

FPGA

Bojan Kuljić

Jan. 2021

(Ver. 1.0)

Contents

- 1. Introduction
- 2. CPU and FPGA synergy
- 3. FPGA in mobile phones
- 4. FPGA in high-performance computing and datacenters
- 5. Conclusions
- 6. References

1. Introduction

Definition of FPGA

FPGA stands for **field programmable gate array**. It is an integrated circuit that can be configured by a designer even after manufacturing – hence the term **field programmable**.

There are two major manufacturers:

- Xilinx – **founded in 1984**
- Altera – **founded in 1983** – acquired by **Intel in 2015**

Other manufacturers include:

- Microchip
 - Microsemi (previously **Actel**) - **acquired by Microchip in 2018**
 - Atmel - **acquired by Microchip in 2016**
- Lattice Semiconductor
 - SiliconBlue Technologies - **acquired by Lattice in 2011**
- QuickLogic
- Achronix

* On **October 27 2020**, AMD announced it would **acquire Xilinx**.

1. Introduction (2)

Era of FPGA started in 1985 with Xilinx XC2064 [1]



- 64 flip flops
- 128 3-LUTs
- 58 IO pins
- 18MHz
- 2um technology

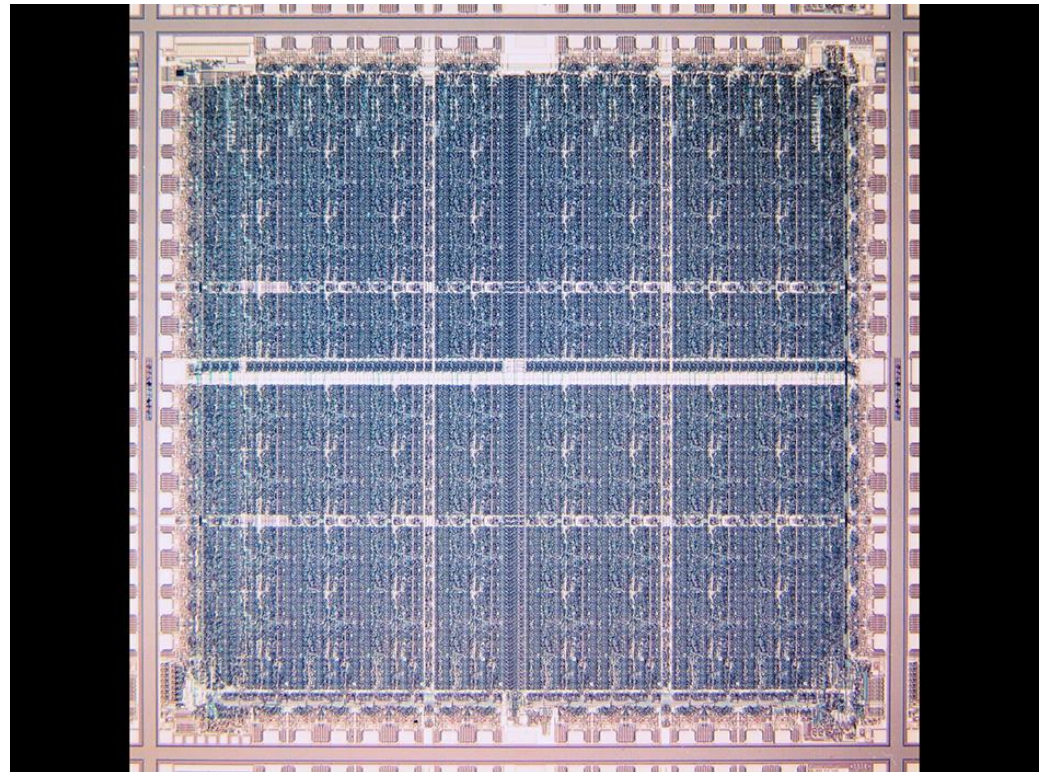


Figure: casing and silicon die of Xilinx first FPGA [1]

1. Introduction (3)

Market size over the years

In 2019 market size reached almost **10 billion dollars!!!**

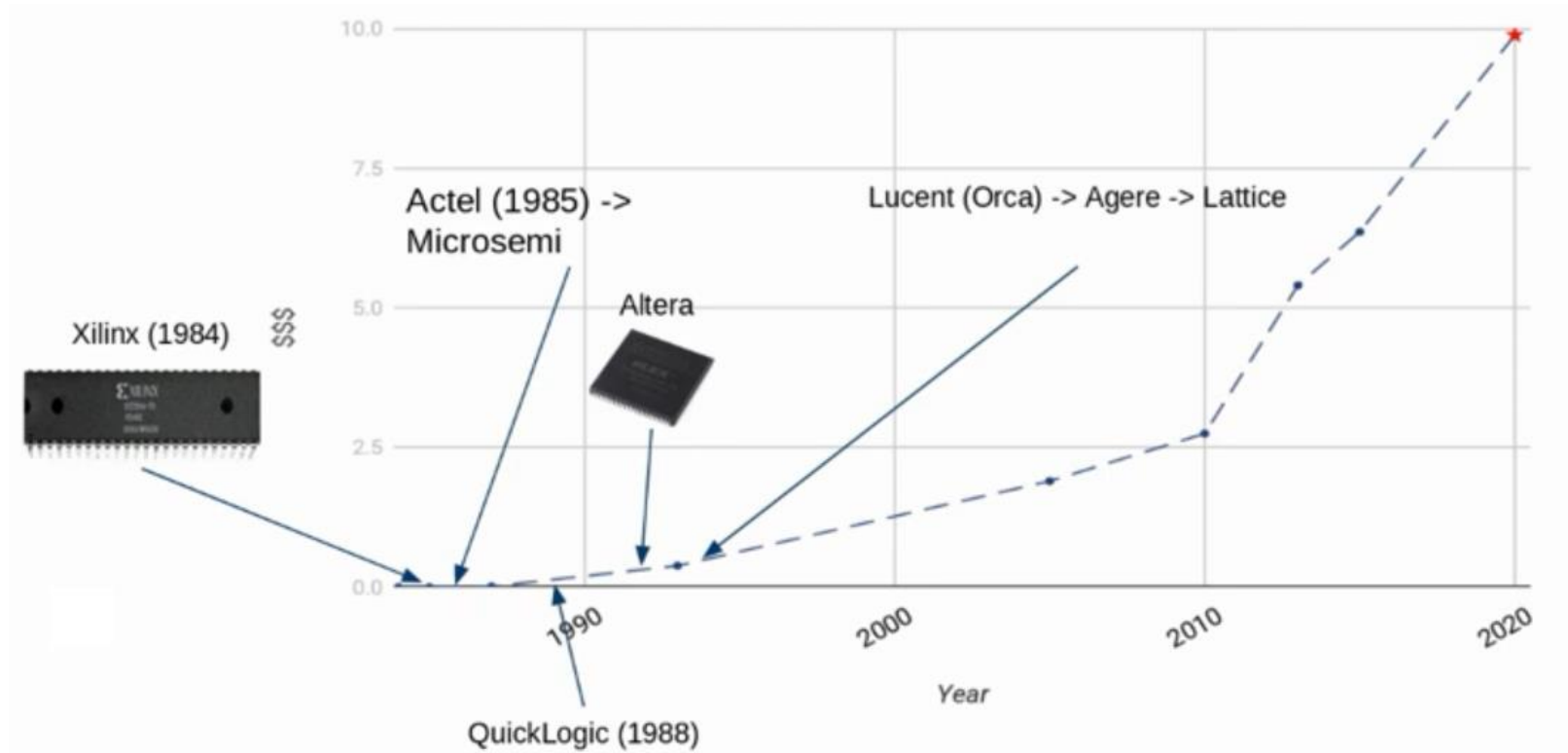


Figure: timeline of FPGA market size

1. Introduction (4)

Future predictions are very favorable for FPGA technology

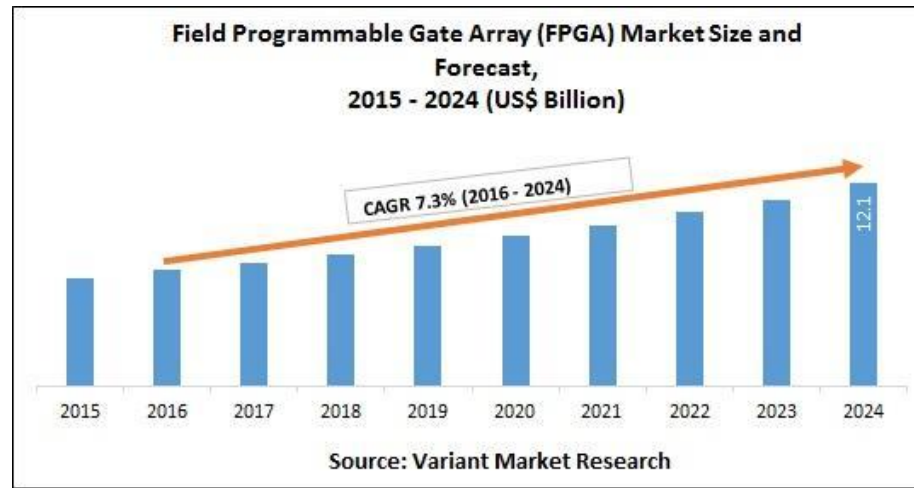


Figure: future FPGA market size prediction [2]

Since the beginning to present day, Xilinx was leader on the market.

Top 5 Company's Sales of FPGA/PLD, in All Applications, Worldwide, (Millions of U.S. Dollars)

Rank 2017	Rank 2018	Company Name	2017	2018	2018 Change	Share of Market
1	1	Xilinx	2,476	2,904	17.3%	51.1%
2	2	Intel	1,858	2,033	9.4%	35.8%
3	3	Microchip (formerly Microsemi)	321	376	17.1%	6.6%
4	4	Lattice Semiconductor	261	285	9.2%	5.0%
		Others	71	85	19.7%	1.5%
Total Market			4,987	5,683	14.0%	100.0%

Source: Gartner (April 2019)

Figure: major companies share of FPGA market size [3]

What happened around 2010 that made FPGA soar on the market?

In the past decade there was a **rise of fields** that required **reconfigurable hardware**:

- Machine learning
- Artificial intelligence research
- Complex simulations
- Biomedical devices
- Networking applications like **routers, servers** etc.

Why didn't FPGA risen on the market earlier?

There are **four reasons** that stand out [4]:

- Most obvious – **price**
- Semiconductor scaling
- Crossover **point between FPGA and ASIC**
- Enhanced software tools

1. Introduction (6)

Semiconductor scaling

In the past 30+ years capacity of the FPGA increased by more than a factor of 10000, speed increased by a factor of 100, while cost and energy consumption per unit function decreased by more than a factor of 1000.

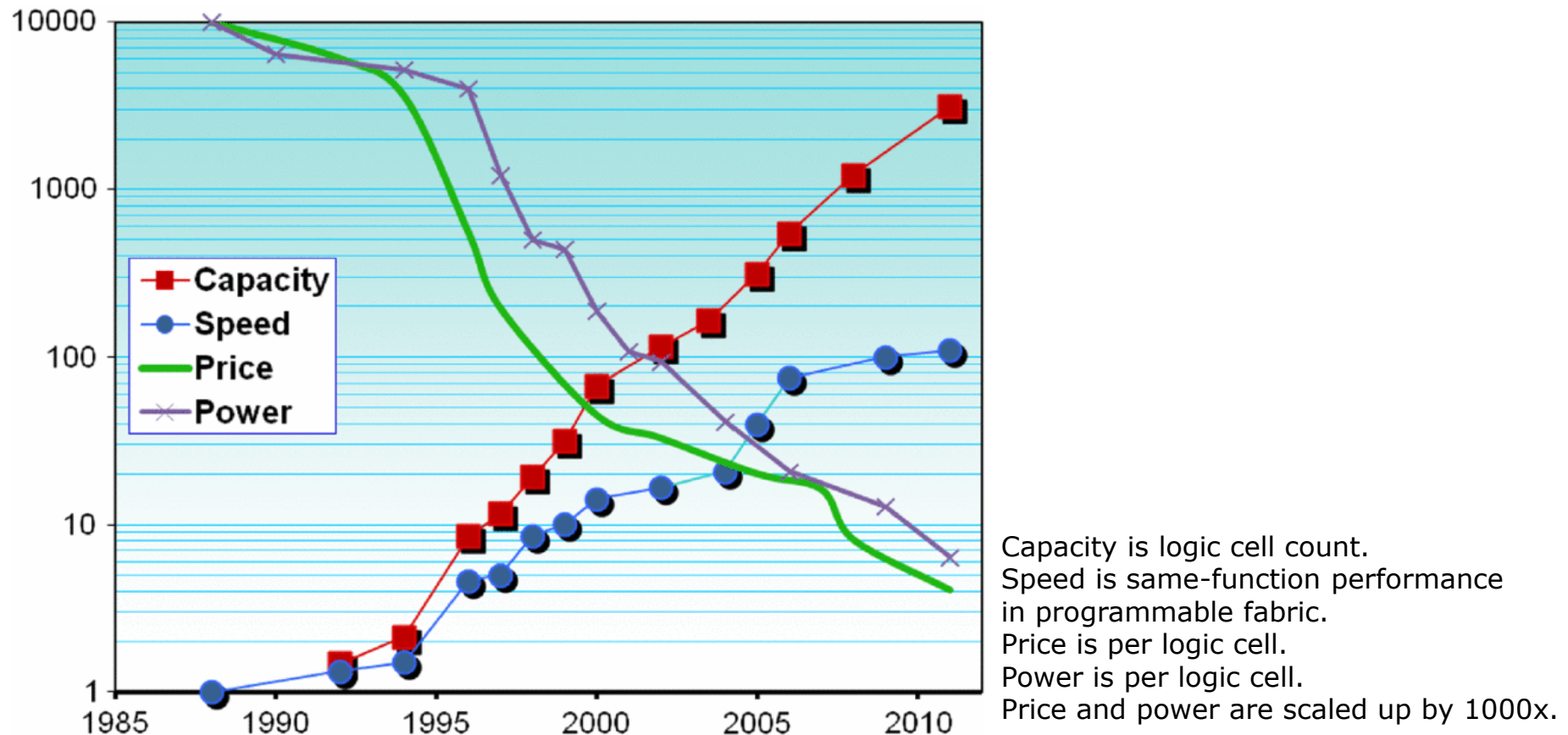


Figure: parameters of FPGA technology through years [5]

Crossover point between FPGA and ASIC

ASIC was always more profitable with large scale production. Up to certain number of units **FPGA it is more profitable**.

With every new generation of FPGAs **crossover point is pushed** more to the right.

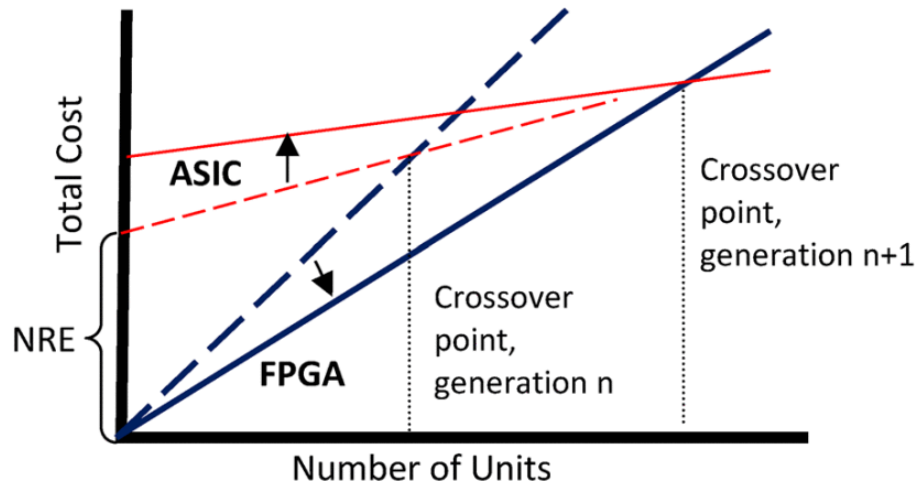


Figure: development cost distribution FPGA vs ASIC [5]

NRE - Non-recurring engineering cost refers to the one-time cost to research, design, develop and test a new product or product enhancement. Even though a company will pay for NRE on a project only once, NRE costs can be prohibitively high and the product will need to sell well enough to produce a return on the initial investment.

1. Introduction (8)

Traditionally many companies used FPGAs during the early design phase and preproduction phases and then switch later to ASIC for volume production. This design practice was for applications whose future commercial success was unknown so the FPGA route offered lower risk.

Over the years advantages shifted towards the FPGA greatly but ASIC still as advantages in some areas.

CHARACTERISTIC	FPGA	ASIC
Time-to-market	Short	Long
High volume unit cost	High	Low
Flexibility after manufacturing	High	None
Performance	Medium	Very high
Density	Medium	Very high
Power consumption	High	Low
Minimum order quantities	None	High
Design flow complexity	Medium	Very high
Complexity of test	Low	High
Turnaround Time	Hours	Months

Table: a qualitative comparison between FPGAs and ASICs

Development of software tools

Today there are **many tools available** for simulation and implementation of different applications in FPGA.

Development of the **software tools had a lot of influence on the popularity** of the FPGA technology in the recent years.

Vendor tools are dominant (synthesis, timing, place & route, bitstreams)

Quartus (Intel), Vivado (Xilinx), Diamond (Lattice)

Vivado's RapidWright is moving toward opening flow, allowing 3rd party place & route (**Yosys** supports this experimentally)

3rd party **EDA tools exist** (synthesis and timing)

Cadence, Mentor, Synopsys

Open Source tools are in **their infancy**

Yosys/nextpnr, Icarus

Kernel interfaces is **developed by Altera and mainly used by Intel**

DFL - Device Feature List

OPAE - Open Programmable Acceleration Engine

Per vendors interfaces – about 20 vendors supported

Example of niche application having a huge influence on FPGA market

Around 2007 in oil and gas implementations a new approach was taken.

The time it took classical computers to simulate the drilling of holes in the earth to find oil was longer than the actual building of a drilling site and the drilling itself.

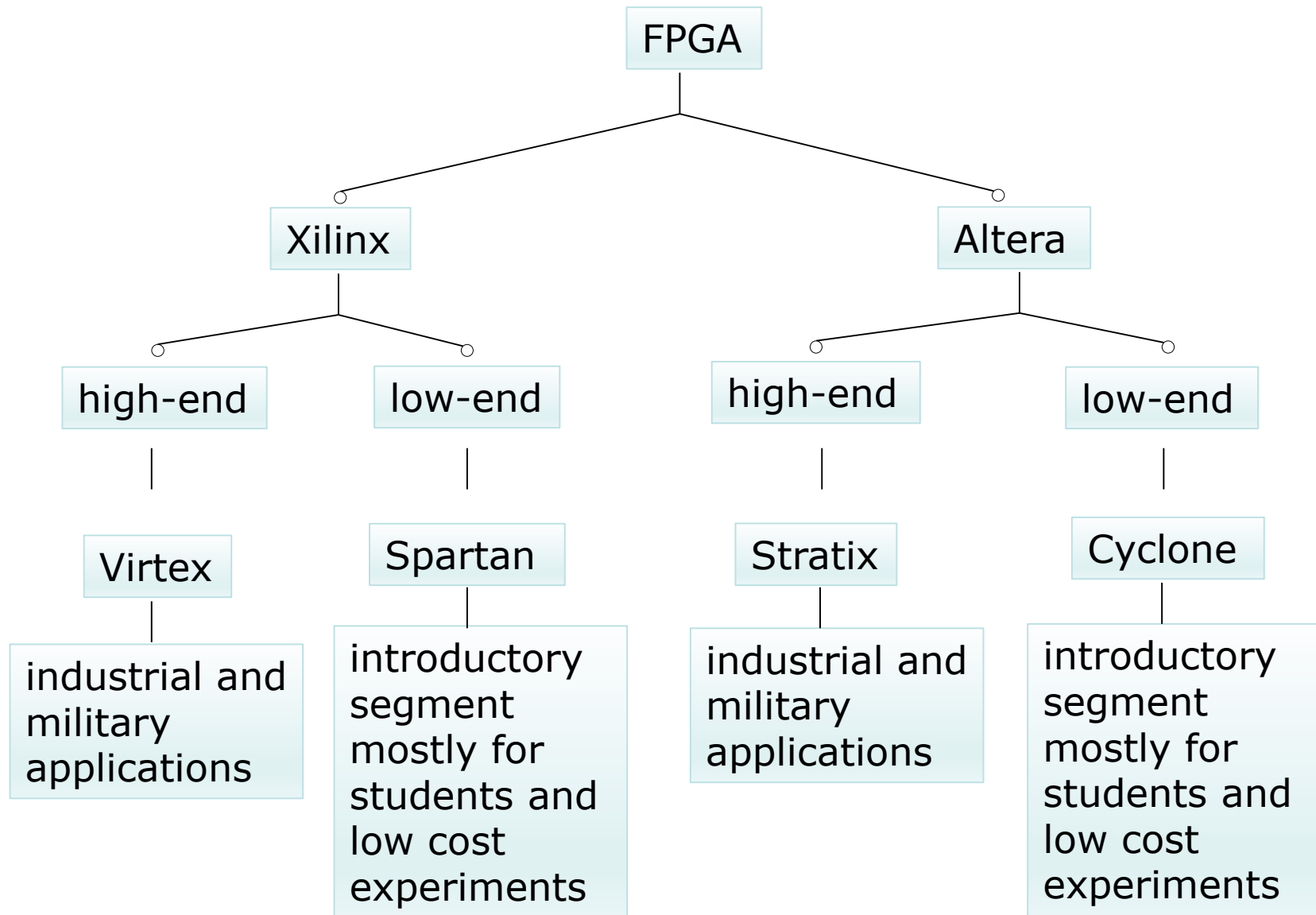
The use of FPGA accelerators dramatically changed this upside-down timing.

