About Faster Technology

- Faster Technology started with focus on government market
  - Signal processing solutions
- Now mostly a commercial focus
- Xilinx Certified Alliance Partner
  - FPGA boards and accessories
  - Design services
- Xilinx Authorized Training Provider

Faster Technology
Operating Beyond the Speed of Technology

Using FPGAs for Computation

- Why
- Hardware
- Software

Why consider FPGAs for HPC?

- Better Power / Performance than GPU/CPU
- 3.8 Tb/s of high speed serial bandwidth
  - PCIe
  - Ethernet
  - Serial Memories
  - CCIX
- 172 GB/s of external DDR4 bandwidth
- 38.3 INT8 TOP/s
  - 12,288 DSP blocks

Faster Technology
Operating Beyond the Speed of Technology

Why consider FPGAs for HPC?

- Multiple memory types / configurations
  - Range of internal memory types
    - Small LUT Rams
    - 11.8 MB of 36Kb Block Rams
    - 45 MB of 36KB Ultra Rams
    - 4 or 8 GB of High Bandwidth Memory (HBM)
  - External parallel memories
    - DDR / QDR / RLDRAM
  - External (multi-ported!) serial memories
    - Micron Hybrid Memory Cube, Mosys Bandwidth Engine

Some FPGA Background

- Two branches to the family tree
  - Traditional FPGAs
    - Columns of configurable resources
    - Configurable routing
  - System on Chip parts
    - Combine a dedicated processor system with traditional FPGA on same die

Traditional FPGAs

- Very wide range of sizes available
  - ~6K to 5M “Logic Cells”
  - Most parts single die
  - Larger multi-die parts
- Current generation arranged into columns of different types of resources and IO

The main building blocks

- Configurable Logic Blocks (CLBs)
  - 8 Look Up Tables (LUTs) for Boolean logic
    - Any function of six inputs
    - Two functions of the same five inputs
    - Small memory
    - Small addressable shift register
  - 16 Registers to hold state
  - Some extra support logic
    - Carry look ahead logic for adders and counters
    - Muxes for wider logic functions

