Interconnect Your Future

Paving the Road to Exascale

Forum Teratec - June 2018

Data is Growing Faster Than Ever

Autonomous vehicle generates 4000GByte per day

CAMERA
~20-40MB Per/sec

SONAR
~10-100KB Per/Sec

GPS
~50KB Per/Sec

RADAR
~10-100KB Per/Sec
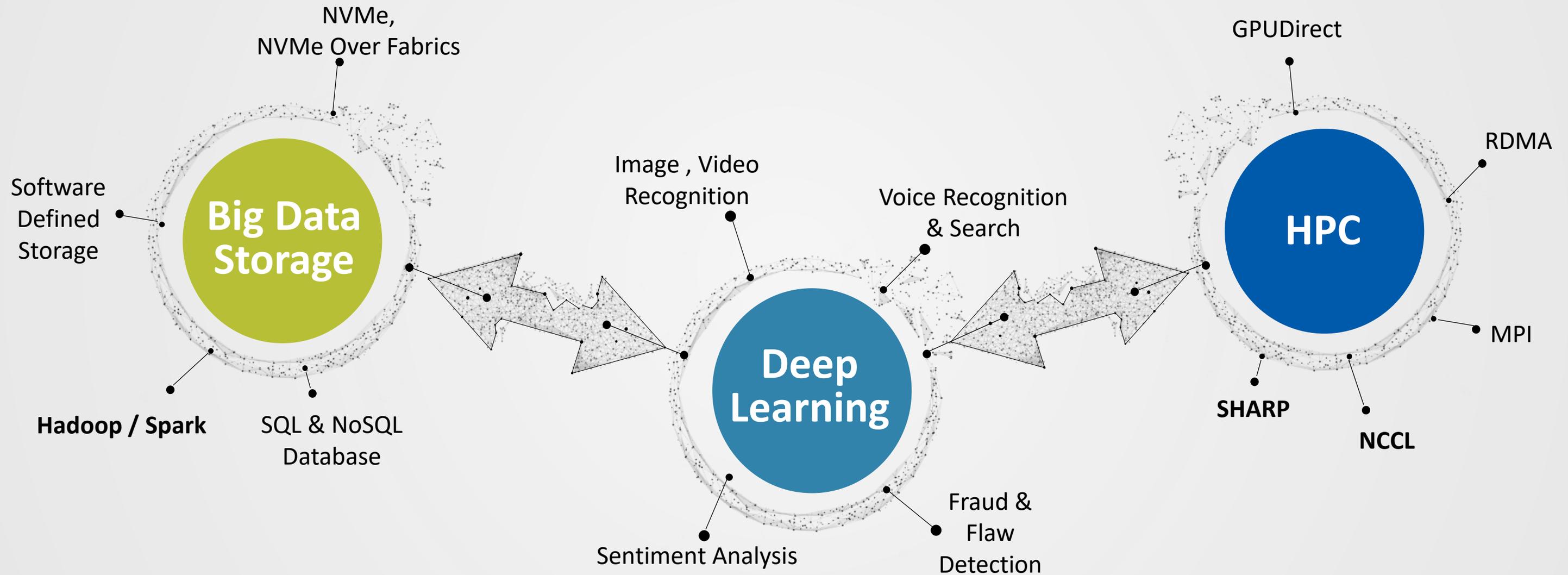
Light Detection & Ranging
~10-70MB Per/Sec

- Data will grow by a factor of 10 over the next decade to 160 Zeta Bytes in 2025 (source: IDC)
- Faster Data processing requires faster Interconnect speeds

HPC and AI Needs the Most Intelligent Interconnect

Same Interconnect Technology Enables a Variety of Applications

NVMe,
NVMe Over Fabrics

GPUDirect

Software
Defined
Storage

**Big Data
Storage**

**Deep
Learning**

HPC

Image , Video
Recognition

Voice Recognition
& Search

RDMA

Hadoop / Spark

SQL & NoSQL
Database

Sentiment Analysis

Fraud &
Flaw
Detection

SHARP

NCCL

MPI

Highest Performance 100, 200 and 400Gb/s Interconnect Solutions

ConnectX·6
Adapters

200Gb/s, 0.6us Latency
200 Million Messages per Second
(10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)

