NetFPGA Summer Course

Presented by:

Noa Zilberman

Yury Audzevich

Technion

August 2 – August 6, 2015
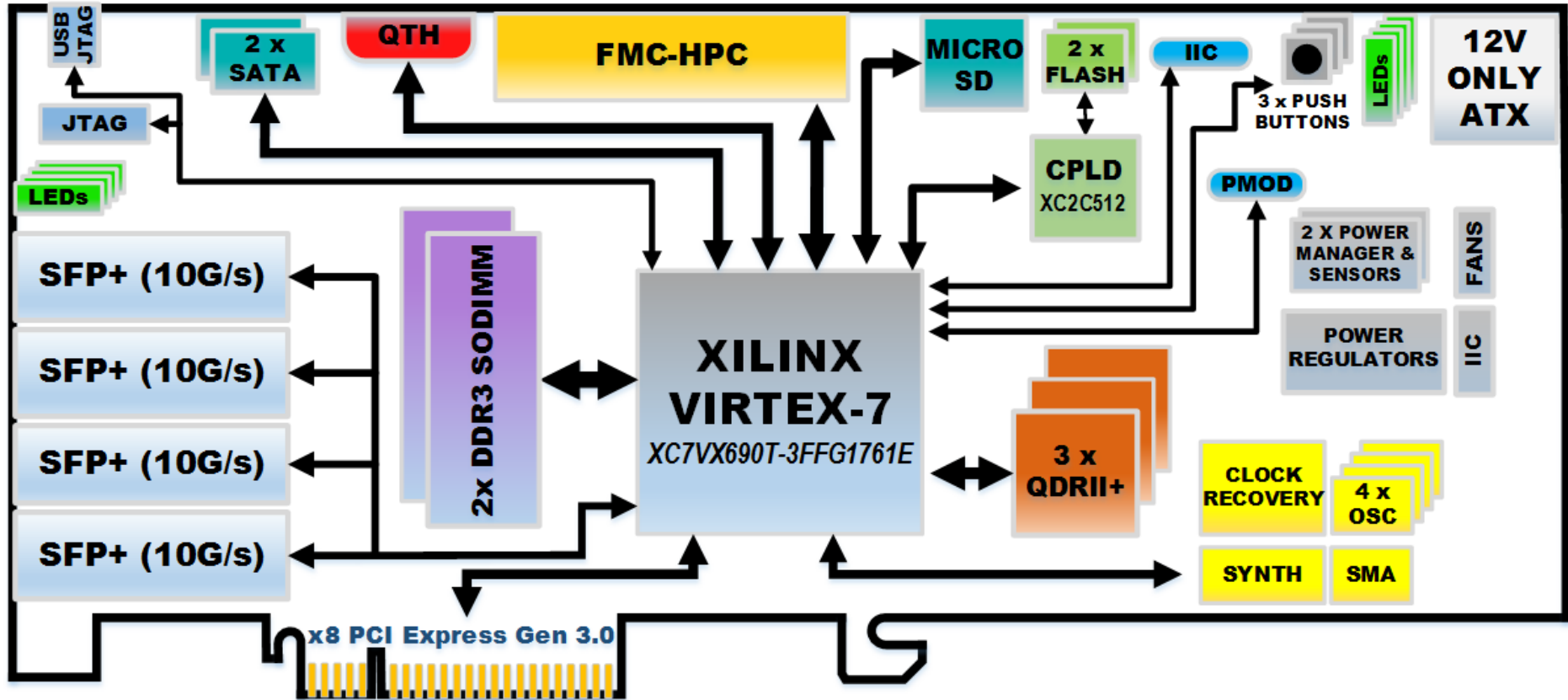
http://NetFPGA.org

NetFPGA SUME HARDWARE

Outline

- **High Level Block Diagram**
- **FPGA**
- **Memory Subsystem**
- **Serial Interfaces**
- **Storage**
- **Configuration**
- **Clocks**
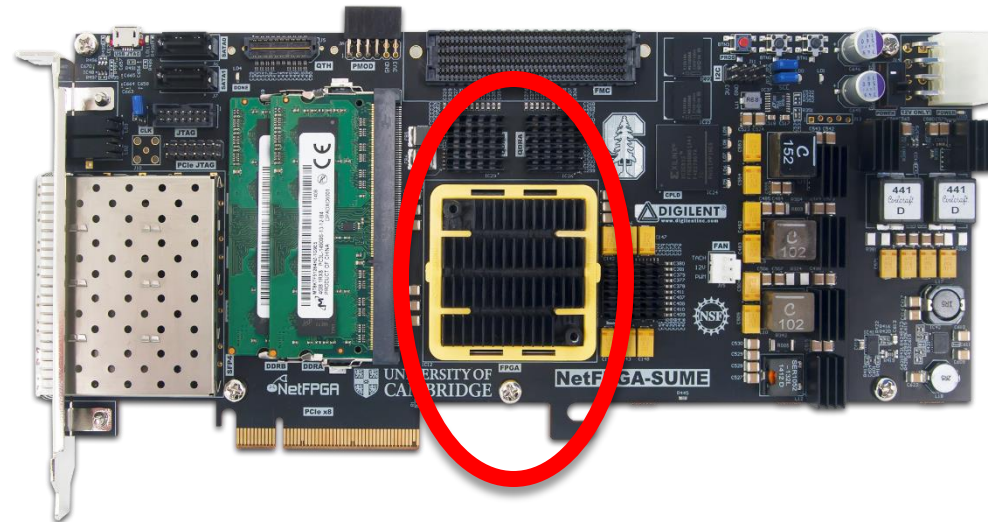- **Status Indications**
- **Power**
- **Misc**

