(intel®)

Intel® FPGAs Accelerate Intel® Xeon® Scalable Processors in Servers and High-End Embedded Systems

Authors

Thomas M.Schulte
Product Marketing Specialist
Intel Programmable Solutions Group

Steve Leibson
Senior Marketing Engineering Manager
Intel Sales and Marketing Group

Abstract

The Intel® Stratix® 10 DX device family provides high-speed hardware acceleration and memory expansion to servers when combined with selected Intel Xeon® Scalable Processors through the lntel Ultra Path Interconnect (UPI) chip-to-chip coherent interface and/or high-speed non-coherent PCI Express* (PCIe*) Gen3 or Gen4 x16 ports, suitable as direct attached FPGA-based hardware accelerators.

Introduction

Hardware acceleration for servers is becoming increasingly common because these accelerators can perform calculations faster and more efficiently than the same computations running in software. The high-performance Intel Stratix 10 DX FPGA family has been designed specifically to provide FPGA-based hardware acceleration and memory expansion for next-generation server systems including servers based on selected Intel Xeon Scalable processors that are compatible with UPI mesh networks. Intel Stratix 10 DX devices, the newest members of the expanding Intel Stratix 10 FPGA family, incorporate a number of specific features that make them ideal for use in these systems, including:

- Multiple coherent UPI ports that support direct, peer-to-peer connection to selected Intel Xeon Scalable processors

- Multiple PCIe Gen4 interface ports (PCIe Gen3-capable as well) with advanced virtualization features

- New memory controller IP for connection to select Intel(r) Optane(tm) DC Persistent Memory accessing up to 4TB of non-volatile memory. High Bandwidth Memory 2 (HBM2) 3D DRAM stacks that deliver extremely high memory bandwidth

These unique features permit Intel Stratix 10 DX FPGAs to act as peers on the UPI network, at the very core of server systems, and to deliver high bandwidth and low latency in both hardware-acceleration and memory-expansion roles through the integrated PCIe Gen4 and UPI interfaces – as explained in this white paper.

Workload acceleration for servers

Diverse server workloads can benefit from hardware acceleration. These specialized workloads include:

- Cloud Disaggregation
- Financial/HPC
- Encryption/Decryption
- Compression/Decompression
- Video Transcoding
- Storage
- Genomics
- Database Acceleration/Data Analytics
- General Cloud-Based Compute Acceleration

FPGAs serve as excellent foundations for hardware-accelerator designs because of their massive parallelism and dynamically reprogrammable datapaths, which can be configured to exactly match the computational needs of specific application workloads. Their versatility enables the provisioning of faster, more power-efficient, lower-latency services, which lower total cost of ownership and maximize compute capacity within a data center's power, space, and cooling constraints. However, the communications between the server processors and the hardware accelerator must be streamlined to extract maximum benefit from the accelerator hardware.

PCIe Gen4: The universal interface for hardware acceleration

The PCIe-SIG formally released the PCIe Gen4 specification in late 2017. First and foremost, PCIe Gen4 ports run at a maximum data rate of 16 GT/s or 2 GBps, so they are twice as fast as PCIe Gen3 ports. Beyond the 2X increase in maximum per-lane transfer rate, PCIe Gen4 provides several additional features that can be used to great advantage to boost system performance. These additional features include:

- Reduced system latency through extended tags and credits for service devices
- Improved I/O virtualization and platform integration
- Bifurcation

PCIe Gen4 faces a tough challenge with respect to effective bandwidth utilization due to the 16 GT/s transfer rate. Two improvements made to the PCIe Gen4 specification help meet this challenge. First, the PCIe Gen4 standard's 10 bit extended tag to increases the number of allowable, outstanding, non-posted requests from 256 to 768.

In addition, the PCIe Gen3 standard's existing flow-control mechanism is limited to a maximum of 127 outstanding header credits and 2,047 outstanding data credits. Credit-based flow control eliminates packet discards caused by receive buffer overflows, but in certain scenarios, the transfer rate of a PCIe Gen3 x16 link can saturate with these credit limits, which can throttle PCIe Gen3 port efficiency. Consequently, the PCI SIG enhanced the flow-control mechanism in the PCIe Gen4 specification by adding scaled flow control, which permits the maximum number of outstanding header and data credits to be scaled. Each PCIe Gen4 flow-control credit can represent 1, 4, or 16 PCIe Gen3 credits based on the programmable scale setting.

