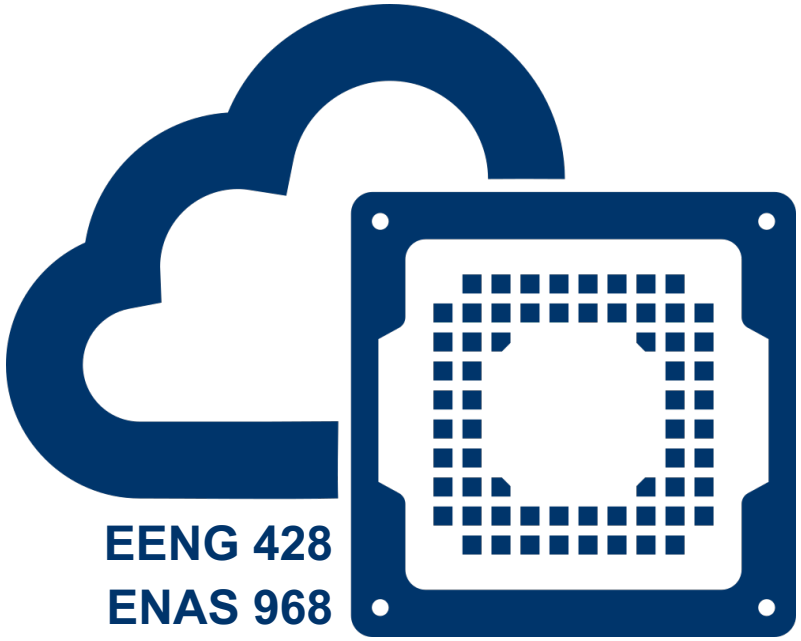
Cloud FPGA



bit.ly/cloudfpga

Lecture: Xilinx FPGAs in Cloud FPGAs

Prof. Jakub Szefer
Dept. of Electrical Engineering, Yale University

EENG 428 / ENAS 968
Cloud FPGA

FPGA Chips Used in Cloud FPGAs

- Different Cloud FPGA providers make various FPGAs available for remote access
  - Major FPGA vendors are: **Xilinx** and **Intel** (Altera)
  - Typically, Cloud FPGAs only provide one type of FPGA board available
  - Directly develop designs in an HDL or use High-Level Synthesis
  - Server-to-FPGA communication via PCIe
  - "Shell" with cloud providers modules, plus interfaces (AXI typically) to user's logic

Nimbix
Alveo

OVH
Arria 10 GX 1150

Baidu Cloud
Kintex UltraScale KU115

Alibaba F2 Cloud
Kintex UltraScale KU115

Huawei Cloud
Virtex UltraScale+ VU9P and Alveo

Alibaba F1 Cloud
Arria 10 GX 1150

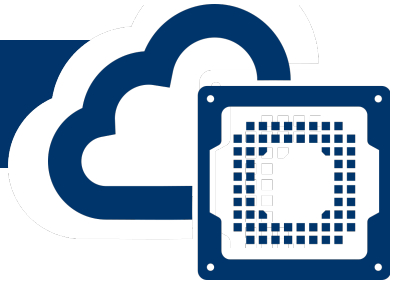
Amazon AWS
Virtex UltraScale+ VU9P

Microsoft Azure
Intel (Altera) Stratix V

Tencent Cloud
Kintex UltraScale KU115

TACC
Intel (Altera) Stratix V

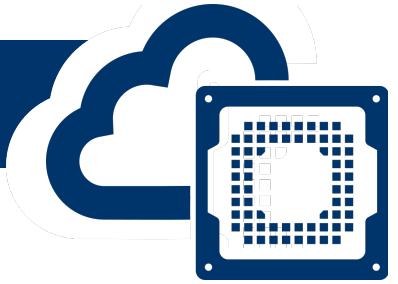
Xilinx FPGA Families



Like most vendors, Xilinx offers a number of FPGA chips targeting different use-cases

- Xilinx 7 Series (28nm):
  - **Spartan-7** – low-end FPGAs, typically for embedded devices
  - **Artix-7** – low-end FPGAs, but with faster I/O
  - **Kintex-7** – low-power FPGAs
  - **Virtex-7** – "baseline" FPGA offering
- UltraScale (20nm):
  - **Kintex UltraScale** and **Virtex UltraScale** – shrink node size, and include super logic regions (SLRs, i.e. multiple FPGA dies per chip)
- UltraScale+ (16nm):
  - **Kintex UltraScale+** and **Virtex UltraScale+** – further shrink node size,
- Other
  - **Zynq SoCs** – include embedded Arm processor
  - **Alveo** – data center accelerator cards, Virtex UltraScale+ with DRAM and other features

Metrics for Comparing FPGAs



Main metrics to consider when comparing FPGAs are:

- **Number of Look-Up Tables** (LUTs)
- **Size of LUTs**, number of inputs
- **Number of Flip-Flops**
- **Size of Block RAM storage**

In Xilinx, a Configurable Logic Block (CLB) typically consists of two slices, each of which contains multiple LUTs and FFs

- **Logic Cells** – a meta-metric for trying to compare among different FPGA, accounting for different LUT sizes and LUT, Flip-Flop configurations

- **Speed grade**
- **I/O resources**
- **Digital Signal Processing (DSP) blocks**
- **Other embedded Hard IP blocks or CPUs**