Block Diagram

FPGA

FPGA- Virtex-7 690T

- **Virtex-7 FPGA introduced in 2012**
- **28nm process**
- **690K Logic cells**
- **866K CLB FF**
- **52Mb RAM**
- **3600 DSP slices**
- **3x PCIe Gen 3 Hard cores**
- **850 I/O**
- **36 GTH transceivers**

Virtex 7 CLB

- **CLB – Configurable Logic Block**
- **The main logic resource**
- **Usually assigned without user intervention**
- **Each CLB contains:**
  - 2 slices
  - 8 LUTs (6 inputs)
  - 16 Flip Flops
  - 2 Arithmetic and carry chains
  - 256b distributed RAM
  - 128b shift registers
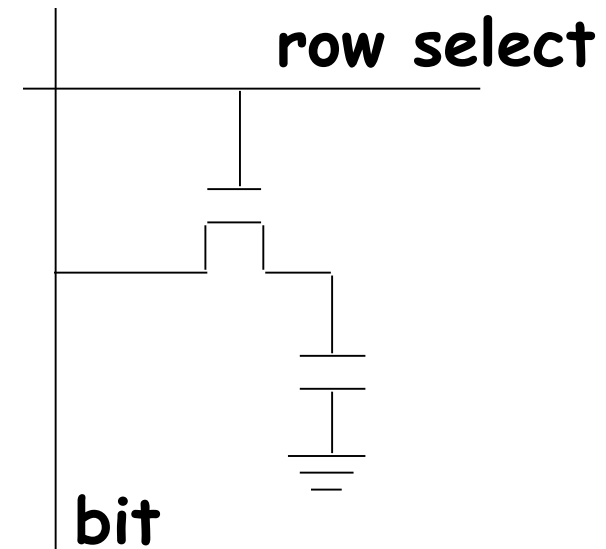- **Refer to Xilinx's UG474**

Memory Subsystem

Memory Interfaces

- **DRAM:
2 x DDR3 SoDIMM
1866MT/s, 4GB
(supports up to
32GB)**
- **SRAM:
3 x 9MB QDRII+,
500MHz**
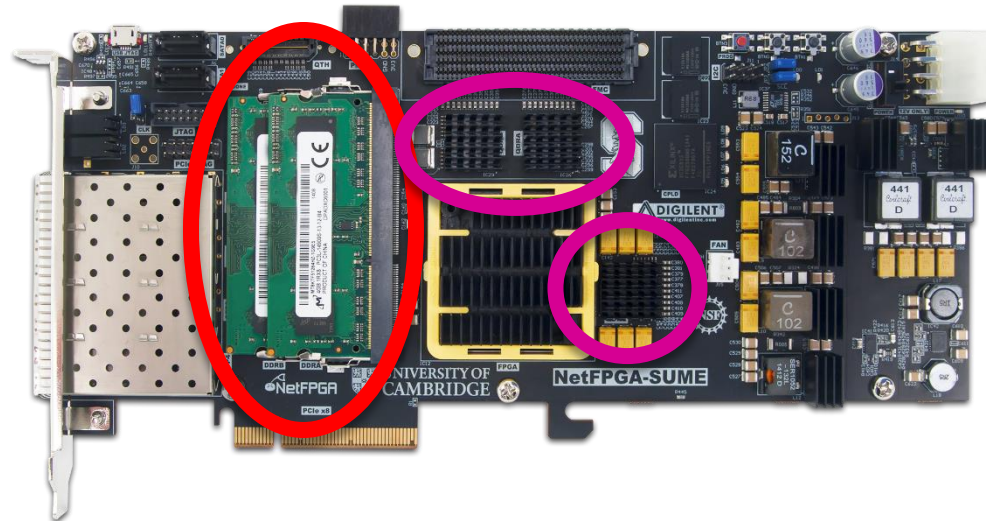
DRAM

- **Dynamic RAM**

- **Based on capacitors, holding charge**

- **SDRAM – Synchronous DRAM**

- **DDR – Double Data Rate**

    – Two data transactions is every clock cycle

        - Rising edge & falling Edge

row select

bit

DDR SDRAM – Prefetch Buffer

- **Fetching a single data word takes time…**
  - Any additional data word on the same row comes with minimal "cost"
- **Idea: with every access, read several adjacent data words**
  - Without individual column request
- **A prefetch buffer holds the fetched words until they are transmitted**
- **Prefetch buffer depth is typically the ratio between core memory frequency and I/O frequency**

11

DDR SDRAM

- ## DDR3 SDRAM - Prefer buffer size is 8n
  - Example:
    - Clock rate is 800MHz
    - Data rate is 1600Mbps x bus width
    - Core rate is 200MHz

- ## DIMM – Dual In-line Memory Module
  - Replaced SIMM – Single In-line Memory module
  - DIMM has separate electrical contacts on each side of the module.