PCIe virtualization using the Intel Stratix 10 DX FPGA

PCIe I/O virtualization allows multiple operating systems running simultaneously within a single server system to natively share PCIe devices. The PCIe Gen4 ports on Intel Stratix 10 DX FPGAs offer several advanced virtualization features that boost the performance of systems running virtual machines (VMs):

- Single-Root I/O Virtualization (SR-IOV) is a new addition to the PCIe specification that allows multiple VMs to share PCIe hardware resources. Prior to the availability of the SR-IOV feature, a system's hypervisor was responsible for creating virtual I/O adapters for the VMs. SR-IOV pushes much of this overhead into the PCIe Gen4 I/O hardware, which removes the hypervisor from the I/O performance path. Consequently, PCIe virtualization makes it easier for server processors to run a larger number of VMs, which reduces the need for more server hardware and thus reduces CAPEX costs, power consumption, and rack-space requirements.

- The Intel Stratix 10 DX FPGA implements eight PCIe Physical Functions (PFs) and 2048 PCIe Virtual Functions (VFs) per PCIe port in hard IP, which means that the Intel Stratix 10 DX FPGA can support more VMs while using significantly fewer programmable logic resources in the FPGA. The Intel Stratix 10 DX FPGA's hard-IP PFs and VFs can be used with VirtIO drivers to relieve the hypervisor's need to emulate physical network and storage devices for each VM, which improves the performance of these virtual I/O devices and eases VM-to-VM migration. VirtIO is part of most standard Linux* libraries.

- Scalable I/O Virtualization (Scalable IOV) is a very efficient hardware-assisted approach to I/O virtualization that has been developed by Intel. It generally provides more I/O scalability (to as many as one million virtual devices) with better latency, while using fewer CPU resources than SR-IOV, and without sacrificing performance. Scalable IOV also addresses limitations associated with direct device assignment such as generational compatibility, live migration, and memory over-commitment.

Additional PCIe features built into the Intel Stratix 10 DX FPGA

The Intel Stratix 10 DX FPGA's PCIe ports support both root-port and endpoint modes. In root-port mode, each of an Intel Stratix 10 DX FPGA's 16-lane PCIe ports supports bifurcation into four PCIe Gen4 x4 or four PCIe Gen 3 x4 ports. In Endpoint mode, the Intel Stratix 10 DX FPGA PCIe ports support bifurcation into two PCIe Gen4 x8 or four PCIe Gen3 x8 ports. Bifurcation allows the PCIe ports on the Intel Stratix 10 DX FPGA to expand the number of available PCIe ports while matching the bandwidth requirements of the link partners at the other end of the ports.

In addition, the Intel Stratix 10 DX FPGA PCIe ports support a hard-IP bypass feature that enables the use of these ports as high-speed PCIe switches. This mode bypasses the PCIe transaction layer (TL) and directly passes packets at the transaction layer to lightweight TL blocks implemented with user-defined logic in the FPGA. This feature can be used, for example, to implement a high-speed PCIe switch.

These bandwidth, latency, virtualization, and bifurcation improvements make PCIe Gen4 interfaces an excellent choice as a universal server interface for hardware accelerators, however UPI ports present an even faster alternative for systems based on Intel Xeon Scalable processors.

UPI for Intel Xeon Scalable Processor hardware acceleration

Intel has developed several shared-memory architectures for multi-processor systems with high-speed, point-to-point, coherent, inter-processor interfaces. One such architecture, called the Intel QuickPath Interconnect (Intel QPI), was introduced in 2008 for Intel Xeon processor families (formerly Haswell and Broadwell) based on the Grantley platform. UPI is the follow-on multiprocessor architecture and interface. Previously, UPI interface ports have only been offered on Intel Xeon processors but Intel Stratix 10 DX FPGAs now incorporate coherent UPI ports as well. UPI makes an ideal interconnect for coupling high-performance hardware accelerators to selected Intel Xeon Scalable processors.

