INAF – Meeting ICT – Trieste, 14-18 nov 2016

High-Performance Computing using FPGAs a proposal to INAF

Massimiliano Belluso, Sergio Billotta e Daniele Gardiol



Summary

- Introduction to FPGAs
 - What is an FPGA?
 -
 - VHDL
- FPGA Vendors
 - List
 - Products
 -
- HPC using FPGAs
 - FPGA Accelerator Cards
- a proposal to INAF
- Q&A



Introduction to FPGAs (1/2)

What is an FPGA?

Field Programmable Gate Arrays (FPGAs) are semiconductor devices that are based around a matrix of configurable logic blocks (CLBs) connected via programmable interconnects.

FPGAs can be reprogrammed to desired application or functionality requirements after manufacturing.

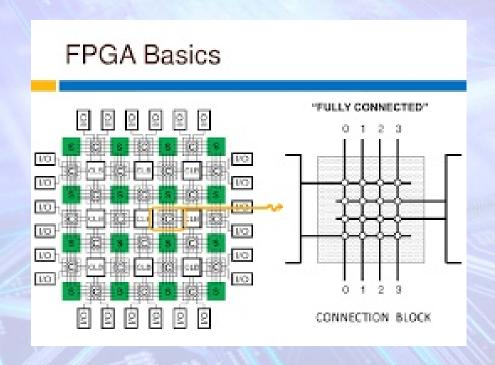






Introduction to FPGAs (2/2)

FPGA Architecture

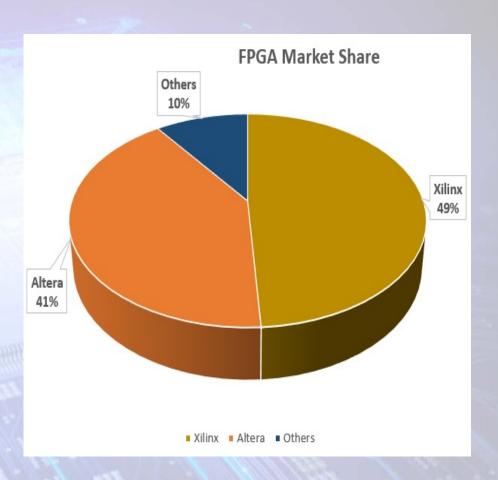




FPGA Vendors

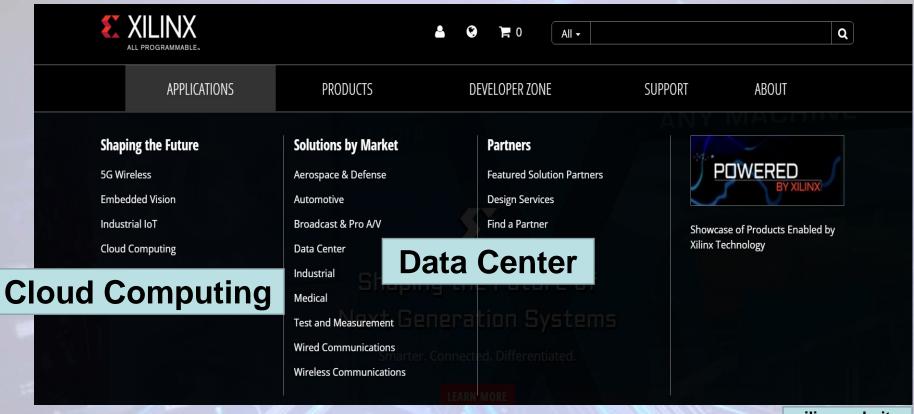
Xilinx Altera

Lattice Semiconductor
Microsemi Corporation
Cypress Semiconductor
QuickLogic
Aeroflex Inc
Tabula





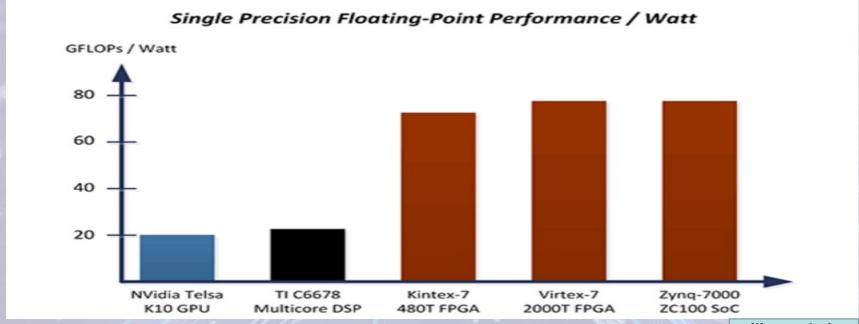
Applications





HPC and Data Storage (1/3)

Xilinx 7-Series FPGAs and All Programmable SoCs deliver power efficient, high-performance processing solutions for aerospace and defense, medical, scientific, oil and gas, financial, communications and life science applications. The parallelism and customizable architecture inherent in the FPGA architecture is ideal for high-throughput processing and software acceleration. These devices are built on a 28nm silicon process that introduces HKMG technology to maximize usable system performance through lower power. All Xilinx devices support long product lifecycles that mitigate obsolescence risk. These factors combine to allow HPC platforms based on Xilinx devices to deliver massive processing performance up to 2 TFLOPS, in a single chip, at a fraction of the power of GPUs and multi-core DSPs.