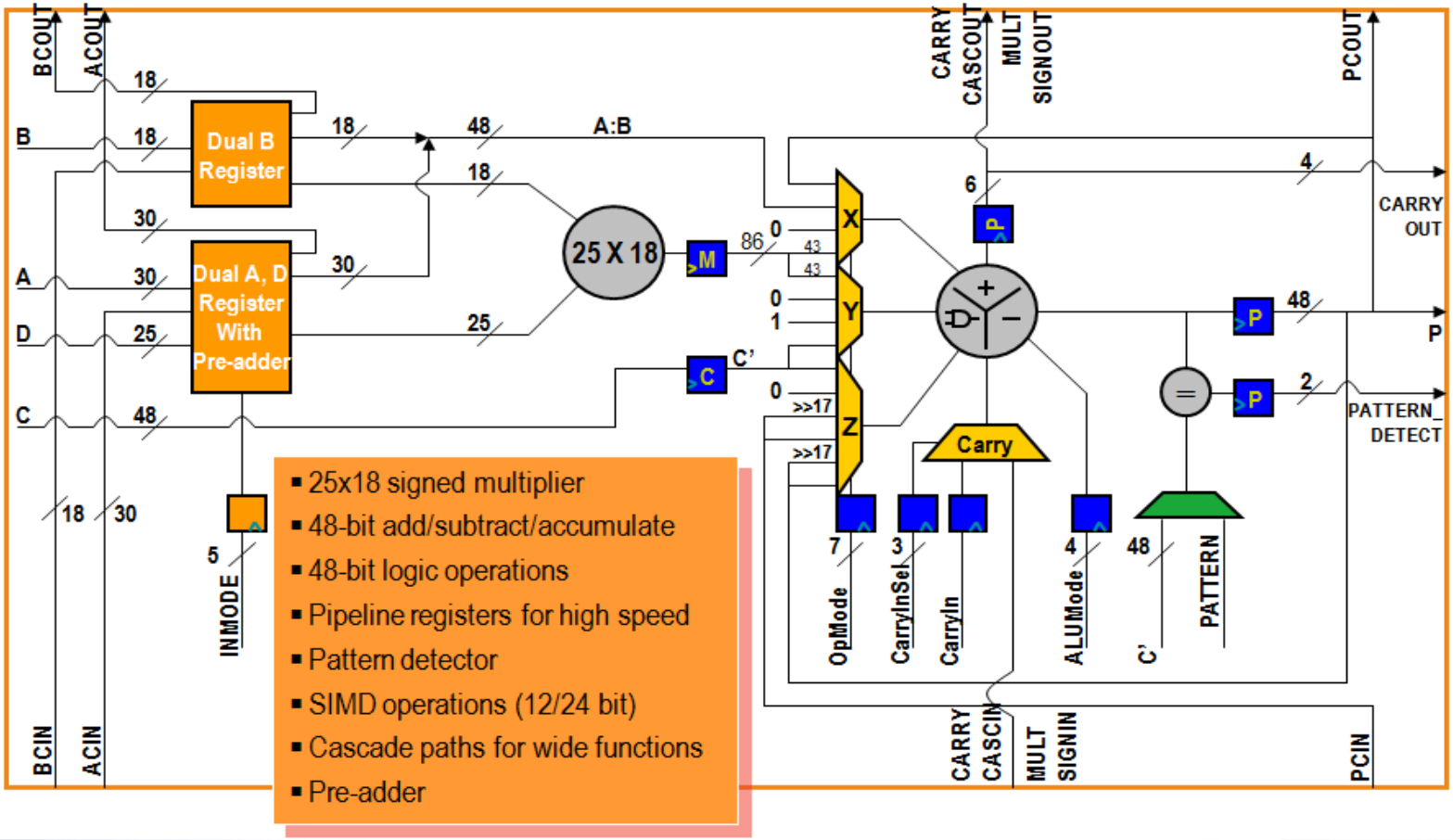
Embedded Memories

- Block RAMs (BRAMs)
  - 36Kb
  - True dual port
  - Can be used as FIFOs
  - Can have ECC, byte write, multiple aspect ratios
- UltraRAMs
  - 36KB
  - Only in some parts

DSP Blocks

- Looks like a bit slice unit
  - Many signals for cycle by cycle control
- 25x18 signed multiplier
- 48 bit accumulator / ALU
- Pre-adder for symmetric FIR filters
- Data path muxes and pipeline registers
- Dedicated cascade routing up the column

Picture of DSP Block



Faster Technology
Operating Beyond the Speed of Technology

External IO

- Banks of IO Blocks (IOBs)
  - High performance
  - High range
  - High density
- Each bank can have a different IO voltage
- Many different supported standards
- Combinatorial and registered data paths
- IO delay lines for timing adjustments
- Simple SERDES

Multi-Gigabit Transceivers

- Grouped into “Quads” of four with some shared resources
- Extremely configurable to support a large range of protocols
- Different types in different FPGAs
  - GTX 12.5 Gb/s
  - GTH 16.3 Gb/s
  - GTY 32.75 Gb/s
  - GTM 58 Gb/s using PAM4

PCIe

- Multiple PCIe blocks in most FPGAs
- End point or root port
  - Root complex with a bit of added logic
- Up to Gen 4x8 in some FPGAs
- Up to Gen 3x16 in most FPGAs
- Can be used with partial reconfiguration to load part of the FPGA design over PCIe

Some other blocks

- ADC
  - System Monitor 1 by 200 KS/s
  - XADC 2 by 1 MS/s
  - Internal temp and voltage sensors
  - Specific external pins
- 100G Ethernet MAC
- 150G Interlaken MAC
- Video Codec
  - H.263 and H.264

