



June 20, 2017

## **Mellanox Interconnect Solutions Scale Deep Learning Platforms to World-Leading Performance**

*Mellanox RDMA Technology Accelerates Leading Deep Learning Frameworks by 2X, Demonstrates Wide Adoption for these Infrastructures*

SUNNYVALE, Calif. & YOKNEAM, Israel--(BUSINESS WIRE)-- Mellanox<sup>®</sup> Technologies, Ltd. (NASDAQ:MLNX), a leading supplier of high-performance, end-to-end interconnect solutions for data center servers and storage systems, announced today that the leading deep learning frameworks such as TensorFlow<sup>™</sup>, Caffe2, Microsoft Cognitive Toolkit, and Baidu PaddlePaddle now leverage Mellanox's smart offloading capabilities to provide world-leading performance and near-linear scaling across multiple AI servers. Mellanox RDMA and In-Network Computing offloads and NVIDIA<sup>®</sup> GPUDirect<sup>™</sup> are key technologies enabling users to maximize their application performance and system efficiencies.

Deep learning is used across industries and the research community to help solve many big data problems such as natural language processing, speech recognition, computer vision, healthcare, life-sciences, financial services and more. Mellanox is enabling these industries into a new era of performance and scalability with the powerful data-centric offload architecture that has been employed by the world's most advanced machine learning platforms.

TensorFlow is an open source software library originally developed by researchers and engineers within Google's Machine Intelligence research group. With the inclusion of RDMA technology in place of traditional TCP, TensorFlow data exchange performance between nodes was accelerated by 2X, enabling faster image processing.

Baidu's PaddlePaddle (Parallel Distributed Deep Learning) is a flexible and scalable deep learning platform. PaddlePaddle supports a wide range of neural network architectures and optimization algorithms, such that it is possible to leverage many CPUs and GPUs to accelerate training. PaddlePaddle leverages RDMA to achieve high throughput and performance, and takes advantage of the more advanced acceleration capabilities of the combined NVIDIA and Mellanox architectures to accelerate deep learning training time by 2X.

"Advanced deep neural networks depend upon the capabilities of smart interconnect to scale to multiple nodes, and move data as fast as possible, which speeds up algorithms and reduces training time," said Gilad Shainer, vice president of marketing at Mellanox Technologies. "By leveraging Mellanox technology and solutions, clusters of machines are now able to learn at a speed, accuracy and scale that push the boundaries of the most demanding cognitive computing applications."

"Developers of deep learning applications can take advantage of optimized frameworks and NVIDIA's upcoming NCCL 2.0 library which implements native support for InfiniBand verbs and automatically selects GPUDirect RDMA for multi-node or NVIDIA NVLink when available for intra-node communications," said Duncan Poole, Director of Platform Alliances at NVIDIA. "NVIDIA NVLink is available in Pascal-based Tesla P100 systems, including the NVIDIA DGX-1 AI supercomputer which has four Mellanox ConnectX<sup>®</sup>-4 100 Gb/s adapters. This allows developers to focus on creating new algorithms and software capabilities, rather than performance tuning low-level communication collectives."

### **Supporting Resources:**

- | Learn more: [Machine Learning](#)
- | Follow Mellanox on: [Twitter](#), [Facebook](#), [Google+](#), [LinkedIn](#), and [YouTube](#)
- | [Join the Mellanox Community](#)

### **About Mellanox**

Mellanox Technologies (NASDAQ:[MLNX](#)) is a leading supplier of end-to-end Ethernet and InfiniBand intelligent interconnect solutions and services for servers, storage, and hyper-converged infrastructure. Mellanox's intelligent interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications and unlocking system performance. Mellanox offers a choice of high performance solutions: network and multicore

processors, network adapters, switches, cables, software and silicon, that accelerate application runtime and maximize business results for a wide range of markets including high performance computing, enterprise data centers, Web 2.0, cloud, storage, network security, telecom and financial services. More information is available at [www.mellanox.com](http://www.mellanox.com).

**Note: Mellanox and ConnectX are registered trademarks of Mellanox Technologies, Ltd. All other trademarks are the property of their respective owners.**

View source version on [businesswire.com](http://www.businesswire.com): <http://www.businesswire.com/news/home/20170619005312/en/>

**Mellanox Technologies, Ltd.**

**Press/Media Contact**

McGrath/Power Public Relations and Communications

Allyson Scott, +1-408-727-0351

[allysonscott@mcgrathpower.com](mailto:allysonscott@mcgrathpower.com)

or

**Israel PR Contact**

Galai Communications Public Relations

Jonathan Wolf, +972 (0) 3-613-52-48

[yonit@galaipr.com](mailto:yonit@galaipr.com)

Source: Mellanox Technologies

News Provided by Acquire Media