FPGA Capability Expansion Despite Technology Challenges

Nui Chong and Xin Wu

XILINX INC, 2100 Logic Drive, San Jose, California, 95124 USA

ABSTRACT

FPGA (Field Programmable Gate Array) has been continually growing – expanding functionality, improving power-performance and enlarging capacity, in past several decades. Moore's Law has been gradually slowing down and technology becomes more expensive in past several nodes. However with continual architecture, design and technology innovations, FPGA's capacity expansion continues, beyond Moore's Law. 3D-IC, embedded processing SOC, HBM integration, RF-SOC and furthermore, coming Versal ACAP (Adaptive Compute Acceleration Platform) family, ensures more and more suitable applications in data centers, machine learning (ML), 5G, automotive and many other applications in coming years.

Keywords: FPGA, beyond Moore's Law, 3D-IC, hetero-integration, Versal ACAP

1. INTRODUCTION

Moore's Law [1] expressed an economical optimism of semiconductor industry forward looking – one could expect half of size, double performance and reduce power, at almost equal cost, in every 12 months (later it modified to 18 months and then 2 years interval). It not only brought confidence to the industry but also continuously guided its 40+ years of growth, till 28nm. The entire industry, and the entire world, benefited from the booming of semiconductor.

It seems since 20nm, due to the usage of multi-patterning lithography, the economic benefit stated by Moore's Law, became more and more challenged – the scaling, improved power-performance still possible with less degree, the "almost equal cost" aspect, is unlikely. Figure 1 shows a trend of processing performance trend along with time [2].

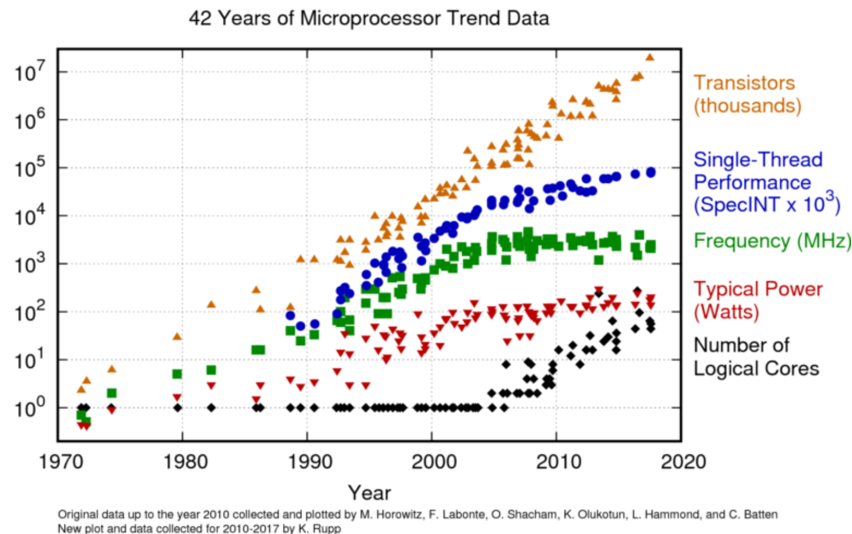

Figure 1. Microprocessor performance trend and showing plateau

Photomask Japan 2019: XXVI Symposium on Photomask and Next-Generation Lithography Mask Technology, edited by Akihiko Ando, Proc. of SPIE Vol. 11178, 1117802 · © 2019 SPIE · CCC code: 0277-786X/19/$21 · doi: 10.1117/12.2537628

2. FPGA CAPABILITY

FPGA's capability, which includes not only its programmable logic count, performance and power, but also functionality, software and applications, has been continually increasing, under Moore's Law and beyond. Figure 2 shows the maximum logic cell (each logic cell roughly equals to ~ 100 logic gates) counts of each generation product family since 0.25um Virtex to the newest 7nm Versal product family. From the chart it can be seen following Moore's Law the maximum logic cell counts (red color) increased from 27,000 in 0.25um node to ~3,700,000 in 16nm (if chose to do so) in a monolithic die, approximately ~137x increase in 9 nodes. Whereas by employing passive 3D-IC (orange color), the maximum logic cell count gains another 2~3x higher than merely following Moore's Law.