The first FPGAs in the datacenter of an oil company, computing seismic images, were built by Maxeler Technologies and delivered to Chevron. [6]

This event had a big influence on high-performance computing (HPC) and datacenter markets:

- In 2017, Microsoft announced its use of Altera FPGAs in the datacenters
- In 2018, Xilinx announced its "Datacenter First" strategy

FPGA families overview for two main competitors



1. Introduction (12)

Overview of Xilinx FPGA families

XC series	
Model	Released
XC2064	1985
XC3020	1988
XC4000	1991
XC3100	1992
XC3200	1992
XC5000	1994
XC8100	1995
XC6200	1995

Artix	
Model	Released
Artix 7	2010

Spartan	
Model	Released
Spartan	1998
Spartan-II	2000
Spartan-3E	2005
Spartan-3A	2007
Spartan-3AN	2008
Spartan-6	2009
Spartan-7	2017

Kintex	
Model	Released
Kintex-7	2010
Kintex UltraScale	2013
KintexUltraScale+	2015

Virtex	
Model	Released
Virtex	1998
Virtex-E	1999
Virtex-EM	2000
Virtex-II	2001
Virtex-IV	2005
Virtex-5	2006
Virtex-6	2009
Virtex-7	2010
Virtex UltraScale	2013
Virtex UltraScale+	2015

Tables: FPGA models without embedded CPU [7]

1. Introduction (13)

Overview of Xilinx FPGA families

Zynq 7000-series 28 nm fabrication process	
Model	Released
Z-7010	2011
Z-7015	2011
Z-7020	2011
Z-7030	2011
Z-7035	2012
Z-7045	2012
Z-7100	2013

Zynq UltraScale+ 16 nm fabrication process		
Model	Sub-models	Released
ZU2	CG, EG	2015
ZU3	CG, EG	2015
ZU4	CG, EG, EV	2015
ZU5	CG, EG, EV	2015
ZU6	CG, EG	2015
ZU7	CG, EG, EV	2015
ZU9	CG, EG	2015
ZU11	EG	2015
ZU15	EG	2015
ZU17	EG	2015
ZU19	EG	2015

Zynq UltraScale+ is available in up to 3 sub-models: CG, EG, and EV - differences among those are in the CPU and GPU configurations.			
	CG	EG	EV
APU	2x Arm A53	4x Arm A53	4x Arm A53
RPU	2x Arm R5	2x Arm R5	2x Arm R5
GPU	-	Arm Mali-400MP2	Arm Mali-400MP2
VCU	-	-	H.264/H.265

Tables: FPGA models with embedded CPU [7]

Overview of Altera FPGA families

Altera		
Model	Released	Description
Flex 8000	1992	Altera's first FPGA
FLEX 10K	1995	FPGA with embedded block RAM
APEX EP20K1500E	1999	More than 1.5 million gates
Excalibur	2000	FPGA with hard embedded ARM processor
Mercury	2001	180nm FPGA with embedded transceivers
Stratix	2002	FPGA with embedded DSP blocks
Cyclone	2002	Low-cost FPGAs
Stratix GX	2003	FPGA with LVDS transceivers
Stratix II	2004	FPGA with larger LUTs
Cyclone II	2005	90-nm low-cost FPGA
Stratix II GX	2006	FPGA with transceivers up to 6 Gbps
Stratix III	2006	65nm refresh of Stratix II
Cyclone III	2007	65nm refresh of Cyclone II
Arria GX	2007	Low-cost transceiver-centric FPGA
Stratix IV	2008	40-nm refresh of Stratix III

Table: Altera FPGA models [8]

Overview of Altera FPGA families (cont.)

Altera		
Model	Released	Description
Stratix IV GT	2009	FPGA with 11.3 Gbps transceiver support
Arria II GX	2009	40-nm refresh of Arria GX
Cyclone III LS	2009	Low-cost FPGA with bitstream protection
Stratix V	2010	28-nm Stratix IV refresh
Cyclone IV	2010	Incremental update of Cyclone series
Arria V	2011	28-nm
Cyclone V	2012	Low-power, low-cost 28-nm FPGA

Table: Altera FPGA models [8]

1. Introduction (16)

Overview of Intel FPGA families

Since acquisition of Altera, Intel has continued to develop all three groups of FPGAs. Current generation of FPGAs is labeled with the number 10.



Introduced in 2013
20 nm technology



Introduced in 2017
20 nm technology



Introduced in 2019
20 nm technology

1. Introduction (17)

Examples of FPGA applications

Low-end FPGAs are traditionally used for [teaching students](#).

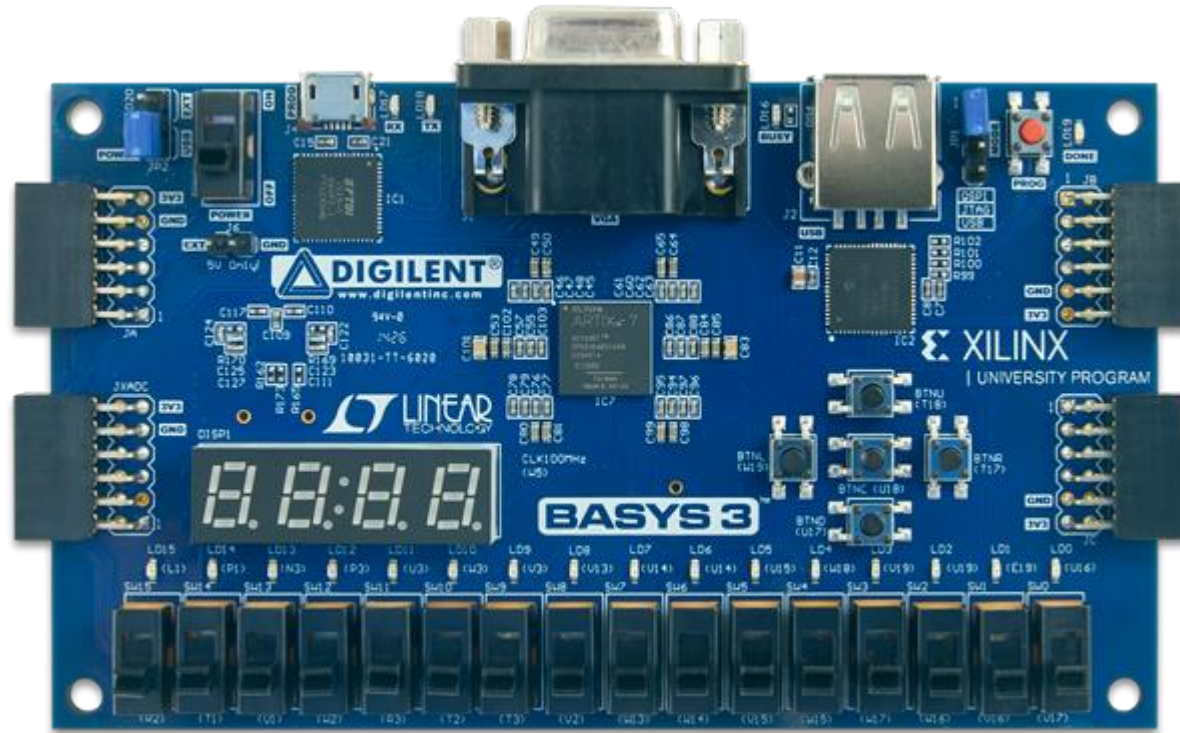


Figure: Basys 3 board from Digilent [9]

1. Introduction (18)

Examples of FPGA applications

Mid-range FPGAs are mostly used for [development and testing](#).

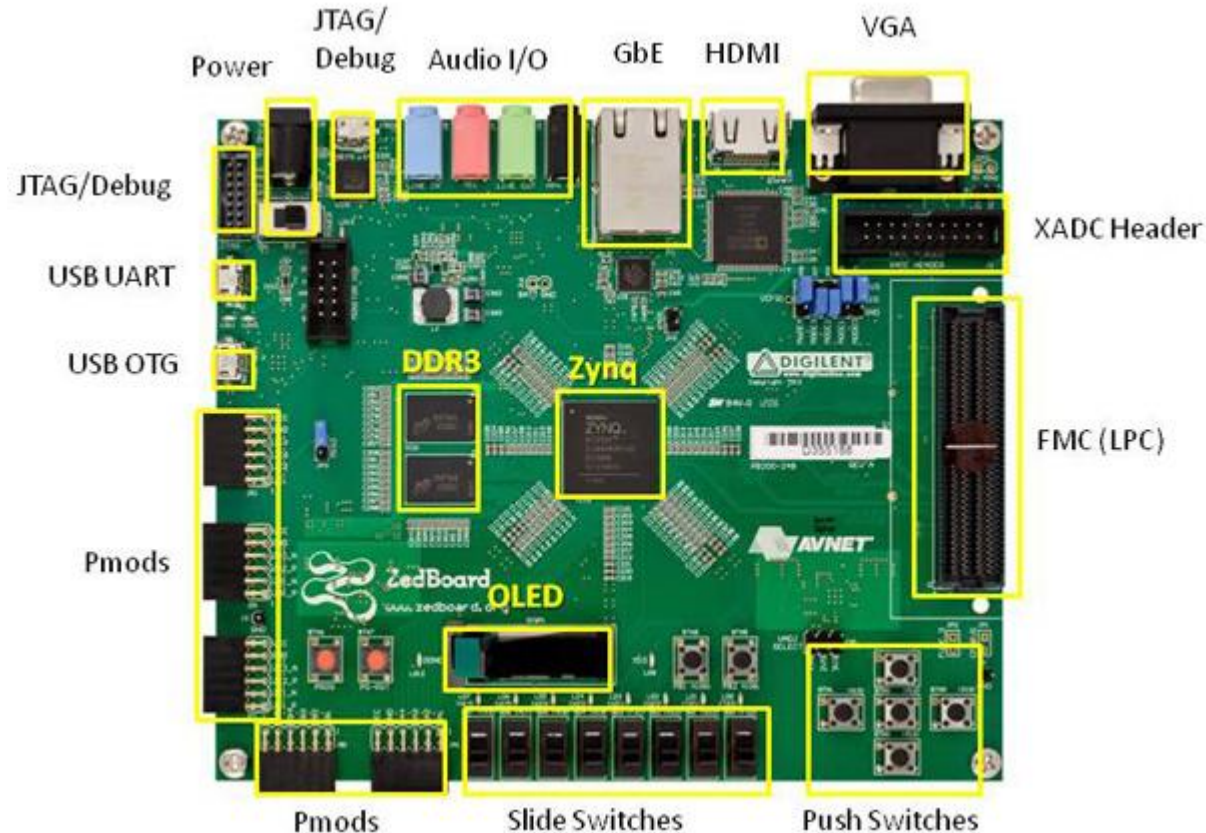


Figure: Zedboard from Digilent and Avnet [10]

Examples of FPGA applications

Hi-end FPGAs are used for state of the art equipment e.g. [high-speed computing and datacenters](#).

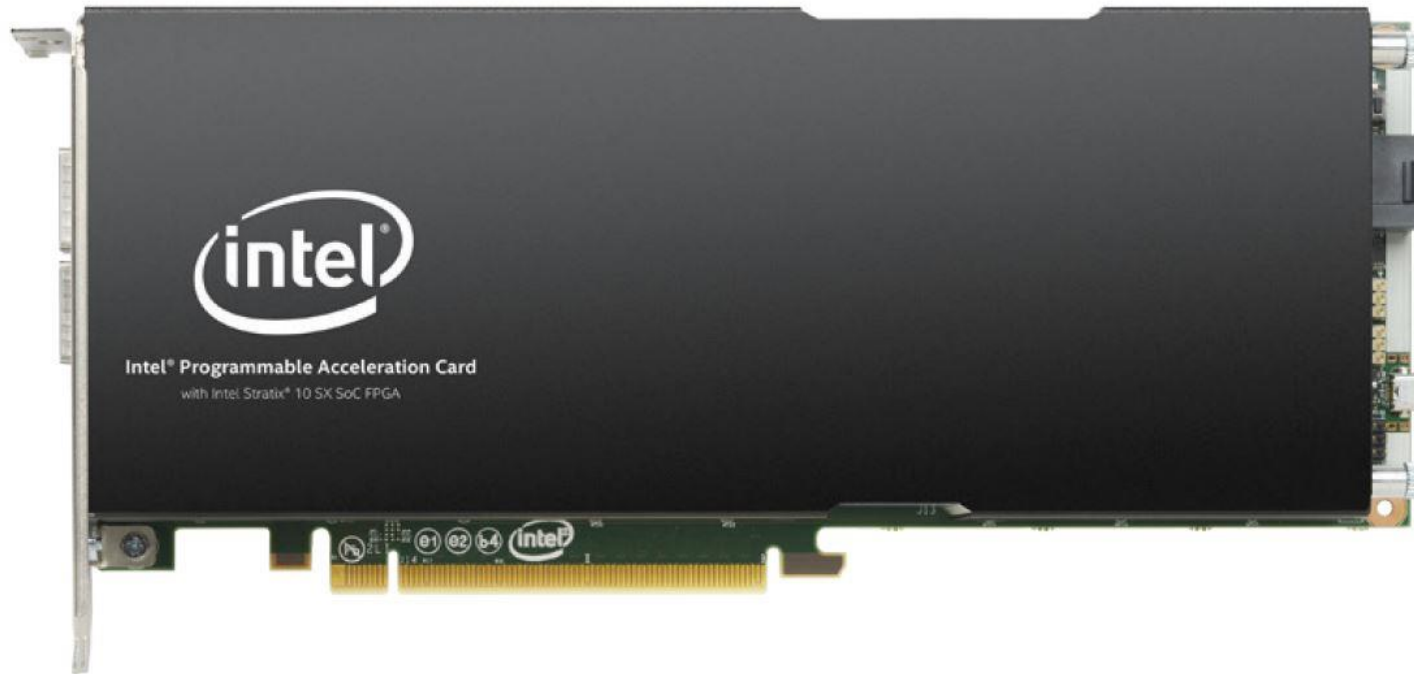


Figure: Intel FPGA PAC D5005 High-end Drop-in Accelerator [11]

1. Introduction (20)

What made FPGA possible?

Wrong analogy is often used in describing FPGA logic, as presented in the figure below.

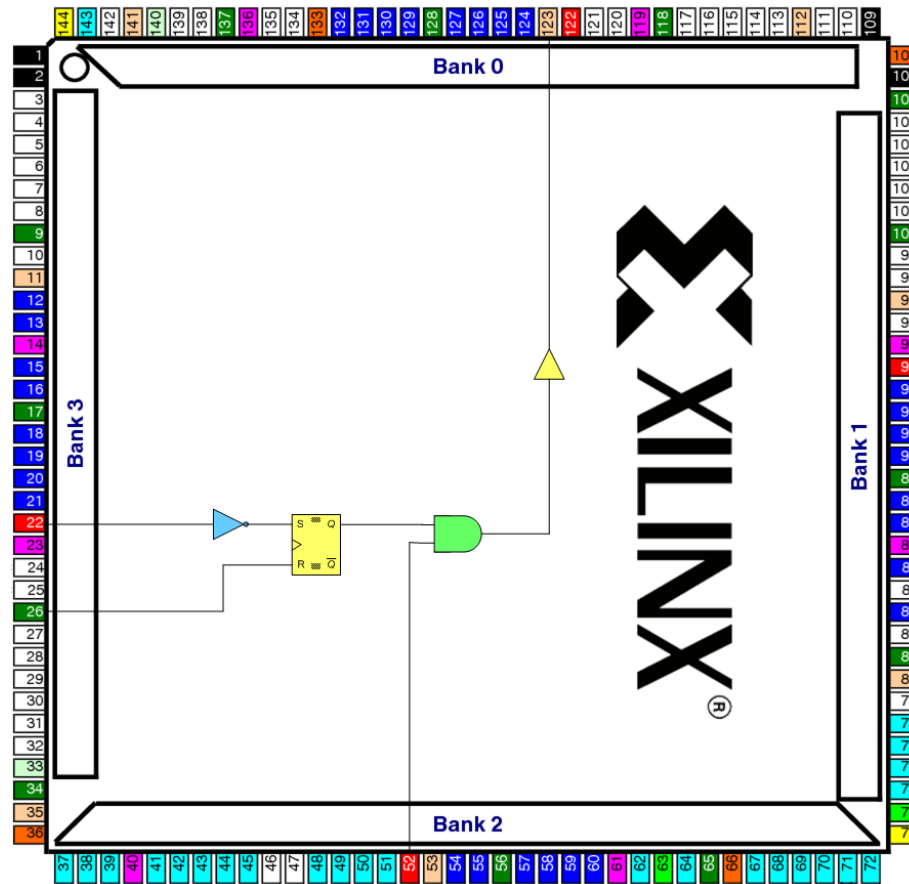


Figure: analogy for representing FPGA logic

There is **nothing technically wrong** with the used **analogy** in the previous figure.

FPGAs are **indeed used for realization of logical functions** in the hardware.

Problem with this analogy is that, as its **direct consequence**, people **automatically** think that FPGA hardware is **consisted of standard logical gates**.

There **aren't any standard** programmable logical gates in FPGA such as **NAND, OR, XOR, NOT...**

FPGAs exist thanks to the discovery of **universal logic cell**.

Famous scientist Claude Shannon laid theoretical **foundations for FPGAs** when he postulated method for writing **any logical function** using **only one electrical circuit**.

Boole's expansion theorem

Boole's expansion theorem, often called **Shannon expansion** or **Shannon decomposition**, is a method by which a **Boolean function** can be represented by the **sum of two sub function** of the original [12].

Shannon developed the idea that **Boolean function can be reduced** by means of the identity:

$$f = x \cdot f_x + \bar{x} \cdot f_{\bar{x}}$$

f - any logical function

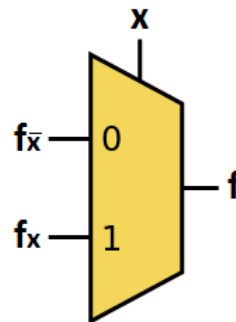
x - variable

\bar{x} - complement of x

f_x - positive Shannon cofactor of f

$f_{\bar{x}}$ - negative Shannon cofactor of f

Electrically, this theorem can be realized with **2:1 multiplexer**:



Example for AND logical function

Truth table		
a	b	F
0	0	0
0	1	0
1	0	0
1	1	1

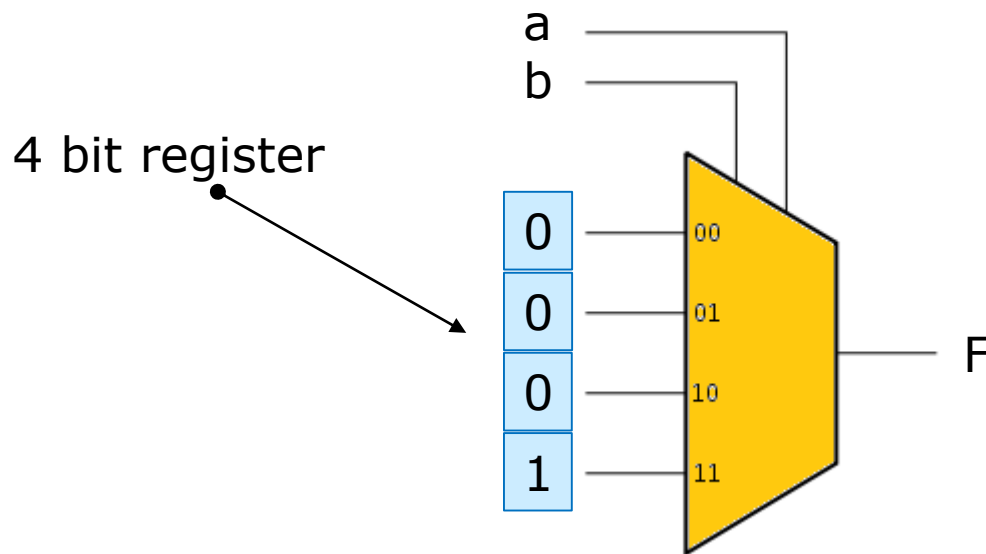
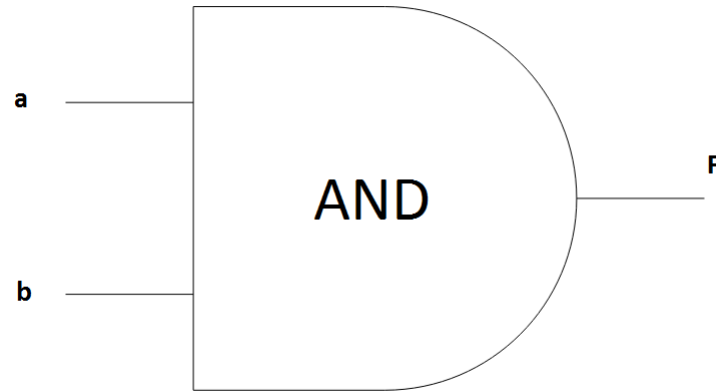


Figure: realization of AND function through universal logic cell

Example for OR logical function

Truth table		
a	b	F
0	0	0
0	1	1
1	0	1
1	1	1

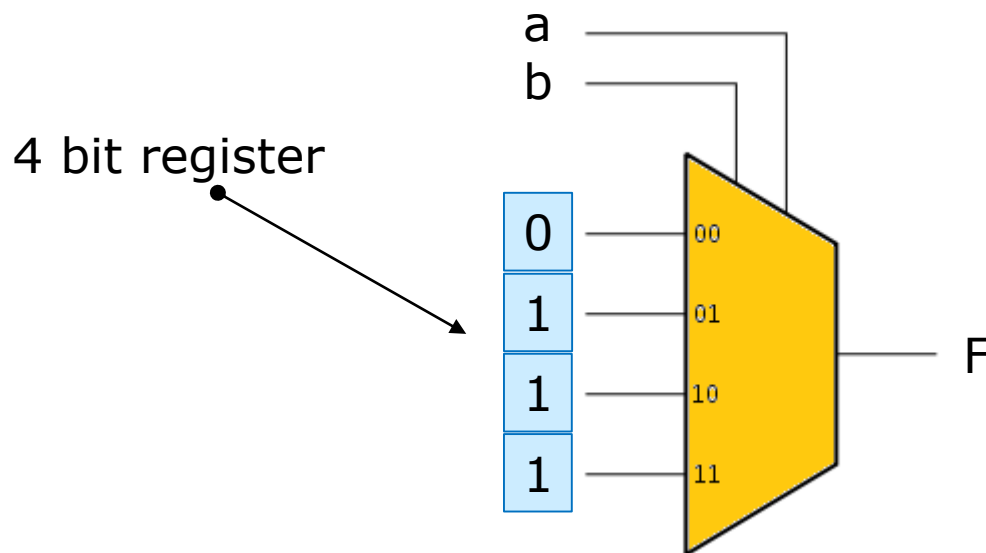
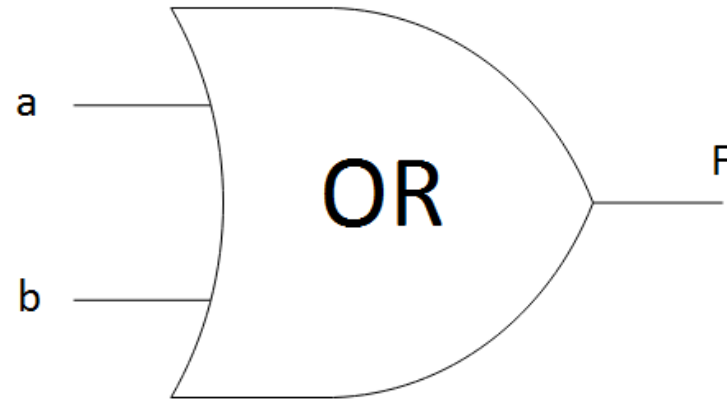


