NVMe PCIe Gen4 Host Controller Core by leveraging Xilinx's UltraScale+ GTY Transceivers

By Design Gateway Co., Ltd.
2020-6-2

Overview The Xilinx UltraScale+™ family features with GTY transceiver which has the capability of supporting a PCI Express® Gen4 interface. Design Gateway's NVMe Host Controller IP core is designed to leverage GTY Transceivers to support the latest NVMe SSD drive PCIe Gen4 technology. This article demonstrates the solution of NVMe Host Controller IP core implementation on Xilinx's Virtex UltraScale+ Evaluation Kit by using Design Gateway's NVMeG4-IP Core which is able to achieve incredibly fast read/write performance more than 4GB/s speed.

Introduction to PCI Express® and technology evolution

PCI stands for Peripheral Component Interconnect. It was introduced back in the early 90's by a group of engineers in Intel, AMD, and other companies and now is a standard connection interface attaching different computer components to a motherboard.

PCI Express® (PCIe) was initially referred to as HSI (for High Speed Interconnect), and underwent a name change to 3GIO (for 3rd Generation I/O) before finally settling on its PCI Special Interest Group (PCI-SIG) name PCI Express.

PCIe provides a point-to-point interconnect solution for communication between two devices. It is scalable, and slots come in different configurations of bidirectional lanes: x1, x4, x8, x16, x32. The number represents the number of lanes in the PCIe slot.

PCIe standard evolved from PCIe 1.0(Gen1), released in 2003 supporting 2.5 gigabit per second (GT/s), to PCIe 5.0(Gen5), released in 2019 supporting 32 GT/s. Figure 1 shows the evolution of the standard and the bandwidth doubling with each generation.

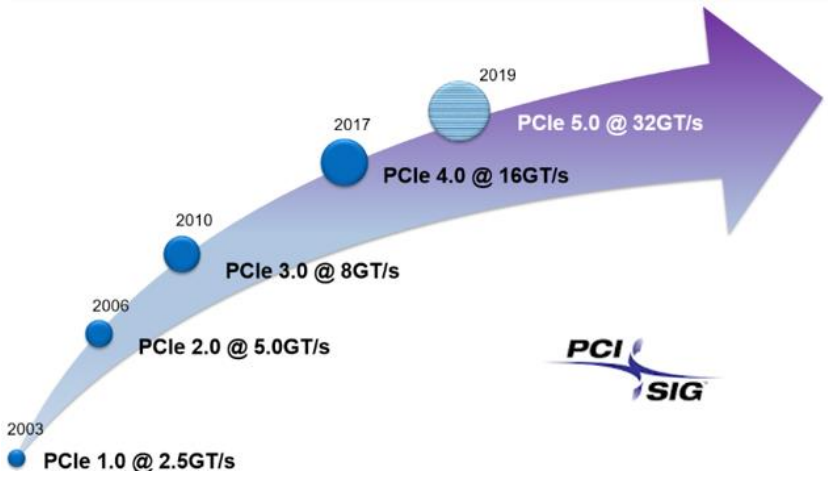

Figure 1: PCI Express® Technology trend/roadmap (Image source: PCI-SIG)

Introduction to NVMe SSD Storage

NVM Express® (NVMe) defines the interface for the host controller to access as SSD by PCIe. NVMe performance evolves together with PCIe performance since PCIe Gen1 to the latest PCIe 4.0(Gen4) which now available in the market. NVMe drives have paved the way for data storage and computing at very high speeds. By using PCIe Gen4 technology, modern NVMe SSD drives can achieve speeds as high as 64 GT/s peak performance.

NVMe optimizes issuing the command and completion process by using only two registers (Command issue and Command completion). Otherwise, NVMe supports parallel operation by supporting up to 64K commands within a single queue. 64K command entries improve transfer performance for both sequential and random access.

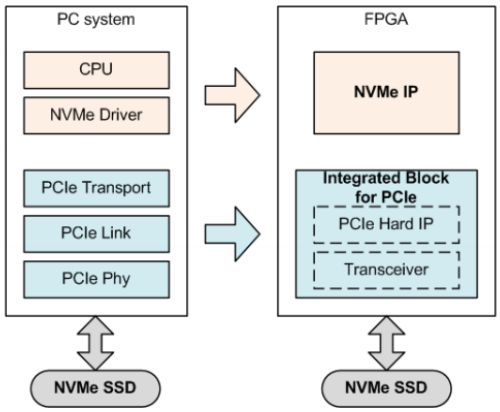

Figure 2: NVMe Protocol layer (Image source: Design Gateway)

To access NVMe SSD, the general system implements NVMe driver running on the processor, as shown in the left side of Figure 2. The physical connection of NVMe standard is PCIe connector which is one-to-one type, so one PCIe host can connect to one PCIe device.