For Intel Xeon processor families that employ QPI, the processor cores, last-level cache (LLC), SDRAM memory controller, I/O controller, and inter-socket QPI ports are connected together using a ring architecture, as shown in Figure 1.
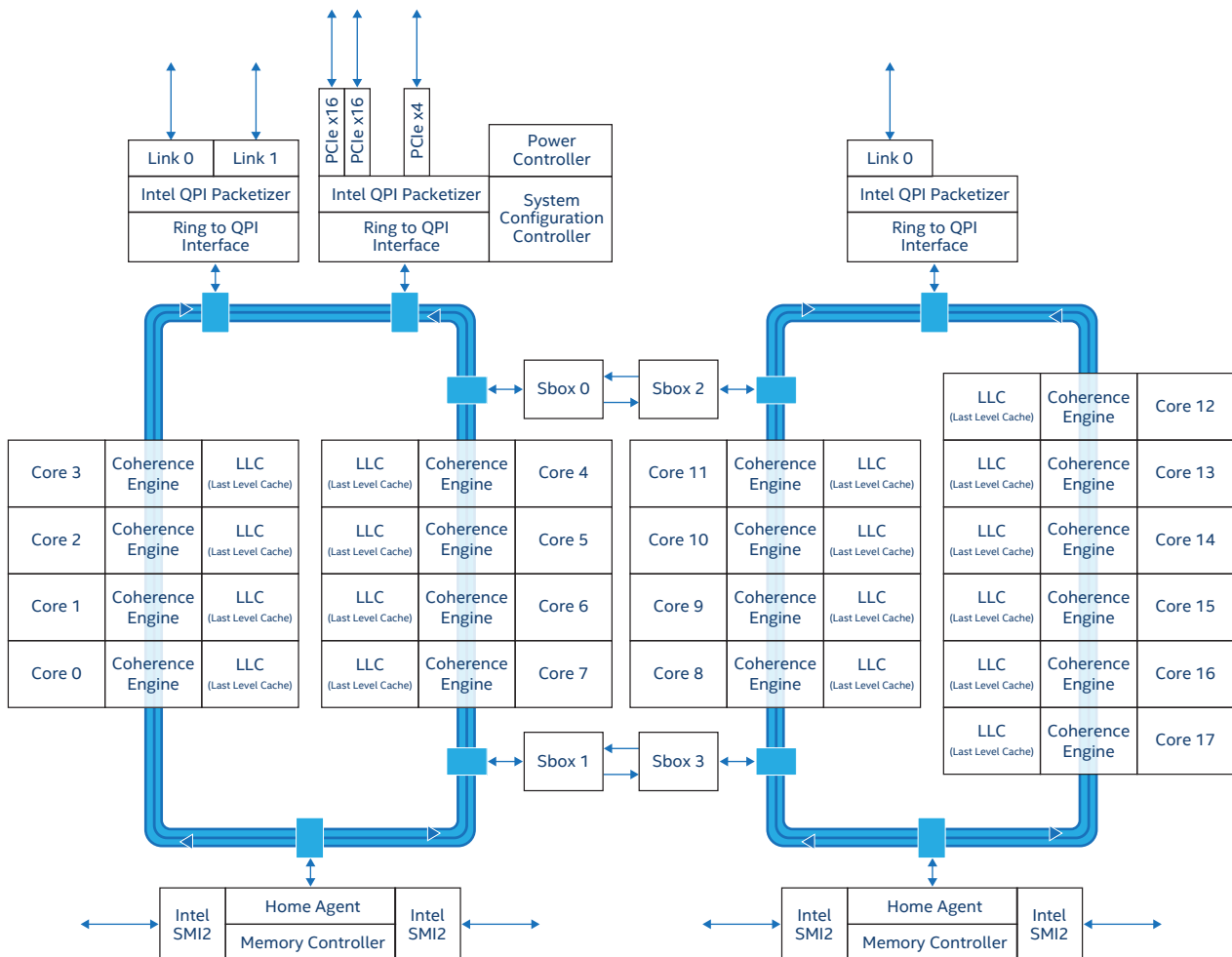

