

### INTELLIGENCE AT THE EDGE

-HIGH PERFORMANCE EMBEDDED COMPUTING TRENDS (HPEC)



### **WE ARE TRITECH**

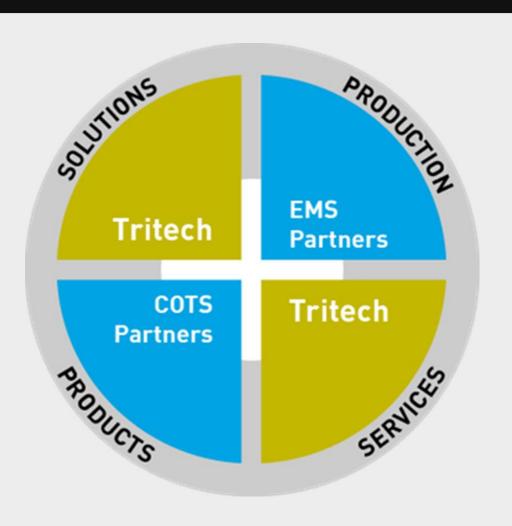
- Embedded products, solutions and engineering services for the intelligent systems and industrial IoT-markets
- Offices in Stockholm, Gothenburg and Helsinki.
- Approx. 125 employees
- Turnover forecast 160 MSEK
- More than 600 products designed for customers since start up
- Partnerships with major technology and COTS vendors







### **INTEGRATED SOLUTIONS BUSINESS MODEL**





# SOLUTIONS BASED ON CUSTOMIZATION OF COTS BUILDING BLOCKS

Sida 4

#### **Embedded COTS system**

#### Custom COTS strategy

Adaption of existing standard products to the requirements of the customer



#### **Modifications**

System configuration including addon hardware

#### **Customer solution**

#### Semicustom design strategy

Custom design by reuse of building blocks



#### System design

- Board development
- Enclosure and thermal design
- System integration

Components and COTS modules

### DRIVERS FOR HIGH PERFORMANCE IN EMBEDDED

- Fog and edge computing
- Autonomous vehicles
- Intelligent machine vision
- Medical imaging
- Real time video applications
- Military

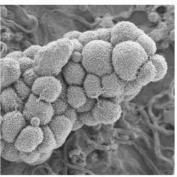


## FROM CLOUD TO VEHICLES

Sida 6

## **DEEP LEARNING EVERYWHERE**











#### INTERNET & CLOUD

Image Classification Speech Recognition Language Translation Language Processing Sentiment Analysis Recommendation