Figure: realization of OR function through universal logic cell

Conclusions from examples so far:

- 4:1 mux can be used for any logical circuit that has 2 inputs and 1 output
- only modification required is modification of the 4 bit register with the values from corresponding truth table for the given function
- there is only delay through the multiplexer
- it is trivial to prove that any multiplexer can be represented with 2:1 multiplexers
- circuitry comprised of SRAM register and multiplexer in FPGA is called LUT (look-up table) and is used for realization of combinatorial logic

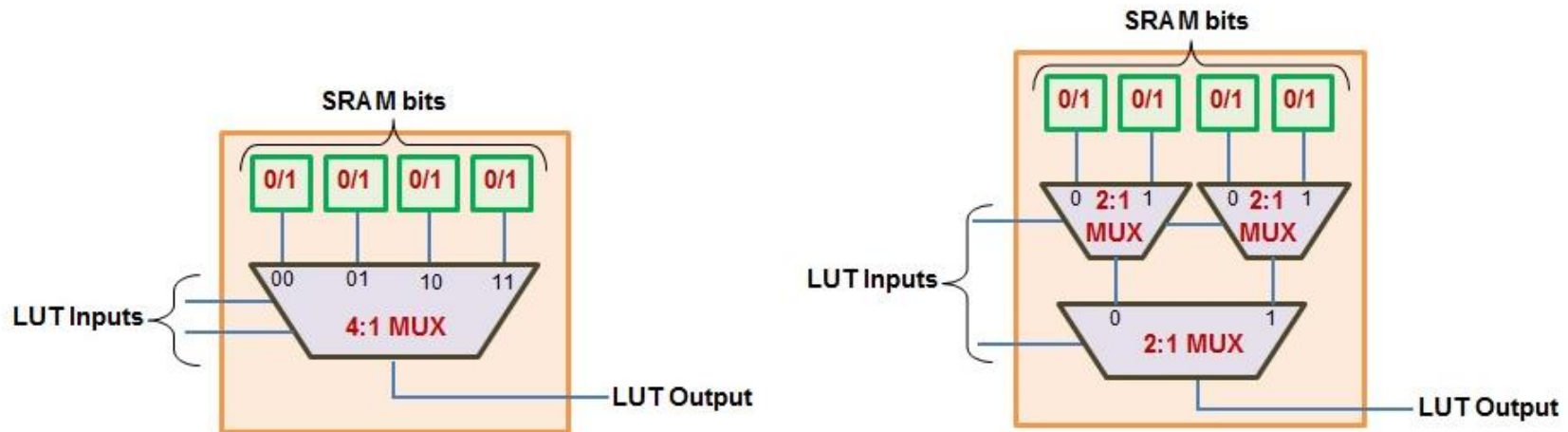


Figure: LUT representation with 4:1 MUX and 2:1 MUX

Number of LUT inputs

For many years scientists tried to calculate optimal number of inputs for the LUT in order to get minimal delay, but unfortunately that heavily depends on the logical function that is being implemented. After various statistical analysis it was determined that optimal number of input lies between 5 and 6. Because of the statistical uncertainty some analysis showed that the number is closer to 5 and others have showed that number is closer to 6 [13].

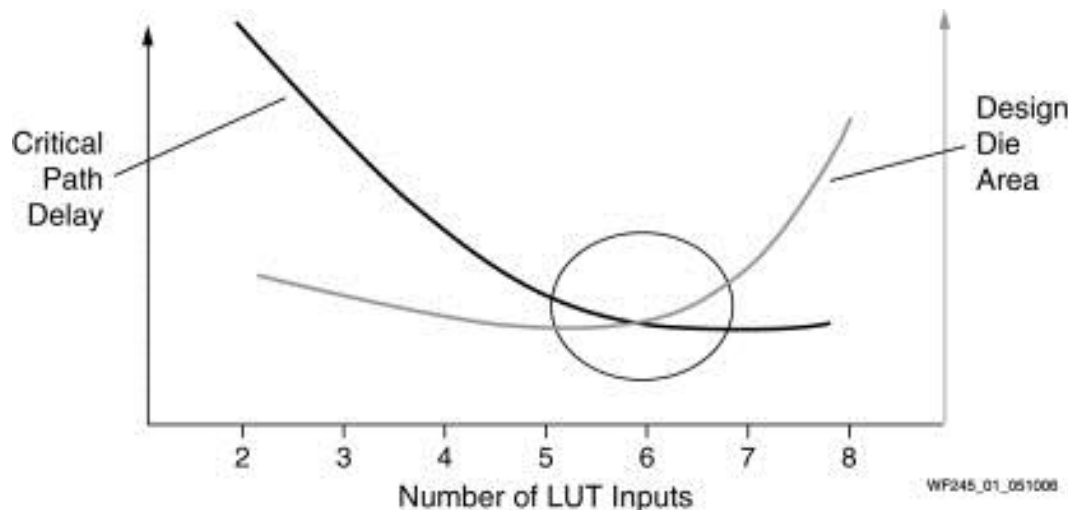


Figure: optimal number of inputs for LUT [13]

Number of LUT inputs

Because of **inconclusive results** of the analysis Xilinx **decided to combine** both options in **Spartan6 LUT**.

Spartan 6 FPGA **has two 5 input** LUTs that are combined into **one 6 input** LUT.

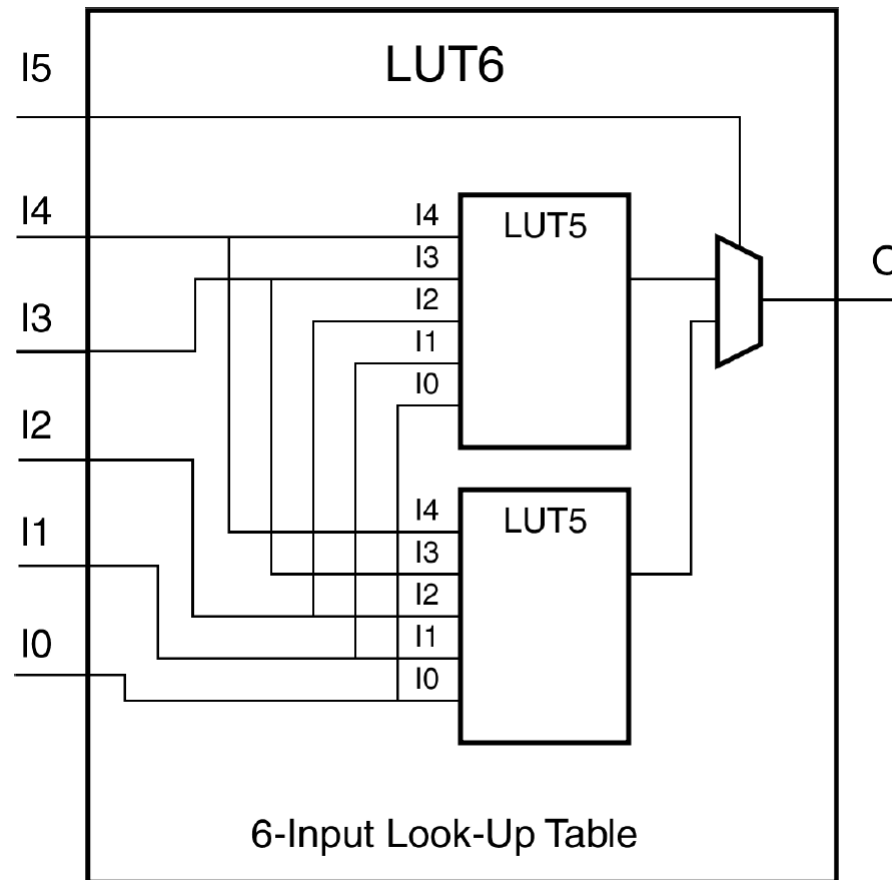


Figure: structure of LUT in Spartan 6 [14]

Configurable logic block

Addition of a flip-flop to the LUT allows realization of combinatorial and sequential logic in FPGA, and is called **configurable logic block (CLB)**.
Xilinx Spartan 3 has CLBs with 4 input LUTs.

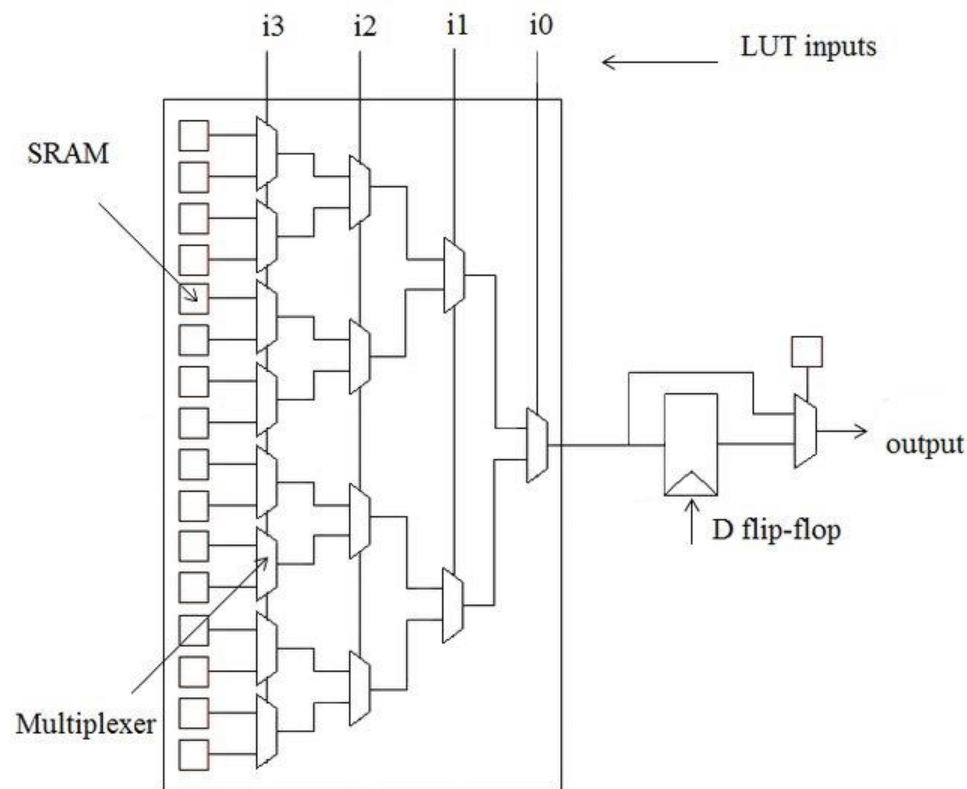


Figure: structure of CLB in Spartan 3 [15]

Configurable logic block

Xilinx Spartan 6 has CLBs with 6 input LUTs.

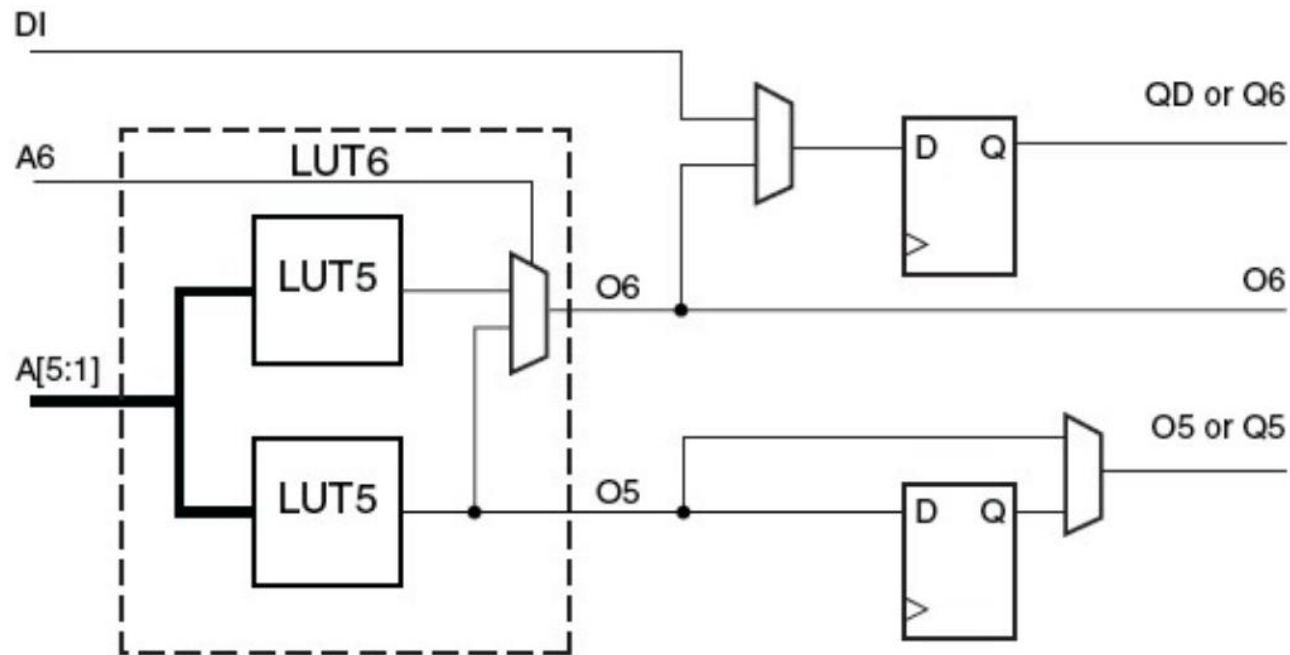


Figure: structure of CLB in Spartan 6 [16]

Routing connections between logic blocks

In order to construct **complex logical functions** it is necessary to interconnect **multiple CLBs** in FPGA.

Xilinx calls these **Switch matrix**, and there is one in **every cross node**.

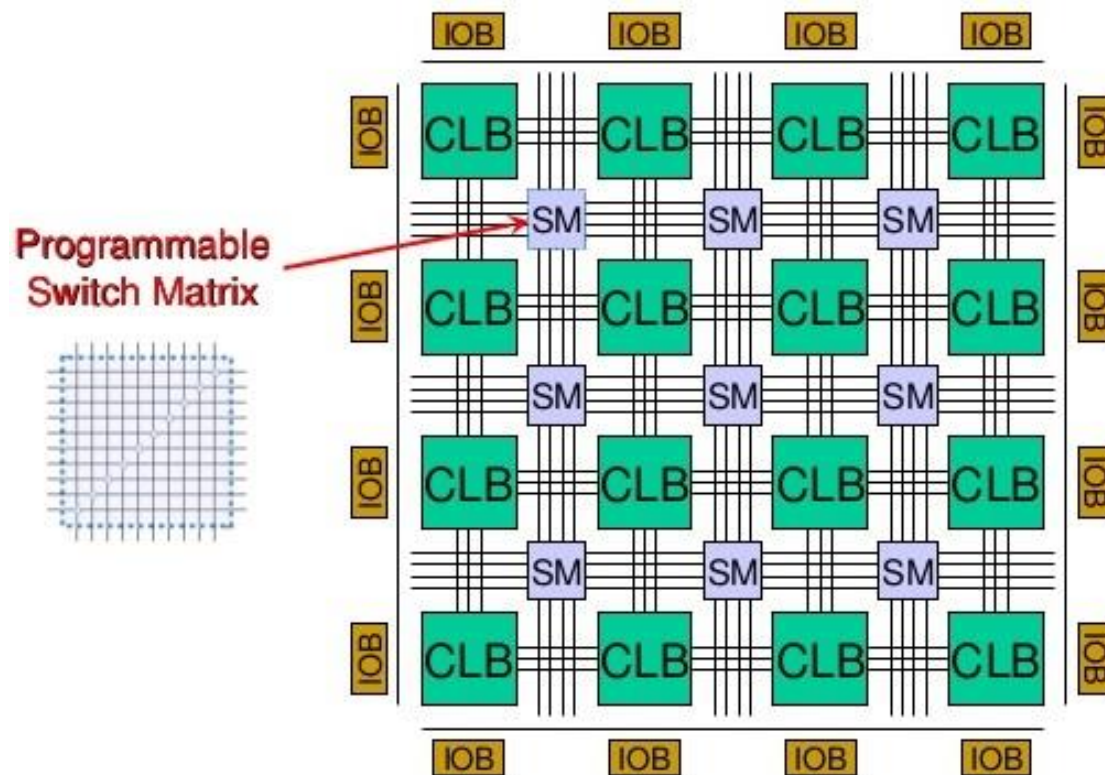


Figure: programmable switch matrix [17]

Routing connections between logic blocks

Every node in switch matrix has 6 switches (transistors) that controls 4 wires which are entering the node.

Switch matrix state is defined through SRAM that is connected to the transistor gates.

SRAM is programmed every time when FPGA boots.

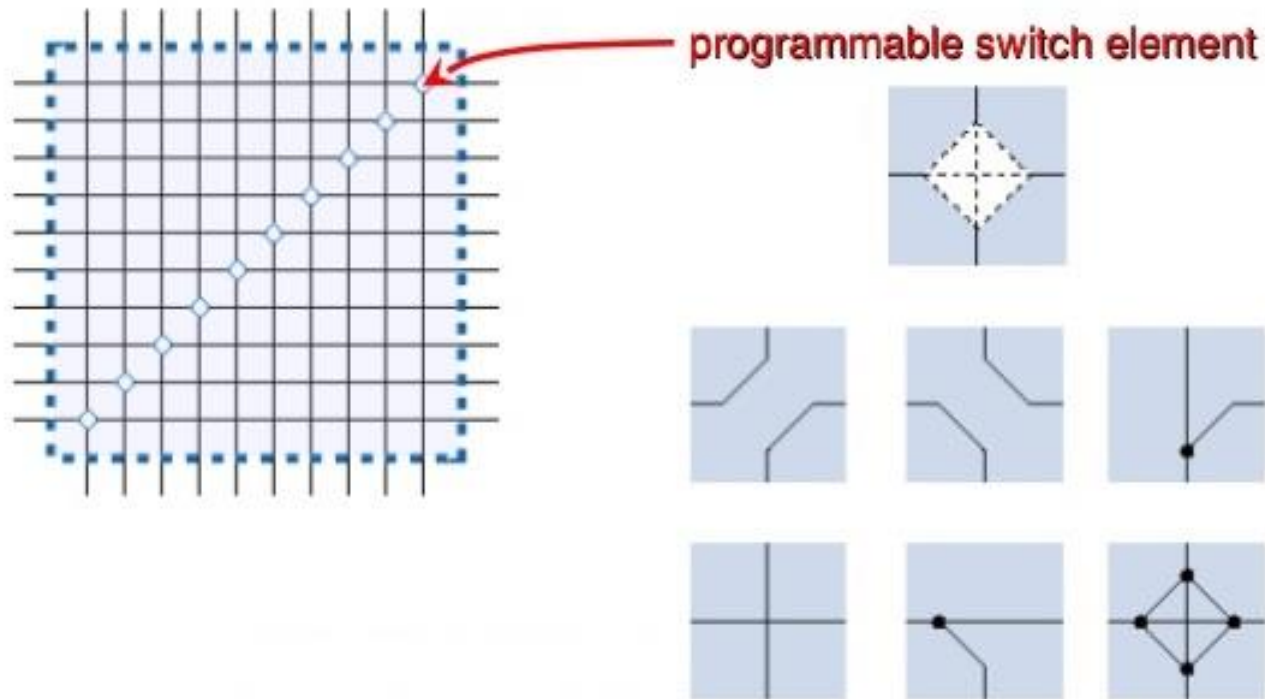


Figure: possible connections through switch matrix [18]

Routing connections between logic blocks

Complex logical functions are realized through **multiple interconnection** using programmable **switch matrix**.

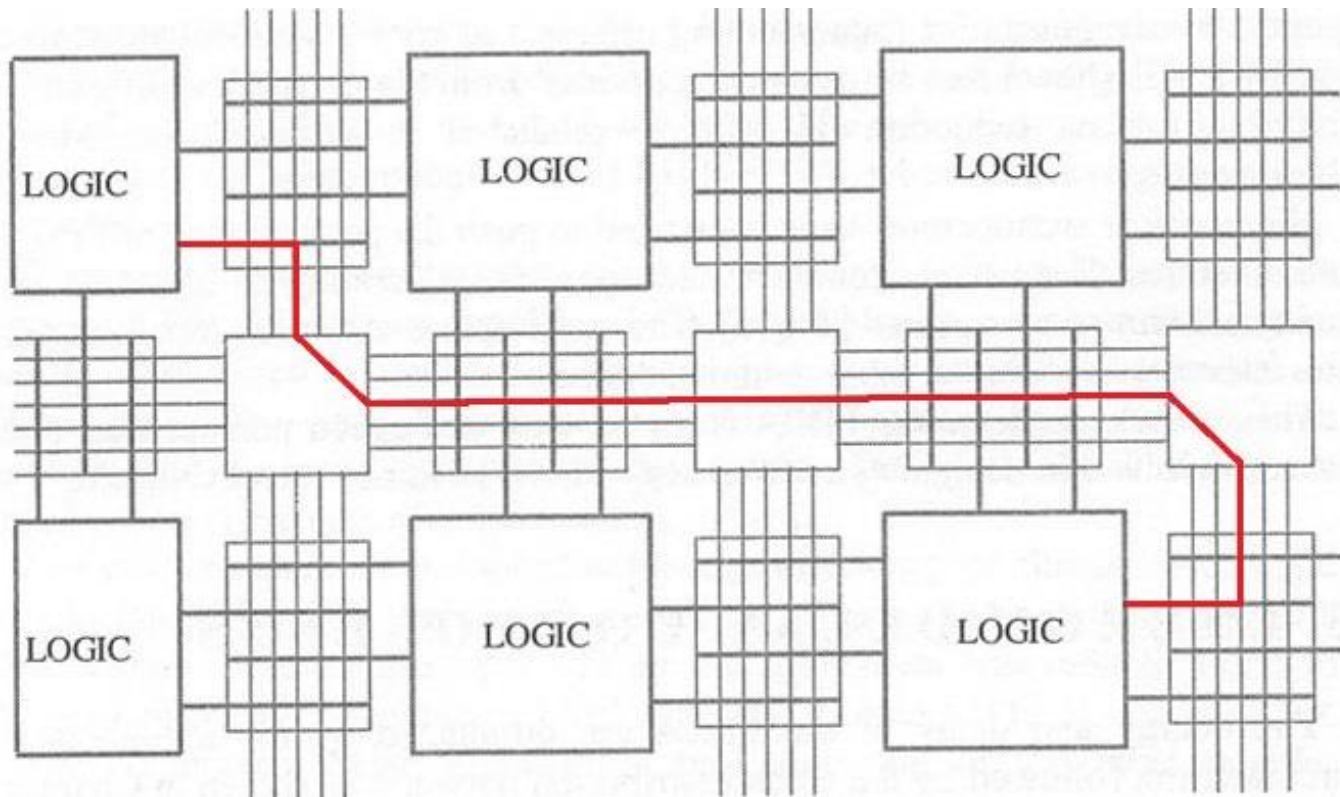


Figure: example of connecting logical blocks through switch matrix [19]

2. CPU and FPGA synergy

Early days

In the beginning FPGA very slowly gain traction on the market because processors had multiple advantages:

- schools and colleges predominantly held courses in programming for processors
- there was a lot of literature on processors
- there was a lot of literature on programming languages for processors
- processors were much cheaper than FPGA
- market was saturated with consumer electronics based on 68000, 6501, x86...