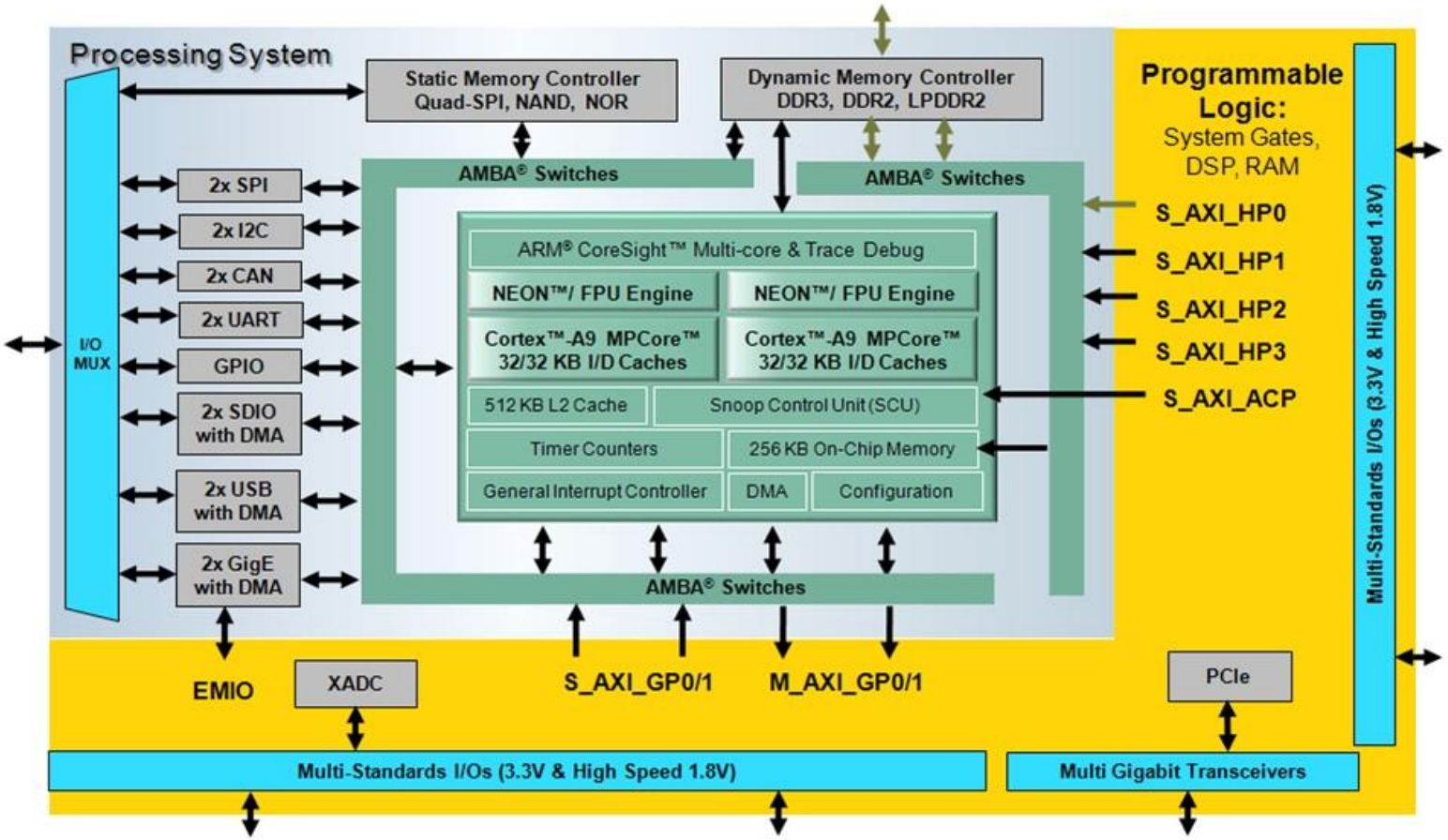
Just Being Introduced

- High Bandwidth Memory
  - Add stacks of HBM DDR in same package
  - 4 or 8 GB
  - 460 GB/S
  - 32 AXI ports to access it
  - Cache Coherent Interconnect (CCIX)
- RFSoC
  - Substitute A2Ds and DACs for an MGT quad
  - For radar and wireless

