layout drawing of two neurons: 150x20 µm^2

- Adaptive Exponential Integrate and Fire (AdEx) Model
- Dedicated circuits in every neuron for:
  - Resting potential
  - Reset potential
  - Threshold potential
  - Reversal potentials
  - Refractory period
  - Membrane time constant
  - Synaptic time constants
  - Adaption and
  - Exponential term

- Accelerated dynamics compared to biological real-time:

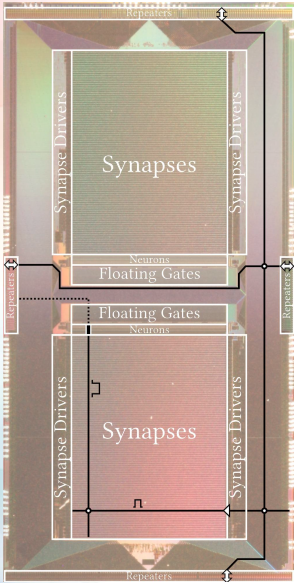


$\tau = C \cdot R, \frac{\tau_{\text{hw}}}{\tau_{\text{bio}}} = 10^{-4}$

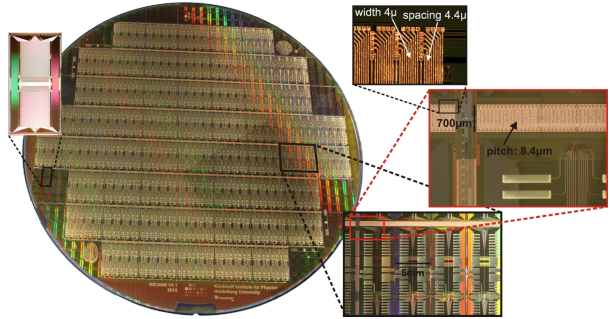
$\tau_{\text{hw}} = 1\,\mu\text{s} \Rightarrow \tau_{\text{bio}} = 10\,\text{ms}$

- Also every synapse is represented by dedicated circuitry matching the speedup

BrainScaleS-1 ASIC

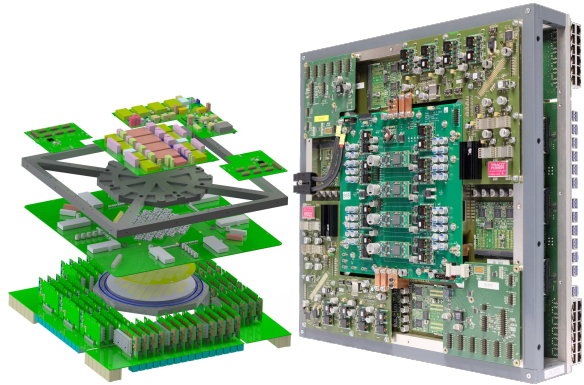


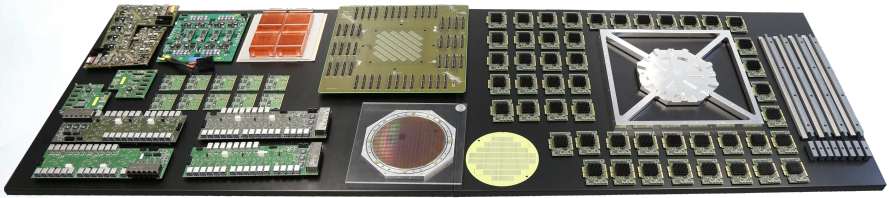
- 512 analog neurons, 110 000 plastic synapses
- Digital communication → mixed-signal system
- Sparse crossbar switches connecting busses → programmable network connections
- Analog parameter storage (floating gates)
- Postprocessing (IZM Berlin) → wafer scale networks