MEDICINE & BIOLOGY

Cancer Cell Detection Diabetic Grading Drug Discovery

#### MEDIA & ENTERTAINMENT

Video Captioning Video Search Real Time Translation

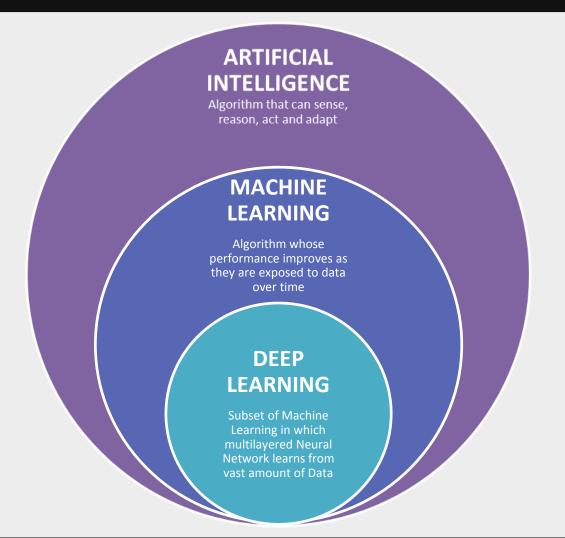
#### SECURITY & DEFENSE

Face Detection Video Surveillance Satellite Imagery

#### **AUTONOMOUS MACHINES**

Pedestrian Detection Lane Tracking Recognize Traffic Sign

#### **ARTIFICIAL DEEP NEURAL NETWORKS - DEEP LEARNING**



## **GROWING SOFTWARE ECOSYSTEM**

# Deep Learning Frameworks





**DEEPLEARNING 4J** 



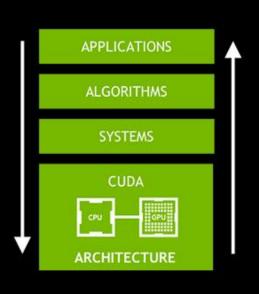


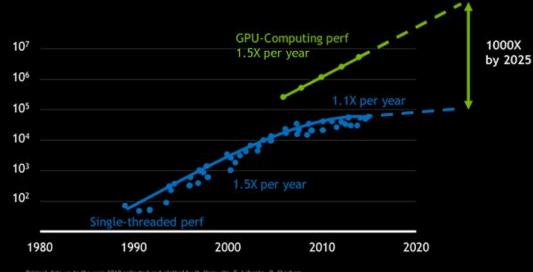




## WHY MOVE TO HETEREGENOUS COMPUTING?

## RISE OF GPU COMPUTING





Original data up to the year 2010 collected and plotted by M. Horowitz, F. Laboote, O. Shacham, K. Olakotun, L. Hammond, and C. Batten New plot and data collected for 2010-2015 by K. Rupp

## **TYPICAL TRAINING AND INFERENCE SCENARIO**

Sida 10



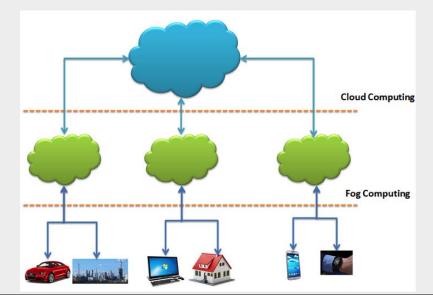
Create your model with Caffe or TensorFlow

Train Your Model in the Cloud

Deploy it in the Field

## FOG AND EDGE COMPUTING

- Industrial, smart city, automotive and media transport applications drive new network requirements.
  - Intelligent endpoints
  - Gateways and edgerouters.
  - Embedded datacenters/microservers.



## WHAT HARDWARE TECHNOLOGY IS NEEDED?

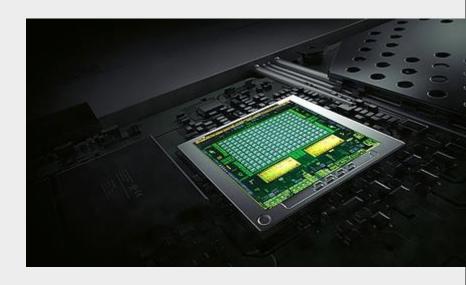
- PC gaming/professional workstation
- Server/datacenter





## EMBEDDING SUCH TECHNOLOGY HAS CHALLENGES, Ida 13

- SWaP considerations
  - Size
  - Weight and
  - Power
- Thermal management
- Temperature range
- Revision control
- Product lifecycle
- Custom-design is expensive and complex



## **HIGH PERFORMANCE COTS OPTIONS**

- Building blocks such as
  - Computer on modules
  - Motherboards
  - Blades/backplanes
  - GPU cards and modules
  - FPGA solutions
  - Storage subsystems

- Application-ready systems such as
  - Rackmount computers
  - Network appliances
  - Box computers