Higher # is faster, e.g., -1 is slowest and -3 is fastest

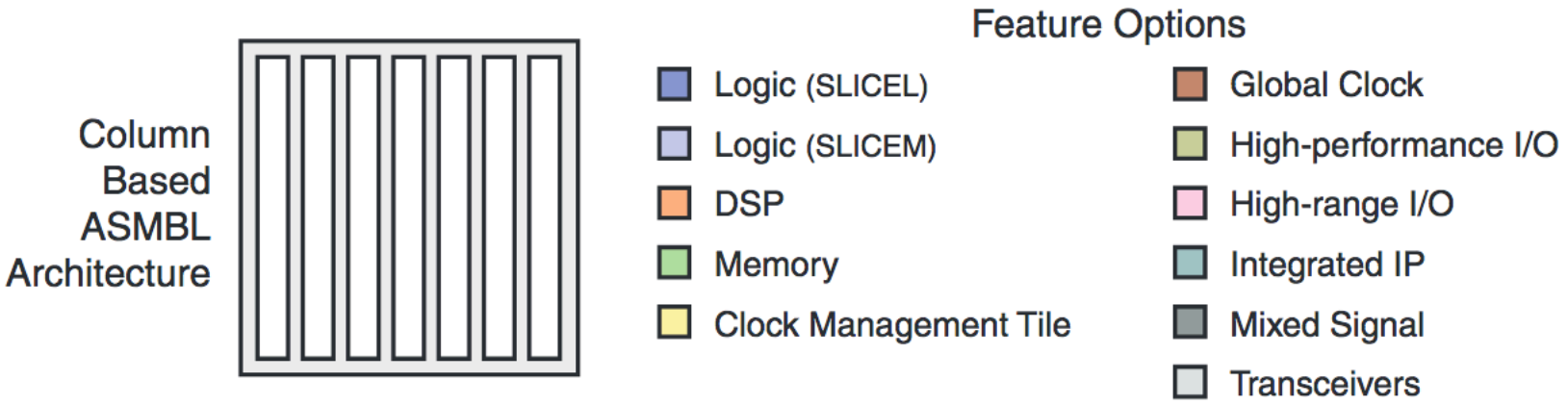
"Hard IP" are digital logic modules realized in the transistors, inside FPGA chip, they can't be re-configured

"Soft IP" are digital logic modules that are realized in the FPGA fabric, often provided by the FPGA vendor

Share: bit.ly/cloudfpga

Xilinx Configurable Logic Block

- Xilinx created the Advanced Silicon Modular Block (ASMBL) architecture instead of the textbook-style matrix-like arrangement of logic blocks
- AMBL is a column-style arrangement:



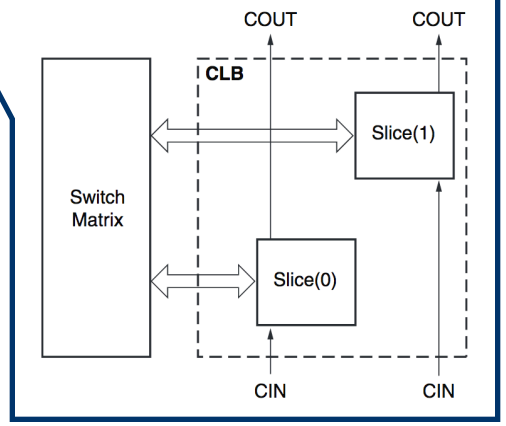
- **`SLICEL`** and **`SLICEM`** are part of Configurable Logic Block (CLBs)
  - **`SLICEL`** has only logic
  - **`SLICEM`** has logic plus it is able to store data using distributed RAM and shift data with 32-bit registers

CLB Structure

- A CLB element contains a pair of slices, and each slice is composed of four 6-input LUTs and eight flip-flops
- Every slice contains:
  - Four LUTs
  - Eight Flip-Flops
  - Carry chains
  - (Shifters)
  - (Distributed RAM)

Either a 6-input LUT with one output, or as two 5-input LUTs with separate outputs

Carry chains bypass the switch matrices



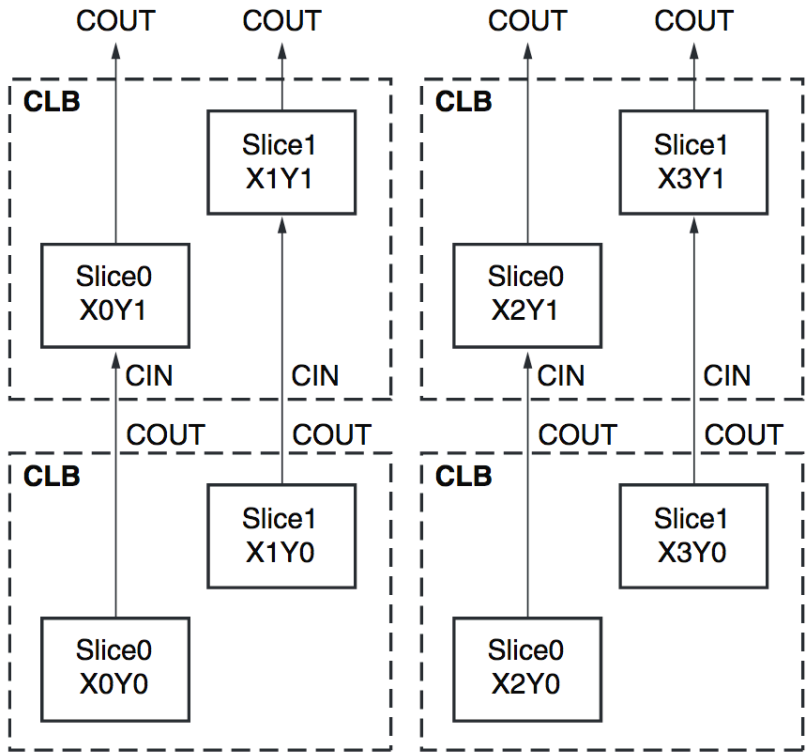
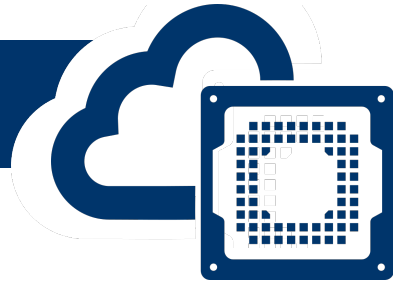