HPC and Data Storage (2/3)

Xilinx High-Performance Compute (HPC) Platforms

Vendor	Platform	Xilinx Device	SP Peak Performance	DDR3 Memory
Xilinx	VC709	Virtex-7 690T	69 GFLOPS	8 GB
Xilinx	VC707	Virtex-7 485T	48 GFLOPS	1 GB
Xilinx	KC705	Kintex-7 325T	48 GFLOPS	1 GB
Xilinx	ZC706	Zynq-7000 7Z045	65 GFLOPS	1 GB



HPC and Data Storage (3/3)

Rapid Prototyping in C/C++ using Vivado HLS

Rapid prototyping with Xilinx delivers uncompromised reliability, Xilinx is known for, with the most raw compute power at the lowest power. Confidently go from concept to market with Xilinx compute acceleration solutions. Vivado HLS provides rapid prototyping design flow software applications developed in single or double precision C/C++. These applications can be compiled to efficient hardware implementations that can be programmed into Xilinx 28nm devices. Vivado HLS is includes as part of Vivado Design Suite: System Edition.

Software Based System Realization with C/C++ and OpenCL

Xilinx is currently working with early customers on a new system level, heterogeneous parallel programming environment that leverage abstractions such as C/C++ and Open Computing Language (OpenCL®), in a comprehensive Eclipse-based development environment.



Data Center

Xilinx offers expertise and Smarter Solutions in three distinct application spaces within the Data Center:

1. Networking

- o Transition from 10G to 40G to 100G Ethernet
- o In-rack routing, e.g, 2 and 3-D Torus
- Ultra Low Latency

2. Storage and Memory

- Solid state drives (SSD)
- Flash memory over fabric
- o Compute near memory
- · High bandwidth stacked memory

3. Parallel Computing

- Deep Learning and Neural Networks
- Video, Image, and Voice Applications
- Security in the Cloud
- o Hadoop, Database, Distributed Search
- Personalized Medicine

Through hardware and software programmability, the power of All Programmable devices accelerates processing and throughput and enables the creation of entirely new classes of servers, storage devices, and network solutions with the flexibility to enable virtualization and services as data center technology continues to evolve.



Altera (INTEL)





Online Newspapers



Newsroom

Top News Sections ▼

News By Category ▼

All News ▼

Search Newsroo

News Release

December 28, 2015

Share this Article









Contact Intel PR

INTEL COMPLETES ACQUISITION OF ALTERA

SANTA CLARA, Calif., Dec. 28, 2015 – Intel Corporation ("Intel") today announced that it has completed the acquisition of Altera Corporation ("Altera"), a leading provider of field-programmable gate array (FPGA) technology. The acquisition complements Intel's leading-edge product portfolio and enables new classes of products in the high-growth data center and Internet of Things (IoT) market segments.



"Altera is now part of Intel, and together we will make the next generation of semiconductors not only better but able to do more," said Brian Krzanich, Intel CEO. "We will apply Moore's Law to grow today's FPGA business, and we'll invent new products that make amazing experiences of the future possible – experiences like autonomous driving and machine learning."

Altera will operate as a new Intel business unit called the Programmable Solutions Group (PSG), led by Altera veteran Dan McNamara. Intel is committed to a smooth transition for Altera customers and will continue the support and future product development of Altera's many products, including FPGA, ARM°-based SoC and power products. In addition to strengthening the existing FPGA business, PSG will work closely with Intel's Data Center Group and IoT Group to deliver the next generation of highly customized, integrated products and solutions.



Online Newspapers





IBM's cloud service will host the Xilinx SDAccel development environment which will allow developers to describe their algorithms in OpenCL, C, and C++ and then compile directly to Xilinx EPGA-based acceleration boards.



This is an open access cloud service, called SuperVessel, which can be used by application developers, system designers, and academic researchers to create, test and pilot their FPGA designs for big data analytic processors and even data gathering IoT node devices.



