


1.8X gains

for CERN's complex generative adversarial network (GAN) model inferencing.¹

2.2X speed up

using multi-stream quantized inference compared to single stream quantized inference.¹

Researchers Demonstrate Significant Deep Learning Inference Speedup

To help address future needs for CERN's Large Hadron Collider (LHC)—the world's largest particle accelerator—researchers at CERN, SURFsara, and Intel have been rethinking approaches for supplying new levels of Monte Carlo based simulations. The team wanted to accelerate a deep learning inferencing workload that held the promise of yielding results much faster than Monte Carlo based simulations. The researchers used the Intel® AI Analytics Toolkit to obtain higher performance for inferencing in 2nd Gen Intel® Xeon Scalable processors with Intel® Deep Learning Boost (Intel® DL Boost). The built-in AI acceleration provided by Intel DL Boost was central to the project's performance gains and shown to accelerate inferencing without sacrificing accuracy.

Products and Solutions

[2nd Gen Intel® Xeon® Scalable Processors](#)

[Intel® Deep Learning Boost](#)

[Intel® AI Analytics Toolkit](#)

Industry

Research

Organization Size

1,001–5,000

Country

Switzerland

Learn more

[Case Study](#)