Zynq SoC Parts

- Zynq-7000
  - 1 or 2 ARM Cortex A9 32 bit processors
  - DDR2/3 memory controller
  - Flash memory controllers
  - AXI interconnections with the FPGA fabric
    - Including Accelerator Coherency Port
  - Supporting IO peripherals
    - Ethernet
    - UART
    - ETC

Zynq-7000 picture

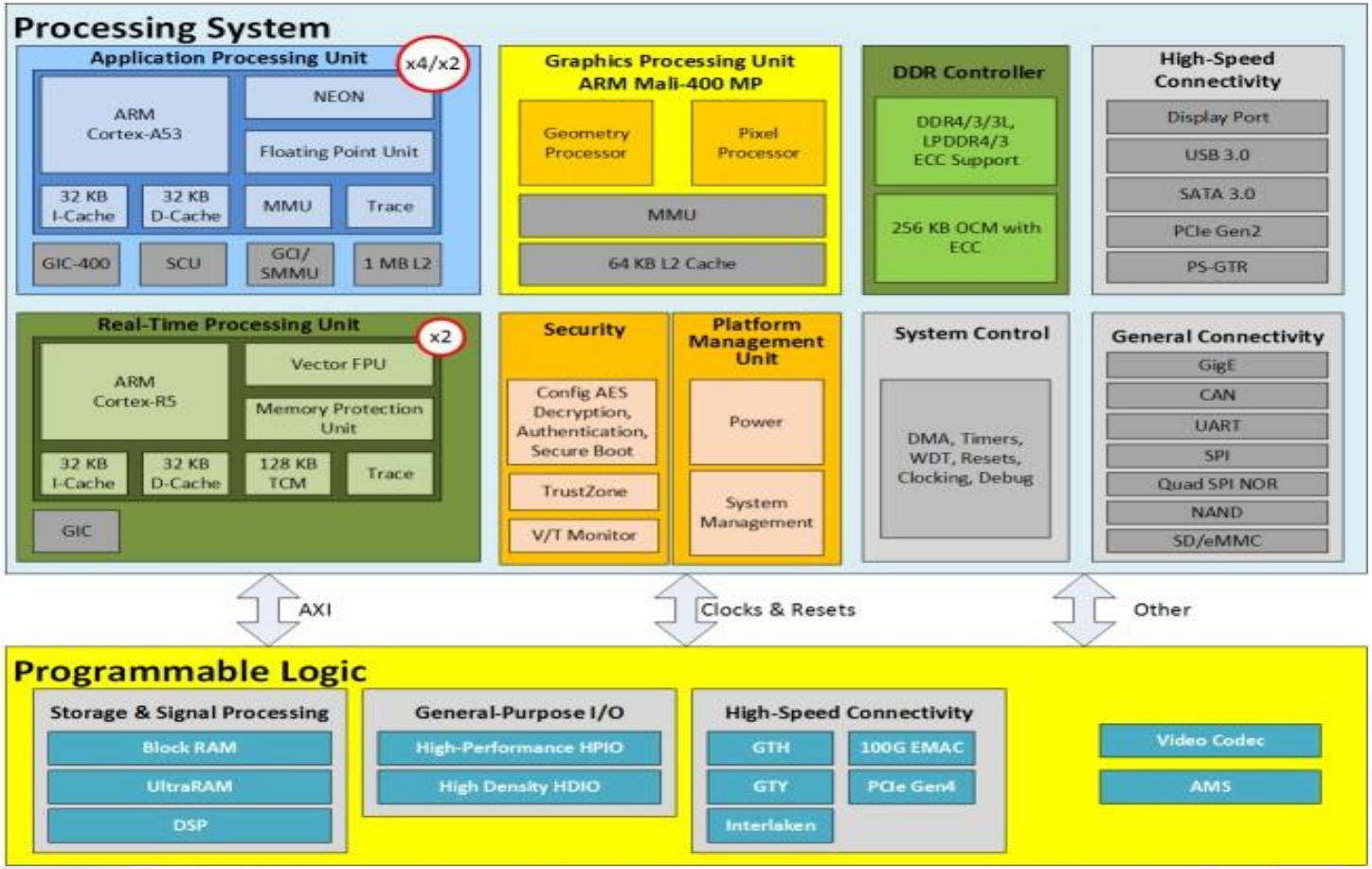


Faster Technology
Operating Beyond the Speed of Technology

Zynq UltraScale+ Parts

- 2 or 4 ARM Cortex A53 64 bit processors
- 2 ARM Cortex R5 32 bit real time processors
- 2 TMR Microblaze support processors
- GPU
- Virtualization and system cache coherency
- DDR3/4 memory controller
- More AXI interconnects with FPGA fabric
- 4 PS MGTs
- More IO peripherals

Zynq MPSoC Picture



Faster Technology
Operating Beyond the Speed of Technology

Zynq PS to PL interconnections

- Processors system and programmable logic connected with AXI
  - Arm eXtensible Interconnect
  - Both master and slave ports to allow either PS or PL to initiate transactions
  - Some ports have special features
- Xilinx Vivado software comes with a catalog of IP building blocks that use AXI for interconnect

The software

- Vivado is top level FPGA design entry tool
  - IP Integrator
- SDK or PetaLinux are embedded software tools
- SDSoC is newer tool that can do it all for using FPGA fabric to accelerate software functions
- FPGA design entry is in a state of transition

How it was/is

- "Programming" FPGAs based on Hardware Description Languages (HDLs)
  - Verilog / System Verilog
  - VHDL
- The FPGA version of assembly language
- Timed languages
  - You have to schedule the operations
  - You deal with binding operations to basic elements in the FPGA
- HDLs are inherently concurrent by default
  - Must add syntax to create sequential blocks of code

HDL Characteristics

- HDL code can be very low level
  - Simple Boolean equations
  - Directly instantiate basic building blocks
- HDL code can be a bit more abstract