Image from [6]

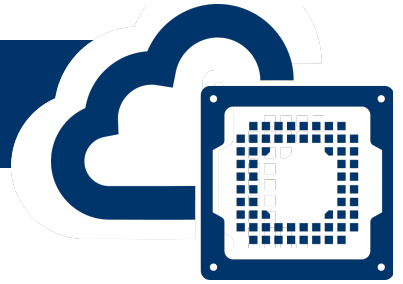
Spartan-7 FPGAs



- Spartan-7 are the lowest-end FPGAs in 7 Series
  - Roughly half the Configurable Logic Blocs (CLBs) compared to Artix-7
  - Block RAM in 100s of Kbits

	Part Number	XC7S6	XC7S15	XC7S25	XC7S50	XC7S75	XC7S100
Logic Resources	Logic Cells	6,000	12,800	23,360	52,160	76,800	102,400
	Slices	938	2,000	3,650	8,150	12,000	16,000
	CLB Flip-Flops	7,500	16,000	29,200	65,200	96,000	128,000
Memory Resources	Max. Distributed RAM (Kb)	70	150	313	600	832	1,100
	Block RAM/FIFO w/ ECC (36 Kb each)	5	10	45	75	90	120
	Total Block RAM (Kb)	180	360	1,620	2,700	3,240	4,320
Clock Resources	Clock Mgmt Tiles (1 MMCM + 1 PLL)	2	2	3	5	8	8
I/O Resources	Max. Single-Ended I/O Pins	100	100	150	250	400	400
	Max. Differential I/O Pairs	48	48	72	120	192	192
Embedded Hard IP Resources	DSP Slices	10	20	80	120	140	160
	Analog Mixed Signal (AMS) / XADC	0	0	1	1	1	1
	Configuration AES / HMAC Blocks	0	0	1	1	1	1
Speed Grades	Commercial Temp (C)	-1,-2	-1,-2	-1,-2	-1,-2	-1,-2	-1,-2
	Industrial Temp (I)	-1,-2,-1L	-1,-2,-1L	-1,-2,-1L	-1,-2,-1L	-1,-2,-1L	-1,-2,-1L
	Expanded Temp (Q)	-1	-1	-1	-1	-1	-1

Artix-7 FPGAs



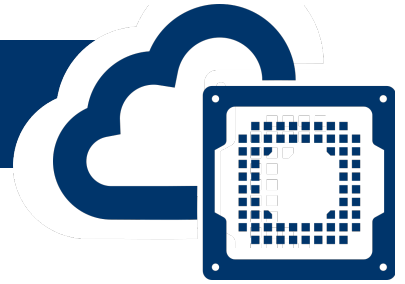
- Artix-7 are the low end FPGAs, but with better I/O and performance
  - Roughly half to quarter fewer CLBs than Kintex-7
  - Block RAM in 1~10 Mbits

	Part Number	XC7A12T	XC7A15T	XC7A25T	XC7A35T	XC7A50T	XC7A75T	XC7A100T	XC7A200T
Logic Resources	Logic Cells	12,800	16,640	23,360	33,280	52,160	75,520	101,440	215,360
	Slices	2,000	2,600	3,650	5,200	8,150	11,800	15,850	33,650
	CLB Flip-Flops	16,000	20,800	29,200	41,600	65,200	94,400	126,800	269,200
Memory Resources	Maximum Distributed RAM (Kb)	171	200	313	400	600	892	1,188	2,888
	Block RAM/FIFO w/ ECC (36 Kb each)	20	25	45	50	75	105	135	365
	Total Block RAM (Kb)	720	900	1,620	1,800	2,700	3,780	4,860	13,140
Clock Resources	CMTs (1 MMCM + 1 PLL)	3	5	3	5	5	6	6	10
I/O Resources	Maximum Single-Ended I/O	150	250	150	250	250	300	300	500
	Maximum Differential I/O Pairs	72	120	72	120	120	144	144	240
Embedded Hard IP Resources	DSP Slices	40	45	80	90	120	180	240	740
	PCIe® Gen2[1]	1	1	1	1	1	1	1	1
	Analog Mixed Signal (AMS) / XADC	1	1	1	1	1	1	1	1
	Configuration AES / HMAC Blocks	1	1	1	1	1	1	1	1
	GTP Transceivers (6.6 Gb/s Max Rate)[2]	2	4	4	4	4	8	8	16
Speed Grades	Commercial Temp (C)	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2
	Extended Temp (E)	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3
	Industrial Temp (I)	-1, -2, -1L	-1, -2, -1L	-1, -2, -1L	-1, -2, -1L	-1, -2, -1L	-1, -2, -1L	-1, -2, -1L	-1, -2, -1L

Share: bit.ly/cloudfpga

Xilinx FPGAs details from [4]

Kintex-7 FPGAs



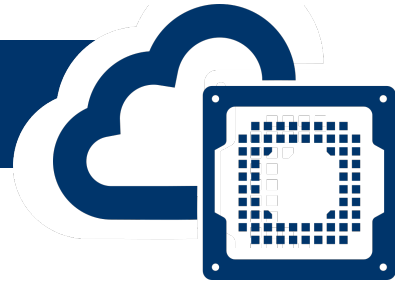
- Kintex-7 are mainly aiming for low-power at high performance
  - Less CLBs compared to Virtex-7
  - Block RAM in 5~30 Mbits