An example of an NVMe storage device is shown in http://www.nvmexpress.org/products/.

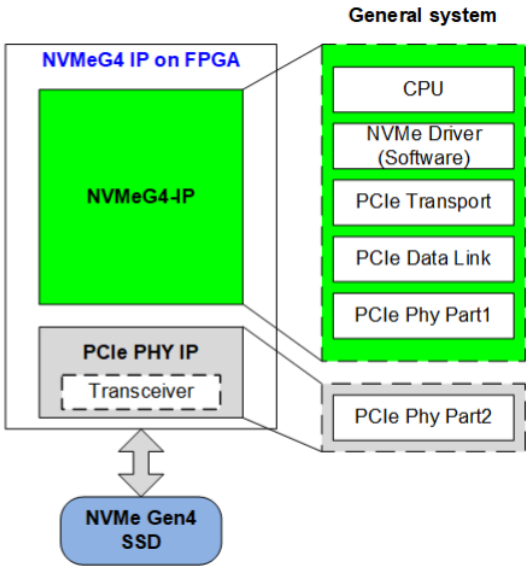

Figure 3: NVMe Implementation. (Image source: Design Gateway)

Conventionally, the NVMe host is implemented by using a Host Processor operating with a PCIe Controller for transferring data to and from the NVMe SSD. NVMe protocol is implemented for device driver communications with the PCIe controller hardware CPU peripheral connected through a very high-speed bus. External DDR memory is required for data buffering and command queue to transfer the data between the PCIe controller and SSD.

Since UltraScale+ features with GTY transceiver which capable of PCIe Gen4 interface support but an PCIe Gen4 integrated Block doesn't available on some device. Design Gateway proposes a solution by using the NVMeG4-IP Core, as shown in Figure 3, to enable a NVMe SSD interface for a UltraScale+ device utilizing GTY transceivers. The NVMeG4-IP is also suitable for building of a multi-channel RAID system with very high performance and the lowest possible FPGA resource consumption.
The NVMeG4-IP core license includes the example reference design that helps designers reduce development time and cost.

Overview of Design Gateway's NVMeG4-IP

NVMe IP Core with PCIe Gen4 Soft IP (NVMeG4-IP) is ideal to access a NVMe SSD without a PCIe integrated block, CPU, and external memory. NVMeG4-IP includes PCIe Gen4 Soft IP and 256 Kbyte memory. This solution is recommended for the application which requires NVMe SSD storage with ultra high speed performance by using a low-cost FPGA which does not contain a PCIe integrated block.