- ## SO-DIMM
  - Small Outline DIMM
  - Usually used in mobile computers

DRAM Modules

- **Consumer and networking applications typically use DRAM devices (components).**

- **Computing applications typically use DRAM modules.**

- **DIMM – Dual In-line Memory Module**

  – Replaced SIMM – Single In-line Memory module

  – DIMM has separate electrical contacts on each side of the module.

- **SO-DIMM**

  – Small Outline DIMM

  – Usually used in mobile computers

DRAM Frequency Errata

- **Xilnx currently has an errata for MIG 7 series DDR3**
- **AR#59167**
- **Triggered by aggressive data patterns (PRBS23)**
- **Caused by loss on the channel and skew with the FPGA**
- **Workaround: max data rate is 1700MT/s**

SRAM

- **Static RAM**

- **Based on transistors (Flip-Flops)**

- **Saving state**

- **Less dense and more expensive than DRAM**

15

QDR SRAM

- **QDR – Quad Data Rate**
  - Synchronous
  - Separate busses for Write and Read
  - Each bus – Double data rate
  - Total of 4 transactions per clock
  - 500MHz $\Rightarrow$ 2000MT/s
  - Constant latency
- **QDR II+**
  - Uses QVLD signal for sampling
    - Rather than free running clock

QDR SRAM – Burst Length

- **Similar concept as DRAM burst length**
- **Valid options: 2 or 4**
  - Part number specific
- **BL=2**
  - Can access a different address every clock
  - Ideal for short queries (e.g. lookups)
- **BL=4**
  - Can change address every 2 clocks
  - Achieves higher frequency
  - Supports half the number of entries of BL=2
    - For the same SRAM density
- **This is a design trade-off**

QDR's Bank Sharing

- **QDR A and QDR B Share Bank 17**
  - For controls
- **Xilinx MIG currently does not support bank sharing**
- **Manual manipulation of the PHY is required in order to use**
  - Calling for a contributed project

DRAM vs. SRAM

	DRAM	SRAM
Density	High	Low
Latency	Variable	Constant
	High	Low
Bandwidth	High	High
Effective bandwidth	Varies, <100%	100%

- **Usage examples:**
  - **Output queues**
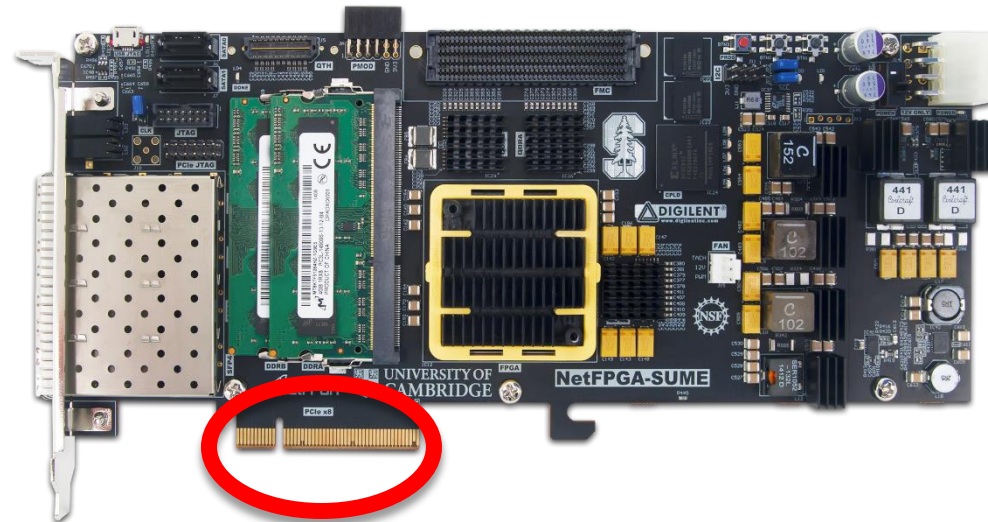 - **Lookup tables**
  - **Storing buffer descriptors**

Serial Interfaces

Serial Interfaces

- **Used for data transfer at high rates**
- **GTH Transceiver (Transmitter/Receiver)**
- **13.1Gb/s**
  - Speed grade: -3
- **FPGA selection**
  - GTH vs. GTZ
    - 13.1Gb/s vs. 28.05Gb/s
  - I/O vs. Serial I/F Rate
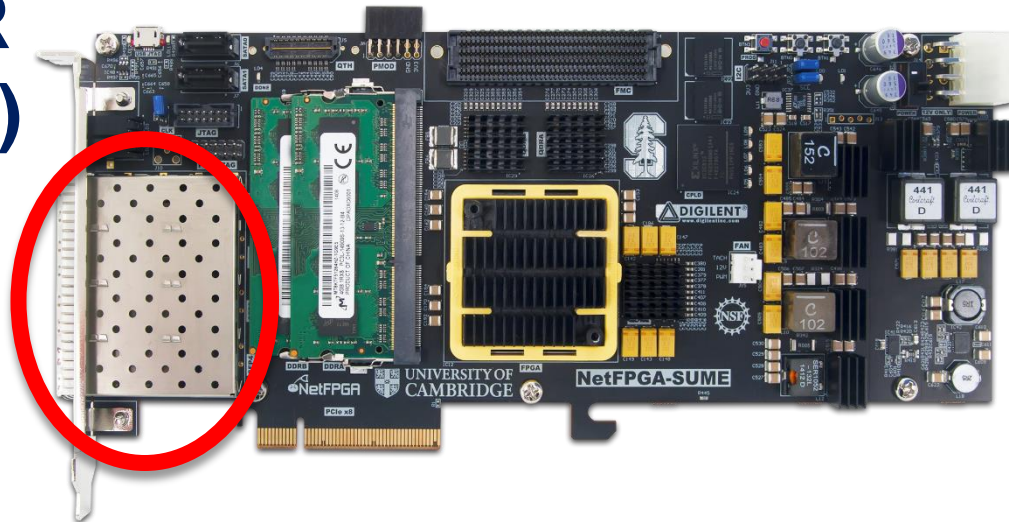   - I/O equals RAM
  - RAM won