Figure 1. The QPI Ring Architecture

Source. https://software.intel.com/en-us/articles/intel-xeon-processor-scalable-family-technical-overview

With even more available cores per processor and much higher memory and I/O bandwidth requirements in the Intel Xeon Scalable processor family, the ring-based QPI architecture could limit system performance. Therefore, the Intel Xeon Scalable processor family introduced a mesh architecture with greatly improved network connectivity, as shown in Figure 2, to improve upon the latencies and bandwidth characteristics of the QPI ring-based architecture. UPI also provides high-speed, low-latency paths among CPU sockets in server designs.
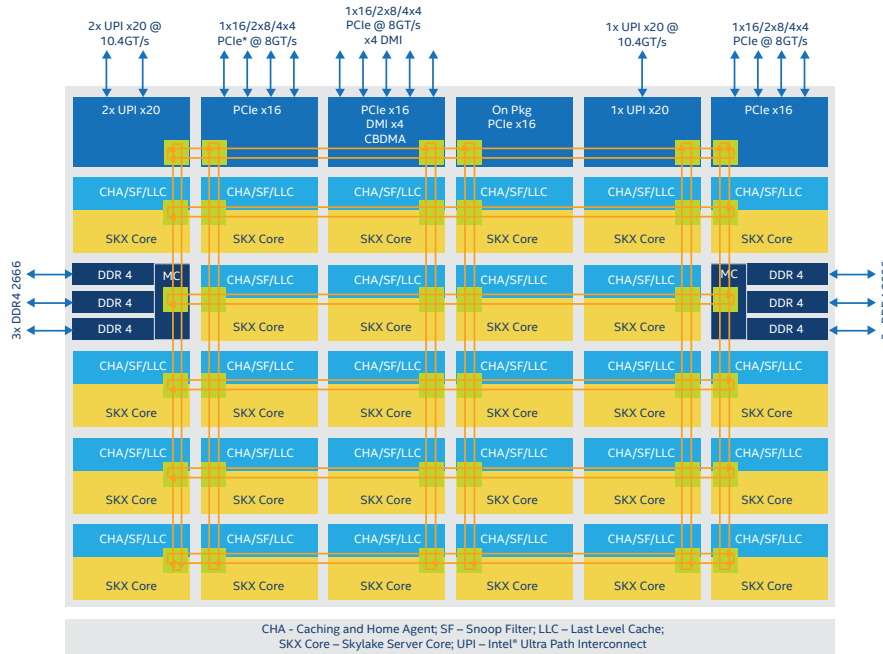

Figure 2. The UPI Mesh

Source. https://software.intel.com/en-us/articles/intel-xeon-processor-scalable-family-technical-overview

Intel Xeon Scalable processors that support UPI provide either two or three external UPI links, used to interconnect multiple Intel Xeon processors using high-speed, low-latency paths to the other CPU sockets. UPI employs a directory-based home snoop coherency protocol, which provides an operational speed as fast as 10.4 GT/s, improves power efficiency through a low-power L0p state, and improves data transfer efficiency over the link using a new UPI packetization format. An improved UPI protocol layer removes some scalability limits previously associated with QPI ports.

Intel Stratix 10 DX devices provide as many as three compatible UPI ports for direct external connection to compatible Intel Xeon Scalable processors in new server designs. Compared to a PCIe Gen3 x16 interface, a UPI port with UPI-Cache Agent IP delivers 2.64x more bandwidth. In addition, UPI delivers 37% lower latency than a non-coherent PCIe Gen4 port[†]. UPI is a coherent interface, which allows with UPI ports to share the Intel Xeon processors' view of global memory.

UPI for memory expansion

Because UPI ports are coherent, Intel Stratix 10 DX device family members that are connected to selected Intel Xeon Scalable processors via coherent UPI links can also provide memory expansion to these same processors. It may not be readily apparent why a cluster of Intel Xeon processors might need memory-expansion capabilities because these server-class processors already incorporate several high-speed DDR4 memory channels. However, Intel Stratix 10 DX FPGAs add a significant amount of system flexibility with respect to controllers for different memory types. In addition to the hardened DDR4 memory controllers incorporated into all Intel Stratix 10 devices, additional soft memory controllers can be added to support essentially any memory type required by a server design including Intel Optane™ DC persistent memory and NAND Flash DIMMs.