BrainScaleS Wafer Module

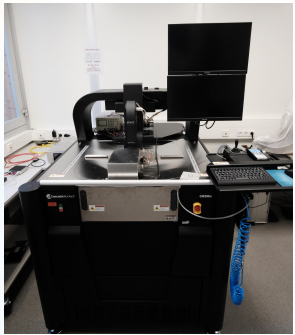
- 20 cm Wafer, 180 nm CMOS
- 14 layer main PCB
- 48 Kintex-7 FPGAs (TU Dresden)
- Only one 48 V external power supply
- Aux. Boards (e.g. for analog readout)
- Connected via ethernet to 48 Gbit/s switch

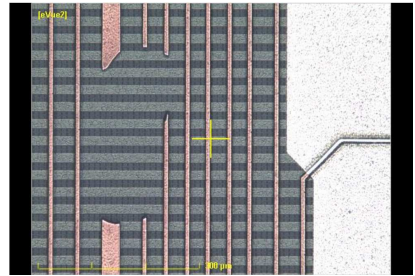


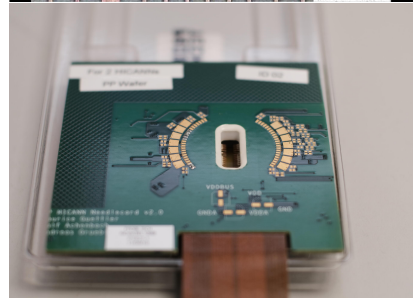


Tests Prior to Assembly

- Optical inspection on wafer prober
  - Find imperfections, e.g., of post processing prior to assembly
- Functional tests with a needle card
  - Differentiate between „inherent“ and system induced problems

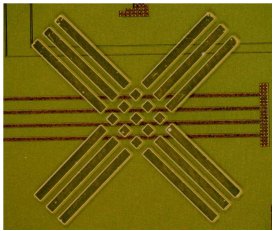


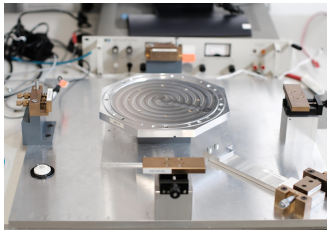


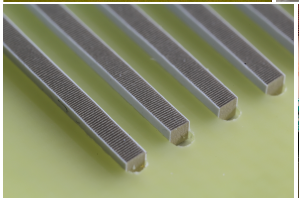


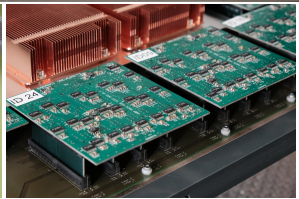
Wafer Main PCB Marriage

- Optical alignment between wafer and main PCB
- Contacts are tested via the FPGA connectors with special boards (Sabanci University)
- Iterative procedure until alignment is sufficient





