Highly Integrated Packet-Based Communication in a Neuromorphic Wafer System

J. Partzsch, C. Mayr, S. Schiefer, S. Hartmann, S. Scholze, R. Schüffny
Endowed Chair for Parallel VLSI Systems and Neural Circuits, Technical University Dresden, Germany

I. Setup of Demonstration

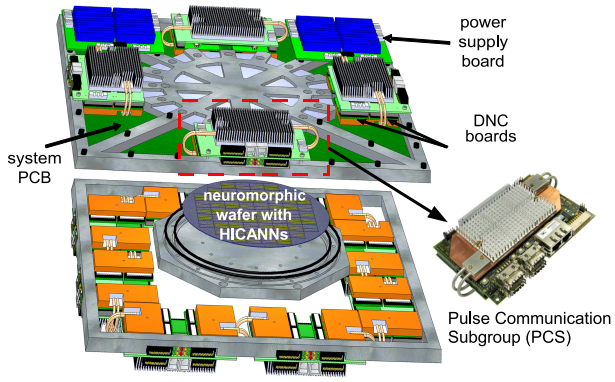

Fig. 1. Overview of one wafer module of the FACETS/BrainScaleS waferscale neuromorphic system [1]

This demonstration is based on the newly developed wafer-scale neuromophic system presented in [1] and [2]. One wafer of this system contains 384 analog network chips (ANC) for a total of 40 M synapses and 200 k neurons. 12 FPGA boards hosting 48 digital network chips (DNC) [3] connect the wafer to outside stimuli by 768 2 Gb/s serial links. As can be seen in Fig. 1, one FPGA board and 4 DNCs form a single unit, which is called a pulse communication subgroup (PCS)[1]. The links realized by the PCS handle the main neuron-neuron communication in the form of quasi-digital pulses. Thus, these communication links are one of the main building blocks required to densely interconnect the neurons between wafers, thus providing biologically realistic connection densities and speeds. They also handle the configuration of the wafer.

We give a live demonstration of the pulse communication subgroup. Pulse packets are routed from the host PC via Gbit Ethernet to the PCS (see Fig. 2), which forwards them to the 4 DNCs on the same board. The DNCs buffer and sort the pulses, implementing 32 2GBit/s Low Voltage Differential Signaling (LVDS) interfaces to the neuromorphic circuits on the wafer. In order to have a low-parasitic-capacitance interface to the wafer and FPGA, the DNC is chip-on-board-bonded on a purpose-designed PCB which is clipped onto the FPGA board. Pulse communication to other wafers is done via an FPGA-FPGA communication using 10Gbit/s Aurora links.

[1]The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007- 2013) under grant agreement no. 269921 (BrainScaleS)

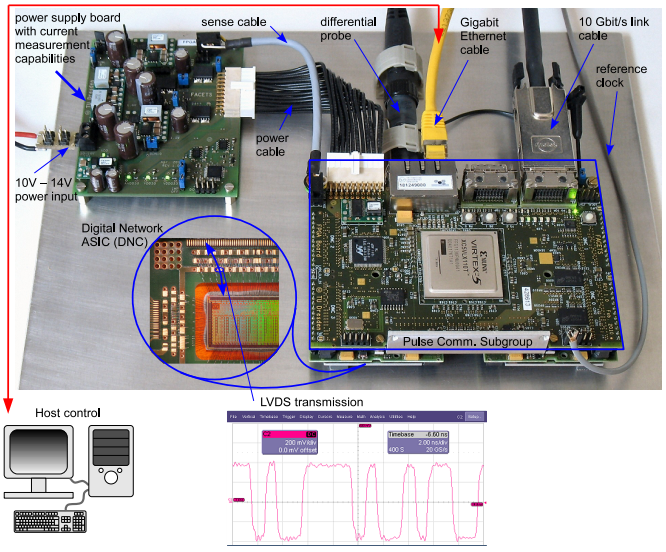

Fig. 2. Pulse communication subgroup: the FPGA/ASIC-based communication infrastructure for the wafer

II. Measurements and Interaction

The communication setup is operated from the host PC via an intuitive Python interface, allowing the user to fully characterise the infrastructure: He can perform bandwidth tests, directly measuring throughput and connection delays. The user can experiment with the spike-train statistics to assess their effect on hardware performance. We also present neuromorphic benchmark simulations (e.g. a synfire chain [4]), which run on PC and route their pulses via the PCS, showing the direct impact of wafer interface characteristics on benchmark performance.

References

[1] J. Schemmel, D. Brüderle, A. Grübl, M. Hock, K. Meier, and S. Millner, "A wafer-scale neuromorphic hardware system for large-scale neural modeling," in *ISCAS 2010*, 2010, pp. 1947–1950.

[2] S. Scholze, S. Schiefer, S. Hartmann, J. Partzsch, C. Mayr, S. Höppner, H. Eisenreich, S. Henker, and R. Schüffny, "VLSI implementation of a 3GEvent/s packet based AER interface with routing and event sorting functionality," *Frontiers in Neuromorphic Engineering*, vol. 5, p. 13, 2011.

[3] S. Scholze, H. Eisenreich, S. Höppner, G. Ellguth, S. Henker, M. Ander, S. Hänzsche, J. Partzsch, C. Mayr, and R. Schüffny, "A 32 GBit/s communication SoC for a waferscale neuromorphic system," *Integration, the VLSI Journal*, vol. 45, no. 1, pp. 61–75, 2011.

[4] J. Kremkow, L. Perrinet, G. Masson, and A. Aertsen, "Functional consequences of correlated excitatory and inhibitory conductances in cortical networks," *Journal of Computational Neuroscience*, vol. 28, no. 3, pp. 579–594, 2010. [Online]. Available: http://dx.doi.org/10.1007/s10827-010-0240-9