

CRAY ACCEL AI DEEP LEARNING

Artificial intelligence (AI) and, in particular, deep learning (DL) are rapidly transforming entire industries and scientific disciplines. If your toe isn't already in the AI/DL water, it might feel like you're already too late. You're not — because Cray can help. How? By giving you the tools and support to guide you to success whether you're just starting out or ready for production AI.

Everything You Need to Go From Pilot to Production

Cray Accel AI offerings turn our supercomputing expertise, technologies and best practices into solutions that advance the adoption of deep learning. These fast-start configurations range from a starter system ideal for AI exploration to a complete, production-level cluster supercomputer for training and inference.

Cray has a long track record supporting AI/DL and other complex computing challenges. The high-performance techniques that make AI possible today have been honed over decades in other technologies such as medical imaging, cybersecurity, climate modeling and seismic processing — technologies driven to success by our supercomputing systems.

So wherever you are on your AI journey, we have the high-powered computing systems and the industry expertise to make AI and DL work for you.

Test, Launch and Grow Your Deep Learning Initiatives

Cray Accel AI solutions are for artificial intelligence applications where machine learning, and especially deep learning, is used. They include Cray® CS-Storm™ GPU-accelerated supercomputers featuring NVIDIA® Tesla® V100 “Volta” GPUs and the NVIDIA® NVLink™ GPU-to-GPU interconnect.

Every Accel AI solution comes with a robust deep learning environment from Bright Computing that includes TensorFlow™, MXNet, Caffe2, Chainer, Microsoft Cognitive Toolkit and more.

Choose from three fast-start configurations: the Accel AI Quick Start for initial deep learning trials, the Accel AI Cluster Starter Kit for deep learning exploration and small proof-of-concept projects, or the Accel AI Deep Learning System for production-level deep learning training and inference.

Accel AI Quick Start

If you're at the tool exploration and model development stage, the Accel AI Quick Start solution delivers all the elements necessary for a small team to get started with deep learning trials. You get a single-chassis, fully configured Cray CS-Storm 500NX 8-GPU server with eight NVIDIA Tesla V100 SXM2 GPUs and Bright Computing's comprehensive deep learning environment.

Accel AI Cluster Starter Kit

The Accel AI Cluster Starter Kit offers the technology you need for application development, initial production and proof-of-concept projects. You get two CS-Storm 500NX 4-GPU servers with the same NVIDIA components and Bright Computing software environment as the Quick Start solution but add a management node and networking allowing you to manage the system and connect it directly to your existing cluster infrastructure.

Accel AI Deep Learning System

When you're ready for a production-level system, choose the Accel AI Deep Learning System. This solution gives you four CS-Storm 500NX 8-GPU servers with the same NVIDIA components and Bright Computing software environment as the Quick Start and Starter Kit solutions, then adds a complete management and networking infrastructure with high-speed Ethernet® and InfiniBand® as well as shared storage.

Complete Deep Learning Software Environment

The comprehensive deep learning software environment from Bright Computing provides the AI frameworks, libraries and tools you need for complex machine learning and deep learning workloads.

Machine and deep learning frameworks include Cafe, Caffe2, Caffe-mpi, Chainer, CNTK, Keras, MXNet, Pytorch, TensorFlow, Theano, Torch

Machine learning libraries include scikit-learn, NVIDIA® CUDA® Deep Neural Network library (cuDNN), cuBLAS linear algebra

Supporting accelerated compute libraries include over 400 MB of Python modules that support the machine learning packages, plus the NVIDIA hardware drivers, CUDA (parallel computing platform API) drivers, CUB (CUDA building blocks), NCCL (library of standard collective communication routines)

Cray Accel AI Solution Components

Configuration	Quick Start	Cluster Starter Kit	Deep Learning System
	For tool exploration and early model development	For deep learning proof-of-concept projects	A complete system for production-level deep learning
AI processing node(s)	1 x Cray® CS-Storm™ 500NX with 8 NVIDIA® Tesla® Volta V100 SXM2 accelerators and 2 Intel® Xeon® processors	2 x Cray CS-Storm 500NX with 4 NVIDIA Tesla Volta V100 SXM2 accelerators and 2 Intel Xeon processors	4 x Cray CS-Storm 500NX with 8 NVIDIA Tesla Volta V100 SXM2 accelerators and 2 Intel Xeon processors
Management node(s)		Cray 2828X 2U servers with dual Intel Xeon processors	2 x Cray 2828X 2U servers with dual Intel Xeon processors
Networking	100 Gb/s InfiniBand® EDR and 10 Gb/s Ethernet	100 Gb/s InfiniBand EDR and 10 Gb/s Ethernet	100 Gb/s InfiniBand EDR and 10 Gb/s Ethernet
Usability Software	DIGITS		
DL Frameworks	Cafe, Caffe2, Caffe-mpi, Chainer, CNTK, Keras, MXNet, Pytorch, TensorFlow, Theano, Torch		
Analytics and Data Preparation	scikit-learn		
Accelerated Compute Libraries	cuDNN, NCCL, cuBLAS, CUB, CUDA, gflags, glog, DDF5, leveldb, lmbd, openCV, Protobuf, sbt, maven, ant, and more		