	Part Number	XC7K70T	XC7K160T	XC7K325T	XC7K355T	XC7K410T	XC7K420T	XC7K480T
	EasyPath™ Cost Reduction Solutions[1]	—	—	XCE7K325T	XCE7K355T	XCE7K410T	XCE7K420T	XCE7K480T
Logic Resources	Slices	10,250	25,350	50,950	55,650	63,550	65,150	74,650
	Logic Cells	65,600	162,240	326,080	356,160	406,720	416,960	477,760
	CLB Flip-Flops	82,000	202,800	407,600	445,200	508,400	521,200	597,200
Memory Resources	Maximum Distributed RAM (Kb)	838	2,188	4,000	5,088	5,663	5,938	6,788
	Block RAM/FIFO w/ ECC (36 Kb each)	135	325	445	715	795	835	955
	Total Block RAM (Kb)	4,860	11,700	16,020	25,740	28,620	30,060	34,380
Clock Resources	CMTs (1 MMCM + 1 PLL)	6	8	10	6	10	8	8
I/O Resources	Maximum Single-Ended I/O	300	400	500	300	500	400	400
	Maximum Differential I/O Pairs	144	192	240	144	240	192	192
Integrated IP Resources	DSP48 Slices	240	600	840	1,440	1,540	1,680	1,920
	PCIe® Gen2[2]	1	1	1	1	1	1	1
	Analog Mixed Signal (AMS) / XADC	1	1	1	1	1	1	1
	Configuration AES / HMAC Blocks	1	1	1	1	1	1	1
	GTX Transceivers (12.5 Gb/s Max Rate)	8	8	16	24	16	32	32
Speed Grades	Commercial Temp (C)	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2
	Extended Temp (E)	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3
	Industrial Temp (I)	-1, -2	-1, -2, -2L	-1, -2, -2L	-1, -2, -2L	-1, -2, -2L	-1, -2, -2L	-1, -2, -2L

Share: bit.ly/cloudfpga

Xilinx FPGAs details from [5]

Virtex-7 FPGAs



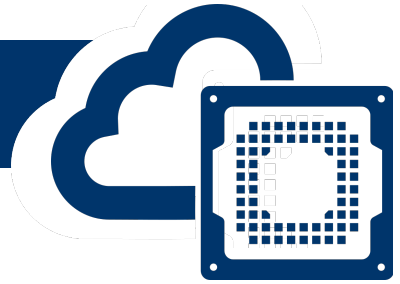
- Virtex-7 are the “baseline” chips with very good performance
  - Roughly 100K slices, 1M flip-flops, 20~50 Mbit Block RAM

	Part Number	XC7V585T	XC7V2000T	XC7VX330T	XC7VX415T	XC7VX485T	XC7VX550T	XC7VX690T	XC7VX980T	XC7VX1140T	XC7VH580T	XC7VH870T
	EasyPath™ Cost Reduction Solutions[1]	XCE7V585T	—	XCE7VX330T	XCE7VX415T	XCE7VX485T	XCE7VX550T	XCE7VX690T	XCE7VX980T	—	—	—
Logic Resources	Slices	91,050	305,400	51,000	64,400	75,900	86,600	108,300	153,000	178,000	90,700	136,900
	Logic Cells	582,720	1,954,560	326,400	412,160	485,760	554,240	693,120	979,200	1,139,200	580,480	876,160
	CLB Flip-Flops	728,400	2,443,200	408,000	515,200	607,200	692,800	866,400	1,224,000	1,424,000	725,600	1,095,200
Memory Resources	Maximum Distributed RAM (Kb)	6,938	21,550	4,388	6,525	8,175	8,725	10,888	13,838	17,700	8,850	13,275
	Block RAM/FIFO w/ ECC (36 Kb each)	795	1,292	750	880	1,030	1,180	1,470	1,500	1,880	940	1,410
	Total Block RAM (Kb)	28,620	46,512	27,000	31,680	37,080	42,480	52,920	54,000	67,680	33,840	50,760
Clocking	CMTs (1 MMCM + 1 PLL)	18	24	14	12	14	20	20	18	24	12	18
I/O Resources	Maximum Single-Ended I/O	850	1,200	700	600	700	600	1,000	900	1,100	600	300
	Maximum Differential I/O Pairs	408	576	336	288	336	288	480	432	528	288	144
Integrated IP Resources	DSP Slices	1,260	2,160	1,120	2,160	2,800	2,880	3,600	3,600	3,360	1,680	2,520
	PCIe® Gen2[2]	3	4	—	—	4	—	—	—	—	—	—
	PCIe Gen3	—	—	2	2	—	2	3	3	4	2	3
	Analog Mixed Signal (AMS) / XADC	1	1	1	1	1	1	1	1	1	1	1
	Configuration AES / HMAC Blocks	1	1	1	1	1	1	1	1	1	1	1
	GTX Transceivers (12.5 Gb/s Max Rate)[3]	36	36	—	—	56	—	—	—	—	—	—
	GTH Transceivers (13.1 Gb/s Max Rate)[4]	—	—	28	48	—	80	80	72	96	48	72
	GTZ Transceivers (28.05 Gb/s Max Rate)	—	—	—	—	—	—	—	—	—	8	16
Speed Grades	Commercial Temp (C)	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2
	Extended Temp (E)[5]	-2L, -3	-2L, -2G	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L, -3	-2L	-2L, -2G	-2L, -2G	-2L, -2G
	Industrial Temp (I)	-1, -2	-1	-1, -2	-1, -2	-1, -2	-1, -2	-1, -2	-1	-1	—	—