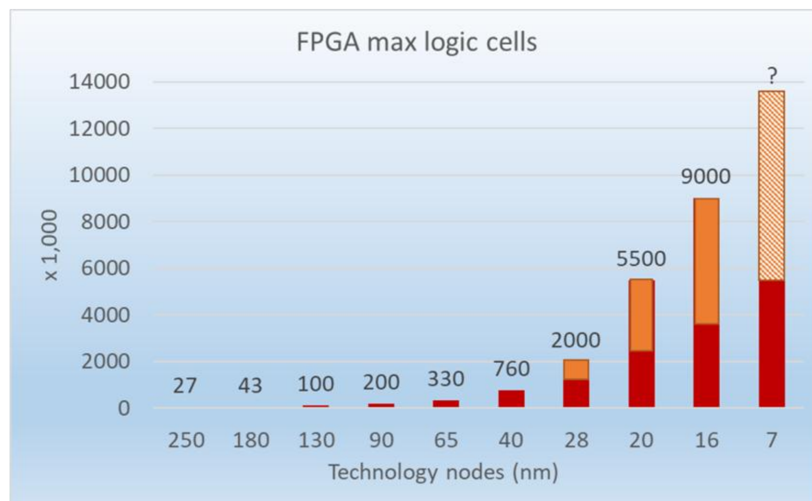

Figure 2. Maximum logic cell counts of FPGA in 10 technology nodes, of monolithic (red) and of passive 3D-IC.

It is noticeable that the logic cell count increase is less than what Moore's Law prediction from node to node. One of the reasons is that not all types of circuits are scaled equally. For example, when normal logic circuits, SRAM in general following technology scaling factor, IOs, analog and RF circuits are usually scaled far less. The other important reason is from one generation to another, FPGA has been continually increasing its functions (hard embedded IPs), which is also counted into die size real estate. Despite more area required by these increases of functionality, they are the key part of FPGA capability improvement, and have very positive impacts of customers' applications.

We will discuss several examples of functionality critical to FPGA applications in the following sections.

2.1 SERDES

Serializer and De-serializer (SERDES) have become popular gradually in past decades. FPGA first employed SERDES in Virtex II Pro family in 2002. It was only a few high-end family members with limited number of SERDES channels of only 3.125Gpbs. Today in 16nm Ultrascale-plus family, SERDES have been much proliferated with higher performance. For example the UV-29Plus has 48 channels of 58Gbps PAM4 and 32 channels of 32Gbps SERDES. The VU-13Plus has 120 channels of 32Gbps SERDES. Both are in a 3D-IC integration. These high speed and large number of SERDES channels become essential in communication industry.

Not only 58Gbps PAM4, Xilinx also demonstrated 112Gbps SERDES [3], and it will be a part of 7nm Versal family offering.

To incorporate SERDES design into FPGA, one key challenge is timing – one needs to be able to bring complicate RF design at early development stage of a technology, not just waiting for its matureness. This requires in-depth

understanding the technology, analyzing existing data and forecasting its RF behavior in RF modeling. Xilinx has spent major effort to master this aspect.

Beyond 112Gbps, industry has not had clear conclusion the roadmap. Xilinx believes optical photonics integration could be one of the options, and has been studying accordingly.

Below Figure 3 lists publications showing the evolution of SERDES in past a decade and half.