Some members of the Intel Stratix 10 DX FPGA family also incorporate high-speed HBM2 DRAM using Intel's Embedded Multi-die Interconnect Bridge (EMIB) technology, which employs small silicon bridges to connect multiple die together in the same package. HBM2 is a 3D, stacked-die DRAM module with a wide parallel interface that delivers significantly more bandwidth than DDR4 SDRAM memory with good power efficiency. HBM2 tiles deliver 4 GB of memory with 256 GBps of bandwidth. Some members of the Intel Stratix 10 DX FPGA family incorporate two HBM2 tiles for a total of 8 GB of DRAM with 512 GBps of bandwidth.

4

CA and HA soft agents for UPI support

Intel's QPI employed separate transaction agents: a Caching Agent (CA) located within each processor core and a Home Agent (HA) located in each memory controller. CAs and HAs perform different functions. A QPI CA initiates transactions into coherent memory on behalf of a processor core and may retain copies of data from those transactions in its own local cache structure. The CA can also provide copies of the coherent memory contents to other, requesting CAs. A QPI HA supervises a portion of the system's coherent memory and is associated with but is not specifically part of a memory controller. The HA must manage and resolve any conflicting transaction requests that arrive from different CAs in the system.

The UPI coherent interconnect merges the functions of these two agents into one Cache and Home Agent (CHA). Each processor core and each distributed last-level cache (LLC) slice within Intel Xeon Scalable processor family incorporate these merged CHAs, which therefore scale with the number of cores and LLC slices in a processor. Each CHA is responsible for tracking requests made by the processor cores and for responding to snoop requests from local and remote agents.

The CHA located at each of the LLC slices maps addresses to a specific LLC bank, to a memory controller, or to an I/O subsystem, and provides the routing information required to reach its destination through the coherent UPI mesh interconnect. CHAs also resolve memory coherency across multiple processor cores.

Separate UPI CA and HA soft agents for Intel Stratix 10 devices

Because Intel Stratix 10 DX devices are not processors, they do not require merged CHAs because they're not generally tasked with all of the duties assigned to processors. Consequently, the UPI implementation for Intel Stratix 10 DX devices includes separate CA and HA IP blocks, which are instantiated in the FPGA fabric as needed by the user's design, according to the tasks assigned to the FPGA. Splitting the CA and HA functions into separate IP blocks reduces the size of the IP blocks and the amount of programmable logic needed to support each UPI port in Intel Stratix 10 DX devices.

The EMIB anatomy of Intel Stratix 10 DX FPGAs

All Intel Stratix 10 FPGAs employ Intel EMIB packaging technology and Intel Stratix 10 DX FPGAs and SoCs are no different. All Intel Stratix 10 FPGAs incorporate a base die, which contains the FPGA's programmable-logic array and interconnect fabric, and multiple specialty tiles that connect to the base die through small silicon EMIB chiplets. These specialty tiles provide a broad assortment of I/O and memory capabilities to the Intel Stratix 10 FPGA family, as shown in Figure 3.
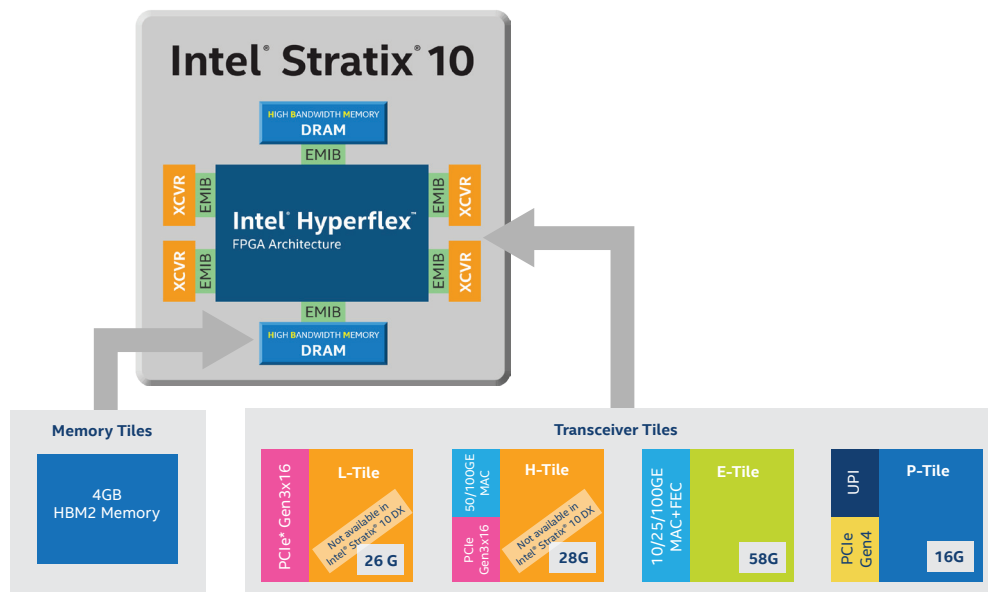