Host Interface

- **PCIe Gen. 3**
- **x8 (only)**
  - x4 requires changes to the clock circuitry
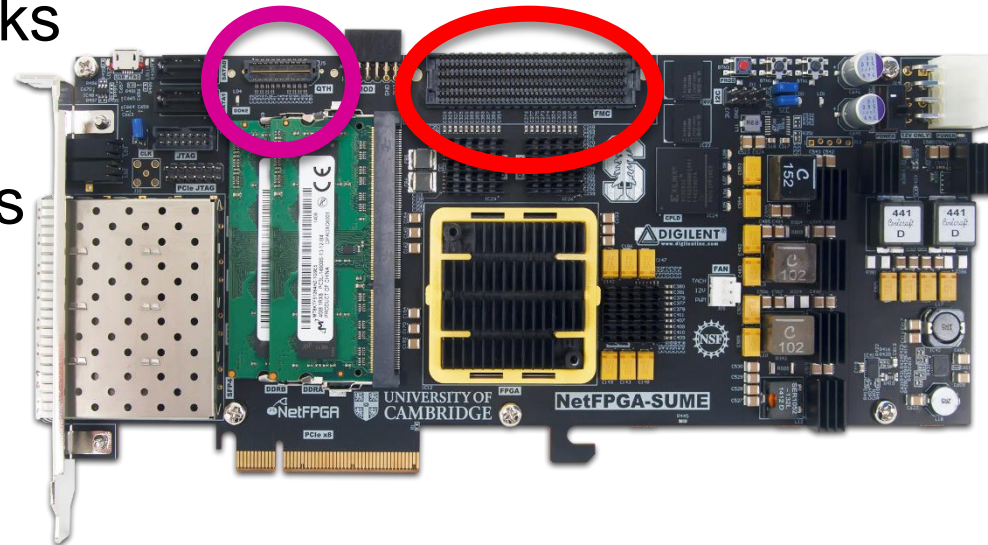- **Hardcore IP**

Front Panel Ports

- **4 SFP+ Cages**
- **Directly connected to the FPGA**
- **Supports 10GBase-R transceivers (default)**
- **Also Supports 1000Base-X transceivers and direct attach cables**

Expansion Interfaces

- **FMC HPC connector**
  - VITA-57 Standard
  - Supports Fabric Mezzanine Cards (FMC)
  - 10 x 12.5Gbps serial links
- **QTH-DP**
  - 8 x 12.5Gbps serial links

- **12.5Gb/s is the validation rate**
- **Actual performance depends on the full channel**
 - Insertion loss, return loss, cross talk,….

Serial Interfaces

- ## Summary:
  - 4 transceivers connect to SFP+
  - 8 transceivers connect to PCIe
  - 10 transceivers connect to FMC
  - 8 transceivers connect to QTH
  - 2 transceivers connect to SATA (see later)
- ## Total: 32
  - 4 transceivers are unused
    - Transceivers are grouped in quads, with shared clocking
    - 2 unused transceivers on SATA quad and 2 on FMC last quad