Quantum
Switch

40 HDR (200Gb/s) Ports
80 HDR100 (100Gb/s) Ports
16Tb/s Throughput, 15.6 Billion msg/sec

Spectrum·2
Switch

16 400GbE Ports, 32 200GbE Ports
64 100GbE Ports, 128 25/50GbE Ports
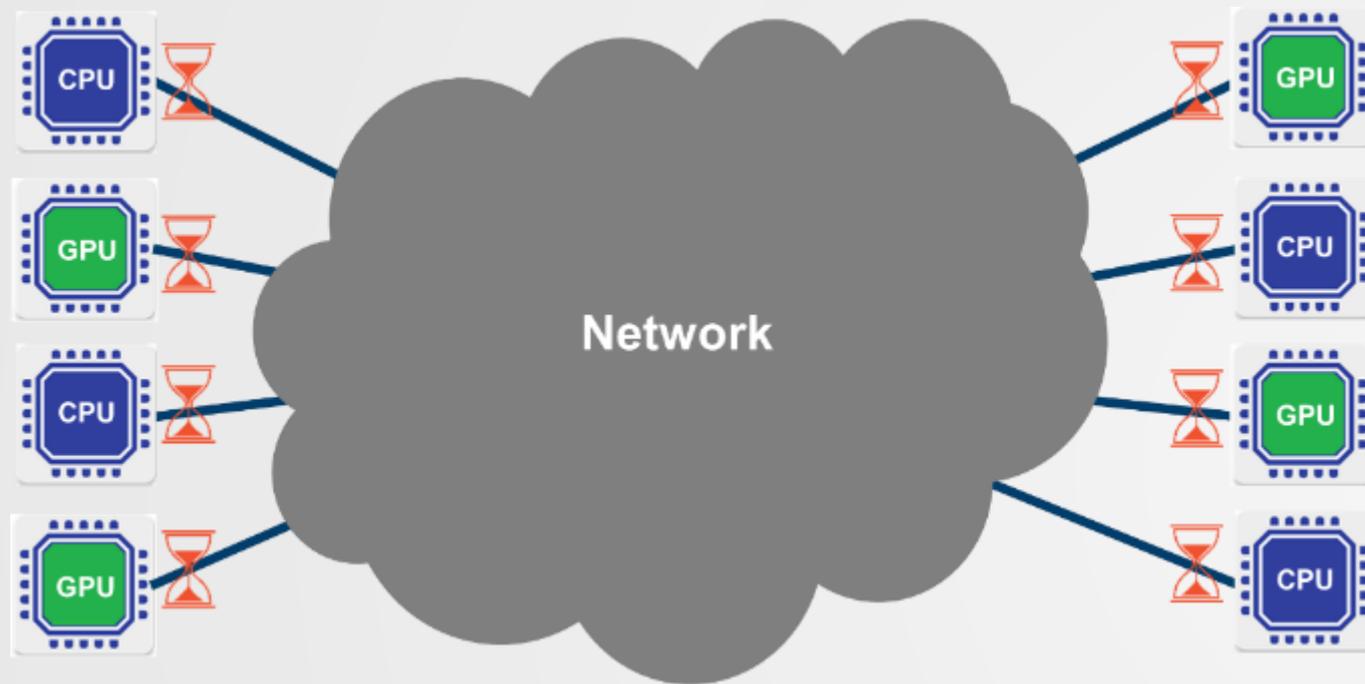Throughput of 6.4Tb/s

