WROKSHOP SUMMARY

Workshop on Fast Realtime Systems and Realtime Machine Learning

Justus–Liebig-University of Giessen

April 8–11, 2024

Is it possible to record data in a scientific experiment with a bandwidth of 1 Terabyte per second, which would correspond to 10 Million smartphones sending data in the same moment? This and other questions of so-called "realtime" data recording and data processing were discussed 48 international participants at a Workshop at Gießen. The workshop was organized with in-person presentations at Gießen campus, but also with remote video contributions from Japan, Switzerland, UK, or US. This so-called "hybrid format" was decided in order to give the participants the possibility to avoid flight travel and thus minimize the CO_2 footprint.

The workshop was supported by the Jennifer2 project and the ErUM–Data–Hub.

Jennifer2 is the second project phase of the *Japan and Europe Network for Neutrino and Intensity Frontier Experimental Research*, as funded by the European Commission within the Horizon 2020 Framework (Marie Skłodowska-Curie Actions MSCA-RISE-2018 - Research and Innovation Staff Exchange). It comprises more than 200 researchers in 15 academic and 2 industrial beneficiary institutions. For additional information see `http://www.jennifer2-project.eu/`.

The ErUM–Data–Hub is the central networking and transfer office for the digital transformation in the exploration of universe and matter, funded by the German Federal Ministry of Education and Research (BMBF) and involves representatives from all ErUM thematic areas: astroparticle physics, elementary Particles Physics, accelerator Physics, research with neutrons, research with synchrotron radiation, research with nuclear probes and ion beams, hadron and nuclear physics and the german observatory council. For additional information see `https://erumdatahub.de/`.

One of the priority questions of the workshop was defined as: would it be possible to detect an *anomaly* in a large data stream in "realtime"? Anomalies are rare objects, such as a rare coin within millions of other coins. A rare bird in a swarm of others, a four-leaf clover in a field with billions of flower or an unexpected new phenomenon in a particle physics experiment. Often such a scientific task is referred to as "finding the needle in the hay stick". The particular challenge of this task, as discussed on the workshop, is given by possible "realtime" performance.

Realtime systems must make decisions very fast. Benchmark results, as presented on the workshops, reach timescales of less 100 nanoseconds. This is the time interval in which the earth at the equator is moving by 0.05 millimeters. It is 0.0001% of the time which is required by the blink of an eye or less than 0.00001% of the duration of a flash in a thunderstorm. There are in fact examples of realtime decisions in daily life, in particular danger-for-life situations such as airbag inflation in a car accident or earthquake pre-warnings by a push message on a phone. As those anomaly detections must be performed on time scales much faster than a human being would be able to react, "intelligent" machines are required. Such a machine requires both dedicated software and dedicated hardware.

For the software, the algorithms are so called *neural networks* and are based upon principles of connected "neurons", in which a neuron represents a mathematical model of a human brain cell. Although such a model is very much simplified (compared to a natural cell), the mathematical model however can be very effective and – as mentioned before – the mathematical operations can be very fast.

For the hardware, very recently, dedicated platforms have been emerging, which have been designed in particular to cope with neural network algorithms. Such hardware platforms are called *Field Programmable Gate Arrays* (FPGA) and are representing large networks of logic cells, which can be configured by the programmer according to her/his particular design wishes. Thus, they are very flexible, but requires specific programming knowledge and implementation techniques. Benchmark results, as presented on the workshop, show that mathematical operations such as 4x4 matrix multiplications per clock cycle are performed by a factor 50 faster compared to a PC. The presently probably most powerful platform is called VERSAL[1] by AMD/Xilinx. Although assigned with a significant price tag, which can be compared to a small car, three groups represented on the workshop were able to acquire such hardware and present first testing results.

More details about the workshop, including pdf of all the presentations, can be found here:
https://indico.belle2.org/event/10782/

[1] https://www.amd.com/de/products/adaptive-socs-and-fpgas/versal.html

WORKSHOP // on // FAST REALTIME SYSTEMS // and // REALTIME MACHINE LEARNING

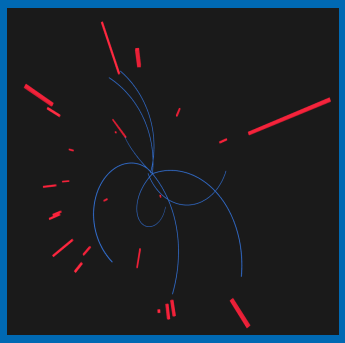
JLU
NEUE WEGE. SEIT 1607.

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

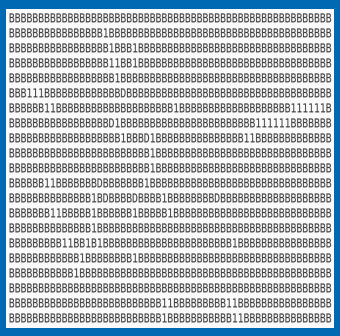
Justus–Liebig-Universität Giessen

April 8–11, 2024

https://indico.belle2.org/event/10782/



https://display.belle2.org/



https://asciiart.club/

ERUM DATA HUB

JENNIFER 2
EU grant n.822070

Bundesministerium für Bildung und Forschung

ErUM-FSP T09 Belle II

Photo: Jens Sören Lange

Photo: Jens Sören Lange