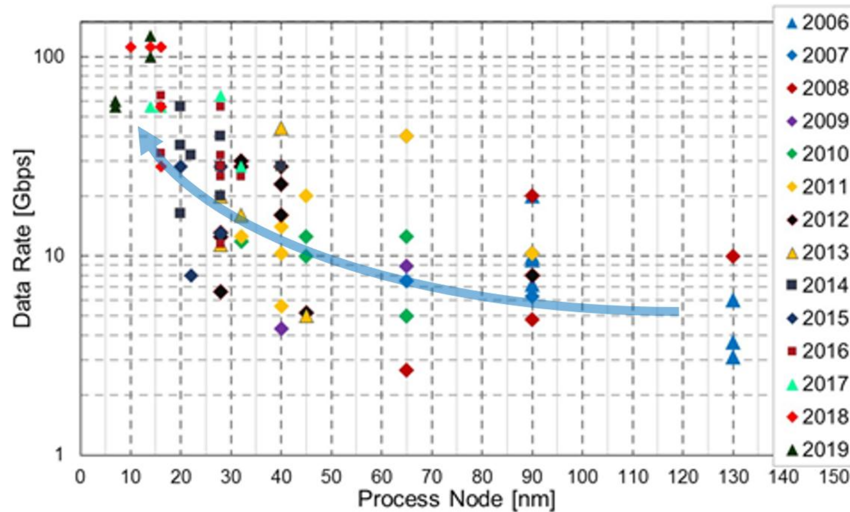

Figure 3: SERDES performance in various technology nodes (ISSCC, VLSI, CICC, ASSCC, etc.)

2.2 Embedded micro-processors, SOC and RFSOC

In the same Virtex-II Pro family in 2002, it was also the first time a micro-processor, PowerPC 405 (PPC405), embedded into certain high-end FPGA family members. A major effort was spent to ensure intimate connections between PPC405 to programmable logic fabric. The embedded micro-processor enabled many new high-end industrial and communication FPGA applications.

However the application usage of embedded micro-processor has still be relatively limited, until the 28nm 7-Series Zynq SOC family in 2012 time frame. The 7-Series Zynq had a much more popular, lower cost ARM A9 processor, with vast effort to build it not only a connected processor but far more capable SOC, which includes memories, IOs and large amount of system software to support applications. The 7-Series Zynq SOC attracted a lot more applications in all range (including very low cost ones).

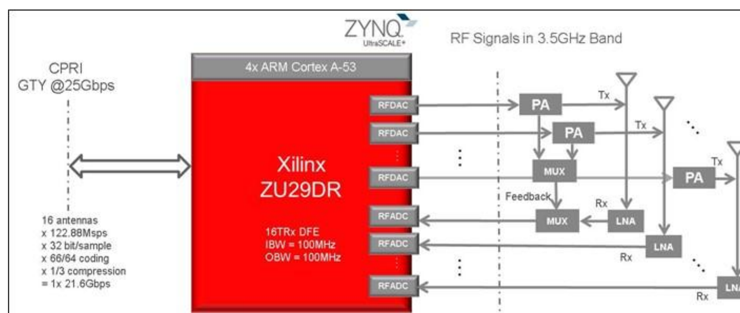

Figure 4. An illustration of using RFSOC in a 5G NR 16 TRX MIMMO 3.5Gbps band design

In 2016 the 16nm Ultrascale-Plus FPGA products, this ARM based SOC has been further upgraded into MPSOC and RFSOC. The MPSOC was an order of magnitude more sophisticate and capable (multiple ARM A53 and AMR R5 with

GPU + video + Codec) than the original 7-Series Zynq SOC; the RFSOC built 4Gbps ADC and DAC (Analog to Digital Convertor and Digital to Analog Convertor) [4]. Combined with MPSOC and RFSOC capability, they become perfect candidate of today's 5G wireless front-haul applications.

Figure 4 is an illustration of a 16 TRX MIMO (Multi-Input Multi-Output) antenna design in a 3.5Gbps 5G NR (New Radio), by using ZU-29DR RFSOC. Because of the integration of ADC and DAC, and the massive capability of MPSOC, it serves beam-forming, DFE (Digital Front End) and ADC/DAC functions. It reduces power, cost and foot-print drastically, and allows system change by its programmability, to suit today's 5G initial deployment which still faces many uncertainties. It is a very successful example of how adding right functionality will expand capability of FPGA and benefit the end user ultimately.

Similar to design and integration of SERDES in FPGA during a not-so-mature technology node, bring up ADC and DAC also requires strong technical capability of predicting and modeling the analog and RF behavior of a new technology.

2.3 HBM integration

Logic and DRAM memory technology development have long been parted. Each optimizes for better technology capability, market needs and cost. Thus integrating DRAM and logic in system in past was mostly through DDR interface. For wide bandwidth DRAM connections, large amount of DDR IOs are needed. Even so the bandwidth is still limited, power and latency still high.