STORAGE

Storage

- **128MB FLASH**

- **2 x SATA connectors**

- **Micro-SD slot**

- **Enable standalone operation**
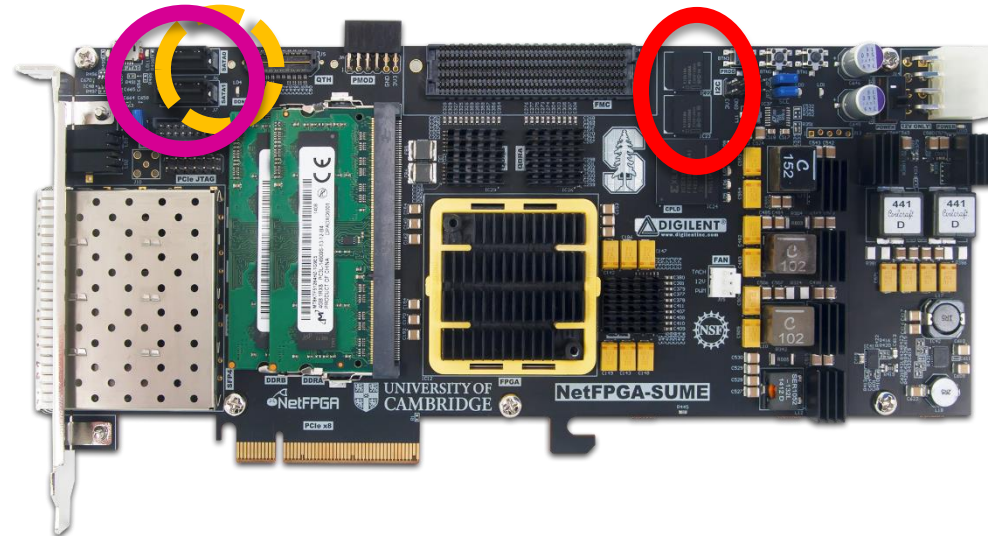
FLASH

- **Non Volatile RAM (NVRAM)**
- **NOR based**
  - High reliability (vs. NAND FLASH)
- **Can read "single" data**
- **Write:**
  - Erase blocks (write '1')
  - Write '0'

FLASH – SUME

- **Using a parallel FLASH**
  - 16 bit wide

- **Used to store the FPGA's image**
  - Loaded upon power up

- **Using 2 FLASH devices in parallel**
  - To achieve PCIe required configuration time

- **Additional storage space available for more bitstream files and user defined purposes**

SATA

- **2 on board SATA connectors**
- **SATA-III compatible (6Gb/s)**
- **Connects to standard HDD/SDD**
  - Use standard SATA cables
- **Uses 2 transceivers**
  - One per connector
- **Enables the stand-alone computing unit operation**

Micro-SD

- **SD – "Secure Digital"**
- **Non volatile memory device**
- **Uses a parallel interface:**
  - 4 bit data
  - 1 bit command
  - …and a protocol
- **Supports UHS-I**
  - But not UHS-II
- **Supports SC, HC and XC class cards**
- **Located at the reverse side (print side) of the board**

NetFPGA

CONFIGURATION

FPGA Configuration

- FPGA configuration data is stored in files called bitstreams
  - Have the .bit file extension.
- Stored in dedicated CMOS Configuration Latches (CCL)
- Defines the FPGA's logic functions and circuit connections
- Remains valid until:
  - Erased
  - Power down

Reset does not affect the FPGA configuration!

