

# Accelerating GAN training using distributed tensorflow and highly parallel hardware

Thursday, October 22, 2020 10:40 AM (20 minutes)

## Abstract

Machine Learning has been used in a wide array of areas and the necessity to make it faster while still maintaining the accuracy and validity of the results is a growing problem for data scientists. This work explores the Tensorflow distributed parallel strategy approach to effectively and efficiently run a Generative Adversarial Network, GAN, model [1] in a parallel environment, as well as benchmarking different types of hardware. More specifically it will use the TensorFlow's Mirrored strategy to parallelize a 3D GAN on multiple GPUs and use a TPU strategy to run it on Google's TPUs. The present work shows two approaches to the Tensorflow mirrored strategy, one approach uses the simplified method of parallelizing the training, where it is specified what each GPU can see, and using the built-in logic from the Tensorflow strategy it can train the model in parallel, and a second approach where it is used a custom training loop by manually setting the training process, this is by manually getting the loss, updating the gradients, and the weights of the GAN, with this, is it is possible to have higher control of the training process as well as add further elements to each GPU work, increasing the overall speedup. For the TPUs we use the TPU distributed strategy present in Tensorflow, applying the same approaches as described for the mirrored strategy. This work is validated by comparing the results obtained by the original 3DGAN model as well as the Monte Carlo simulated data obtained from Geant4. It shows the run times and speed-ups obtained in both types of hardware comparing both approaches.

## References

[1] G. R. Khattak, S. Vallecorsa, F. Carminati and G. M. Khan, "Particle Detector Simulation using Generative Adversarial Networks with Domain Related Constraints," 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Boca Raton, FL, USA, 2019, pp. 28-33, doi: 10.1109/ICMLA.2019.00014.

**Primary authors:** DA COSTA CARDOSO, Renato Paulo (Universidade de Lisboa (PT)); VALLECORSIA, Sofia (CERN)

**Presenter:** DA COSTA CARDOSO, Renato Paulo (Universidade de Lisboa (PT))

**Session Classification:** Workshop

**Track Classification:** 6 ML infrastructure : Hardware and software for Machine Learning