  - Allow the synthesis tools to infer what you want from the code
- HDL code can be structural
  - Looks like a hierarchy of function calls

How it is/will be

- The FPGA equivalent to programming in C is: Programming in C!
  - C is an untimed language, the tools deal with scheduling and binding of operations.
  - Vivado HLS is the C design entry tool.
    - A point tool that will create a module to use in a higher level Vivado HDL / IPI based design
    - Will compile C, C++, SystemC, OpenCL kernel code into Verilog or VHDL.
    - Can package it as IEEE IP-XACT with AXI interfaces for the arguments

HLS

- One HLS project = one IP core
- User specifies algorithm in C, HLS performs scheduling and binding
- User controls implementation results using directives
  - Default implementation is as simple as possible
  - Unroll loops
  - Pipeline functions
  - Implement arrays / memories differently
  - Use directives remove bottlenecks in the critical path to achieve desired performance
- Package results as IP-XACT

SDSoC and SDAccel

- Alternative top level tools
  - Sit on top of Vivado, SDK, PetaLinux
  - All design entry is in C,C++, SystemC, OpenCL
- SDSoC for Zynq parts
- SDAccel for PCIe FPGA cards
- Board vendor supplies a board support package
  - Reference design with infrastructure
  - User design merged into this

SDSoC and SDAccel

- User marks functions for acceleration
  - Each function turned into an HLS project
  - User creates HLS directives
- Design refactored with accelerated functions moved to FPGA fabric
  - Data flow analysis performed on accelerated functions
  - Appropriate data path and DMA engines wrapped around HLS generated IP cores
  - FPGA configuration files produced by SDx using Vivado, SDK, etc

SD software environments

- reVISION
  - Machine vision stack
- BNN
  - Binary Neural Networks
- DSP
  - FFTs, FIRs, Wireless algorithms
- Networking

Faster Technology
Operating Beyond the Speed of Technology