NetFPGA

FPGA Configuration

- **Multiple ways to configure the FPGA:**
1. **Through the JTAG chain, using USB-JTAG**
   - J16, labelled PROG
2. **Through the JTAG chain, using 14-pin JTAG header**
   - J9
3. **From the FLASH**
   - Loading one of four possible bitstream files
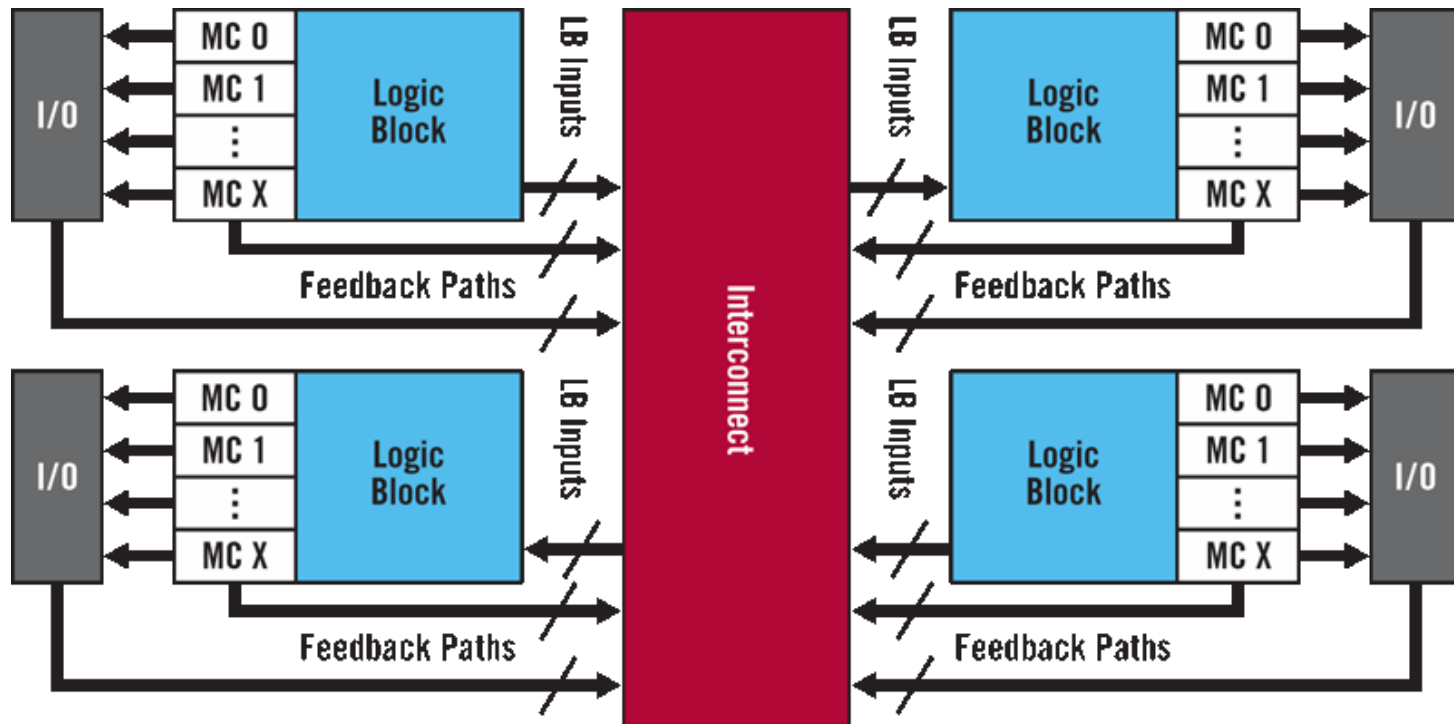
FPGA Configuration
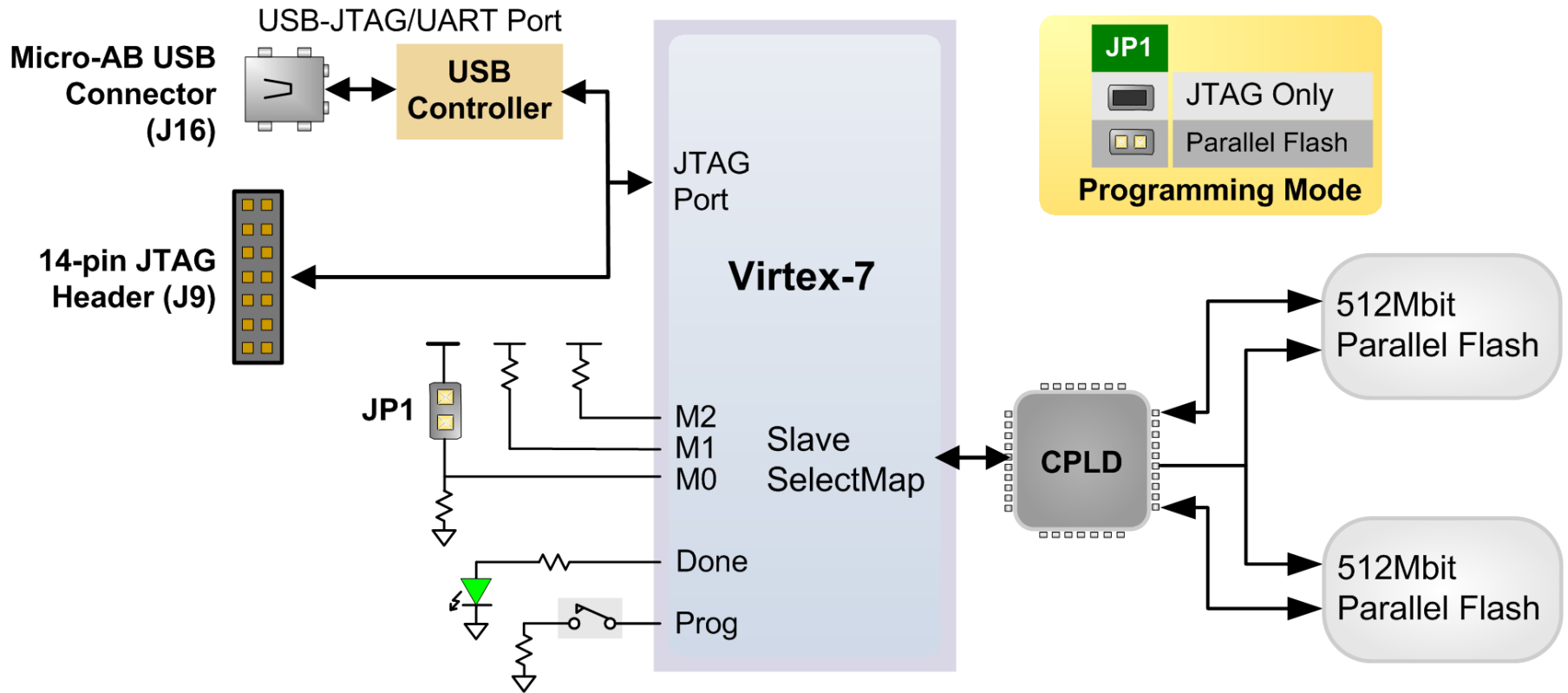
CPLD

- CPLD – Complex Programmable Logic Device
- Non-volotile
- A combination of
  - Programmable AND/OR array
  - Macrocells
- Macrocells
  - Functional blocks
  - Perform combinatorial or sequential logic
  - True or complement
  - Varied feedback path

CPLD

- **An example of Xilinx CPLD (block diagram):**

Source: http://www.xilinx.com/cpld/

CPLD

- CoolRunner II XC2C512
  - 512 macro cells
- Same JTAG chain as the FPGA
- Used as an interface converter between FLASH and FPGA
  - CPLD $\Leftrightarrow$ 2xFLASH: Master BPI (16 bit)
  - CPLD $\Leftrightarrow$ FPGA: 32bit SelectMap
- See UG470
- Goal: Respond to PCI enumeration commands within 200 milliseconds of power up