Early days

Very early users started designing software processors:

- many users had existing programs compiled for processors that could easily be reused
- standard software libraries
- there were much more experts on the market that were proficient in programing for processors

2. CPU and FPGA synergy (3)

Software processor

Example of **software processor** that is **very modest** in design:

- IR - **instruction register**, width **16bits**
- PC - **program counter**, loadable counter, width **8bits**
- ACC - **accumulator**, a general purpose data register, width **8bits**
- ALU - **arithmetic and logic unit**, performs: add, sub and bitwise_and, width **8bits**
- CONTROL_LOGIC - **decodes opcode field**, generating required **internal control signals** for each instruction
- ZERO - **zero detect**, produces a logic 1 output when ACC is zero
- DATA MUX - selects **immediate or memory data**, width **8bits**
- ADDR MUX - selects **program counter or absolute address**, width **8bits**

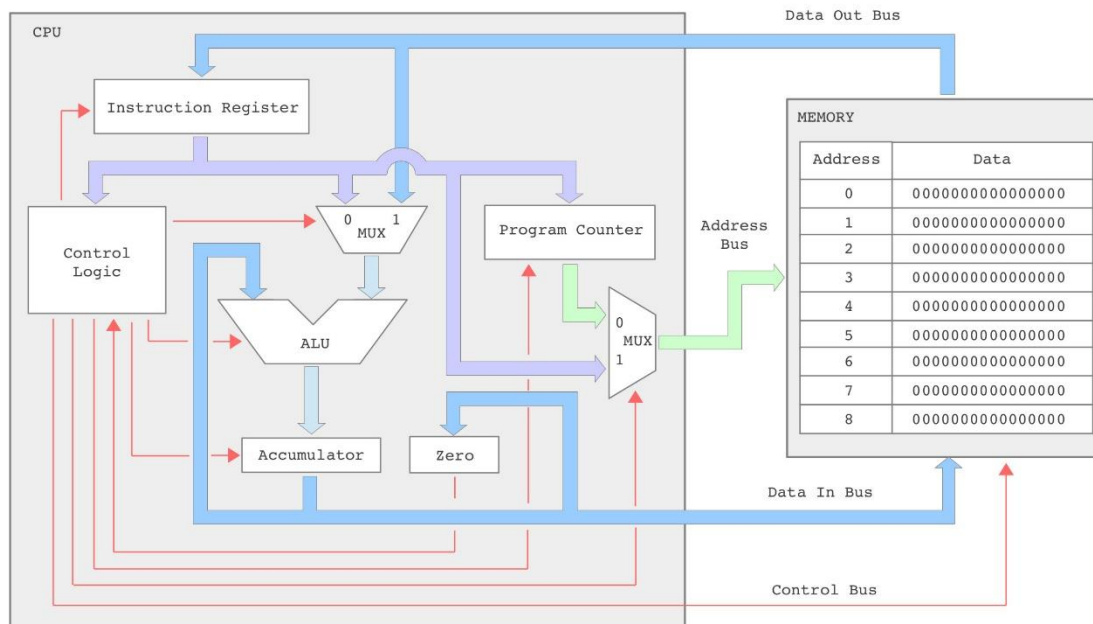


Figure: example of simple software processor

Xilinx simple 8 bit software processor - Picoblaze

Xilinx quickly recognized this growing need on the market and came up with a simple design. Author was Ken Chapman.

Main features:

- 16 byte-wide general-purpose data registers
- 1K instructions of programmable on-chip program store, automatically loaded during FPGA configuration
- Byte-wide Arithmetic Logic Unit (ALU) with CARRY and ZERO indicator flags
- 64-byte internal scratchpad RAM
- 256 input and 256 output ports for easy expansion and enhancement
- Automatic 31-location CALL/RETURN stack
- Predictable performance, always two clock cycles per instruction, up to 200 MHz or 100 MIPS in a Virtex-II Pro FPGA
- Fast interrupt response; worst-case 5 clock cycles
- Optimized for Xilinx Spartan-3 architecture—just 96 slices and 0.5 to 1 block RAM
- Support in Spartan-6, and Virtex-6 FPGA architectures
- Assembler, instruction-set simulator support

2. CPU and FPGA synergy (5)

Xilinx simple 8 bit software processor - Picoblaze

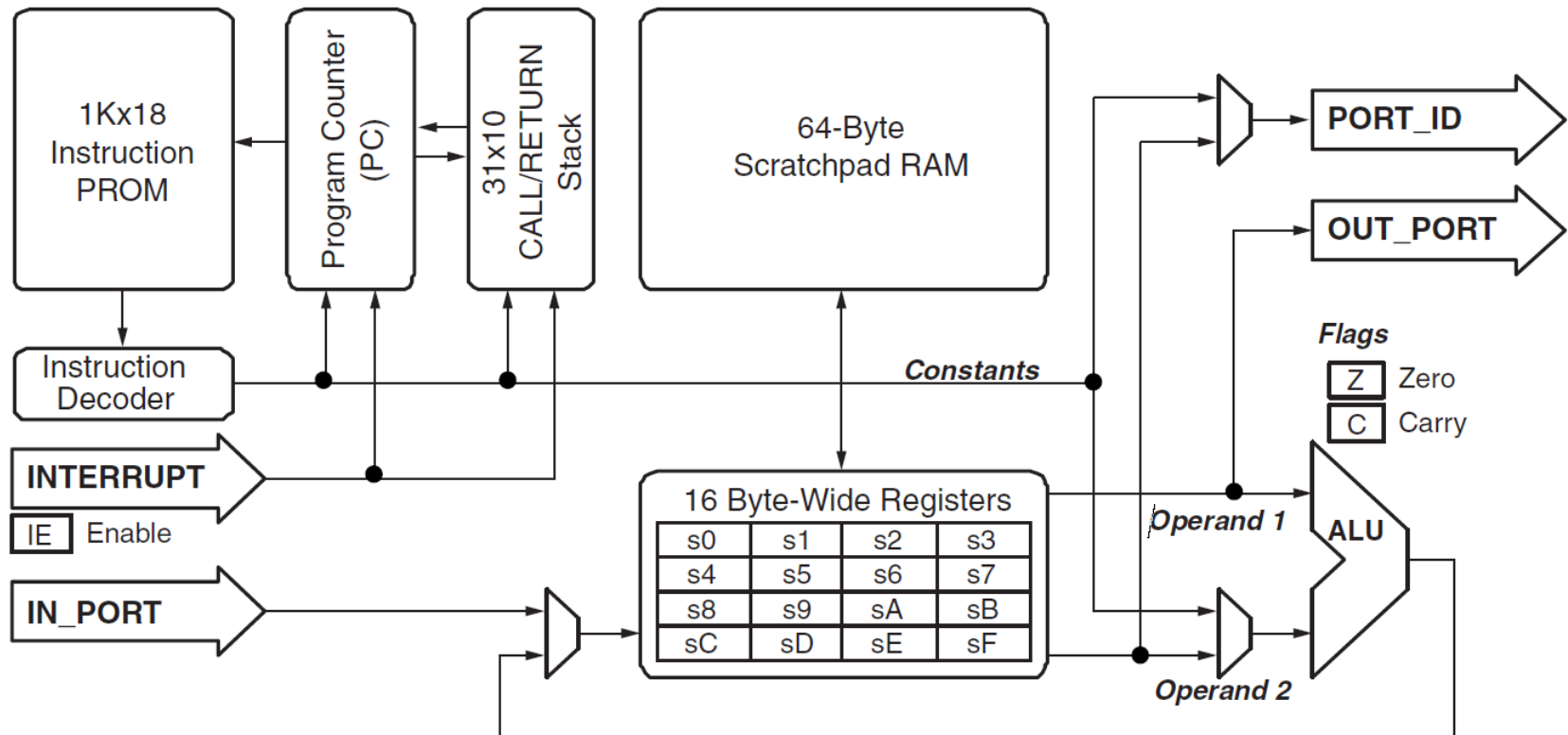


Figure: Picoblaze block diagram [20]

Xilinx 32 bit software processor - Microblaze

After [success with Picoblaze](#), Xilinx decided to publish [32 bit RISC](#) software processor.

Features (every section can be [separately configured or completely turned off](#)):

- Clock up to 300 MHz
- DDR3 memory support
- Multichannel DMA
- LVDS I/O Performance 1.25Gb/s
- Transceiver Performance: 6.25Gb/s
- Ethernet subsystem
- Controller Area Network
- Streaming FIFO
- HDMI Camera/Display Interface
- MIPI-CSI, MIPI-DSI
- Video DMA
- Timer / Watchdog
- Mutex / Mailbox
- UART
- USB 2.0
- Quad SPI
- General Purpose I/O (GPIO)
- Pulse-Width Modulator (PWM)

2. CPU and FPGA synergy (7)

Xilinx 32 bit software processor - Microblaze

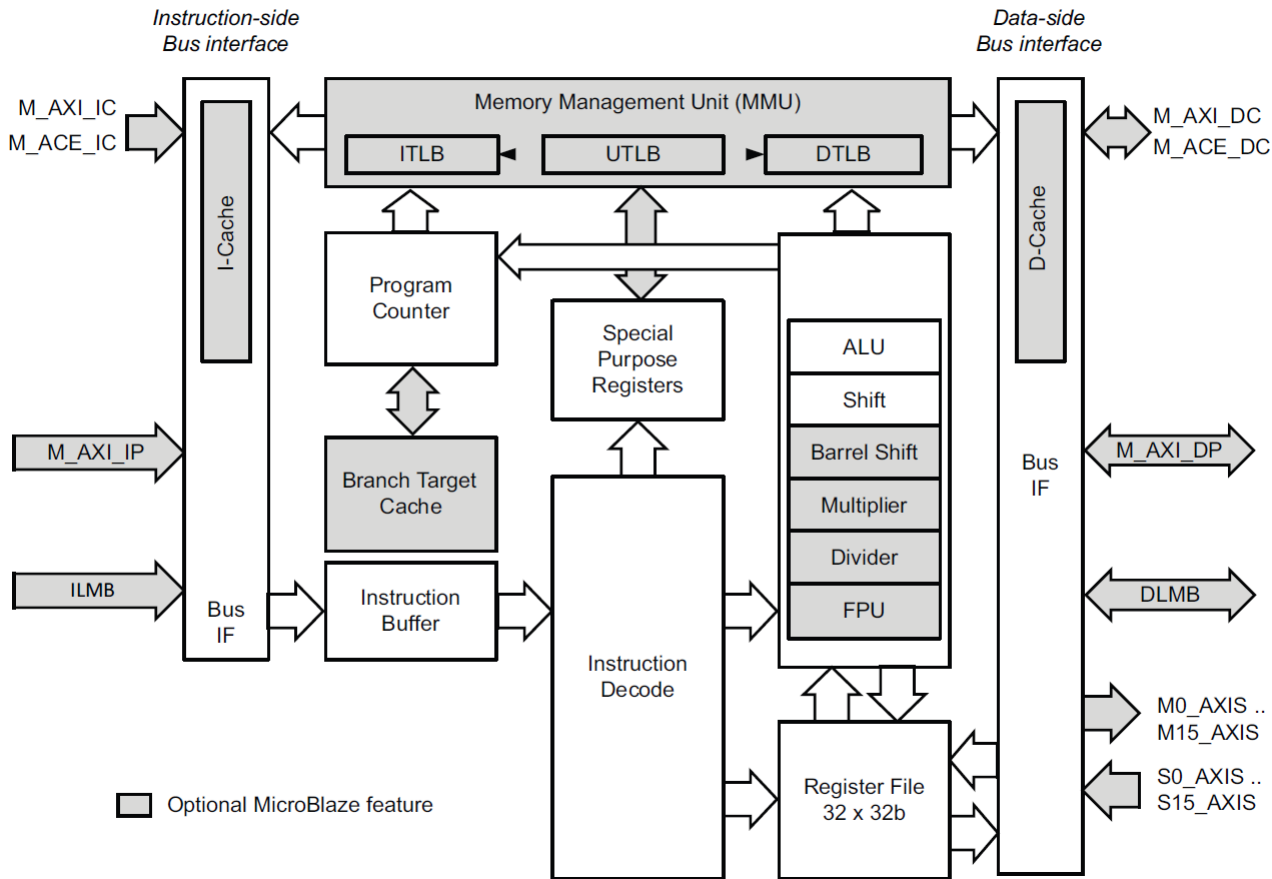


Figure: Microblaze block diagram [21]

Intel (Altera) 32 bit software processor – Nios II

The [Nios II architecture](#) defines the following functional units:

- Register file
- Arithmetic logic unit (ALU)
- Interface to custom instruction logic
- Exception controller
- Internal or external interrupt controller
- Instruction bus
- Data bus
- Memory management unit (MMU)
- Memory protection unit (MPU)
- Instruction and data cache memories
- Tightly-coupled memory interfaces for instructions and data
- JTAG debug module

2. CPU and FPGA synergy (9)

Intel (Altera) 32 bit software processor – Nios II

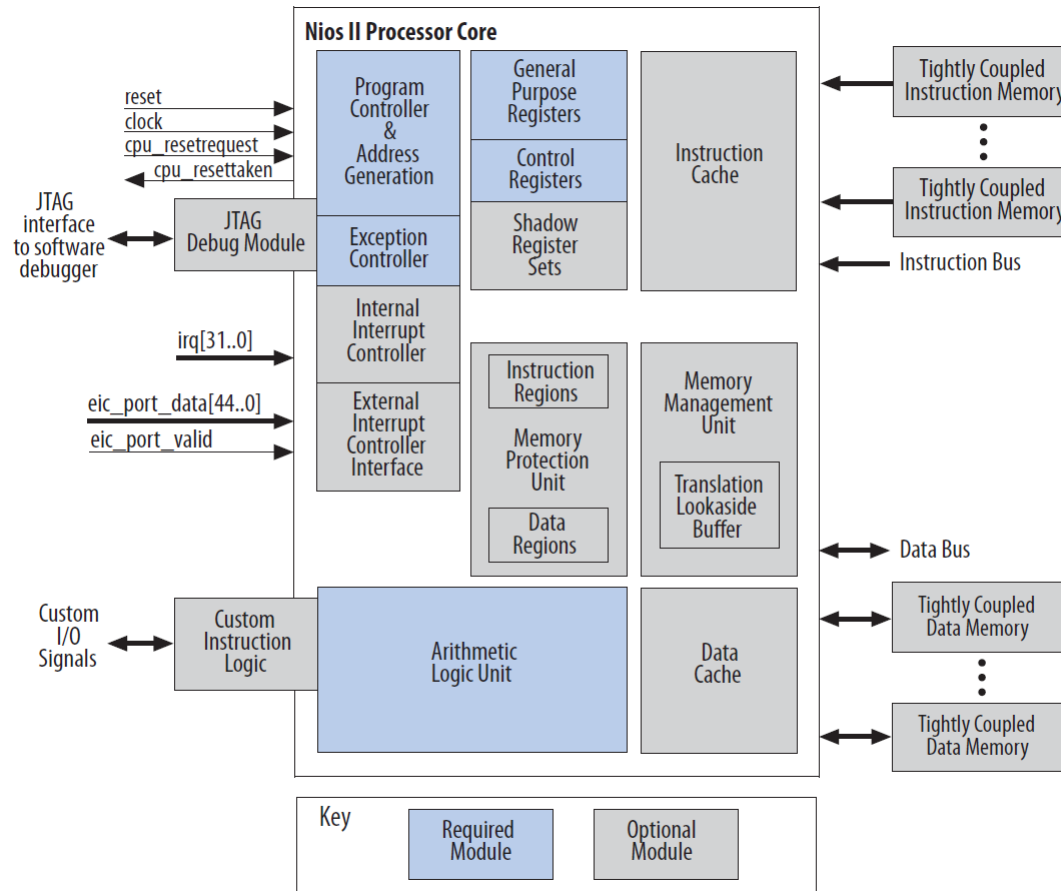


Figure: Nios II block diagram [22]

Migrating from soft to hard processors

After **success with software processors** FPGA manufacturers decided to **implement hard processors** on the same silicon die and **integrate** it with FPGA fabric. **Instead designing their own processor** FPGA manufacturers opted to **buy licenses** for already established processor brands.

Intel (**Altera**) and **Xilinx** acquired the same processor architectures through licensing: **ARM, MIPS, PowerPC**.

With **new architecture** came **new paradigm** in embedded systems:

- accent is on **software ecosystem** where multiple software platforms are interconnected
- moving from **system on board** to **system on chip concept**
- Introducing **high level** programming languages and **GUI oriented** frameworks

2. CPU and FPGA synergy (11)

System on a Board

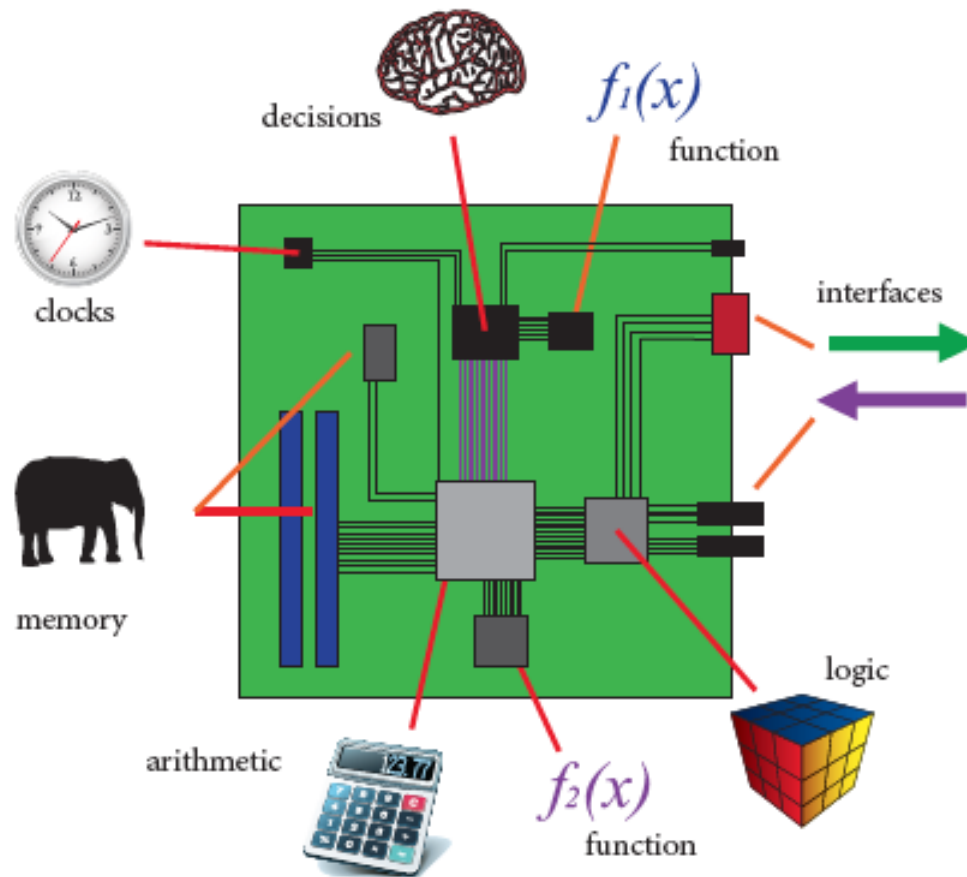


Figure: traditional electronics system [23]

2. CPU and FPGA synergy (12)

System-on-Chip (SoC)

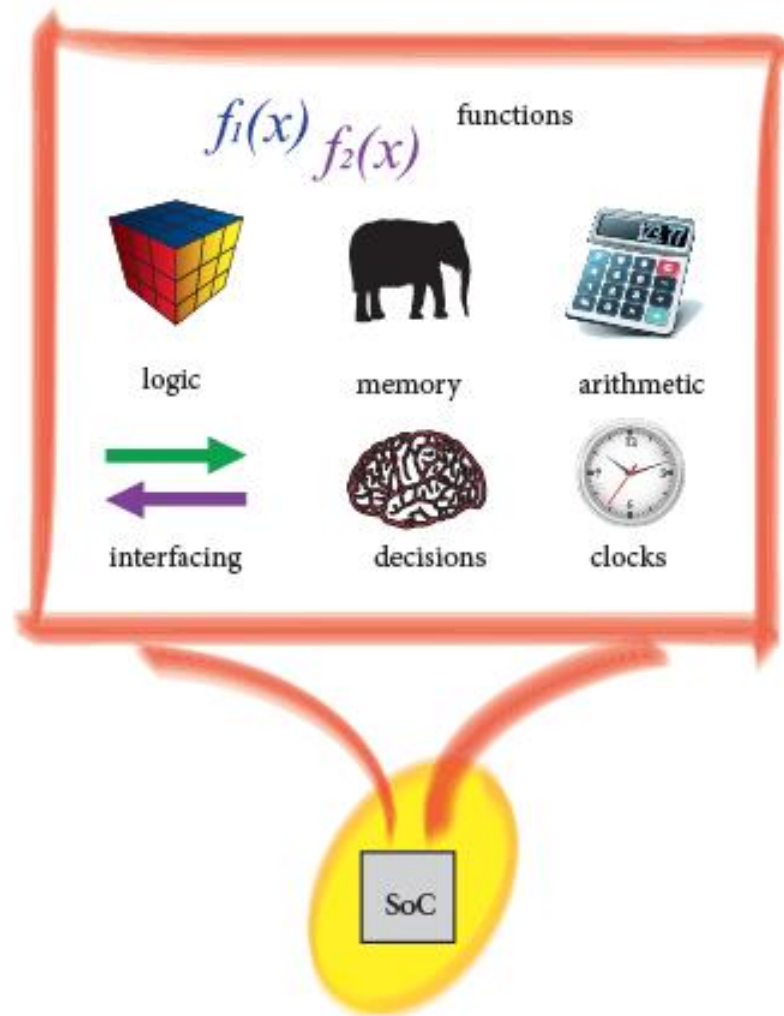


Figure: new SoC concept [23]

2. CPU and FPGA synergy (13)

Examples of Soc with different manufacturers

In recent years **ARM processors** became **preferred option** with many manufacturers.