Figure 3. Intel Stratix 10 FPGAs Employ Intel EMIB Technology

As Figure 3 shows, with the introduction of the Intel Stratix 10 DX FPGAs, there are now four basic transceiver tiles and one memory tile used in the manufacture of Intel Stratix 10 FPGAs. The four transceiver tiles provide SERDES transceivers capable of different data rates, as fast as 58 Gbps in the case of the E-Tile. In addition, some members of the Intel Stratix 10 FPGA family incorporate high-speed HBM2 3D DRAM stacks, which also take advantage of Intel EMIB interconnect technology.

Previously announced members in the Intel Stratix 10 FPGA family portfolio include the Intel Stratix 10 GX, SX, TX, and MX FPGAs. Each of these Intel Stratix 10 FPGA types use the same base FPGA die, combined with different specialty I/O and memory tiles, as shown in Figure 4.
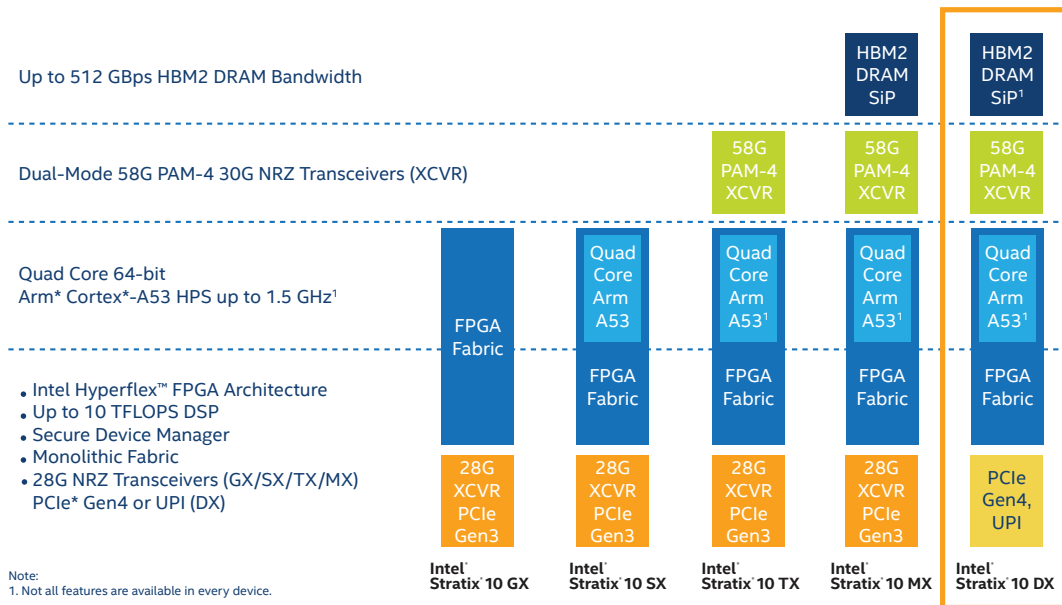