Share: bit.ly/cloudfpga

Xilinx FPGAs details from [5]

Zynq-7000 SoCs and Zynq with UltraScale+



- Zynq SoC combine an Arm processor core with FPGA chip all in one package
  - **PS** is the Processing System, the CPU
  - **PL** is the Programmable Logic, the FPGA
- Different versions combine different CPUs with various sizes of FPGAs
- Zynq with UltraScale+ also exists and provides a bigger FPGA with the same benefit of Arm processors

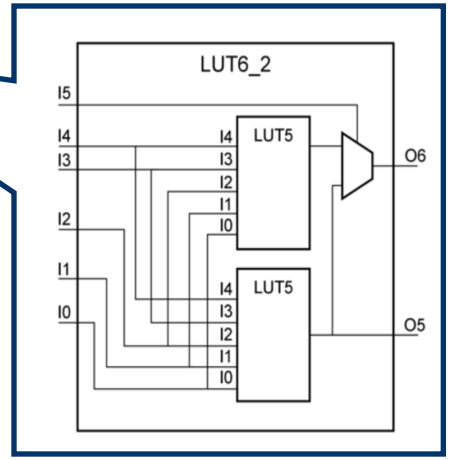
		Cost-Optimized Devices						Mid-Range Devices			
	Device Name	Z-7007S	Z-7012S	Z-7014S	Z-7010	Z-7015	Z-7020	Z-7030	Z-7035	Z-7045	Z-7100
	Part Number	XC7Z007S	XC7Z012S	XC7Z014S	XC7Z010	XC7Z015	XC7Z020	XC7Z030	XC7Z035	XC7Z045	XC7Z100
Processing System (PS)	Processor Core	Single-Core Arm® Cortex®A9 MPCore™ Up to 766MHz			Dual-Core Arm Cortex-A9 MPCore Up to 866MHz			Dual-Core Arm Cortex-A9 MPCore Up to 1GHz[1]			
	Processor Extensions	NEON™ SIMD Engine and Single/Double Precision Floating Point Unit per processor									
	L1 Cache	32KB Instruction, 32KB Data per processor									
	L2 Cache	512KB									
	On-Chip Memory	256KB									
	External Memory Support[2]	DDR3, DDR3L, DDR2, LPDDR2									
	External Static Memory Support[2]	2x Quad-SPI, NAND, NOR									
	DMA Channels	8 (4 dedicated to PL)									
	Peripherals	2x UART, 2x CAN 2.0B, 2x I2C, 2x SPI, 4x 32b GPIO									
	Peripherals w/ built-in DMA[2]	2x USB 2.0 (OTG), 2x Tri-mode Gigabit Ethernet, 2x SD/SDIO									
	Security[3]	RSA Authentication of First Stage Boot Loader, AES and SHA 256b Decryption and Authentication for Secure Boot									
	Processing System to Programmable Logic Interface Ports (Primary Interfaces & Interrupts Only)	2x AXI 32b Master, 2x AXI 32b Slave 4x AXI 64b/32b Memory AXI 64b ACP 16 Interrupts									
Programmable Logic (PL)	7 Series PL Equivalent	Artix®-7	Artix-7	Artix-7	Artix-7	Artix-7	Artix-7	Kintex®-7	Kintex-7	Kintex-7	Kintex-7
	Logic Cells	23K	55K	65K	28K	74K	85K	125K	275K	350K	444K
	Look-Up Tables (LUTs)	14,400	34,400	40,600	17,600	46,200	53,200	78,600	171,900	218,600	277,400
	Flip-Flops	28,800	68,800	81,200	35,200	92,400	106,400	157,200	343,800	437,200	554,800
	Total Block RAM (# 36Kb Blocks)	1.8Mb (50)	2.5Mb (72)	3.8Mb (107)	2.1Mb (60)	3.3Mb (95)	4.9Mb (140)	9.3Mb (265)	17.6Mb (500)	19.2Mb (545)	26.5Mb (755)
	DSP Slices	66	120	170	80	160	220	400	900	900	2,020
	PCI Express®	—	Gen2 x4	—	—	Gen2 x4	—	Gen2 x4	Gen2 x8	Gen2 x8	Gen2 x8
	Analog Mixed Signal (AMS) / XADC[2]	2x 12 bit, MSPS ADCs with up to 17 Differential Inputs									
	Security[3]	AES & SHA 256b Decryption & Authentication for Secure Programmable Logic Config									
	Speed Grades: Commercial	-1			-1			-1			-1
	Speed Grades: Extended	-2			-2,-3			-2,-3			-2
	Speed Grades: Industrial	-1, -2			-1, -2, -1L			-1, -2, -2L			-1, -2, -2L

Share: bit.ly/cloudfpga

Programming 7 Series FPGAs

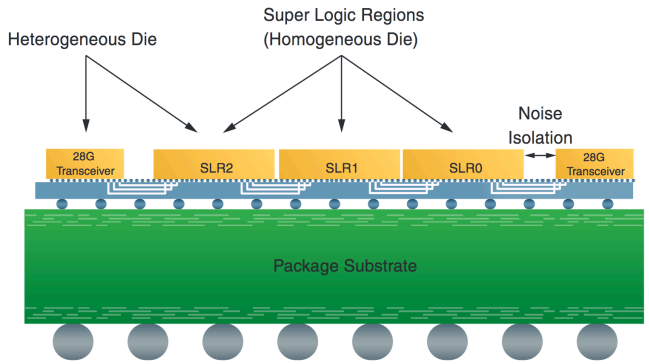
- Vivado tools compile HLD hardware descriptions into bitstreams so they can be run on Xilinx devices
  - Automatically optimize design
    - Map memories to block RAMs
    - Use fast carry chains where possible
    - Pack logic into same LUT
    - Etc.

