Stand alone NVMeOF Acceleration Solution

U50 based 100Gb ethernet NVMe Storage JBOF with Application Acceleration

SOLUTION BRIEF

- 100Gb/s NVMeOF Alveo U50 based solution
- Support for up to 24 NVMe SSDs
- Six U50s in a single chassis
- 100GbE line rate performance
- Ability to add Hardware Accelerators

INTRODUCTION

Unlike most of the JBOF designs in today's storage marketplace, this Xilinx NVMe-Over-Fabrics (NVMeOF™) reference was created with the idea of adding computational storage into next generation networked storage solutions. Using Remote Direct Memory Access (RDMA) this design provides a low latency, high performance, industry standard interconnect for up to 24 NVMe SSDs. This platform provides the flexibility to define custom acceleration functions within an NVMeOF compliant environment and eliminates the need for an external processor or Network-Interface-Card (NIC) thus enabling a highly integrated and cost-effective JBOF storage solution.

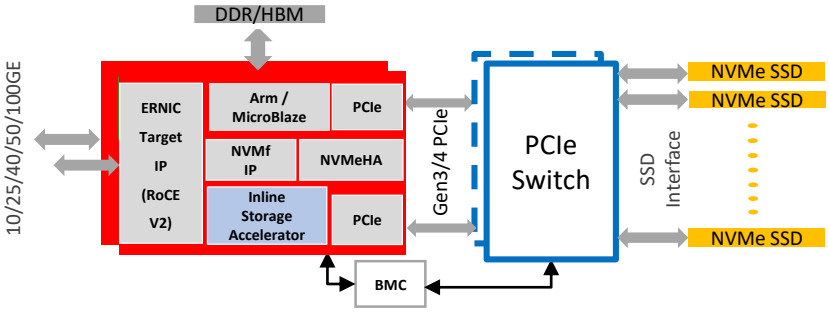

Figure 1

SOLUTION OVERVIEW

This design transports NVMe packets over a standard RDMA Ethernet fabrics allowing the creation of remote storage namespaces, which enables the efficient pooling and sharing of storage resources across datacenter servers. These remote namespaces can dramatically reduce the storage cost, footprint, and power within datacenters. With additional capability for hardware acceleration services via namespace management, the Alveo U50 solution becomes a true disaggregated computational storage accelerator.

This Xilinx solution provides reliable transport of NVMe frames with low latency, high throughput and massive scalability to remote hosts. A block diagram of a typical system solution is depicted above in Figure 1. The Xilinx NVMeOF reference design implements the NVM express Over Fabric protocol and the RDMA NIC protocol in the single highly integrated Xilinx FPGA contained in an Alveo U50 add in card with a significant amount of programable logic remaining for use as computational storage accelerators.

The key data transfer commands in the NVMeOF protocol are offloaded to hardware, while the embedded CPU in the Xilinx device processes the control plane commands giving this Xilinx solution significant performance advantages over processor-only implementations. The Xilinx ERNIC IP integrated into this reference design provides reliable transport, flexibility in network interconnect and the performance to support line speed bandwidth. An implementation supporting 24 drives can be accomplished using just over 200K LUTs leaving the rest of the U50's resources for customizable acceleration engines or other types of differentiation.

Stand alone NVMeOF Acceleration Solution

U50 based 100Gb Ethernet NVMe Storage JBOF with Application Acceleration

SOLUTION DETAILS

Feature Overview	Description
Supported RDMA Protocol	RoCEv2
Network Side Interface	Up to 100Gb Ethernet 100GbE, 50GbE, 40GbE, 25GbE and 10GbE
SSD side Interface	Up to one PCIe Gen3 x16 or two Gen4 x8 interfaces
Send and Receive Queue-Pairs (QPs)	Up to 32
Completion Queues (CQ) Queue Depth	64 entries per queue
Latency	Hardware design limited only
Performance (Single U50, 4KB, 8 SSDs, OIO=64)	2.5M IOPs
Management Interfaces	SMBus, NVMe-MI
Future support	NVMe 1.4+ NVMe 1.1+ NVMe/TCP
Inline Accelerator Examples	Storage services: • (De)Compression • (De)Encryption • Data protection Database Acceleration: • Scan • Filter • Aggregate
Resource Utilization	The full solution resources include NVMF, ERNIC, CMAC, NVMeHA, AXI-DMA and DDR-MIG IPs. The resource utilization depends on selected configuration and additional accelerators.

TAKE THE NEXT STEP:

Register for solution access at https://www.xilinx.com/member/nvmeof-bitfile.html

Learn more about Alveo accelerators https://www.xilinx.com/products/boards-and-kits/alveo.html