Figure 4. Intel Stratix 10 FPGA Portfolio

The I/O tile that differentiates Intel Stratix 10 DX devices from other members in the Intel Stratix 10 device families is the P-Tile, which incorporates the PCIe Gen4 and UPI interface blocks. The Intel Stratix 10 DX devices' PCIe port supports PCIe Gen3 or Gen4 x16 operations, with transfer rates as fast as 16 GT/s. Intel Stratix 10 DX FPGAs support coherent UPI transfer rates as fast as 11.2 Gbps using 20 lanes. All Intel Stratix 10 DX FPGAs also incorporate one E-Tile, which allows these devices to support multiple 100GE communications channels using multiple SERDES transceivers capable of operating at 58 Gbps using PAM-4 modulation.

Members of the Intel Stratix 10 DX Device Family

As shown in Figure 5, there are three members in the Intel Stratix 10 DX device family:

- The Intel Stratix 10 DX 1SD110 device, which incorporates one PCIe Gen4 x16 port, four 100G Ethernet ports, and a Hard Processor System consisting of a 1.5-GHz, Quad-core Arm* Cortex-A53 processor and peripherals

- The Intel Stratix 10 DX 1SD210 device with three PCIe Gen4 x16 or three UPI ports, four 100G Ethernet ports, and 8 GB of HBM2 DRAM

- The Intel Stratix 10 DX 1SD280 device with four PCIe Gen4 ports or three UPI ports, and two 100G Ethernet ports

Table 1 illustrates the features of each of these three Intel Stratix 10 DX devices and Figure 5 illustrates three possible use cases for these devices.

INTEL® STRATIX®10 DX DEVICE FAMILY[1]														
FPGA CORE		**MEMORY**			**DSP**			**GPIO**	**XCVR**	**HARD IP**				**HBM2**
INTEL STRATIX 10 DX DEVICE	**LOGIC ELEMENT (KLE)**	**M20K Bits (Mbits)**	**MLAB (Mbits)**	**eSRAM (Mbits)**	**18x19 MULTS**	**PEAK TFLOPS**	**PEAK TMACS**	**MAX GPIO**	**MAX XCVR**	**PCIe* GEN4**	**UPI**	**100GE**	**HPS[2]**	**DRAM STACK**
1SD110	1,325	107	7	-	5,184	4.1	10.4	528	32	1	-	4	Yes	-
1SD210	2,073	134	11	90	7,920	6.3	15.8	612	84	3 or 3	4	-	-	2
1SD280	2,753	229	15	-	11,520	9.2	23.0	816	84	4 or 3	2	-	-	-

Notes:
[1] Preliminary and subject to changes
[2] HPS = Quad-core Arm* Cortex*-A53 (1.5 Ghz) + peripherals

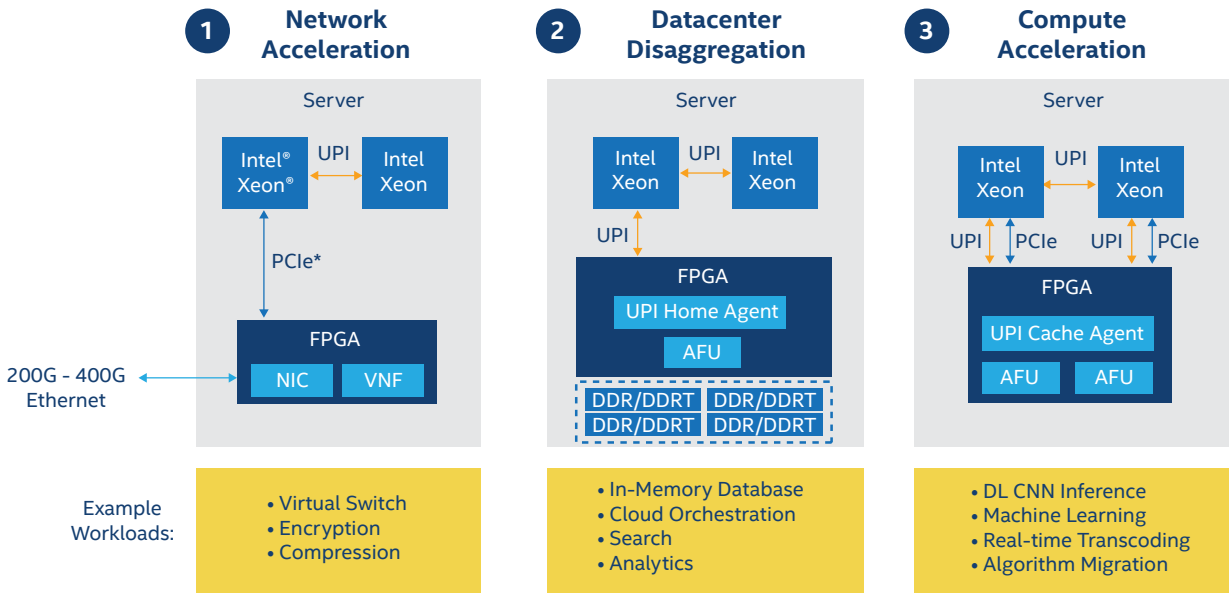**Table 1.** Intel Stratix 10 DX Device Family Members