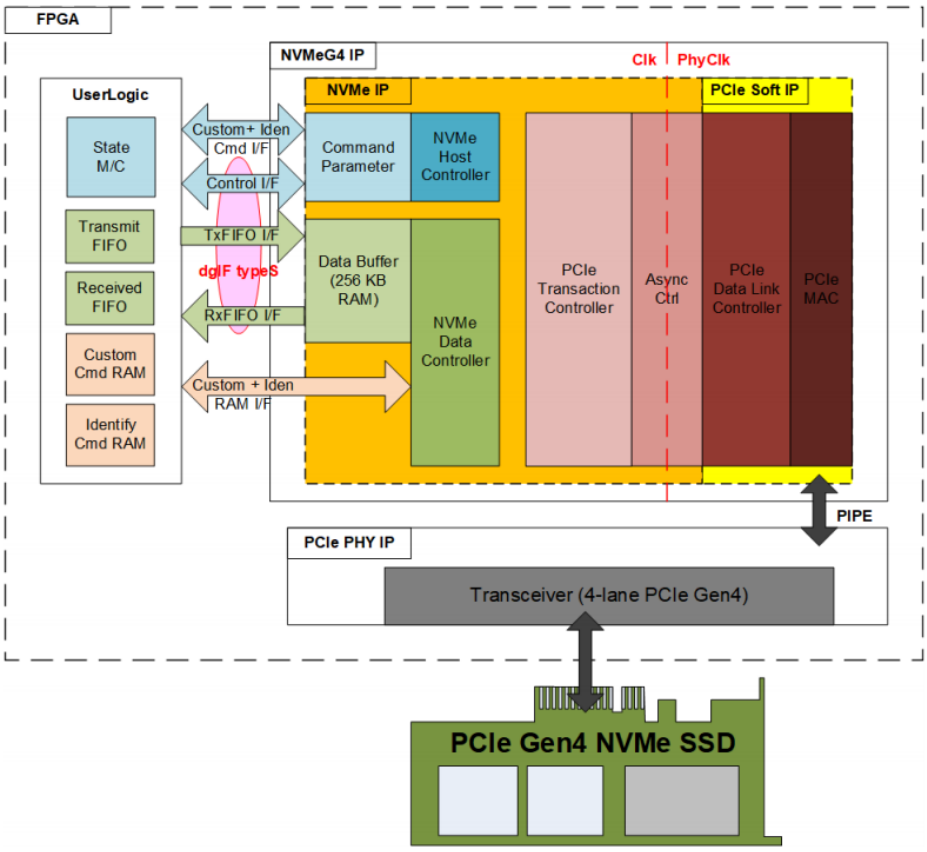

Figure 4: NVMeG4-IP block diagram. (Image source: Design Gateway)

NVMeG4-IP's Features

NVMeG4-IP has many features, some of which are highlighted below:

- Implement application layer, transaction layer, data link layer, and some parts of the physical layer to access the NVMe SSD without CPU usage
- Operate with Xilinx PCIe PHY IP configured as a 4-lane PCIe Gen4 (256-bit bus interface)
- Includes 256 Kbyte RAM data buffer
- Simple user interface via dgIF typeS
- Supports six commands, i.e. Identify, Shutdown, Write, Read, SMART, and Flush (support additional command as optional)
- Supported NVMe device:
  - Base Class Code:01h (mass storage), Sub Class Code:08h (Non-volatile), Programming Interface:02h (NVMHCI)

- MPSMIN (Memory Page Size Minimum): 0 (4Kbyte)
- MDTS (Maximum Data Transfer Size): At least 5 (128 Kbyte) or 0 (no limitation)
- LBA unit: 512 byte or 4096 byte

- User clock frequency must be more than or equal to PCIe clock (250MHz for Gen4)
- Available reference design:
  - ZCU102 with AB17-M2FMC adapter board
  - KCU105 with AB18-PCIeX16/AB16-PCIeXOVR adapter board.
  - VCU118 with AB18-PCIeX16 adapter board

Design Gateway developed the NVMeG4-IP to run as a NVMe host controller for accessing a NVMe SSD. The user interface and standard features are designed for ease of usage without needing knowledge of the NVMe protocol. The additional feature of NVMeG4-IP is the built-in PCIe soft IP which implements the Data link layer and some parts of the Physical layer of the PCIe protocol by pure logic. So, NVMeG4-IP can run in an FPGA, which does not have a PCIe integrated block, by using built-in PCIe soft IP and Xilinx PCIe PHY IP core. Xilinx PCIe PHY IP is a free IP core available which includes a transceiver and a logic equalizer.

NVMeG4-IP supports six NVMe commands, i.e. Identify, Shutdown, Write, Read, SMART, and Flush. 256 Kbyte Memory Block (UltraRAM and BlockRAM) is integrated in the NVMeG4-IP to act as a data buffer. The system does not need a CPU and external memory. More details of the NVMeG4-IP are described in its datasheet which can be downloaded from our website: https://dgway.com/products/IP/NVMe-IP/dg_nvmeg4_ip_data_sheet_xilinx_en.pdf

FPGA resource usages on the XCVU9P-FLGA2104-2L FPGA device are shown in Table 1 below

As shown in Table 2, the main advantage of NVMeG4 IP is the performance which operates at PCIe Gen4. Also, the maximum number of connected SSDs is not limited by the number of PCIe hard IP but

Table 1: Example Implementation Statistics for Ultrascale/Ultrascale+ device

Family	Example Device	Fmax (MHz)	CLB Regs	CLB LUTs	CLB	IOB	BRAM Tile	URAM	Design Tools
Virtex-Ultrascale+	XCVU9P-FLGA2104-2L	300	19639	21029	4397	-	12	8	Vivado2019.1

limited by the number of transceivers and the resource utilization. However, the disadvantage of NVMeG4 IP is the resource utilization which is large for implementing the PCIe soft IP. Also, there is limitation that NVMeG4 IP can support only 4-lane PCIe Gen4 SSD.