- Users can also use Xilinx-specific modules in HLD to have greater control of how the hardware inside the FPGA is programmed
  - E.g. **`LUT6_2`**, a six input, two output LUT
  - E.g. **`BRAM_SINGLE_MACRO`**, a single ported Block RAM
  - E.g. **`ROM64X1`**, 64 entry Read-Only Memory with 1-bit words
  - E.g. **`PLLE2_BASE`**, a Phase Locked Loop block

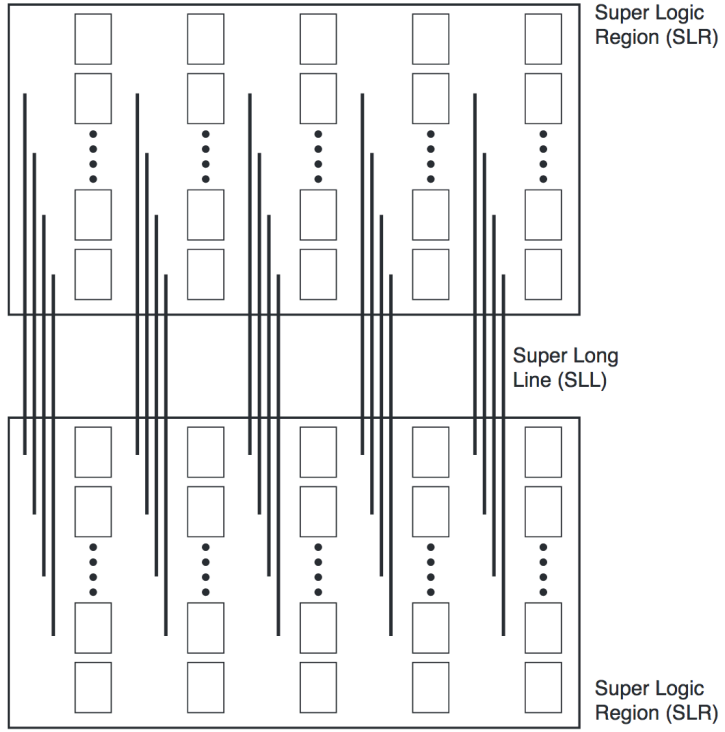


3D Integration and Super Logic Regions

- Many Xilinx (and other) FPGA vendors are moving to 3D integration
- In Xilinx, multiple smaller FPGA dies, Super Logic Regions, are placed on same interposer



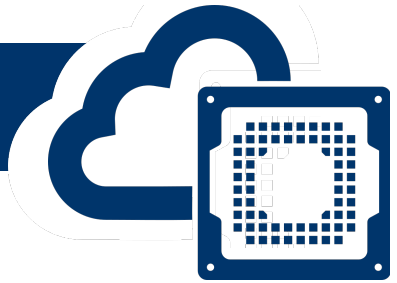
- The smaller FPGA dies are connected by Super Long Lines, effectively extending the size of the columns in the FPGAs and connecting CLBs from different dies



Share: bit.ly/cloudfpga

SLL image from [6], SLR die diagram from [7]

UltraScale and UltraScale+ FPGAs



Comparison of UltraScale and UltraScale+ Chips

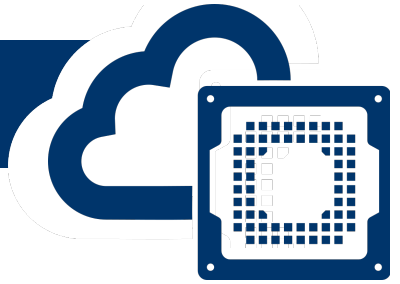
- UltraScale and UltraScale+ provide bigger and faster FPGAs, leveraging mostly same basic CLB design as 7 Series devices
  - UltraScale+ further adds features such as High Bandwidth Memory (HBM, DRAM on same chip) and UltraRam blocks

	Kintex UltraScale FPGA	Kintex UltraScale+ FPGA	Virtex UltraScale FPGA	Virtex UltraScale+ FPGA	Zynq UltraScale+ MPSoC	Zynq UltraScale+ RFSoC
MPSoC Processing System					✓	✓
RF-ADC/DAC						✓
SD-FEC						✓
System Logic Cells (K)	318–1,451	356–1,143	783–5,541	862–8,938	103–1,143	678–930
Block Memory (Mb)	12.7–75.9	12.7–34.6	44.3–132.9	23.6–94.5	4.5–34.6	27.8–38.0
UltraRAM (Mb)		0–36		90–360	0–36	13.5–22.5
HBM DRAM (GB)				0–16		
DSP (Slices)	768–5,520	1,368–3,528	600–2,880	2,280–12,288	240–3,528	3,145–4,272
DSP Performance (GMAC/s)	8,180	6,287	4,268	21,897	6,287	7,613
Transceivers	12–64	16–76	36–120	32–128	0–72	8–16
Max. Transceiver Speed (Gb/s)	16.3	32.75	30.5	58.0	32.75	32.75
Max. Serial Bandwidth (full duplex) (Gb/s)	2,086	3,268	5,616	8,384	3,268	1,048
Memory Interface Performance (Mb/s)	2,400	2,666	2,400	2,666	2,666	2,666
I/O Pins	312–832	280–668	338–1,456	208–2,072	82–668	280–408

Share: bit.ly/cloudfpga

Comparison table from [2]