Xilinx Zynq

Zynq-7000 All Programmable
SoCs with Cortex-A9 MPCore



Altera Arria V & Cyclone V

Hard processor system (HPS)
with Cortex-A9 MPCore



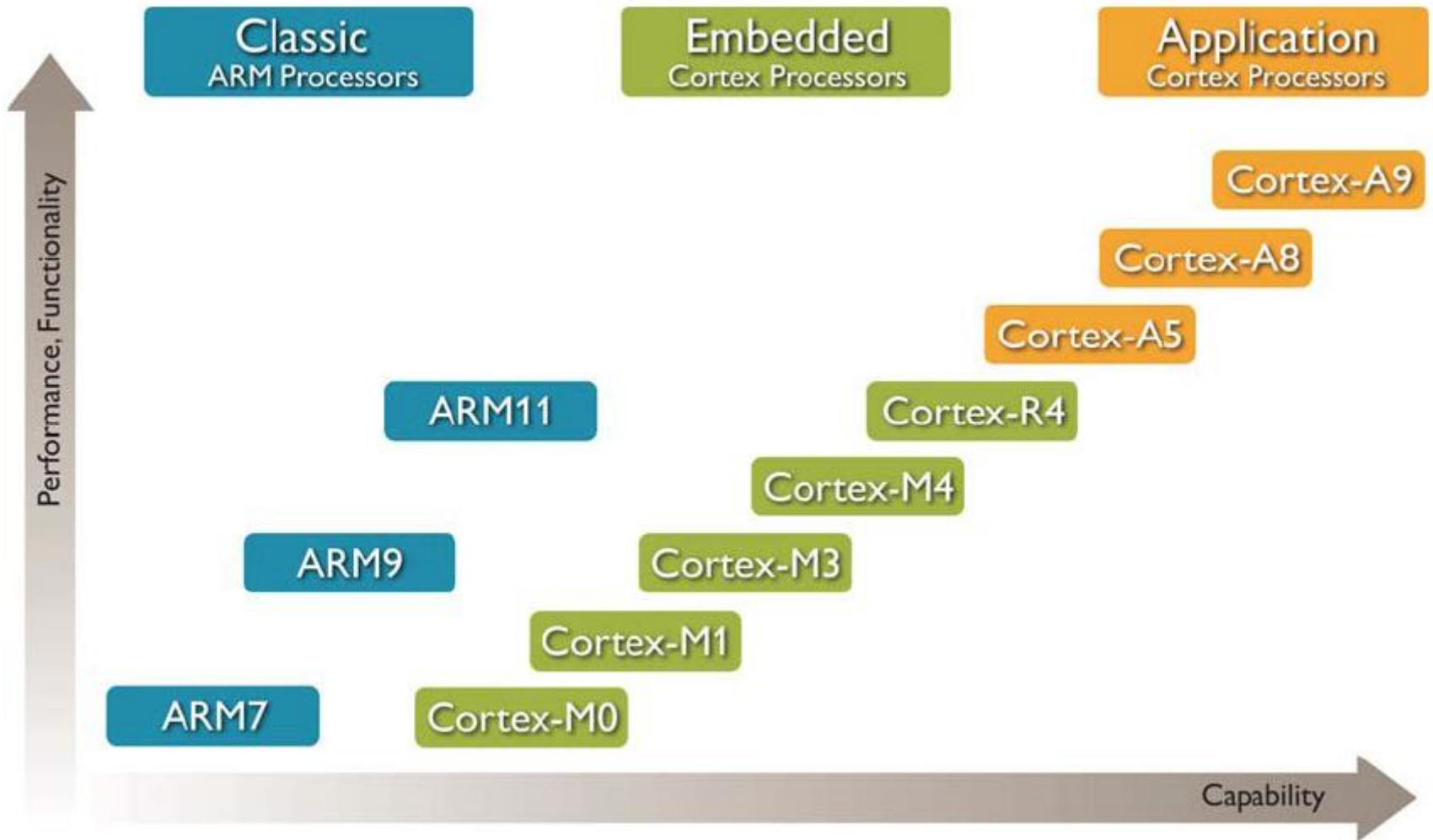
Microsemi Smartfusion2

Cortex M3



2. CPU and FPGA synergy (14)

ARM processor roadmap



WP369_03_042110

Figure: different ARM families [24]

Zynq architecture

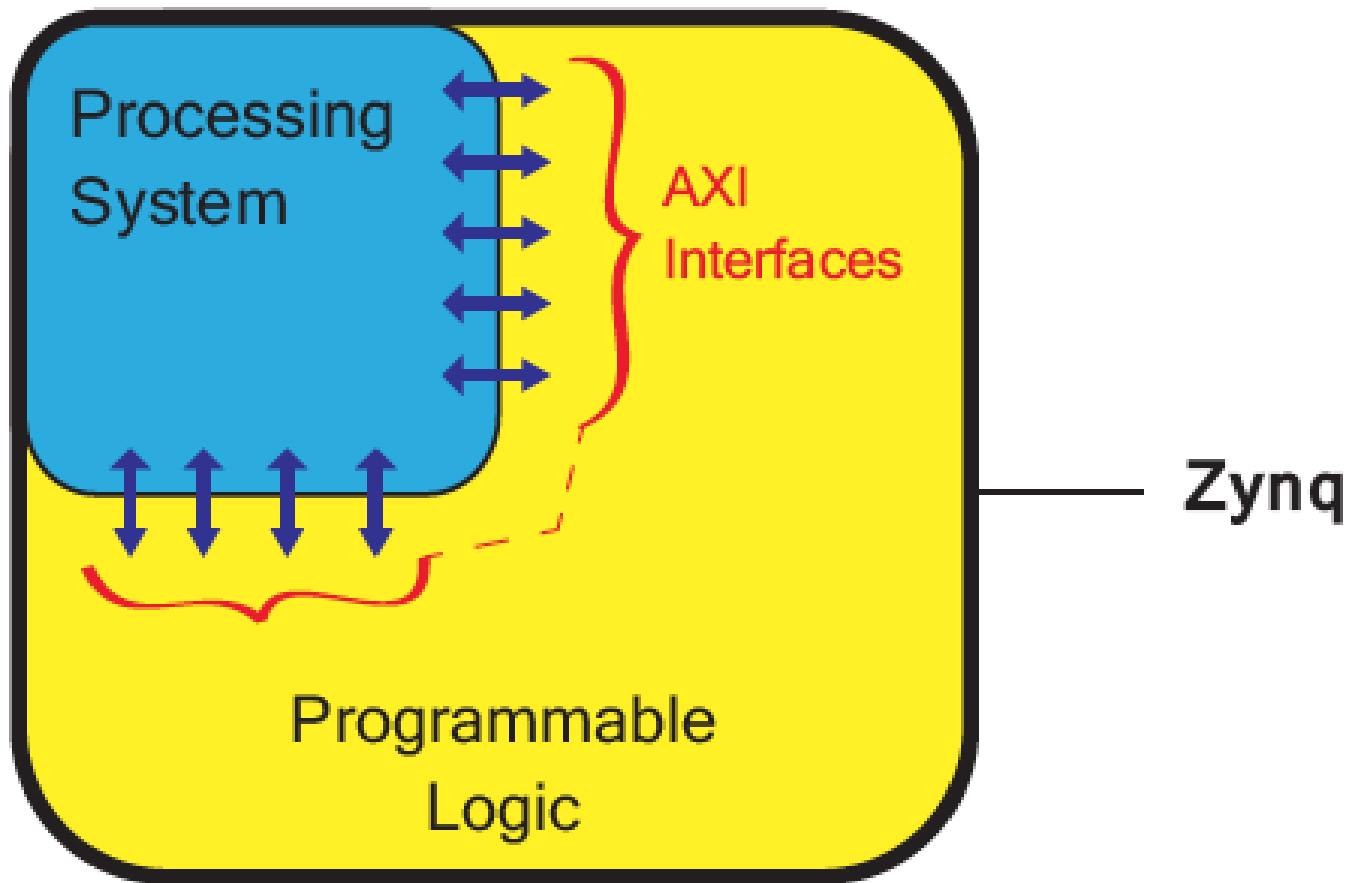


Figure: connection between processor and FPGA logic over AXI bus [23]

Zynq embedded SoC architecture

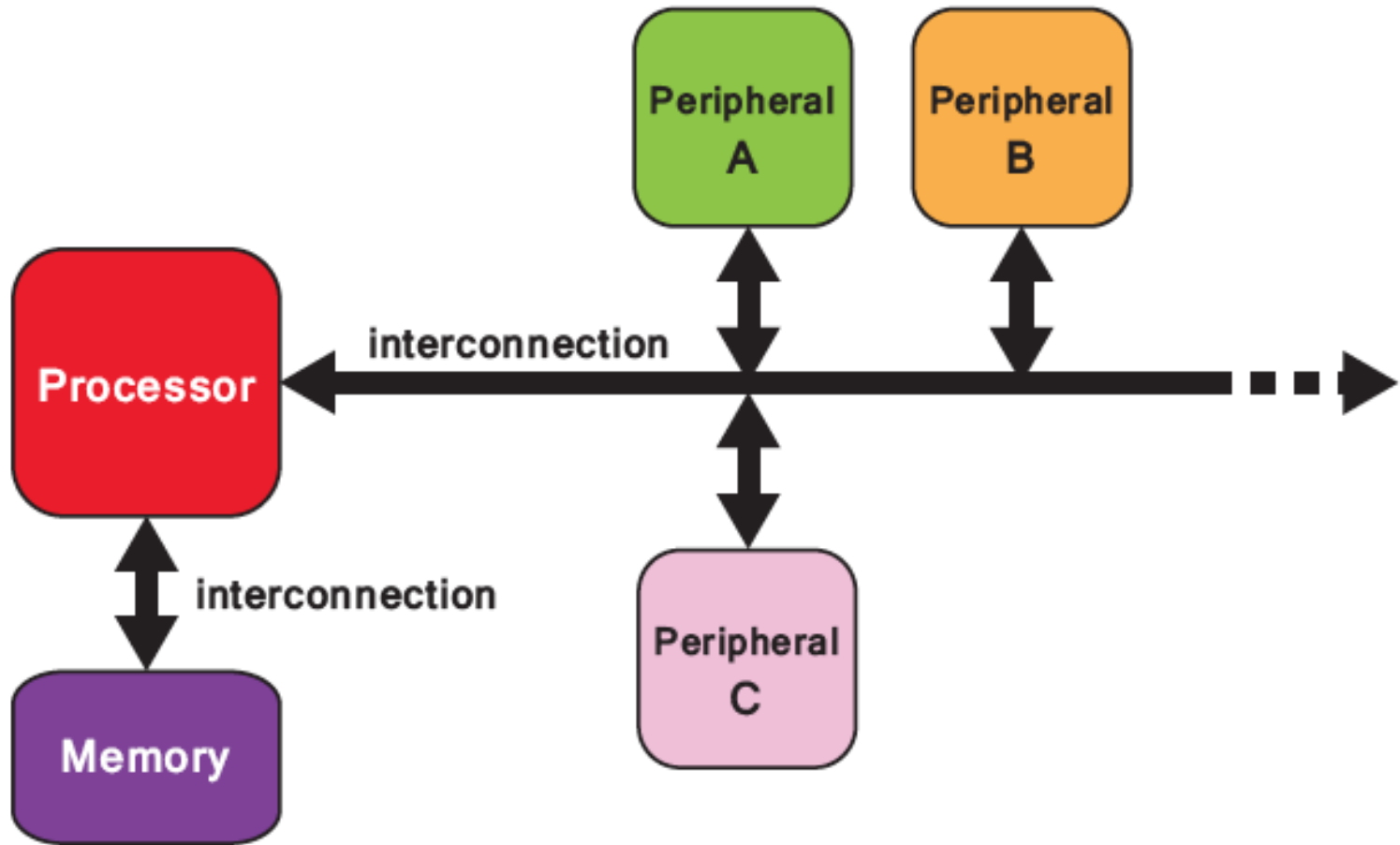


Figure: structure of interconnections in Zynq [23]

2. CPU and FPGA synergy (17)

Implementing embedded SoC on Zynq

Main accent today is on development of software stack.

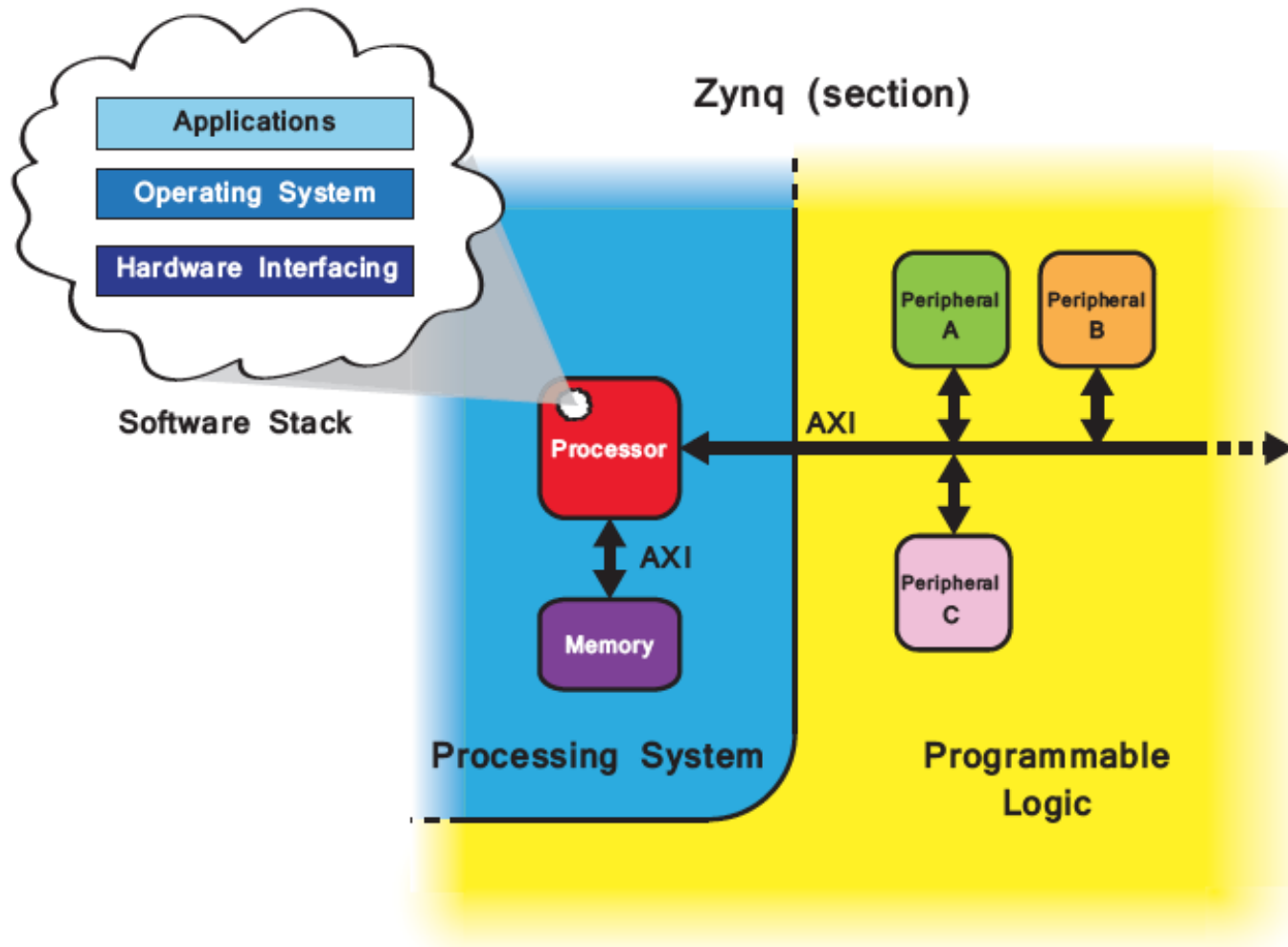


Figure: place of software stack in Zynq [23]

2. CPU and FPGA synergy (18)

Zynq SoC Ecosystem

In order to support complex applications Zynq hardware has [various platforms in synergy](#).

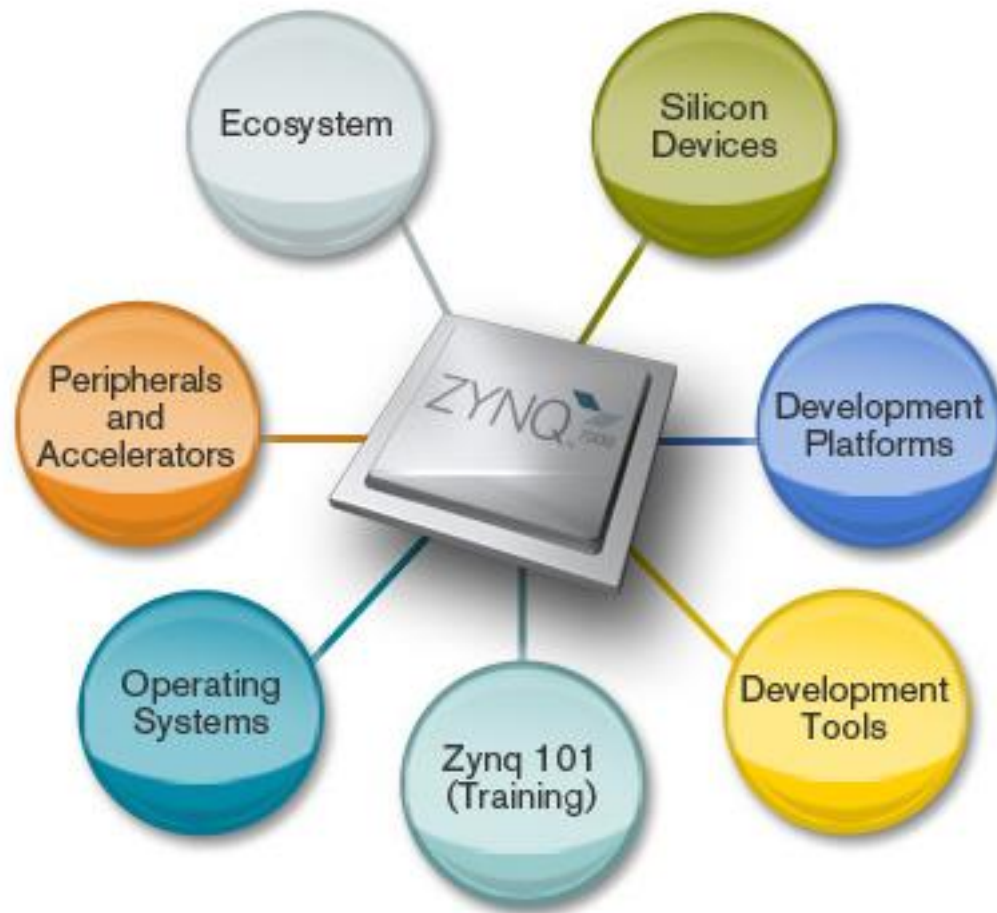


Figure: complex ecosystem surrounding Zynq [23]

2. CPU and FPGA synergy (19)

Zynq SoC Ecosystem

Ecosystem connect not only different software but also different companies.