Feature	NVMe IP	NVMeG4 IP
PCIe Interface	128-bit AXI4 Stream	256-bit PIPE
Xilinx PCIe IP	Integrated Block for PCIe (PCIe Hard IP)	**PCIe PHY IP (Transceiver and equalizer)**
PCIe Hard IP	Necessary	**Not use**
PCIe Speed	Gen3 or lower, depending on Xilinx IP	**Support only 4-lane PCIe Gen4**
User Interface	dgIF typeS	dgIF typeS
FPGA resource	**Smaller (859 CLBs)**	Larger (3294 CLBs)
Memory resource	66 BRAMTile (URAM support as optional)	**12 BRAMTiles and 8 URAM blocks**
Maximum SSD	Depend on the number of PCIe Hard IPs (6 PCIe Gen3x16 = 6 SSDs on VU9P)	**Depend on the number of transceivers** **(120 GTY = 30 SSDs on VU9P)**
SSD Performance	Write : Up to 2300 MB/s Read : Up to 3400 MB/s	**Write : Up to 4300 MB/s** **Read : Up to 4700 MB/s***

Table 2: DG's NVMe IP Core comparison between NVMeG4-IP and NVMe-IP

Implementation and performance on VCU118

Figure 5 shows the overview of the reference design based on the VCU118 to demonstrate NVMeG4-IP operation. The NVMeG4IPTest module in the demo system includes with following modules: TestGen, LAxi2Reg, CtmRAM, IdenRAM and FIFO.

For more detail of NVMeG4-IP reference design, please refer to the NVMeG4-IP reference design document provided on Design Gateway's website here: https://dgway.com/products/IP/NVMe-IP/dg_nvmeg4ip_refdesign_xilinx_en.pdf

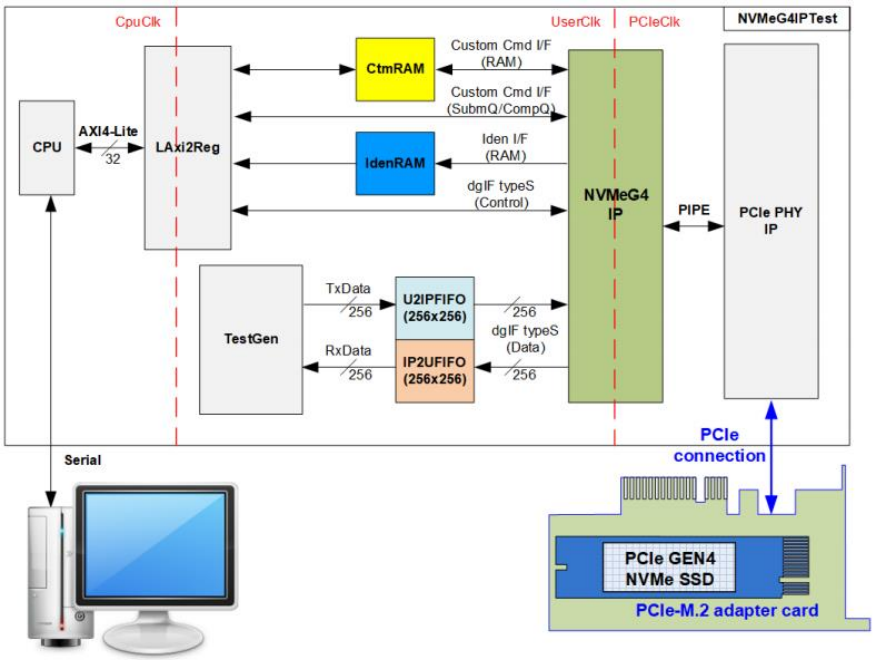

Figure 5: NVMeG4-IP reference design overview. (Image source: Design Gateway)

The demo system is designed to write/verify data with the NVMe SSD on the VCU118. The user controls the test operation through a Serial console. For the NVMe SSD to interface with the VCU118, an AB17-M2FMC adapter board is required as shown in Figure 5.

Figure 6 shown demo hardware set up on VCU118 board. For more detail of the board setting and environment set up to operate this demo, please see the NVMeG4-IP demo instruction document provided on Design Gateway's website here:
https://dgway.com/products/IP/NVMe-IP/dg_nvmeg4ip_instruction_xilinx_en.pdf