### **HPEC TRENDS FOR "SEMI-CUSTOM" DESIGN**

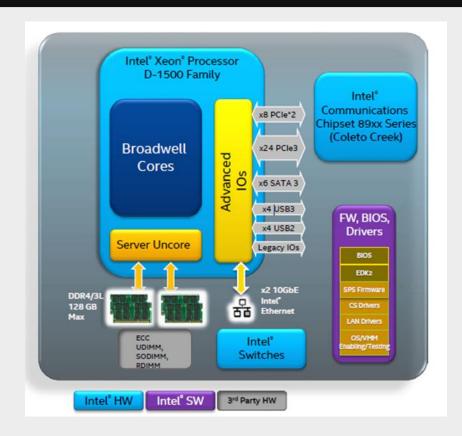
- New COM Express standard for servertechnology, type 7
- Type 6 still mainstream
- Faster storage with NVMe
- Embedded discrete GPUs -MXM modules
- Mini PCIe/mSATA moving to M.2 connector
- Early adoption of 10 Gb ethernet
- Widespread adoption of USB
   3.0 and PCIe 3.0





#### INTEL XEON D FITS DEMANDING FOG APPLICATIONS

- CPU
  - Up to 16C BDW Intel® Xeon® XMT (14 nm)
  - Targeted TDP ~19-65 W
  - Intel® Xeon® Features
  - Industrial temp options
  - Functional safety IEC 61508 support option
- Memory
  - Two Memory Channels, two DIMMS/channel
  - DDR3L/DDR4
  - 128 GB Max Capacity
- Integrated IOs
  - x24 PCle 3.0, x8 PCle 2.0
  - x6 SATA3
  - x4 USB 3.0, x4 USB 2.0
    - x2 10 GbE Intel® Ethernet
- Platform highlights
  - Intel® Advanced Vector Extensions 2 (AVX2)
  - Intel® Cache Monitoring Technology (CMT)
  - Intel® Cache Allocation Technology (CAT)
  - Memory Bandwidth Monitoring (MBM)
  - Intel® Virtualization Technology
  - Intel® Advanced Encryption Standard New Instructions (Intel® AES-NI)
  - Intel® Secure Key Instructions
- Storage/Network Features
  - Storage/Network Environment (Reliability, Temp, Availability)
  - Validated with external Crypto accelerator (Coleto Creek)
  - Non-Transparent Bridging (NTB), Asynchronous DRAM self-refresh (ADR), Intel® QuickData Technology



### **COM EXPRESS TYPE 7 – SERVER ON MODULE**

- Xeon D and Atom (Denverton) incl. industrial temp options
- Changes when compared to Type 6
  - No graphics, no audio
  - 4x 10 Gigabit Ethernet with complete set of sideband signals (including SDT pins to enable hardware based precision timing)
  - Introduce eSPI as future replacement for LPC
  - 32 PCI Express Lanes
  - Reduction to 2x SATA in favors of PCIe for NVMe storage interface
  - Reduction to 4x USB 3.0 and no additional 4x UBS 2.0 ports

- Added interfaces
  - 8 PCIe (32 lanes now)
  - 4x 10G Ethernet
    - with SDP
  - NC-SI
  - eSPI shared with LPC



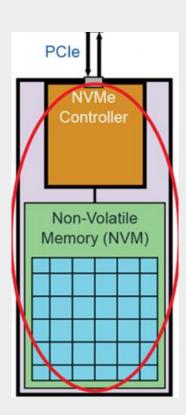


## **TYPE 7 VS TYPE 6 OVERVIEW**

COM Express Type 6			COM Express Type 7	
Gigabit Ethernet	USB 3.0 0-3		Gigabit Ethernet	USB 3.0 0-3
LPC			LPC / eSPI	
CATA 0.2	PCle 6-7		SATA 0-1	10GBaseKR 0-3
SATA 0-3			USB 2.0 0-3	
HDA	DDI 0-2			
USB 2.0 0-7		PCIe 0-15		
ExpressCard				
PCIe 0-5	PEG			
GPIO / SDIO				
LVDS / eDP			GPIO / SDIO PCIe 16-31 SER 0-1 / CAN	PCIe 16-31
SER 0-1 / CAN				
SPI & I2C			SPI & I2C	
Power	Power		Power	Power