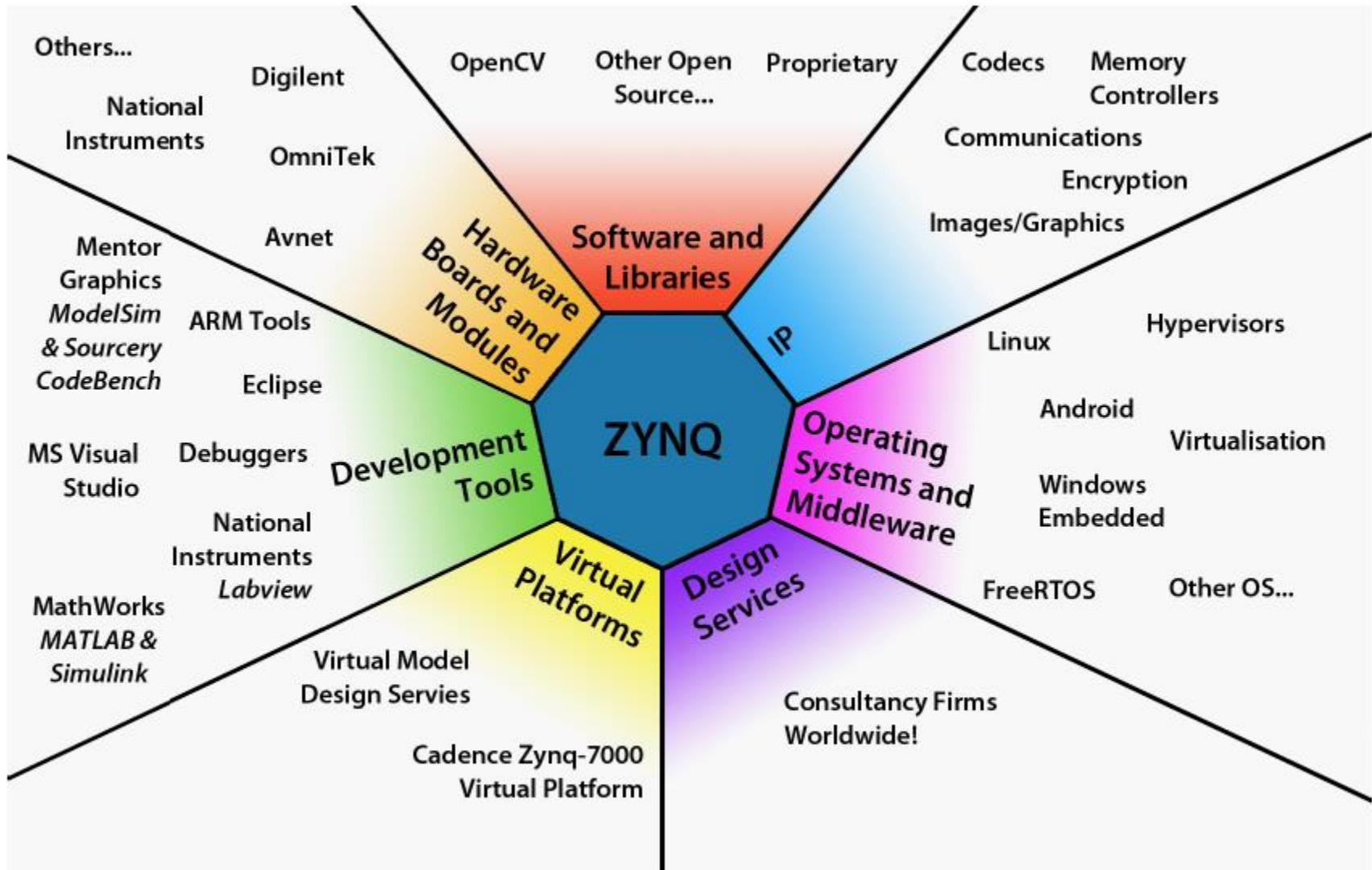
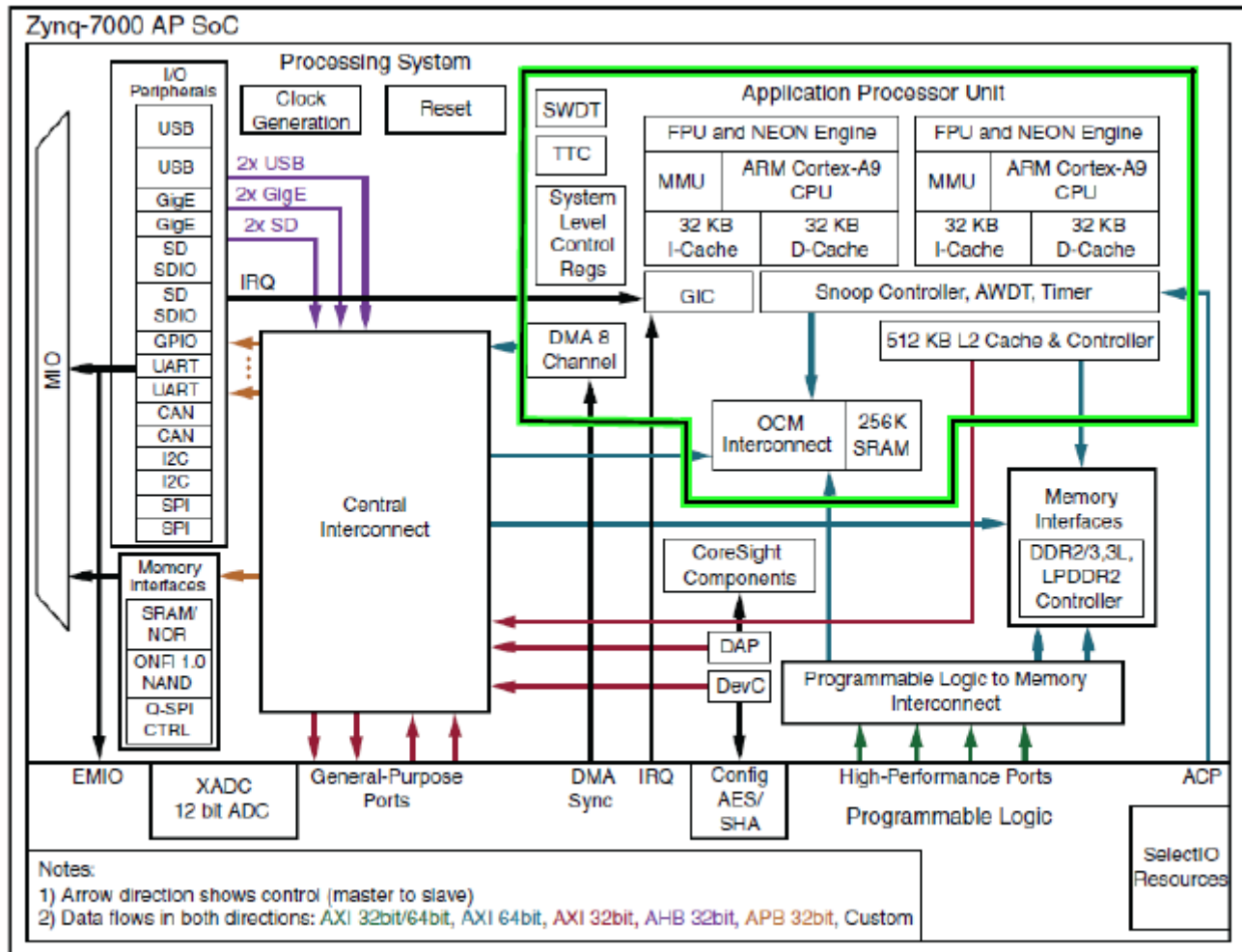


Figure: Zynq ecosystem extends from hardware and software to the companies [23]

2. CPU and FPGA synergy (20)

Zynq processing system



DS100_01_00713
© Xilinx

Figure: block schematic of Zynq-7000 [23]

2. CPU and FPGA synergy (21)

Options between various implementation platforms

	Total System Cost	Flexibility	Differentiation	Time-to-Market	Cost of Derivatives	Risk
Zynq SoC	Low + best value	Most flexible: HW and SW programmable + programmable I/O	Highest degree of programmability, HW/SW co-design	Fastest for integrated HW & SW differentiation	Lowest due to HW & SW programmability	Predictably low risk
ASSP + FPGA	Higher than Zynq SoC (system dependent)	Highly flexible but ASSP I/O limited compared to Zynq SoC	HW and SW programmable, ASSP-dependent	Fastest if ASSP requires HW differentiation	Low to high depending on FPGA vendor	Low to high depending on FPGA vendor
ASSP	Lowest if SW-only programmability is sufficient	Good but SW-programmable only	Limited to SW programmable only - easy cloning	Fastest if SW-only differentiation required	Lowest if SW-only derivatives needed	Can be Lowest if SW-only programmability is sufficient
ASIC	High to prohibitive	Once manufactured only limited SW flexibility	Best HW differentiation but limited SW differentiation	Lowest & riskiest	Highest	Terrible (respins)

Figure: options with different platforms in regards to risk and development time [25]

Zynq advantages and features

Lowest NRE, Best Risk Mitigation	Greatest Flexibility & Differentiation	Streamlined Productivity & Fast TTM	Lowest Cost of Derivatives & Highest Profitability
<ul style="list-style-type: none">✓ Already manufactured silicon✓ Negligible development & design tool costs✓ Xilinx IP library + third-party IP✓ Extensive development boards	<ul style="list-style-type: none">✓ All Programmable HW, SW & I/O✓ Anytime field programmable✓ Partial reconfiguration✓ System Secure (encryption)	<ul style="list-style-type: none">✓ Instant HW/SW co-development✓ All Programmable Abstractions (C, C++, OpenCV, OpenCL, HDL, model-based entry)✓ Vivado Design Suite, Vivado HLS, IP Integrator & UltraFast Methodology✓ Broad and open OS & IDE support (Open-source Linux & Android, FreeRTOS, Windows Embedded, Wind River, Green Hills, & many others)	<ul style="list-style-type: none">✓ IP standardized on ARM AMBA AXI4✓ Reuse precertified code (ISO, FCC, etc.)✓ Reuse & refine code & testbenches✓ Volume silicon, power circuitry, PCBs & IP licensing

Figure: list of advantages and flexibility of Zynq platform [25]

3. FPGA in mobile phones

3. FPGA in mobile phones (1)

FPGA technology feasibility for mobile phones

For a very long time FPGA technology **wasn't viable for mobile phones**:

- Relatively **low profit** on high production series
- NRE cost
- ASIC **outperforms** FPGA
- ASIC is **usually much smaller** than FPGA
- ASIC solutions are much **harder to reproduce (reverse engineering)** by rival companies

In recent years there was a change in **extremely important factor** of the mobile phone manufacturing:

RELEASE SCHEDULE.

3. FPGA in mobile phones (2)

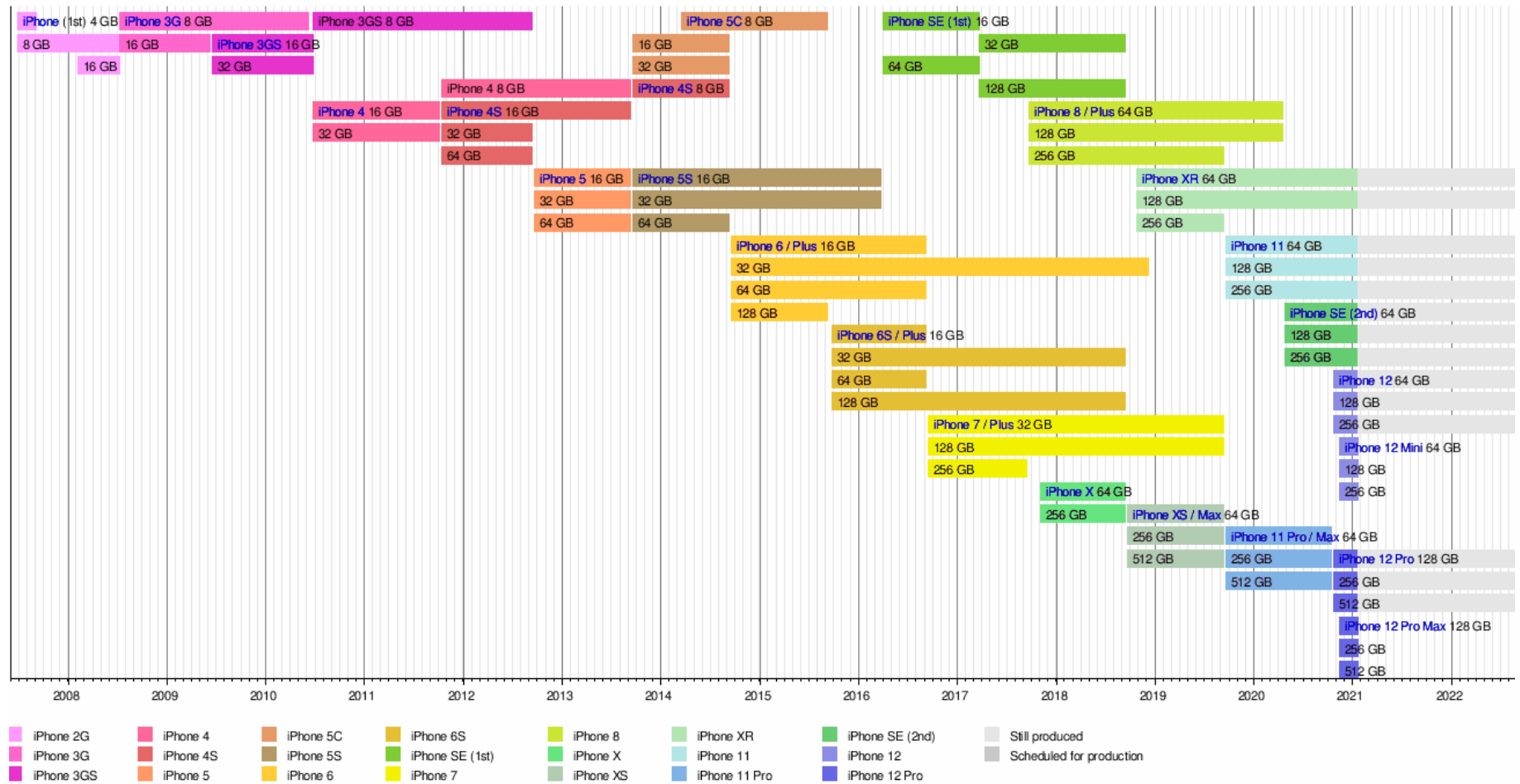
Mobile phones release schedule

Mobile phone development is **complex task an it use to take years** to release new phone model. In recent years companies, such as Apple and Samsung, started **pushing for release dates of 12 months** for their **flagship phone models**.

Such a **short time period is often not enough** for the engineers to finish ASIC development, thus **new market was opened** for FPGA technology.

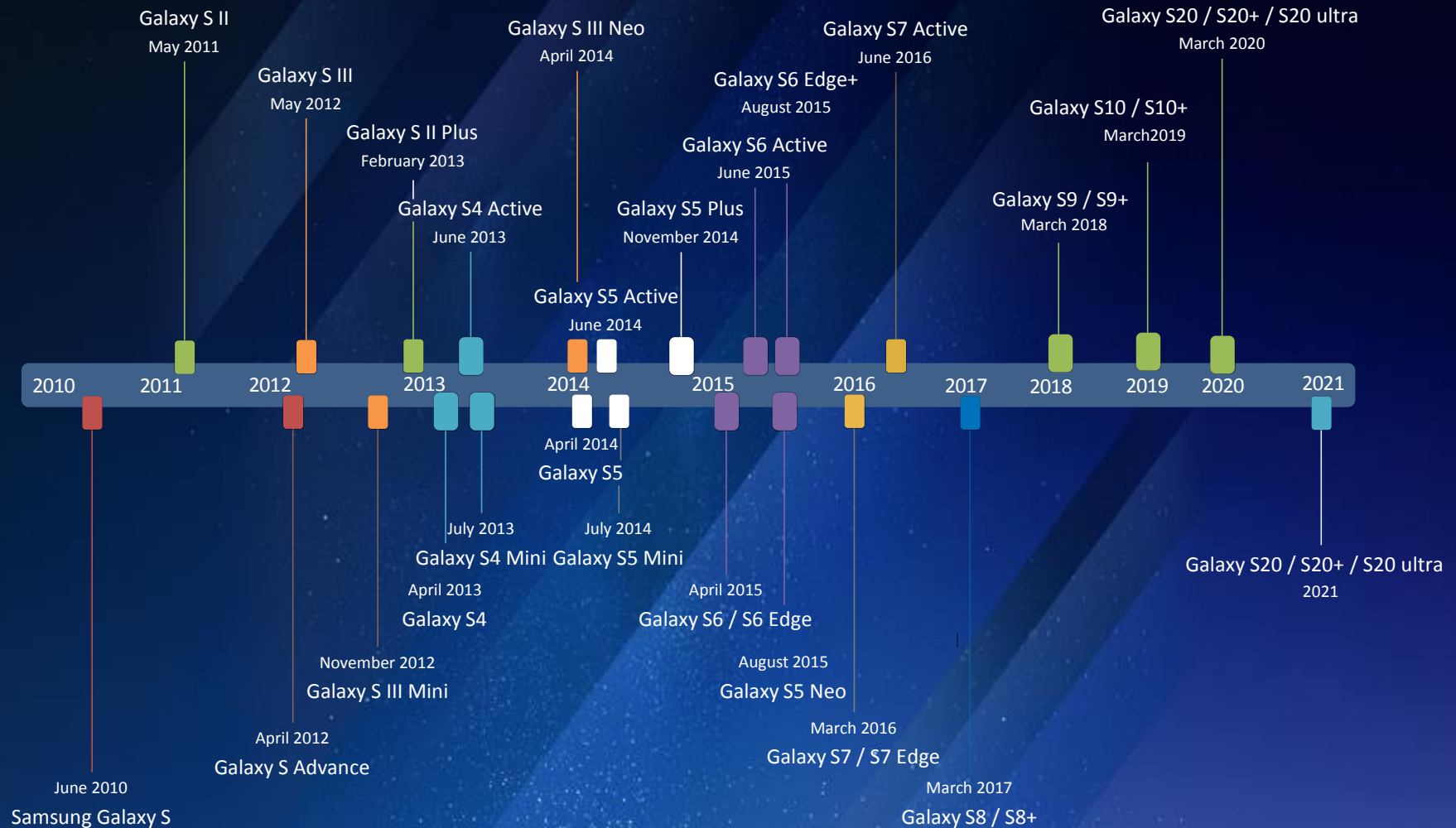
3. FPGA in mobile phones (3)

Apple iPhone release dates [26]:



3. FPGA in mobile phones (4)

Samsung Galaxy S series release dates [27]:



Advantages of FPGA technology in recent years

With time to market constantly decreasing FPGA gained more areas for application:

- used as “glue logic” for communication between different areas of the phone
- suitable for offloading main processor
- introduction of machine learning and AI in phones
- easy to upgrade firmware because FPGAs are reprogrammable
- FPGA chip are constantly getting cheaper

3. FPGA in mobile phones (6)

FPGA in Apple flagship phones

During teardown of iPhone 7, company Chipworks found a Lattice Semiconductors ICE5LP4K FPGA [28].

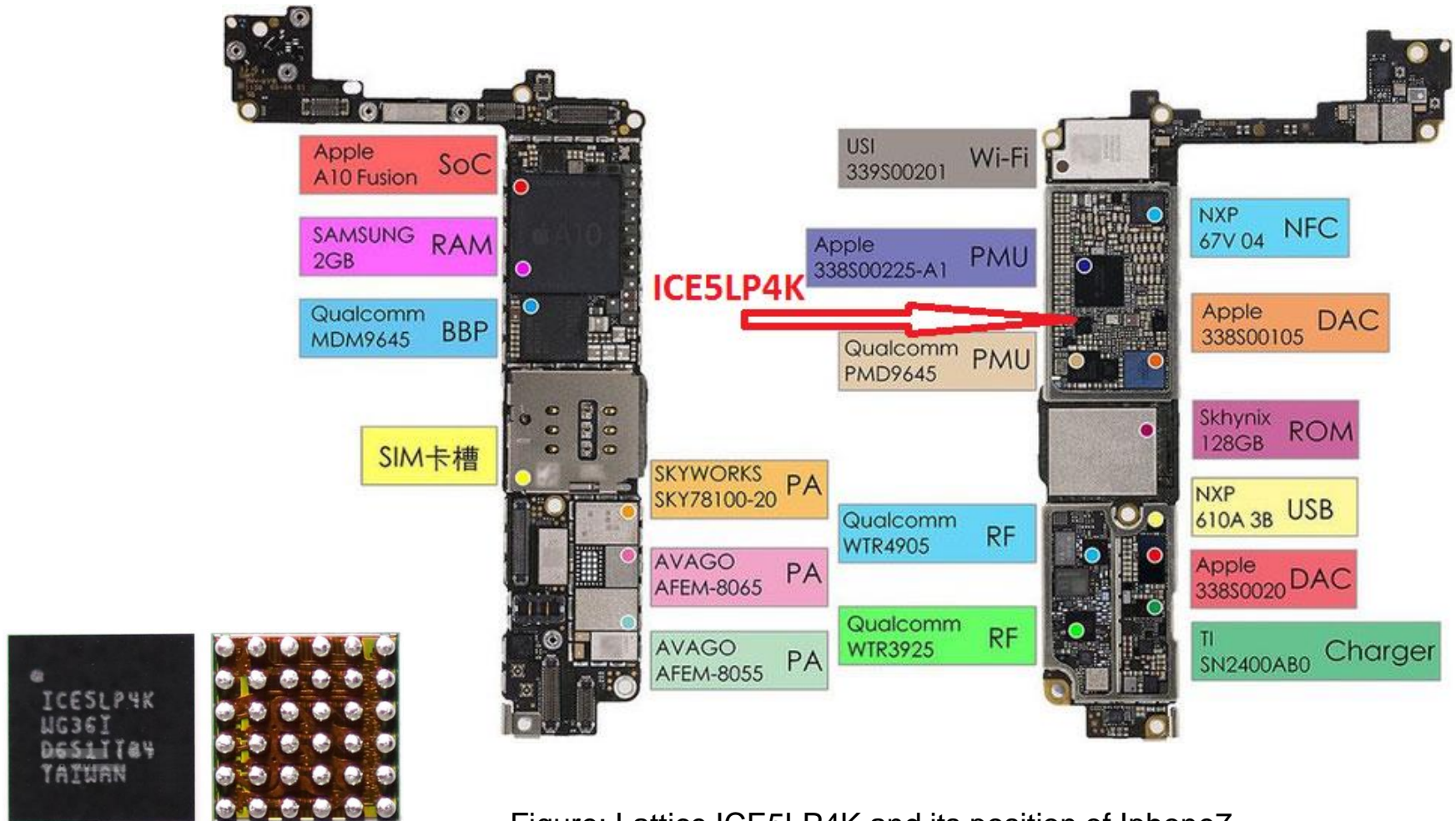


Figure: Lattice ICE5LP4K and its position of Iphone7

3. FPGA in mobile phones (7)

FPGA in Apple flagship phones

Apple **never officially commented usage** of the FPGA on iPhone 7 but there has been many **speculation** among the engineers.

Possible usages:

- encryption code **that cannot be accessed or modified** by malware from iOS and which can be updated in the future as needed
- “**glue logic**” between different systems on PCB
- machine learning algorithms

With each new iPhone, **Apple is embedding more and more artificial intelligence**. The advanced camera capabilities in the iPhone 7, for example, come from computer-vision algorithms running on a new image signal processor, an in-house Apple design. The company **is trying to differentiate from rivals** like Google, **which rely heavily on the cloud** (rather than the device) for **access to artificial intelligence**. Apple's argument is that **it's more secure and private** to do some of the **computing on the device** instead having to route every piece of data to the cloud.