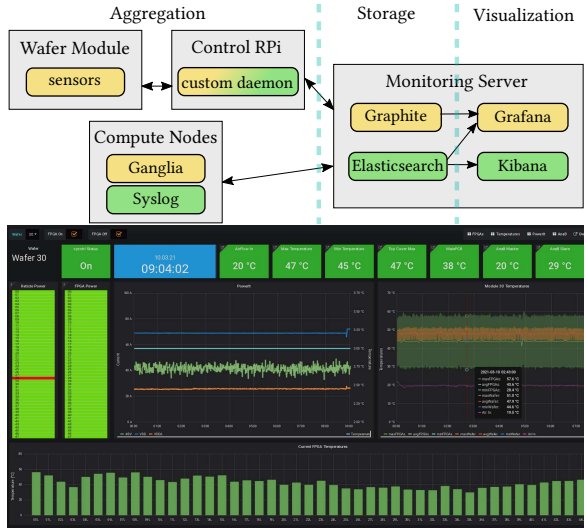




Machine Room with 20 BrainScaleS Wafer Modules

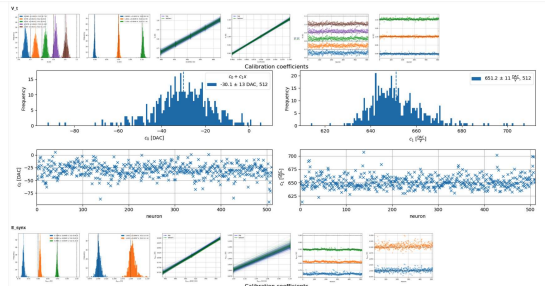
Monitoring

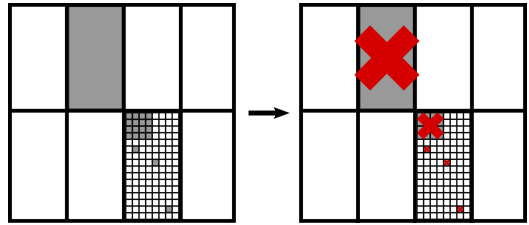
- Machine room and wafer modules are monitored
- About 1200 time-series data sources within one wafer module
- Important sensors like wafer temperature are read out every few seconds.
- Events for powering parts on/off
- Alerts are generated and reported via mail and chat



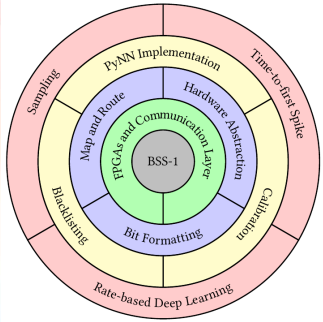
Calibration and Availability Management

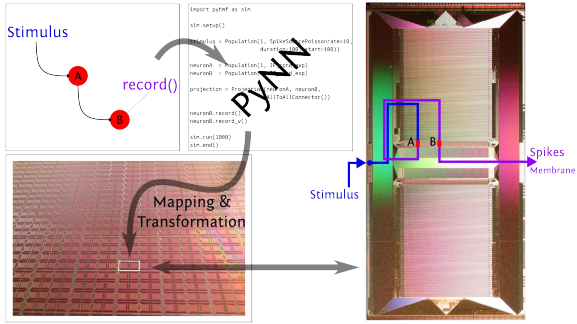
- Calibration of analog parameters
  - 111 calibration steps
- Hardware tests:
  - Communication link tests (JTAG, highspeed)
  - Digital memory checks of 42 MiB per wafer
- Availability management:
  - Derived resource availability from calibration and digital memory checks
  - Experiment specific resource availability flexibly adjustable via cli
- Tests and calibration are fully automated and executed periodically to track hardware state





From Network to Hardware

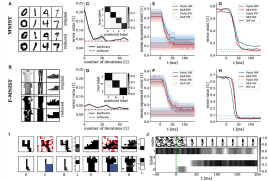




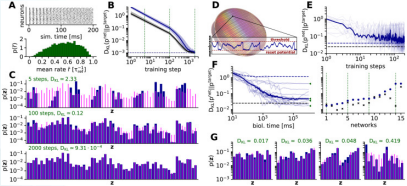
- The BrainScaleS operating system converts from a biological network description in pyNN to a hardware configuration
- Spike input and output via the FPGAs and analog recordings are managed
- For details → E. Müller: “BrainScaleS: Development Methodologies and Operating System” Friday, 19 March 2021

Experiments on BrainScaleS

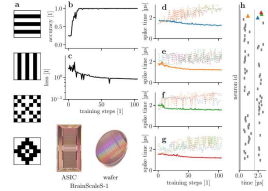
- “Accelerated physical emulation of Bayesian inference in spiking neural networks”



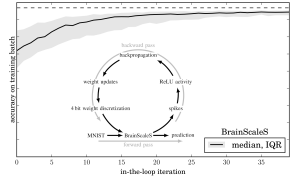
- “Stochasticity from function - why the Bayesian brain may need no noise”



- “Fast and deep neuromorphic learning with time-to-first-spike coding”