Clocks

Clocks

Clock Name	Frequency	Common Usage
FPGA_SYSCLK	200MHz	General Purpose, Shared with source with QDR devices
QDRII_SYSCLK	200MHz	Used by MIG for QDRIIA, QDRIIB
QDRIIC_SYSCLK	200MHz	Used by MIG for QDRIIC
SATA_SYSCLK	150MHz	Used by SATA transceivers
SFP_CLK	156.25MHz (Configurable)	Shared by transceivers of SFP+ and QTH
DDR3_SYSCLK	233.33MHz	Shared by MIG for DDR3A, DDR3B
PCIE-CLK	100MHz	Used by PCI-Express Core
USER_CLK	156.25MHz (Configurable)	User defined purposes
FMC_GBT_CLK0	-	Used by FMC transceivers (From FMC)
FMC_GBT_CLK1	-	Used by FMC transceivers (From FMC)
FMC_CLK0	-	Used by FMC card (From FMC)
FMC_CLK1	-	Used by FMC card (From FMC)

Clocks

- **Special clocks:**
- **Configurable – through I2C:**
  – USER clock
  – SFP (& QTH) clock
- **External Clock**
  – PCIe clock
  – FMC clocks
- **Recovered Clock**
 – SFP (& QTH) clock

Miscellaneous

Status and Indications

- 10G ports indication LEDs
  - 4x Green
  - 4x Yellow
  - Programmable
- 5 Programming status LEDs
  - Controlled by CPLD
    - LD5 – No valid bitstream in indicated section
    - LD6 – Fallback LED (load bitstream from default boot section)
    - LD7-LD9 – Indicate current boot section
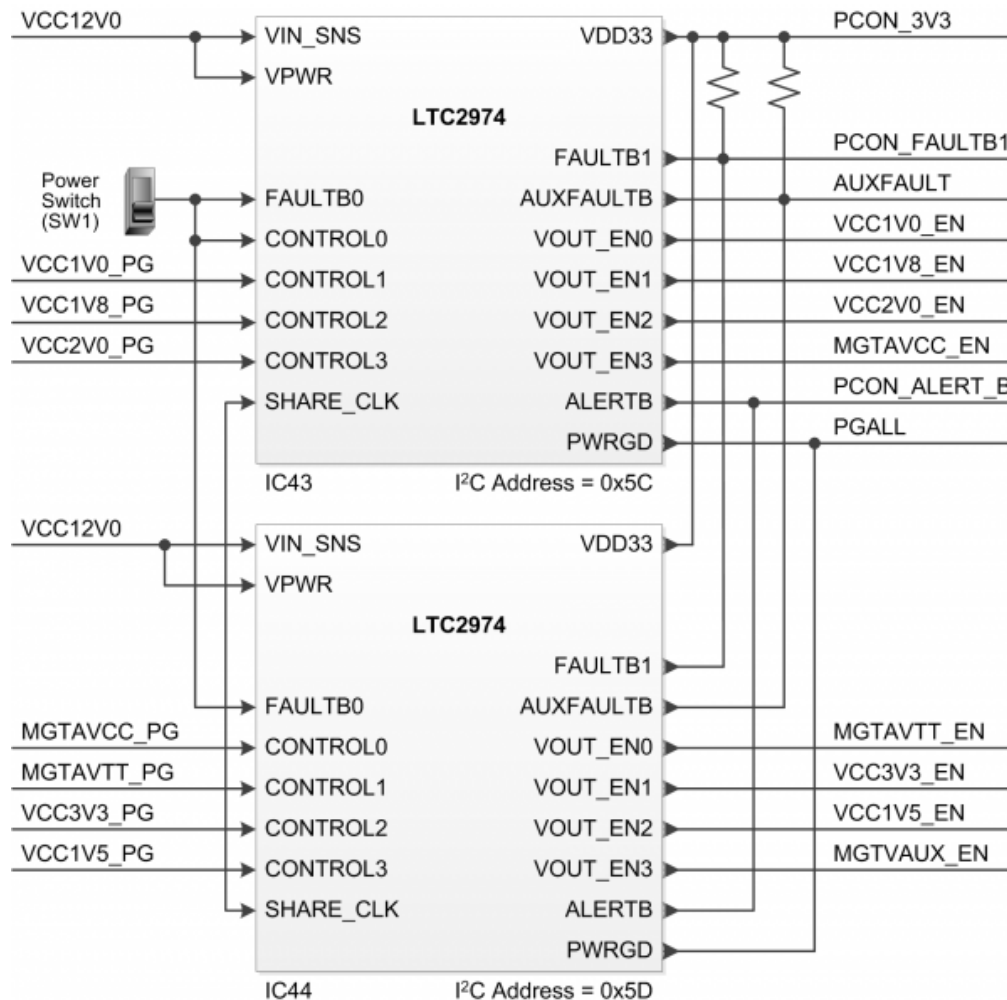- 2 General purpose LEDs

Power

- **Power Supply**
  - 12V
  - From a dedicated power connector
    - Not through PCIe connector
- **On-Off connector**
- **6 major power rails**
  - Different voltage: 1.0, 1.2V, 1.5V, 1.8V, 3.3V
  - Several dedicated, isolated, power rails
- **Maximum power consumption ~150W**
  - Typical power consumption is much lower