3. FPGA in mobile phones (8)

Lattice ICE5LP4K capabilities

Even [without comments from Apple](#), it is [obvious from feature list](#) of ICE FPGA family that this technology [can be utilized](#) for different advanced algorithms:

- flexible logic architecture with **up to 3,520 4 input LUTs**, up to 26 I/O pins for customized interfaces and up to **80 Kbits** of embedded distributed memory
- ultra-low power advanced process with sleep current as **low as 35uA** and **1-10 mA** active current for most applications
- high performance signal processing using **DSP blocks** with multiply and accumulate functions
- hardened **SPI and I2C** blocks to interface to variety of [sensors and peripherals](#)

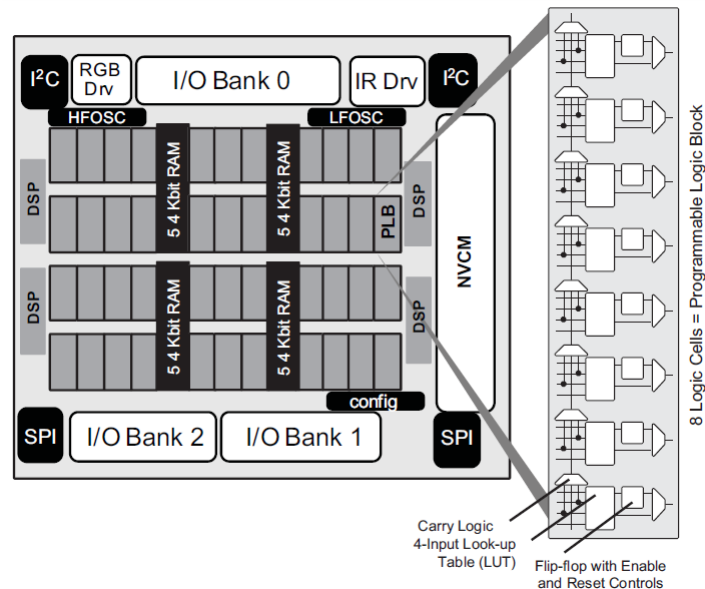


Figure: Lattice ICE5LP4K block schematic [29]

3. FPGA in mobile phones (9)

FPGA technology in Samsung phone

Samsung used Lattice's LP1K9D (Low-power FPGA) in Galaxy S5 phone.

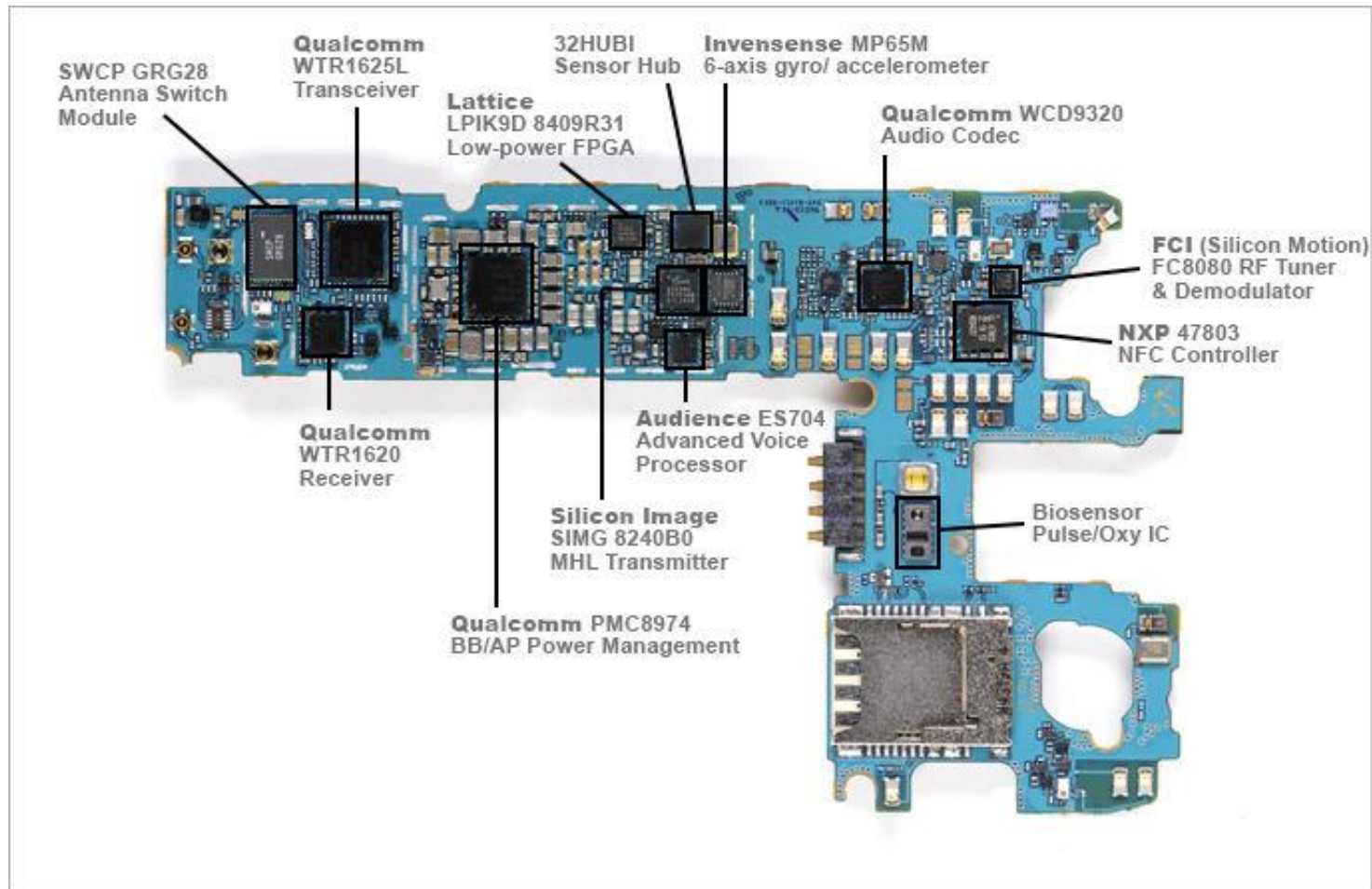


Figure: Samsung Galaxy S5 PCB with Lattice's LP1K9D [30]

3. FPGA in mobile phones (10)

FPGA technology future in mobile phones

FPGA **has minor role** in the phones from **major manufacturers** but open hardware and open source communities **are putting great efforts** to give FPGA **much more important role**. Because of many **concerns about privacy and security** regarding users data, open communities are trying to **develop completely open mobile phone**. On December 16, 2020 project Precursor managed to secure 186% of required manufacturing investment through open crowd funding [31].



Figure: Precursor mobile phone

3. FPGA in mobile phones (11)

Full specifications of Precursor

- FPGA - **Xilinx XC7S50** primary System on Chip (SoC) FPGA using -L1 speed grade for longer battery life; tested with 100 MHz VexRISC-V, RV32IMAC + MMU, 4k L1 I/D cache
Lattice Semi iCE40UP5K secondary Embedded Controller (EC) FPGA managing power, standby, and charging functions; tested with 18 MHz VexRISC-V, RV32I, no cache
- System Memory - **16 MB external SRAM**
- Storage - **128 MB flash**
- Display - **536 x 336 black & white LCD** with 200ppi, backlight
- Audio - 0.7 W notification speaker, vibration motor, 3.5 mm headset jack
- Connectivity - **802.11 b/g/n WiFi** via sandboxed Silicon Labs WF200C chipset for battery conservation
- USB - 1x **USB 2.0 Type-C** port for data and charging
- Expansion - Flex PCB breakout for **8x FPGA GPIO** via the battery compartment
- Security - **Dual hardware TRNG**
- Anti-tamper features - **User-sealable metal** can for trusted components, dedicated real-time clock (RTC) with basic clock integrity monitoring, power monitors trip reset in case of power glitches, always-on accelerometer/gyro to detect movement in standby, **support for instant secure erase via battery-backed AES key** and self-destruct circuit
- Battery - Replaceable 1,100 mAh Li-Ion battery giving ~100 hours standby with Wi-Fi + embedded controller + static display enabled, or 5.5 hours continuous use.

3. FPGA in mobile phones (12)

Another open source smartphone?

Precursor have some **things in common** with other open source devices, but it differs in **the decision to host the SoC on an FPGA**.

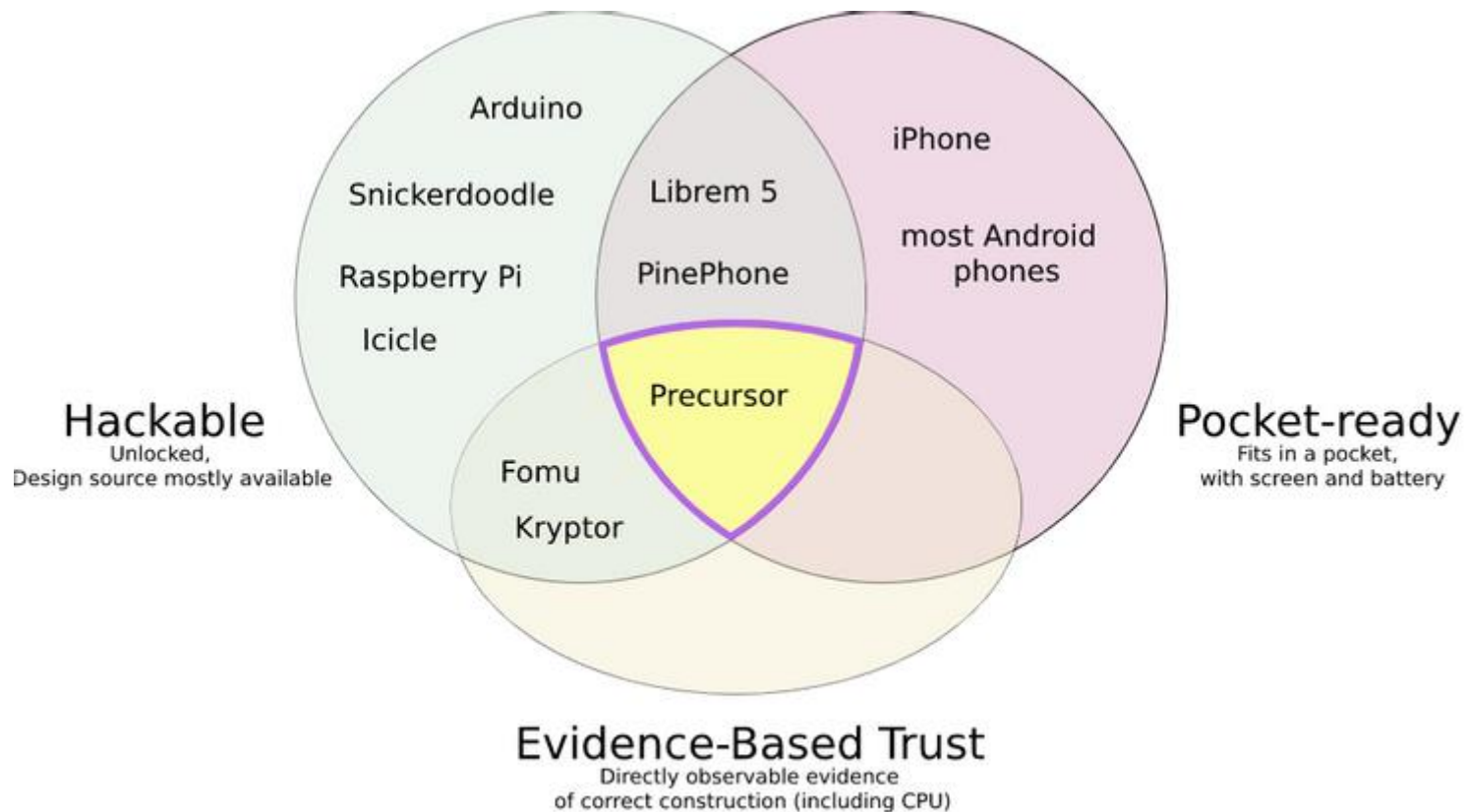


Figure: Venn diagram displaying relations between different devices

3. FPGA in mobile phones (13)

Why FPGA and not CPU?

A processor is essentially a tiny, complex circuit user can interact with using instruction based architecture. User have no control over what is actually inside it. User can give it calculations to perform using an instruction set provided by the manufacturer. When it comes to the security, user simply have to take the chip creators on their word when they say they are secure.

This frequently gets proved wrong, as it did in the case of the critical vulnerabilities found in AMD Ryzen chips some time ago [32].

FPGAs are integrated circuits that can be reconfigured using code. This might not sound all that different on the surface, but rather than giving the FPGA instructions like in the case of a regular processor, the circuitry itself is being configured.

This is where the "evidence-based trust" idea central to the Precursor project comes from. User will be able to know, down to the very last logic gate in the software CPU, how the data is being handled.

4. FPGA in high-performance computing and datacenters

4. FPGA in high-performance computing and datacenters (1)

Big companies using FPGA technology

Cloud based services are priority in today's IT sector. Big companies are using servers to provide services to their clients, and as time progresses need for processing power grows. In recent years some companies turned to FPGA technology in order to achieve higher performances through parallelism and reconfigurability.

Big companies that are using FPGA in their cloud services:

- Microsoft uses FPGA to accelerate services Bing, Azure and Office365
- Intel bought Altera and produces FPGA cards for integration with servers
- Amazon created Amazon EC2 F1 which uses Xilinx UltraScale+ VU9P FPGAs for accelerating computationally intensive algorithms

4. FPGA in high-performance computing and datacenters (2)

Microsoft's beginning with FPGA technology

In 2012 Doug Burger (computer chip researcher) had a meeting with Steve Ballmer (CEO of Microsoft) where he tried to pitch new idea about architecture transformation in processing data on Microsoft servers [33]. Burger wanted to create specialized chips that the company could reprogram for particular tasks and equip all of Microsoft's servers.

This idea didn't sit well with Ballmer. In Ballmer mind, Microsoft had spent 40 years building PC software like Windows, Word, and Excel. Internet was still a place where Microsoft was only beginning to gain footing and Ballmer didn't want to take risk with unknown technology. Luckily Qi Lu was present on the meeting, and he was in charge of running Bing, Microsoft search engine. He was interested in the idea and thus project Catapult was brought to life.

Burger's team managed to build successful prototype which convinced Qi Lu to give Burger the money to build and test over 1,600 servers equipped with FPGAs. After almost a year of developing the final test showed that Bing's "decision tree" machine-learning algorithms ran about 40 times faster with the new chips.

Bing dominated Microsoft's online ambitions until 2015 when the company got two other massive online services: the business productivity suite Office 365 and the cloud computing service Microsoft Azure. Microsoft executives realized that the only efficient way of running a growing online empire is to run all services on the same foundation.

If Project Catapult was going to transform Microsoft, it had to work inside Azure and Office 365 as well as with Bing.

4. FPGA in high-performance computing and datacenters (3)

Microsoft's beginning with FPGA technology

Integrating all three services on the same platform posed big challenge. The problem was that developed prototype was design to accelerate machine learning algorithm that was used by the Bing search engine.

The traffic bouncing around Azure's data centers was growing so fast, the service's CPUs couldn't keep pace. The Catapult could help with this too - but not the way it was designed for Bing. Azure and Office365 needed programmable chips right where each server connected to the primary network, so they could process all that traffic before it even got to the server.

This meant that Burger's team had to redesign FPGA solution one more time. It took them more than a year to achieve this, but Microsoft finally had a unified platform for all online services.

4. FPGA in high-performance computing and datacenters (4)

Microsoft's beginning with FPGA technology



Figure: Catapult team members Adrian Caulfield, Eric Chung, Doug Burger, and Andrew Putnam [34]

4. FPGA in high-performance computing and datacenters (5)

Microsoft's future in FPGA technology

After [immensely successful project Catapult](#), Burger's team continued with innovations and in 2017 they released new FPGA board, codenamed [Project Brainwave](#) [35].

The [Microsoft Brainwave mezzanine card](#) extends each server with an [Intel Altera Stratix 10 FPGA accelerator](#), synthesized to act as a “Soft DNN Processing Unit,” or DPU, and a fabric interconnect that enables datacenter-scale persistent neural networks.



Figure: The Microsoft Brainwave mezzanine card [36]

4. FPGA in high-performance computing and datacenters (6)

Microsoft's future in FPGA technology

Project Brainwave is a **scalable acceleration platform for deep learning**, which can provide real time responses for **cloud-based AI services**.

Microsoft's Project Brainwave consists of **three components**:

- a **high-performance systems architecture** that pools accelerators for datacenter-wide services and scale - by linking their accelerators across a high bandwidth, low-latency fabric, Microsoft can **dynamically allocate these resources to optimize their utilization while keeping latencies very low**
- a **“soft” DNN (deep neural network) processor** (DPU – DNN processing unit) that is programmed, or synthesized, on 14nm class Altera FPGAs
- a **compiler and run-time environment** to support efficient deployment of trained neural network models using CNTK (Cognitive Toolkit) - Microsoft's DNN platform.

4. FPGA in high-performance computing and datacenters (7)

Microsoft's future in FPGA technology

A great example of the **advantage of FPGAs in machine learning** is the ability to **customize the level of precision** required for a particular layer in a deep neural network.

Microsoft's DPU can be programmed to process calculations for virtually **any precision required** by the neural network being used, **delivering excellent performance**. Also, Microsoft can reprogram (synthesize) these chips in a matter of weeks for a different use case.

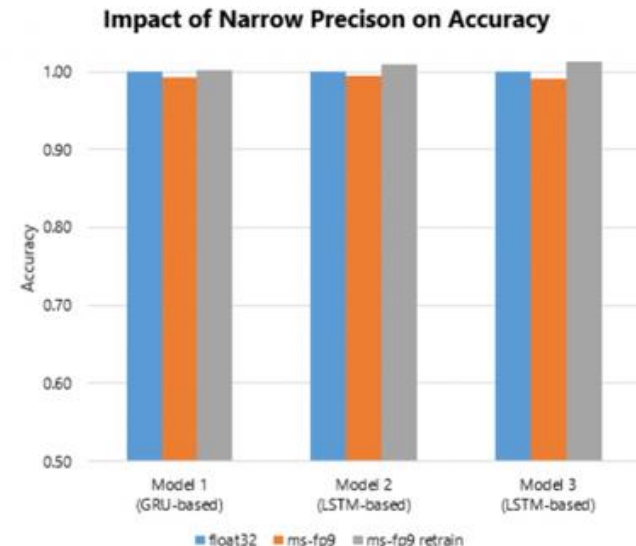
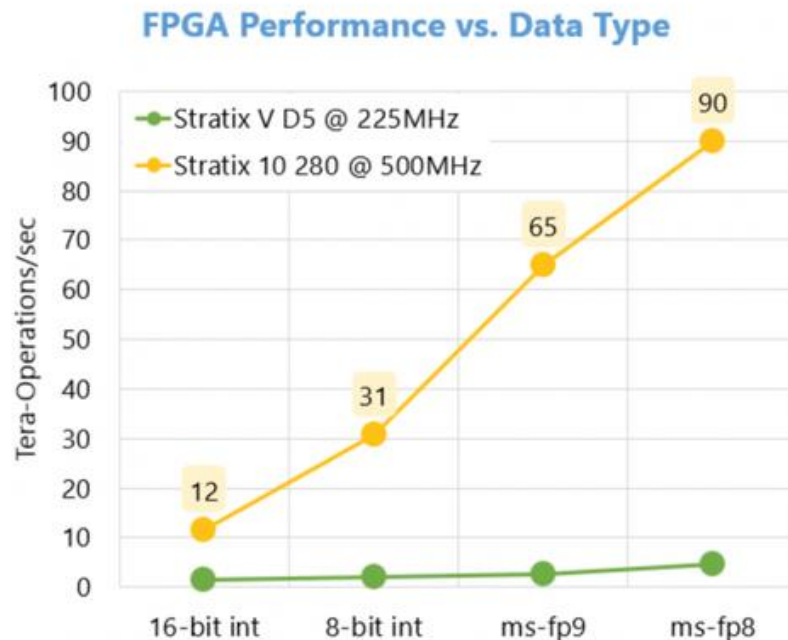


Figure: Comparison between performance and precision on Brainwave card [37]

4. FPGA in high-performance computing and datacenters (8)

Intel's expansion to FPGA market

For decades Xilinx and Altera **were in constant battle for supremacy** in the FPGA acceleration game and everybody were convinced that the **next-generation Stratix battling Virtex will decide the winner**.

Then, in 2015, **Intel bought the Altera** and everything changed.

While engineers were still **debating LUTs, SerDes, Moore's law etc**, **Intel packed Arria 10 chips onto PCIe cards** and placed them in **Xeon-powered servers**.

In 2018 Intel announced that **top-tier OEMs including Dell and Fujitsu are rolling out servers pre-equipped with Intel programmable acceleration cards (PACs) containing Arria 10 GX FPGAs** [38].

That meant servers which were **available for volume shipment with ready-to-run**, results-proven acceleration for key workloads including financial risk analytics and database acceleration. **Results on both applications were beyond compelling**. On financial risk analysis, there's an **850% per-symbol algorithm speedup** and a greater than **2x simulation time speedup** compared with traditional "Spark" implementation. On database acceleration, Intel claims 20X+ faster real-time data analytics, 2x+ traditional data warehousing, and 3x+ storage compression.

4. FPGA in high-performance computing and datacenters (9)

Intel's expansion to FPGA market

In 2017, Intel has announced its first full-fledged FPGA card for accelerating datacenter workloads. Known as the Intel Programmable Acceleration Card, or PAC for short, the device was powered by the Arria 10 GX-1150 FPGA [39].

PAC used PCIe Gen3 to hook into the server and was equipped with 8 GB of DDR4 memory, along with 128 MB of flash. As a half-length, half-height card it could fit unobtrusively into most standard servers and consumed just 60 watts of power.



Figure: Intel Arria 10 GX-1150 FPGA card [39]

Intel's expansion to FPGA market

The Arria 10 GX-1150 FPGA is aimed at a data-intensive applications:

- deep learning inference
- database acceleration
- video transcoding
- financial analytics
- HPC applications in genomics
- oil & gas simulations

To help developers ease the hardships of programming, managing, and running these reconfigurable chips there is Intel's Open Programmable Acceleration Engine (OPAE) technology, which is a programming interface and toolset for Altera silicon. OPAE also includes useful things like simulators, support for virtualization, code samples, and command-line utilities.

