



# APPROACHES TO HIT FINDING AND TRIGGER GENERATION

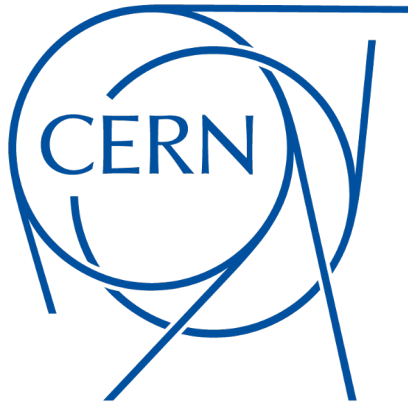
CERN openlab lightning talk

August 15, 2019

Niklas Böhm

Paola Sala & Manuel Rodriguez

University of Tübingen



Why should we care about neutrino research?

Why should we care about neutrino research?

→ Neutrinos are in a quantum superposition.

Why should we care about neutrino research?

- Neutrinos are in a quantum superposition.
- Neutrinos don't interact with almost anything.

## PROBLEM STATEMENT

We want to have a fast inference algorithm for neutrino classification.

$$f: \mathbb{R}^{500 \times 500 \times 3} \rightarrow \{0, \dots, 12\}$$

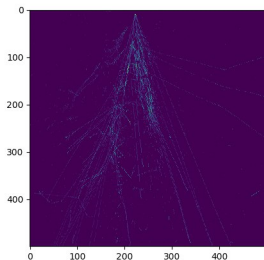


Figure: A sample image of a simulation of a neutrino interacting with argon.

## THE STATE AT THE START

- Two FPGA boards
- Clean, simulated training data
- A trained model

## THE STATE AT THE START

- Two FPGA boards
- Clean, simulated training data
- A trained model

Now get it running on the FPGA.

**Problem** Issues with Keras export to ONNX.<sup>1</sup>

---

<sup>1</sup>Open Neural Network Exchange Format: <https://onnx.ai/>.



**Problem** Issues with Keras export to ONNX.<sup>1</sup>

**Solution** Use Pytorch instead.

---

<sup>1</sup>Open Neural Network Exchange Format: <https://onnx.ai/>.

**Problem** Need a GPU for training.

**Problem** Need a GPU for training.

**Solution** Ask for it

**Problem** Need a GPU for training.

**Solution** Ask for it (that was easy).

**Problem** Need a GPU for training.

**Solution** Ask for it (that was easy).

**Problem** All available GPUs are occupied.

**Problem** Need a GPU for training.

**Solution** Ask for it (that was easy).

**Problem** All available GPUs are occupied.

**Solution** Use a different cluster and adapt code.

**Problem** Need a GPU for training.

**Solution** Ask for it (that was easy).

**Problem** All available GPUs are occupied.

**Solution** Use a different cluster and adapt code.

**Problem** Training needs time.

**Problem** Need a GPU for training.

**Solution** Ask for it (that was easy).

**Problem** All available GPUs are occupied.

**Solution** Use a different cluster and adapt code.

**Problem** Training needs time.

**Solution** Take care of other problems.



## INFERENCE SPEED

We now have an implementation that gives the same results on both GPU and FPGA.

We now have an implementation that gives the same results on both GPU and FPGA.

**Problem** The FPGA is too slow.

We now have an implementation that gives the same results on both GPU and FPGA.

**Problem** The FPGA is too slow.

**Solution** Use other, faster FPGA.

**Problem** Output probabilities (classification) is very different from reference; due to “overflows”.

**Problem** Output probabilities (classification) is very different from reference; due to “overflows”.

**Solution** Investigate weights and find the spot where overflow occurs.

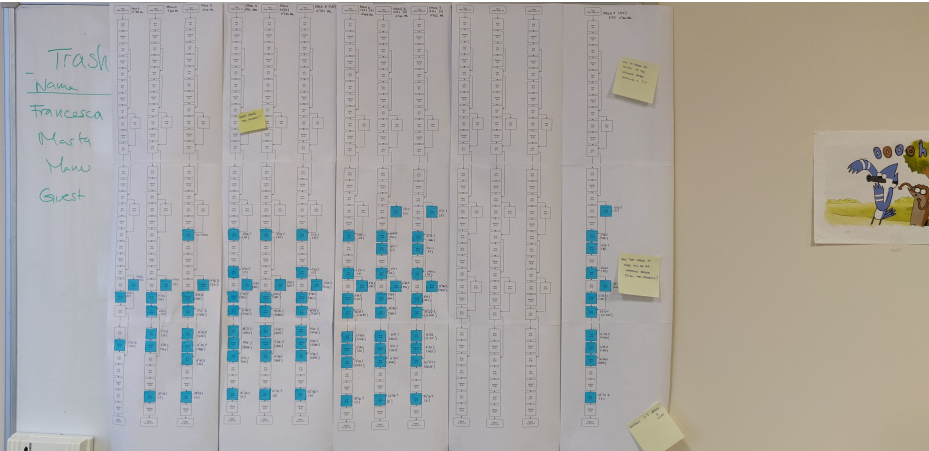


Figure: Printout of overflows in layers over epochs. Made by Manuel Rodriguez.

**Problem** Weights are not the issue.

## MORE PROBLEMS

**Problem** Weights are not the issue.

**Solution** Implement Software checks. Ask vendor for further information since it's a proprietary library.



## MORE PROBLEMS

**Problem** Weights are not the issue.

**Solution** Implement Software checks. Ask vendor for further information since it's a proprietary library.

**Alternative Solution** Change training procedure and pray it prevents overflows from happening (use **Batch Normalization**, gradient clipping, gradient normalization).

## MORE PROBLEMS

**Problem** Weights are not the issue.

**Solution** Implement Software checks. Ask vendor for further information since it's a proprietary library.

**Alternative Solution** Change training procedure and pray it prevents overflows from happening (use **Batch Normalization**, gradient clipping, gradient normalization).

It did not.

**Problem** The FPGA keeps dying.

## WHILE WAITING FOR AN ANSWER

**Problem** The FPGA keeps dying.

**Solution** Stress test it.

## WHILE WAITING FOR AN ANSWER

**Problem** The FPGA keeps dying.

**Solution** Stress test it.

**Problem** The FPGA does not survive the stress test.

## WHILE WAITING FOR AN ANSWER

**Problem** The FPGA keeps dying.

**Solution** Stress test it.

**Problem** The FPGA does not survive the stress test.

**Solution** Make it less stressful.

## WHILE WAITING FOR AN ANSWER

**Problem** The FPGA keeps dying.

**Solution** Stress test it.

**Problem** The FPGA does not survive the stress test.

**Solution** Make it less stressful.

**Problem** It still dies.

## WHILE WAITING FOR AN ANSWER

**Problem** The FPGA keeps dying.

**Solution** Stress test it.

**Problem** The FPGA does not survive the stress test.

**Solution** Make it less stressful.

**Problem** It still dies.

**Solution** Also ask the vendor for help.



## AN ANSWER FROM MICRON

There was an error in the batch normalization implementation in the library. It gets fixed.

There was an error in the batch normalization implementation in the library. It gets fixed.

**Problem** The FPGA does not run anymore.

There was an error in the batch normalization implementation in the library. It gets fixed.

**Problem** The FPGA does not run anymore.

**Solution** Uhm...

To be continued.