LinkX
Interconnect

Transceivers
Active Optical and Copper Cables
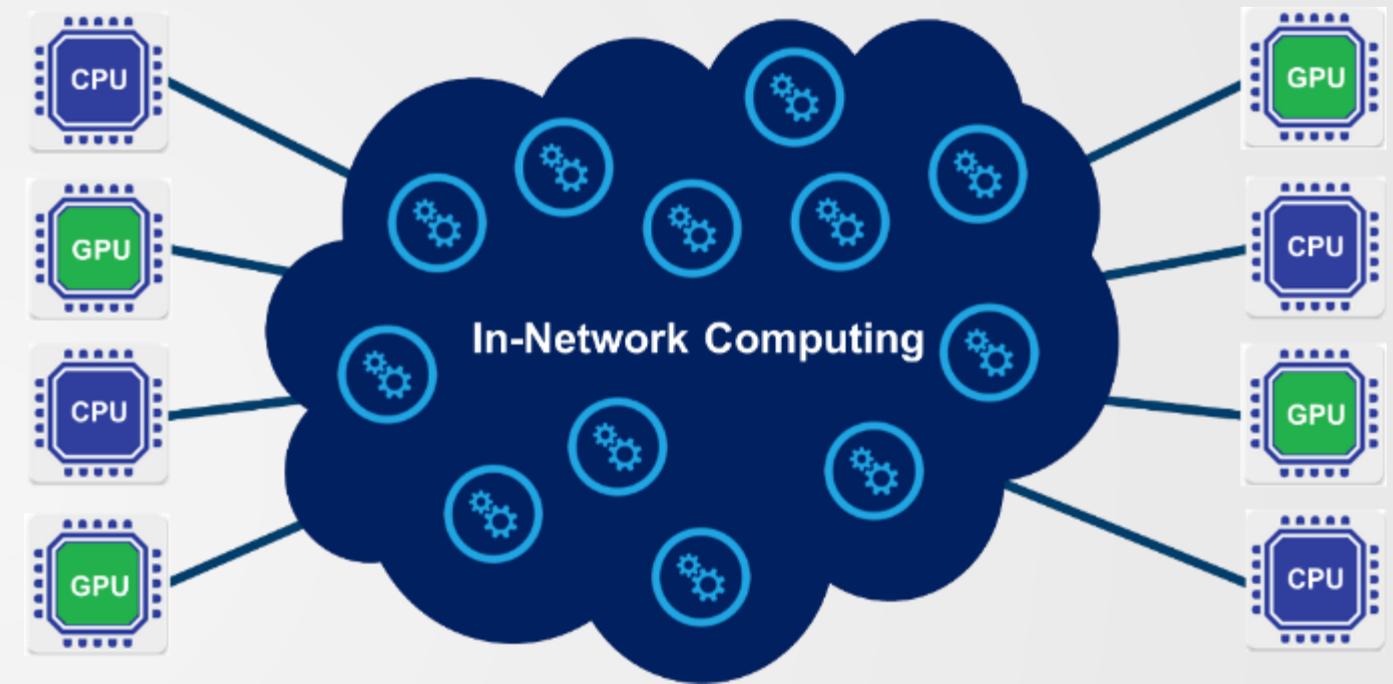(10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)
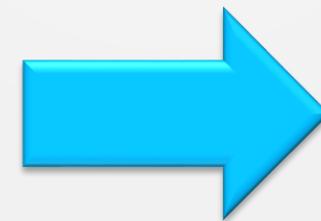
The Need for Intelligent and Faster Interconnect

CPU-Centric (Onload)

Data-Centric (Offload)

Network

In-Network Computing

Must Wait for the Data
Creates Performance Bottlenecks

Analyze Data as it Moves!

Faster Data Speeds and In-Network Computing Enable Higher Performance and Scale

In-Network Computing Delivers Highest Performance

SHARP
In-Network Computing → **10X** Performance Acceleration
Critical for HPC and Machine Learning Applications

MPI
In-Network Computing → **35X** Performance Acceleration
Delivers Highest Application Performance

SHIELD
Self-Healing Technology → **5000X** Faster Network Recovery
Unbreakable Data Centers

GPUDirect™ RDMA
GPU Acceleration Technology → **10X** Performance Acceleration
Critical for HPC and Machine Learning Applications

Mellanox Accelerates Next Generation HPC and AI Systems

Oak Ridge National Laboratory → Summit CORAL System — U.S. DEPARTMENT OF ENERGY

Lawrence Livermore National Laboratory → Sierra CORAL System — U.S. DEPARTMENT OF ENERGY

AIST — National Institute of Advanced Industrial Science and Technology (AIST) → Fastest Supercomputer in Japan
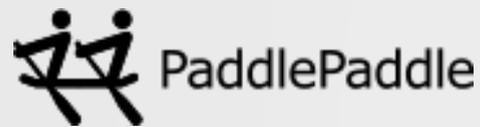
University of Toronto → Fastest Supercomputer in Canada
InfiniBand Dragonfly+ Topology — SciNet

bp → Highest Performance with InfiniBand

An Intelligent Network Unlocks the Power of Artificial Intelligence

RDMA Supercharges Leading AI Frameworks

Microsoft Cognitive Toolkit | **Apache Spark** | **Caffe2**

PaddlePaddle | **torch** | **TensorFlow**

Mellanox Supercharges Leading AI Companies

facebook | **Microsoft** | **Yahoo!** | **NVIDIA**

PayPal | **Alibaba Group** | **Baidu** | **Tencent 腾讯**

60%
Higher ROI

50%
Lower CapEx & OpEx

Thank You