4. FPGA in high-performance computing and datacenters (11)

Intel's expansion to FPGA market

In February, 2019 Intel presented **FPGA Programmable Acceleration Card N3000** for Networking. This is a **PCIe 3.0 x16 card** with both networking and an Intel Arria 10 FPGA onboard. The dual Intel XL710 NICs support **either 8x 10GbE or 4x 25GbE networking** per Intel's spec. Card is **targeting the 5G infrastructure market**.



Figure: Intel FPGA PAC N3000 [40]

4. FPGA in high-performance computing and datacenters (12)

Intel's expansion to FPGA market

In August, 2019 Intel released [FPGA PAC D5005 High-end Drop-in Accelerator](#) with a Stratix 10 FPGA onboard. The card overall is a full height, 3/4 length double-slot, PCIe Gen3 x16 device carrying a 215W TDP. It, therefore, [needs additional power via the server](#).



Figure: Intel FPGA PAC D5005 On HPE ProLiant DL380 Gen10 [41]

4. FPGA in high-performance computing and datacenters (13)

Intel's expansion to FPGA market

On the devices are **two QSFP28 transceivers and 32GB of DDR4 ECC memory**. The **Intel MAX 10 FPGA** BMC supports IPMI 2.0 and out-of-band management. At the center of the design is an **Intel SX280HN2F43E2VG FPGA**.

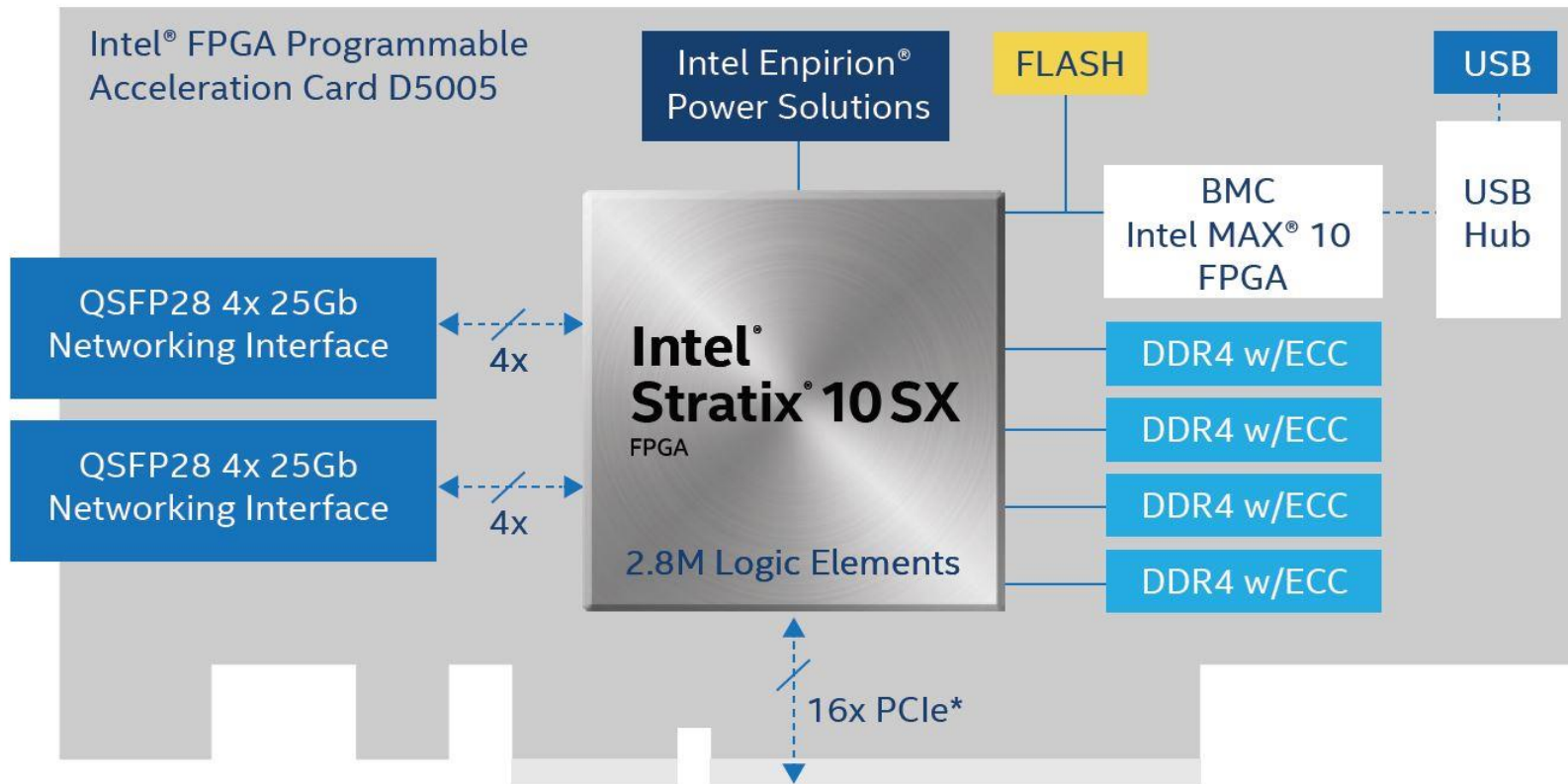


Figure: Intel FPGA PAC D5005 Diagram [41]

4. FPGA in high-performance computing and datacenters (14)

Amazon services with FPGA technology

Microsoft and Intel based their platforms on Altera FPGA chips, but Xilinx found a partner in Amazon. Amazon has different platforms that are in use for cloud computing through Amazon Web Services (AWS). Platform that uses FPGA chips is called Amazon Elastic Compute Cloud (EC2) F1.

This platform is suitable for:

- genomics research
- search/analytics
- image and video processing
- network security
- electronic design automation (EDA)
- image and file compression
- big data analytics

AWS has recognized necessity of incorporating younger generations so Amazon founded AWS Educate which is available to accredited educational institutions, professors, and students free of charge to access the cloud computing services. Professors, teaching assistants, educators, and students receive free access to AWS technology, open source content for their courses, training resources, demos, special on-campus programs and a community of cloud evangelists.

4. FPGA in high-performance computing and datacenters (15)

Amazon services with FPGA technology

Amazon EC2 F1 specifications:

- High frequency [Intel Xeon E5-2686 v4](#) (Broadwell) processors
- NVMe SSD Storage
- Xilinx [Virtex UltraScale+ VU9P](#) FPGAs
- 64 GiB of ECC-protected memory on 4x DDR4
- Dedicated PCI-Express x16 interface
- Approximately [2.5 million logic elements](#)
- Approximately [6,800 Digital Signal Processing](#) (DSP) engines

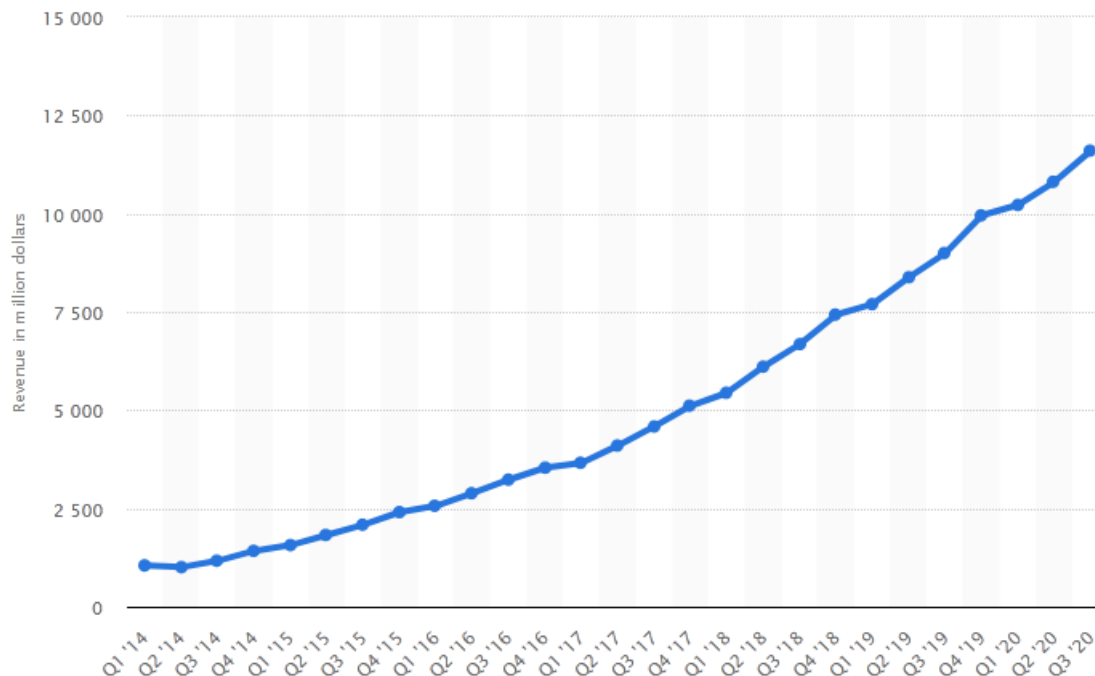


Figure: Quarterly revenue of Amazon Web Services from 1st quarter 2014 to 3rd quarter 2020 [42]

4. FPGA in high-performance computing and datacenters (16)

Amazon services with FPGA technology



Figure: XUPP3R PCIe Accelerator Board based on the Xilinx Virtex UltraScale+ VU9P FPGA [43]

4. FPGA in high-performance computing and datacenters (17)

Amazon services with FPGA technology

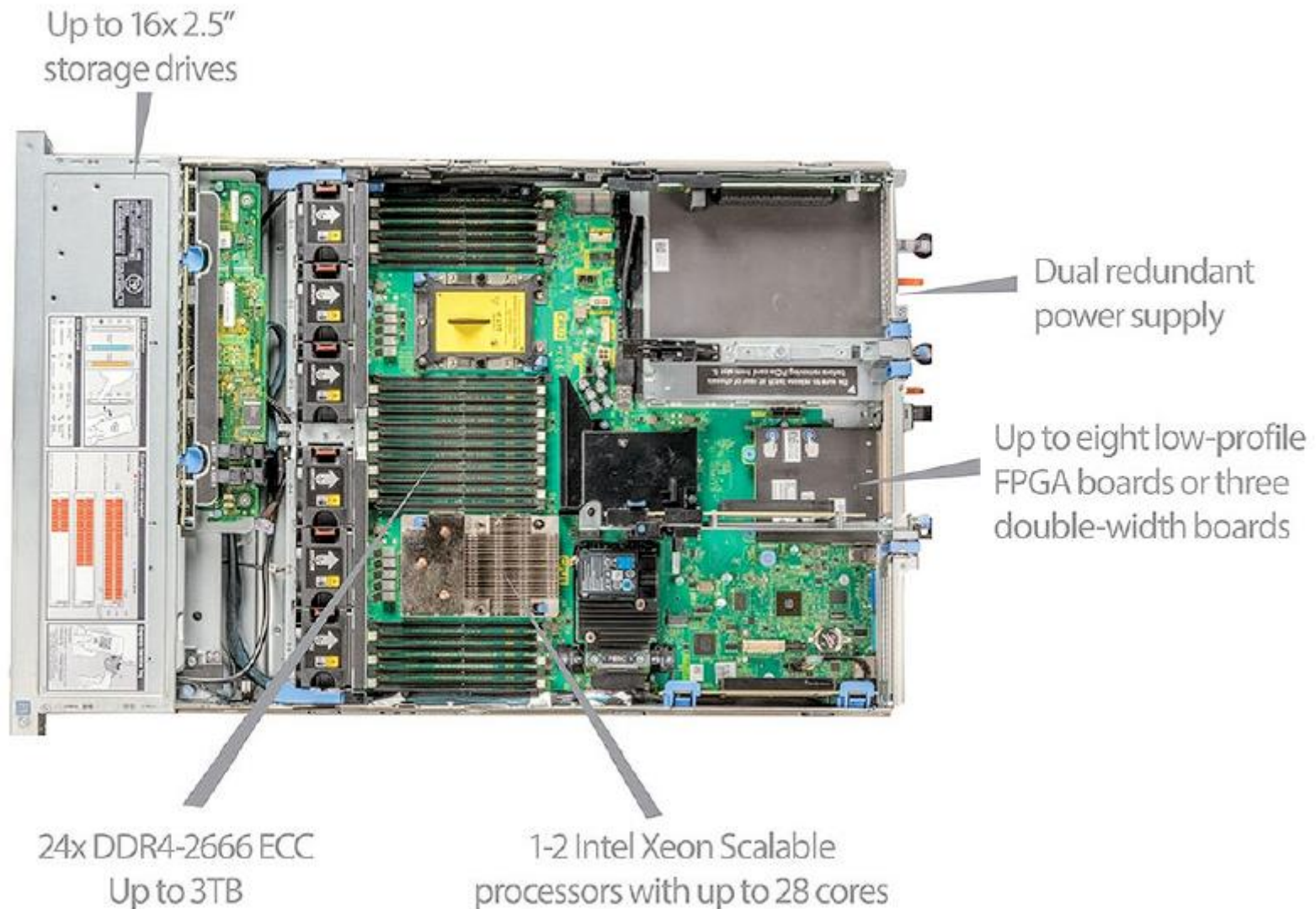


Figure: TeraBox 2000D 2U FPGA chassis combines one or two Intel Xeon processors with as many as eight of FPGA-based PCIe cards [43]

5. Conclusions

5. Conclusions

FPGA technology traveled a great length from emerging newcomer who was struggling in the shadow of much more popular CPU technology, all the way to the extremely viable candidate to replace CPU technology in datacenter servers.

It is obvious that FPGA technology has not yet reached its full potential from the fact that all major FPGA manufacturers have been acquired by the biggest companies on the market (Intel acquired Altera and AMD has announced plans to acquire Xilinx).

Since FPGA technology entered data server market, maybe in near future it will be common to have FPGA chips in desktop computers.

6. References

6. References

- [1]: Chip Hall of Fame: Xilinx XC2064 FPGA
<https://spectrum.ieee.org/tech-history/silicon-revolution/chip-hall-of-fame-xilinx-xc2064-fpga>
- [2]: Field Programmable Gate Array Market Size, Share & Trends Analysis Report By Technology (SRAM, Antifuse, Flash), By Application (Military & Aerospace, Telecom), By Region, And Segment Forecasts, 2020 – 2027
<https://www.grandviewresearch.com/industry-analysis/fpga-market>
- [3]: Xilinx Says Its New FPGA is World's Largest
<https://www.enterpriseai.news/2019/08/21/xilinx-says-its-new-fpga-is-worlds-largest/>
- [4]: S. M. Trimberger, "Three Ages of FPGAs: A Retrospective on the First Thirty Years of FPGA Technology," in *Proceedings of the IEEE*, vol. 103, no. 3, pp. 318-331, March 2015, doi: 10.1109/JPROC.2015.2392104
- [5]: S. M. Trimberger, "Field Programmable Gate Array Technology", Springer /b S Publication (1 January 2008), ISBN-13: 978-8181286031
- [6]: Nemeth, T., Stefani, J., Liu, W., Dimond, R., Pell, O., Ergas, R. An implementation of the acoustic wave equation. In *Proceedings of the 78th Society of Exploration Geophysicists Meeting*, (Las Vegas, NV, 2008)
- [7]: List of Xilinx FPGAs with and without integrated CPU
https://en.wikipedia.org/wiki/List_of_Xilinx_FPGAs

6. References

- [8]: Altera Parts History, compiled by John Lazzaro
<https://www-inst.eecs.berkeley.edu/~cs294-59/fa10/resources/Altera-history/Altera-history.html>
- [9]: Basys 3 board based on the Artix-7™ Field Programmable Gate Array from Xilinx
<https://reference.digilentinc.com/reference/programmable-logic/basys-3/start>
- [10]: ZedBoard Xilinx Zynq-7000 complete development kit for designers
<http://zedboard.org/product/zedboard>
- [11]: Doug Black, "Xilinx vs. Intel: FPGA Market Leaders Launch Server Accelerator Cards", August 6, 2019
<https://www.hpcwire.com/2019/08/06/xilinx-vs-intel-fpga-market-leaders-launch-server-accelerator-cards/>
- [12]: Boole's expansion theorem
https://en.wikipedia.org/wiki/Boole%27s_expansion_theorem
- [13]: R.C. Cofer, Benjamin F. Harding, in Rapid System Prototyping with FPGAs, 2006
<https://www.sciencedirect.com/topics/computer-science/simple-programmable-logic-device>
- [14]: Spartan-6 Libraries Guide for HDL Designs, UG615(v14.7) October2,2013
www.xilinx.com
- [15]: H. Mehri and B. Alizadeh, "An analytical dynamic and leakage power model for FPGAs," *2014 22nd Iranian Conference on Electrical Engineering (ICEE)*, Tehran, 2014, pp. 300-305, doi: 10.1109/IranianCEE.2014.6999552

6. References

- [16]: Spartan-6 FPGA Configurable Logic Block User Guide, UG384 (v1.1) February 23, 2010
www.xilinx.com
- [17]: EET 3350 Digital Systems Design Textbook: John Wakerly Chapter 9: 9.6 FPGAs
- [18]: Patrick Francis, "CSET 4650 Field Programmable Logic Devices"
- [19]: Clive MaxMaxfield, "FPGA Architectures"
<https://doi.org/10.1016/B978-0-7506-8974-8.00002-8>
- [20]: Laszlo Bako, Szabolcs Hajdu, Fearghal Morgan, "Evaluation and Comparison of Low FPGA Footprint, Embedded Soft-Core Processors", MACRo 2015,
DOI: 10.1515/macro-2017-0003
- [21]: Bikram Adhikari, Deepak Gurung, Giresch Singh Kunwar, Prashanta Gyawali,
"FPGA Control of a Mobile Inverted Pendulum Robot", Journal of the Institute of
Engineering 2011, DOI: 10.3126/jie.v8i1-2.5111
- [22]: Mouna Baklouti, Mohamed Abid, "Multi-Softcore Architecture on FPGA",
International Journal of Reconfigurable Computing, vol. 2014, Article ID 979327,
13 pages, 2014. <https://doi.org/10.1155/2014/979327>
- [23]: Louise Crockett, Ross Elliot, Ross Elliot, Robert Stewart, "The Zynq Book: Embedded
Processing with the Arm Cortex-A9 on the Xilinx Zynq-7000 All Programmable Soc",
Strathclyde Academic Media 2014, ISBN-13: 978-0992978709
- [24]: Brianne Miles, "ECE 699: Lecture 2 Introduction to Zynq"
<https://slideplayer.com/slide/10437954/>

6. References

- [25]: Xcell Journal, no. 88, Q3 2014
www.xilinx.com
- [26]: Timeline of iPhone models
https://en.wikipedia.org/wiki/Template:Timeline_of_iPhone_models
- [27]: Samsung Galaxy S Series Timeline
<https://www.officetimeline.com/blog/samsung-galaxy-s-series-timeline>
- [28]: John Martellaro, "Thoughts About Apple's Secret iPhone 7 Chip"
<https://www.macobserver.com/columns-opinions/editorial/apple-secret-iphone-7-chip/>
- [29]: iCE40 Ultra Family Data Sheet, DS1048 Version 1.8, June 2015
<https://www.latticesemi.com/Products/FPGAandCPLD/iCE40Ultra>
- [30]: M. Alarcon, R. Fontaine, D. James, R. Krishnamurthy, J. Morrison, D. Yang and C. Young,
"Samsung Galaxy S5 Teardown", April 11, 2014
<https://www.techinsights.com/blog/samsung-galaxy-s5-teardown>
- [31]: Sutajio Kosagi - Precursor, Mobile, Open Hardware, RISC-V System-on-Chip (SoC)
Development Kit
<https://www.crowdsupply.com/sutajio-kosagi/precursor>
- [32]: Gavin Phillips, "The New AMD Ryzen Vulnerabilities Are Real: What You Need to Know",
Mar 19, 2018
<https://www.makeuseof.com/tag/amd-cpu-vulnerability/>

6. References

- [33]: Robert McMillan, "Microsoft Supercharges Bing Search With Programmable Chips", 2014
<https://www.wired.com/2014/06/microsoft-fpga/>
- [34]: Cade Metz, "Microsoft Bets Its Future on a Reprogrammable Computer Chip", 2016
<https://www.wired.com/2016/09/microsoft-bets-future-chip-reprogram-fly/>
- [35]: Doug Burger, "Microsoft unveils Project Brainwave for real-time AI", 2017
<https://www.microsoft.com/en-us/research/blog/microsoft-unveils-project-brainwave/>
- [36]: Allison Linn, "Real-time AI: Microsoft announces preview of Project Brainwave", 2018
<https://blogs.microsoft.com/ai/build-2018-project-brainwave/>
- [37]: Karl Freund, "Microsoft: FPGA Wins Versus Google TPUs For AI", 2017
<https://www.forbes.com/sites/moorinsights/2017/08/28/microsoft-fpga-wins-versus-google-tpus-for-ai/?sh=7a77015f3904>
- [38]: Kevin Morris, "Accelerating Mainstream Servers with FPGAs Intel Puts FPGAs Inside", Electronic Engineering journal, 2018
<https://www.eejournal.com/article/accelerating-mainstream-servers-with-fpgas/>
- [39]: Michael Feldman, "Intel Unveils Programmable Acceleration Card for Servers", 2017
<https://www.top500.org/news/intel-unveils-programable-acceleration-card-for-servers/>
- [40]: Cliff Robinson, "Intel FPGA Programmable Acceleration Card N3000 for Networking", 2019
<https://www.servethehome.com/intel-fpga-programmable-acceleration-card-n3000-networking/>

6. References

- [41]: Patrick Kennedy, "Intel FPGA PAC D5005 High-end Drop-in Accelerator Launched", 2019
<https://www.servethehome.com/intel-fpga-pac-d5005-high-end-drop-in-accelerator-launched/>
- [42]: Quarterly revenue of Amazon Web Services from 1st quarter 2014 to 3rd quarter 2020,
Published by Statista Research Department, Jan 12, 2021
<https://www.statista.com/statistics/250520/forecast-of-amazon-web-services-revenue/>
- [43]: Ron Huizen, "FPGAs in the Cloud: Should you Rent or Buy FPGAs for Development and Deployment?", BittWare Whitepaper, 2015
https://www.bittware.com/wp-content/uploads/sites/4/2015/03/BMC_whitepaper.pdf