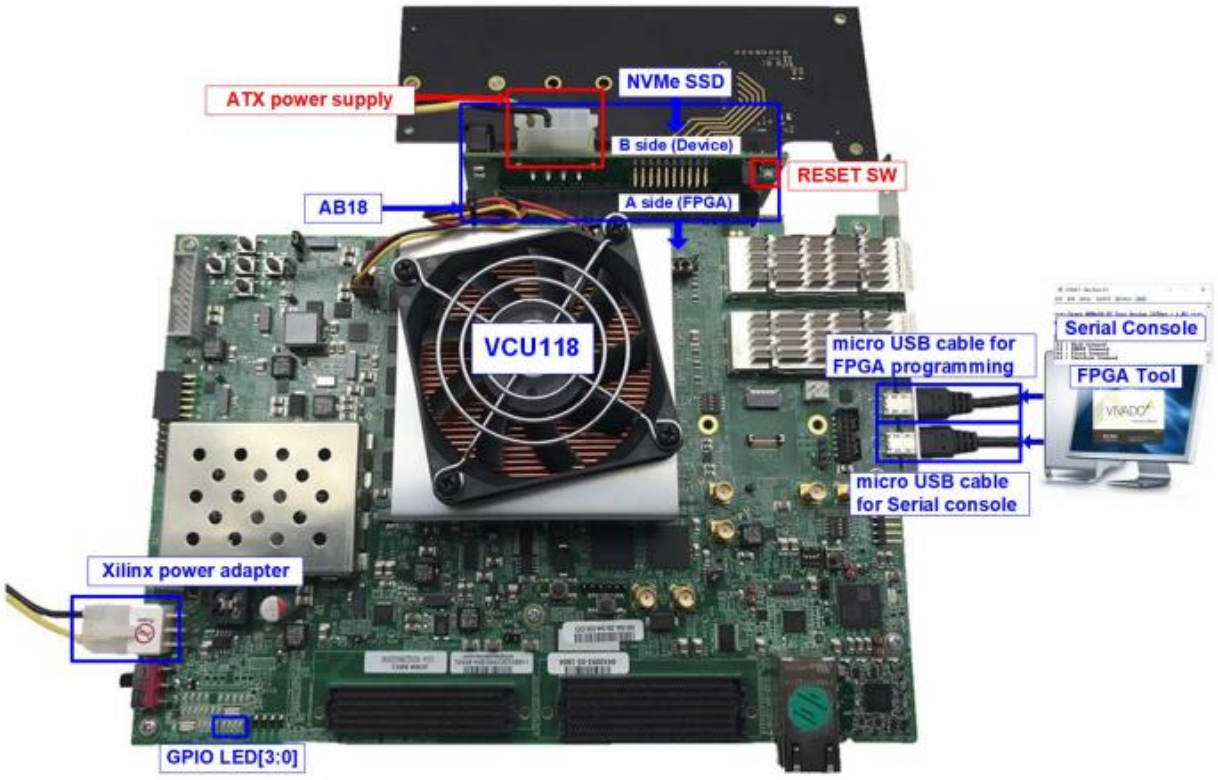

Figure 6: NVMeG4-IP demo environment set up on VCU118. (Image source: Design Gateway)

The example test result when running the demo system on the VCU118 while using the 1 TB Gigabyte AORUS NVMe PCIe Gen4 SSD is shown in Figure 7.

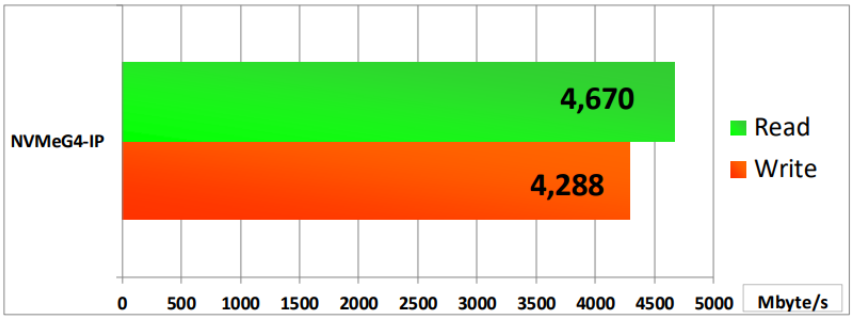

Figure 7: NVMe SSD read/write performance on VCU118 by using Gigabyte AORUS NVMe PCIe Gen4 SSD *SSD. (Image source: Design Gateway)*

Conclusion

NVMeG4-IP Core provides a solution to enable NVMe PCIe Gen4 SSD interface on the VCU118 evaluation kit and also the solution for Xilinx®'s UltraScale+™ device family features with GTY transceiver but a PCIe Gen4 integrated block doesn't available. NVMeG4-IP is designed with the goal of achieving the highest possible performance with the lowest possible FPGA resource usage for NVMe SSD access without requiring a CPU. It's very suitable for high performance NVMe storage without CPU invention and able to implement multiple NVMe SSD interfaces by utilizing GTY transceivers with no limitations from the number of available PCIe integrated blocks on the FPGA device.

For more detail of NVMeG4-IP and available reference design, please visit Design Gateway's website at https://dgway.com/NVMeG4-IP_X_E.html.