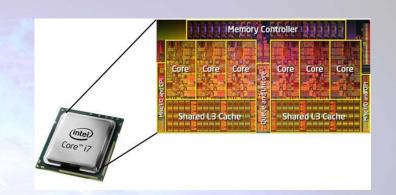
HPC Requirements

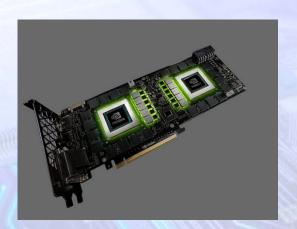
- Compute Elements
 - CPU (Multi-core,)
 - GPU FPGA (News)
- Programming Model
 - Efficient, Open,







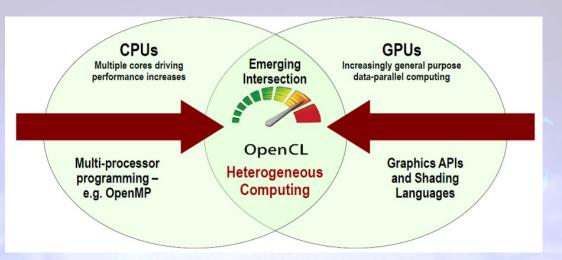




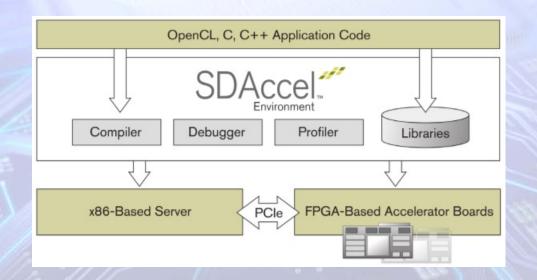


HPC Requirements













Accelerator card based on FPGAs



Solutions

Support

About

Contact



Nallatech's server qualified PCIe accelerator cards and integrated rack-mount servers are compatible with the Altera Software Development Kit (SDK) for OpenCL.

Nallatech's accelerators allow customers to combine the OpenCL FPGA programming model with Altera's massively parallel FPGA architecture. This combination enables dramatic acceleration intensive applications while reducing power consumption and total cost of ownership.

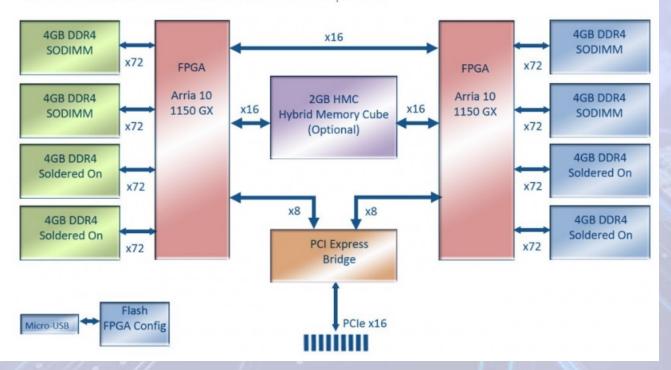




Accelerator card based on FPGAs

Datacenter Co-Processor FPGA Compute Acceleration Card

- GPU Form Factor Card with (2) Arria 10 10A1150GX FPGAs
 - o Dual Slot Standard Configuration
 - o Single Slot width possible, if user design fits within ~100W power footprint
- PCle Gen3 x 16 Host Interface
- 290 GBytes/s Peak Aggregate Memory Bandwidth
 - o 85GB/s Peak DDR4 Memory Bandwidth per FPGA (4 Banks per FPGA)
 - o 30GB/s Write + 30GB/s Read Peak HMC Bandwidth per FPGA





Accelerator card based on FPGAs

TeraDeep's real-time video analytics run on (gasp) FPGA-based Micron/Pico Computing AC-510 platform

by **₹ XILINX sleibso** on 10-18-2016 03:43 PM

(3,853 Views)

Real-time video analytics that can recognize 1000 object classes at frame-rate speeds need heavy-duty processing, so TeraDeep pulled a really big gun from the rack: a <u>Micron Pico Computing AC-510 accelerator card</u> based on a <u>Xilinx Kintex UltraScale KU060 FPGA</u> connected to Micron's Hybrid Memory Cube (HMC). TeraDeep's <u>press release</u> says:

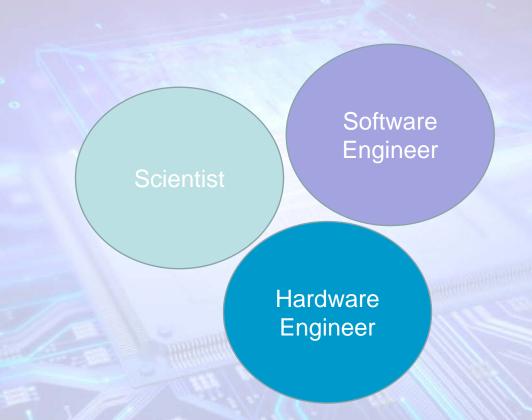
"For low-latency applications such as video analytics, where quick recognition and tracking of fast-moving objects is critical, the graphical processing units (GPUs) used in conventional systems are at a disadvantage. TeraDeep instead uses an FPGA-based architecture that offers faster analytics at half the power, making it an ideal candidate for on-premise appliances.

The first version of the company's solution is an FPGA-based PCIe board that achieves a four-time lower latency compared with the latest GPUs."

TeraDeep will be demonstrating this technology in Micron's booth at the Supercomputing 2016 Conference on November 14-17 in Salt Lake City, Utah.



Power to the People!





..... a proposal to INAF