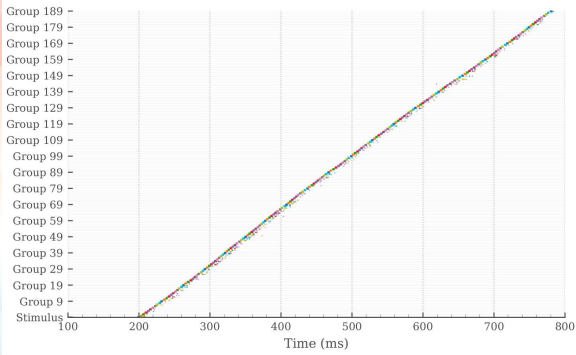


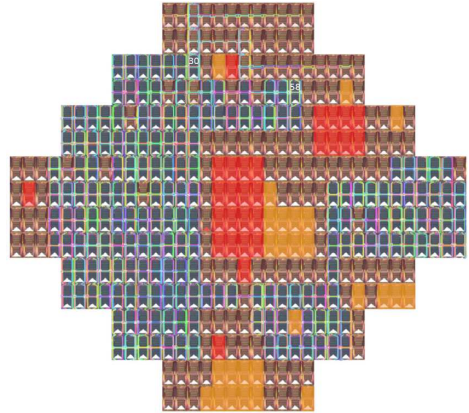
- “Neuromorphic Hardware In The Loop: Training a Deep Spiking Network on the BrainScaleS Wafer-Scale System”



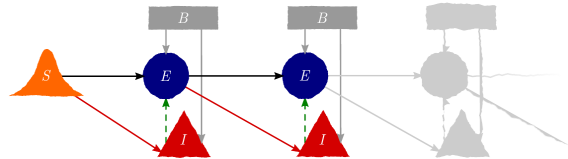
- More results can be found in about 60 theses (PhD plus undergraduate)

Wafer Scale Feed-Forward Chain



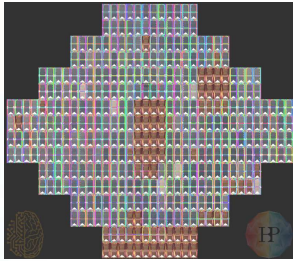


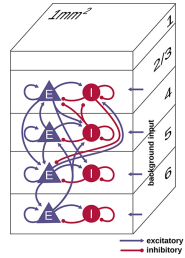
- Wafer Scale Feed-Forward Chain
- 230 chips used:
  - 19000 logical neurons (76000 neuron circuits)
  - 1.4 million synapses



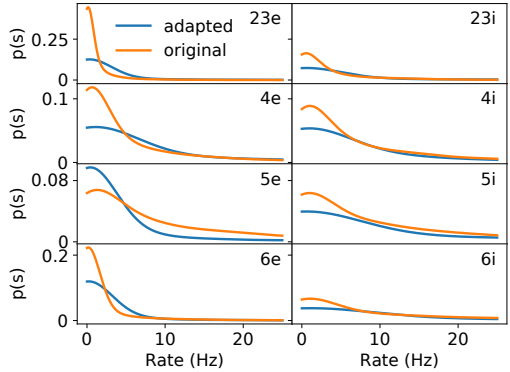
Towards Wafer Scale Full-Scale Cortical Microcircuit Model

- Towards emulation of the “Cortical Microcircuit Model” (Potjans/Diesmann (2014))
- Full model consists of 80000 neurons
  - Downscaling to 10% to fit on wafer





Simulation:



also:

- CUBA $\rightarrow$ COBA synapses
- Reduced external drive
- Simulation studies of influence of paramater variations

Summary and Outlook

- BrainScaleS demonstrates “From Clean Room to Machine Room” and “From Network to Hardware” of accelerated mixed-signal wafer scale neuromorphic computing
- In the pipeline:
  - On-module analog digitization $\rightarrow$ a fully self-contained module
  - Horizontal configuration complying with open compute rack standard
  - Next-generation interim solution until BrainScaleS-2 wafer scale is done
  - Interconnection via EXTOLL (cf. poster by T. Thommes)