SLRs and UltraScale and UltraScale+ Chips

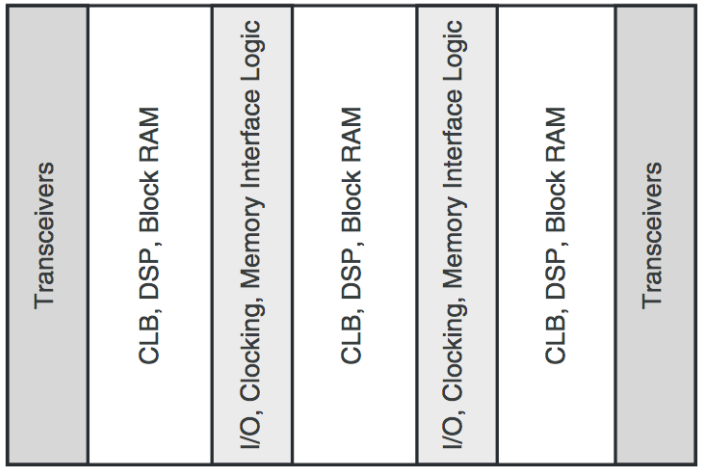


- UltraScale and UltraScale+ chips use Super Logic Regions
- SLR size and number depends on the chip

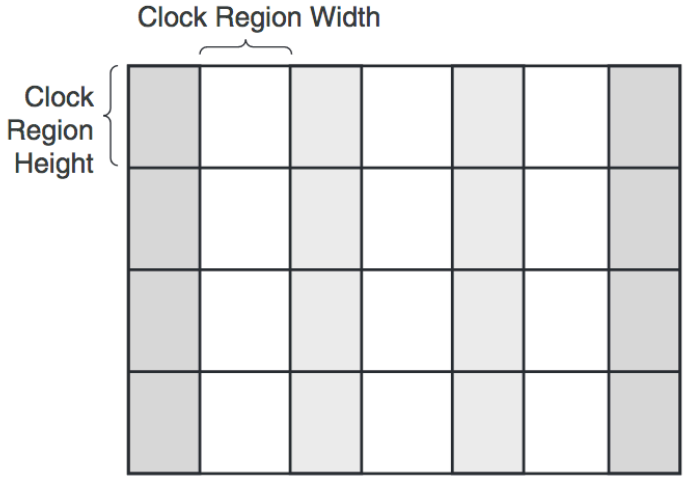
	Kintex UltraScale		Virtex UltraScale				Virtex UltraScale+								
Device	KU085	KU115	VU125	VU160	VU190	VU440	VU5P	VU7P	VU9P	VU11P	VU13P	VU31P	VU33P	VU35P	VU37P
# SLRs	2	2	2	3	3	3	2	2	3	3	4	1	1	2	3
SLR Width (in regions)	6	6	6	6	6	9	6	6	6	8	8	8	8	8	8
SLR Height (in regions)	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4

UltraScale and UltraScale+ Device Layout

- The devices follow column-style organization
  - CLBs and other blocks are organized in columns
  - Interconnection is done through routing matrix with varying length wires



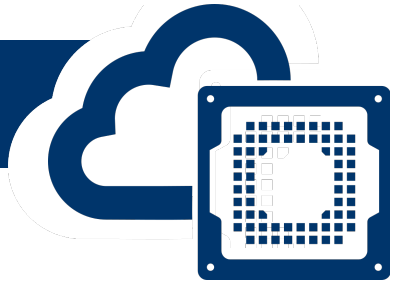
- The FPGA is divided into multiple clock regions
  - Height of clock region equals 60 CLBs
  - Clock regions are used to distribute clock signals with minimal delay
    - Support multiple clocks per region



Share: bit.ly/cloudfpga

Images from [2]

Some Features of UltraScale and UltraScale+ FPGAs

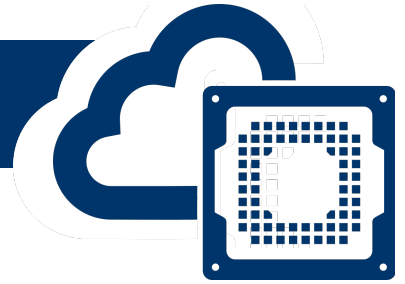


- **Configurable Logic Block**
  - Follows the 7 Series design
    - Configurable Logic Block (CLB) in the UltraScale architecture contains 8 LUTs and 16 flip-flops
    - The LUTs can be configured as either one 6-input LUT with one output, or as two 5-input LUTs with separate outputs but common inputs
    - Each LUT can optionally be registered in a flip-flop

- **Interconnect**
  - Vertical and horizontal routing wires in the UltraScale architecture that span 1, 2, 4, 5, 12, or 16 CLBs
  - Super Long Lines are used for crossing SLR boundaries

- **UltraRAM**
  - Faster alternative to Block RAM
  - Distributed in the FPGA fabric like Block RAM