HBM (High Bandwidth Memory) is a stacked DRAM cube using 3D-IC with uBump (micro-bump) and TSV (Through Silicon Via). Because its wide (1024 channel in HBM-2) and high speed (2~3 Gbps in HBM-2), its bandwidth, power and latency are un-surpassable. Integrating HBM to a logic IC also needs to employ passive 3D-IC technologies, such as CoWoS (Chip on Wafer on Substrate) and others.

In the 16nm Ultrascale-Plus FPGA family, HBM integration becomes available, which largely expands on-chip data processing capability because of ultra-wide bandwidth with low power and low latency. Figure 5 showed an illustration of FPGA integrated with 2 HBMs [5].

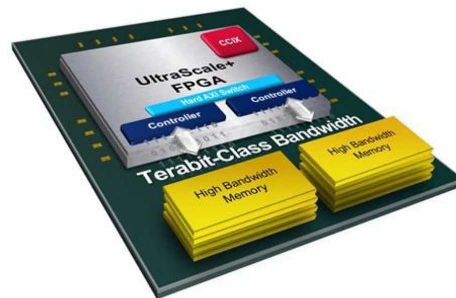

Figure 5. An illustration of an Ultrascale-Plus FPGA integrated with 2 HBMs

The key to succeed in HBM integration relies on Xilinx' long and successful experiences of 3D-IC SSIT (Stacked Silicon Integration Technology) since 2011, as well as industry's volume production learning.

3. NEXT, VERSAL THE ACAP

Because of the rapid growth in datacenter, explosion of AI (Artificial Intelligence), ML (Machine Learning), CNN (Convoluted Neuron Network), as well as other rapid evolutions in industry such as 5G wireless initial deployment, automotive ADAS (Advanced Driver Assistance System) and development of fully AD (Auto Drive), etc., the next

generation of FPGA, Versal, has been built in 7nm technology and set up an ACAP (Adaptive Compute Acceleration Platform) [6]. Figure 6 is a function block diagram of Versal.

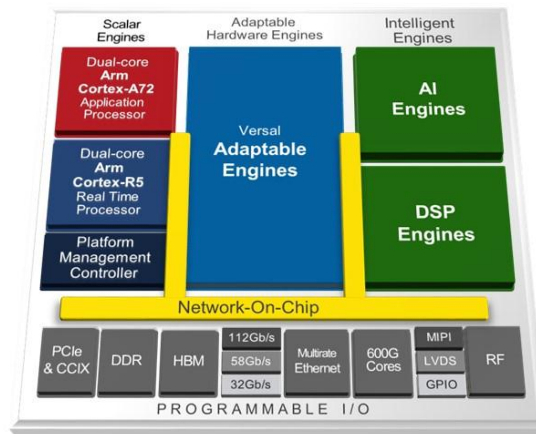

Figure 6. A functional block diagram of 7nm Versal FPGA

Several new functional blocks, such as AI Engine, Network-on-chip, etc. will be created, and many other functional blocks, will be upgraded. All these hard blocks will serve as foundation to establish an adaptive acceleration platform including both hardware and software, as illustrated in Figure 7.