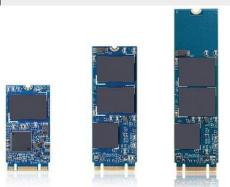
## STORAGE MOVING FROM SATA TO NVME

- SATA speed limited to 6Gbit/s.
- With PCI Express the max. theoretical bandwidth of the host interface is increased to 4GB/s with an x4 Gen3 link.
- Initial adoption in server/datacom and video.
- 3DXPoint (Intel Optane/Micron QuantX) in early stage
- SATA performance still good enough for most embedded applications.
- Native NVMe is supported in the latest Intel Core BIOS.



## M.2 (NGFF) REPLACING MINI PCIE AND MSATA

- Smaller and more flexible
- Defines several pinouts and dimensions
- 2242, 2280 mainstream for storage
- 67 pins with 0,5 mm pitch different socket types, keys, for:
  - Connectivity
  - WWAN / SSD / Others
  - SSD Drives
- Both NVMe and SATA widely available for storage





### DISCRETE GPU TECHNOLOGY FOR EMBEDDED





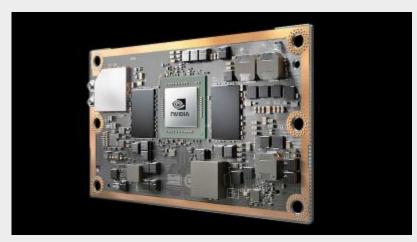
- MXM (3.1) modules are suitable for carrierboard/COM module projects
- VPX/XMC/Compact PCI serial variants are available
- Up to Nvidia GTX1080 available
- Extended lifecycle and revision control supported
- Wide temp options on some SKUs
- Conduction cooling solutions possible on some SKUs

### **HPEC BOX COMPUTER EXAMPLE**

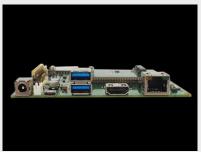


- High performance 7th gen Xeon and Core i 7 CPUs, C236 chipset
- Nvidia GTX 1050 GPU
- 10 GbE plus multiple PoE/PoE+ ports
- USB 3.0, PCle 3.0 supported
- 32 isolated DIO
- 3 SIM for WiFi/3G/4G/ LTE/GPS/GPRS/UMTS
- Vehicle power with 80V surge protection
- -20 + 60C op. temp range
- Anti shock, anti vibration

## **ENTRY EDGE AI SOLUTIONS – NVIDIA JETSON**



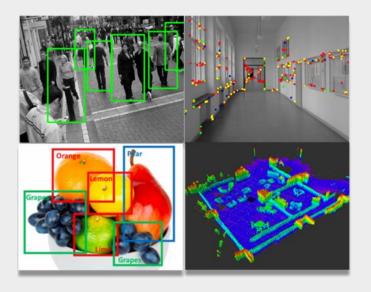




- Nvidia SOM with Tegra TX1 or TX2 SOC
- Multicore ARM CPU with Nvidia GPU
- Off the shelf carriers and boxes available
- Custom carriers designed based on Nvidias SOM spec
- Ideal solution for lower performance requirements while still using CUDA and Nvidia SW ecosystem

### **ENTRY EDGE AI SOLUTIONS – AAEON AI CORE**





- Mini PCIe module based on Movidius/Intel Myriad 2 VPU
- Verified with Aaeon SBCs and box PCs, and UP maker boards
- Works in any mini PCIe slot in x86 based systems
- Low cost and low TDP (1W) solution for smart vision applications
- Supports Caffee and Tensorflow.

### TRITECH PARTNERS FEATURED

Sida 25



congared







- COM Express type 7 and 6 modules
- Industrial, embedded GPU cards and modules, Jetson Tegra TX1/2 solutions
- Embedded, rugged. wide temp box computers for HPEC applications incl. GPGPU
- Al Core VPU module, SBCs, box PCs