Virtex UltraScale+ VU9P



- Chip used in Amazon F1 and Huawei Cloud
  - About 25,000$ per chip

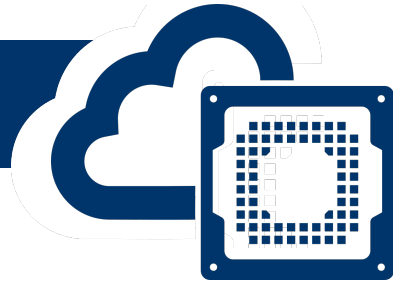
70,000$ on DigiKey, per chip

	VU3P	VU5P	VU7P	VU9P	VU11P	VU13P	VU19P	VU27P	VU29P
System Logic Cells	862,050	1,313,763	1,724,100	2,586,150	2,835,000	3,780,000	8,937,600	2,835,000	3,780,000
CLB Flip-Flops	788,160	1,201,154	1,576,320	2,364,480	2,592,000	3,456,000	8,171,520	2,592,000	3,456,000
CLB LUTs	394,080	600,577	788,160	1,182,240	1,296,000	1,728,000	4,085,760	1,296,000	1,728,000
Max. Distributed RAM (Mb)	12.0	18.3	24.1	36.1	36.2	48.3	58.4	36.2	48.3
Block RAM Blocks	720	1,024	1,440	2,160	2,016	2,688	2,160	2,016	2,688
Block RAM (Mb)	25.3	36.0	50.6	75.9	70.9	94.5	75.9	70.9	94.5
UltraRAM Blocks	320	470	640	960	960	1,280	320	960	1,280
UltraRAM (Mb)	90.0	132.2	180.0	270.0	270.0	360.0	90.0	270.0	360.0
HBM DRAM (GB)	–	–	–	–	–	–	–	–	–
CMTs (1 MMCM and 2 PLLs)	10	20	20	30	12	16	40	16	16
Max. HP I/O(1)	520	832	832	832	624	832	1,976	520	676
Max. HD I/O(2)	–	–	–	–	–	–	96	–	–
DSP Slices	2,280	3,474	4,560	6,840	9,216	12,288	3,840	9,216	12,288
System Monitor	1	2	2	3	3	4	4	4	4
GTY Transceivers 32.75Gb/s(3)	40	80	80	120	96	128	80	32	32
GTM Transceivers 58.0Gb/s(3)	–	–	–	–	–	–	–	48	48
100G / 50G KP4 FEC	–	–	–	–	–	–	–	24/48	24/48
Transceiver Fractional PLLs	20	40	40	60	48	64	40	40	40
PCIe Gen3 x16	2	4	4	6	3	4	0	1	1
PCIe Gen3 x16 / Gen4 x8 / CCIX(4)	–	–	–	–	–	–	8	–	–
150G Interlaken	3	4	6	9	6	8	0	8	8
100G Ethernet w/RS-FEC	3	4	6	9	9	12	0	15	15

Share: bit.ly/cloudfpga

Xilinx FPGAs details from [2]

Kintex UltraScale KU115



- Chip used in Baidu, Tencent, and Alibaba Clouds
  - About 6,000$ per chip

	KU025[1]	KU035	KU040	KU060	KU085	KU095	KU115
System Logic Cells	318,150	444,343	530,250	725,550	1,088,325	1,176,000	1,451,100
CLB Flip-Flops	290,880	406,256	484,800	663,360	995,040	1,075,200	1,326,720
CLB LUTs	145,440	203,128	242,400	331,680	497,520	537,600	663,360
Maximum Distributed RAM (Mb)	4.1	5.9	7.0	9.1	13.4	4.7	18.3
Block RAM Blocks	360	540	600	1,080	1,620	1,680	2,160
Block RAM (Mb)	12.7	19.0	21.1	38.0	56.9	59.1	75.9
CMTs (1 MMCM, 2 PLLs)	6	10	10	12	22	16	24
I/O DLLs	24	40	40	48	56	64	64
Maximum HP I/Os[2]	208	416	416	520	572	650	676
Maximum HR I/Os[3]	104	104	104	104	104	52	156
DSP Slices	1,152	1,700	1,920	2,760	4,100	768	5,520
System Monitor	1	1	1	1	2	1	2
PCIe Gen3 x8	1	2	3	3	4	4	6
150G Interlaken	0	0	0	0	0	2	0
100G Ethernet	0	0	0	0	0	2	0
GTH 16.3Gb/s Transceivers[4]	12	16	20	32	56	32	64
GTY 16.3Gb/s Transceivers[5]	0	0	0	0	0	32	0
Transceiver Fractional PLLs	0	0	0	0	0	16	0

Share: bit.ly/cloudfpga

Table from [2]

Alveo Cards

- Xilinx Alveo U200/U250/U280 Accelerator Cards
  - Designed for data center acceleration
  - Combine FPGA with DRAM chips on same PCB board
  - 800,000 to 1,000,000 LUTs
  - Thermal Design Power (TDP) of 225W
  - Passive or active cooling
  - Up to PCIe 4.0 and DDR4 and QSFP networking

Summary

- Different Cloud FPGA vendors give access to different types of FPGAs
- Amazon F1 and many others use Xilinx
- Typically used FPGAs are high-end
  (but not biggest or fastest)

Share:
bit.ly/cloudfpga

References

1. “Vivado Design Suite 7 Series FPGA and Zynq-7000 SoC Libraries Guide”. Available at: https://www.xilinx.com/support/documentation/sw_manuals/xilinx2018_3/ug953-vivado-7series-libraries.pdf.
2. “UltraScale Architecture andProduct Data Sheet: Overview”. Available at: https://www.xilinx.com/support/documentation/data_sheets/ds890-ultrascale-overview.pdf.
3. Xilinx Alveo, Adaptable Accelerator Cards for Data Center Workloads. Available at: https://www.xilinx.com/products/boards-and-kits/alveo.html.
4. “Cost-Optimized Portfolio Product Tables and Product Selection Guide”. Available at: https://www.xilinx.com/support/documentation/selection-guides/cost-optimized-product-selection-guide.pdf.
5. “All Programmable 7 Series Product Selection Guide”. Available at: https://www.xilinx.com/support/documentation/selection-guides/7-series-product-selection-guide.pdf.
6. “7 Series FPGAs Configurable Logic Block”, Available at: https://www.xilinx.com/support/documentation/user_guides/ug474_7Series_CLB.pdf.
7. “Xilinx Stacked Silicon Interconnect Technology Delivers Breakthrough FPGA Capacity, Bandwidth, and Power Efficiency”. Available at: https://www.xilinx.com/support/documentation/white_papers/wp380_Stacked_Silicon_Interconnect_Technology.pdf.