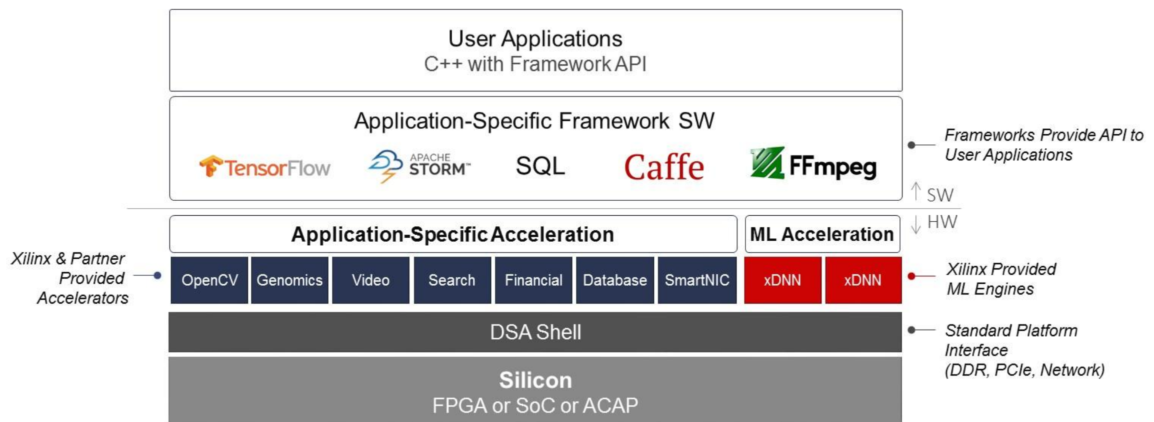

Figure 7. A stack of SW and HW application adaptive acceleration platform of Versal

With 7nm technology, as a first order calculation, one would be able to estimate approximately ~50% of scaling, ~20% performance gain at equal power vs. 16nm. However with all these innovations, software and hardware stacked platforms, the actual application will see much higher gain in overall performance, as shown in Figure 8.

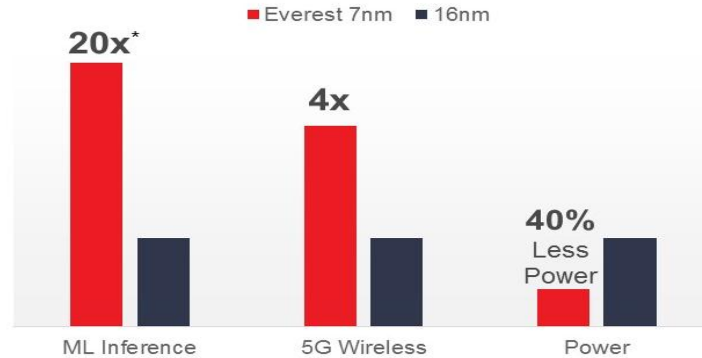

Figure 8. Versal application performance and power comparison vs. 16nm

3.1 EUV readiness in 7nm and possible application

In TSMC's 7nm technology, critical layers are all using immersion multi-patterning with uni-directional design. TSMC's 7-Plus platform, will adopt several EUV lithography.

Potential use of EUV for better uniformity thus better power-performance (for example in MEOL (Mid-End of Line) to reduce resistance), will be assessed.

4. CONCLUSION

FPGA continually increases its capability which ultimately benefits end customer applications, despite in past several technology nodes more challenges becoming obvious. The increase of capability has been achieved not only by technology scaling, but also by multiple "more-than-Moore" techniques in integration as well as continual architecture, design, software innovations in functionality advancement.

Author would like to express their thanks to many in Xilinx of their help on this publication.

REFERENCES

[1] Moore, Gordon E. (1965-04-19). "Cramming more components onto integrated circuits". Electronics. Retrieved 2016-07-01.

[2] K. Rupp, "42 Years of Microprocessor Trend Data", https://www.karlrupp.net/2018/02/42-years-of-microprocessor-trend-data/ (2018)

[3] KeeHian Tan ; Ping-Chuan Chiang ; Yipeng Wang ; Haibing Zhao ; Arianne Roldan ; Hongyuan Zhao ; Nakul Narang ; Siok Wei Lim ; Declan Carey ; Sai Lalith Chaitanya Ambatipudi ; Parag Upadhyaya ; Yohan Frans ; Ken Chang, "A 112-GB/S PAM4 Transmitter in 16NM FinFET", VLSI Symposium 2018, page 45-46 (2018)

[4] Xilinx RFSOC product datasheet https://www.xilinx.com/publications/product-briefs/rfsoc-product-brief.pdf (2019)

[5] Suresh Ramalingam, "HBM package integration: Technology trends, challenges and applications", 2016 IEEE Hot Chips 28 Symposium (HCS) (2016)

[6] Victor Peng, "Adaptable Intelligence", 2018 IEEE Hot Chips 30 Symposium (HCS) Key Note Speech, (2018)