Power
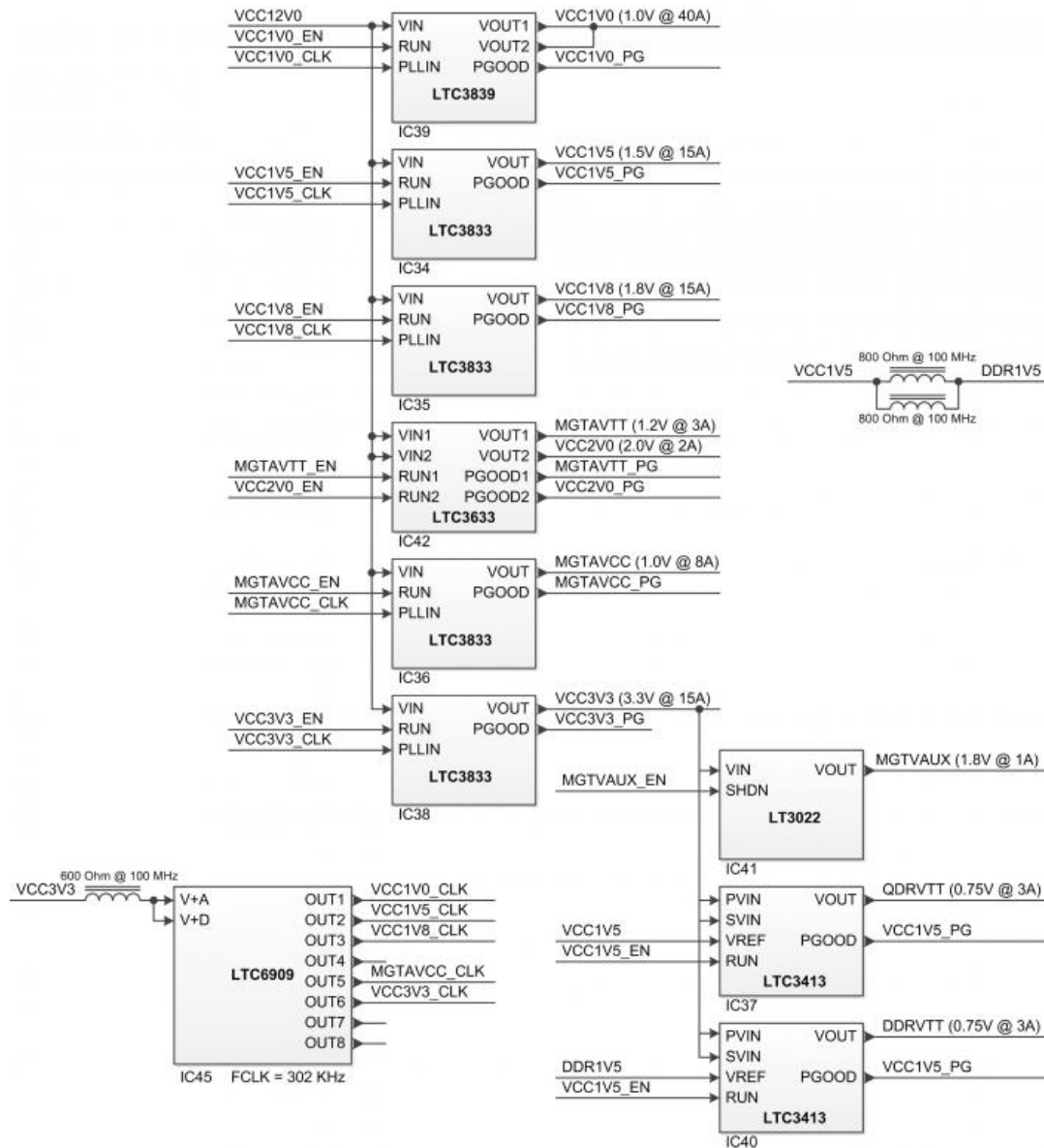
Power Control

- **On board power sequencing and supervision**

Sensors

- Per power rail:
  - Output voltage
  - Current consumption
  - Temperature
- Can set alerts (fault or warning):
  - Over/under voltage
  - Over/under current
  - Over temperature
  - Timeout reaching or exceeding a certain voltage
- Using I2C (PMBus)

Conclusion

NetFPGA

Acknowledgments (I)

NetFPGA Team at University of Cambridge (Past and Present):

Andrew Moore, David Miller, Muhammad Shahbaz, Martin Zadnik
Matthew Grosvenor, Yury Audzevich, Neelakandan Manihatty-Bojan,
Georgina Kalogeridou, Jong Hun Han, Noa Zilberman, Gianni Antichi,
Charalampos Rotsos, Marco Forconesi, Jinyun Zhang, Bjoern Zeeb

NetFPGA Team at Stanford University (Past and Present):

Nick McKeown, Glen Gibb, Jad Naous, David Erickson,
G. Adam Covington, John W. Lockwood, Jianying Luo, Brandon Heller, Paul
Hartke, Neda Beheshti, Sara Bolouki, James Zeng,
Jonathan Ellithorpe, Sachidanandan Sambandan, Eric Lo

All Community members (including but not limited to):

Paul Rodman, Kumar Sanghvi, Wojciech A. Koszek,
Yahsar Ganjali, Martin Labrecque, Jeff Shafer, Eric Keller ,
Tatsuya Yabe, Bilal Anwer, Yashar Ganjali, Martin Labrecque,
Lisa Donatini, Sergio Lopez-Buedo

Kees Vissers, Michaela Blott, Shep Siegel, Cathal McCabe

Acknowledgements (II)