Figure 5. Use cases for Intel Stratix 10 DX Devices

Example 1 in Figure 5 shows an Intel Stratix 10 DX FPGA connected to a host Intel Xeon Scalable processor using a PCIe Gen4 link to create a network accelerator. Example workloads for such an accelerator include a virtual switch, encryption/decryption, and compression/decompression. An FPGA-based hardware network accelerator such as the one labeled as Example 1, frees CPU cycles that can be better utilized for high-value, revenue-generating tasks in the data center.

Example 2 in Figure 5 shows an Intel Stratix 10 DX FPGA connected to a pair of Intel Xeon Scalable processors using a coherent UPI link. The Intel Stratix 10 FPGA is attached to and controls a large amount of local SDRAM and/or Optane DC persistent memory (labeled DDRT in Figure 5). The FPGA can use this large bank of attached memory to perform in-memory database operations including searches and data analytics using processing engines instantiated in the FPGA's programmable logic. Even faster operation is possible through the use of multiple FPGAs, each processing a portion of the database in parallel using its local memory.

Example 3 in Figure 5 shows an Intel Stratix 10 DX FPGA connected to a pair of Intel Xeon Scalable processors through multiple PCIe Gen4 and coherent UPI ports. In this configuration, the Intel Stratix 10 DX FPGA can operate either as a full peer with the Intel Xeon Scalable processors or as a peripheral device over the multiple PCIe Gen4 ports. Applications for this system configuration include deep learning, machine learning, real-time video transcoding, and general algorithmic offload through multiple accelerator functional units (AFUs).

Intel Quartus® Prime software support for PCIe Gen4 and UPI

The Intel® Quartus® Prime sofware includes connectivity and programming support for the Intel Stratix 10 DX devices' PCIe Gen4 support. Access to UPI support including the HA and CA IP blocks requires Intel approval and a special license.

Conclusion

Meeting high-bandwidth and low-latency requirements for hardware accelerators in high-performance servers presents a significant design challenge. Applications for such accelerators include diverse applications such as cloud disaggregation, financial/HPC, encryption/decryption, compression/decompression, video transcoding, storage, genomics, database acceleration/data analytics, and general compute acceleration.

The design of these accelerators must start with a solid technology foundation, which includes high-speed interfaces that support the system's bandwidth and latency requirements. Two such high-speed interfaces are UPI, which specifically supports Intel Xeon Scalable processors, and PCIe Gen4. Both of these interfaces are available in new Intel Stratix 10 DX FPGA family members, which can be used to develop hardware accelerators for any of the applications listed above.

References

- Intel Xeon Processor Scalable Family Technical Overview

- PCI-SIG SR-IOV Primer

- SR-IOV for NFV Solutions: Practical Considerations and Thoughts

- Intel Scalable I/O Virtualization Technical Specification

- Intel's View of the Chiplet Revolution

Where to Find More Information

For more information about Intel and the Intel Stratix 10 FPGAs, visit www.intel.com/content/www/us/en/products/programmable/fpga/stratix-10.html

♻ Please Recycle

WP-01288-1.0