Flash Memory Summit

Advancing the Verification of SSDs over Fabrics

Dhruv Garg

Siemens EDA

Agenda

- High Performance Remote Storage Access
  - ❑Challenges
  - ❑NVMe-oF (NoF) Features
  - ❑NoF Variants
  - ❑Performance Metrices
- NoF QVIP Solution
  - ❑Basic Characteristics
  - ❑NoF QVIP Features
  - ❑Recommendations for trouble free debugging

High Performance Remote Storage Access

Challenges

- ❑Increase in devices accessing plethora of content available online
- ❑Rise in real time applications
- ❑Rise in need for fast data processing
- ❑Requirement for *real time data interoperability*

NVMe-oF (NoF) Features

- Low latency efficient storage protocol over fabrics to provide large scale sharing of storage over distance
- Designed to work with the range of storage networking fabrics
  - NoF using RDMA
  - NoF using Fibre Channel
  - NoF using TCP/IP
- Uses multi-queue model of NVMe resulting in low latency design

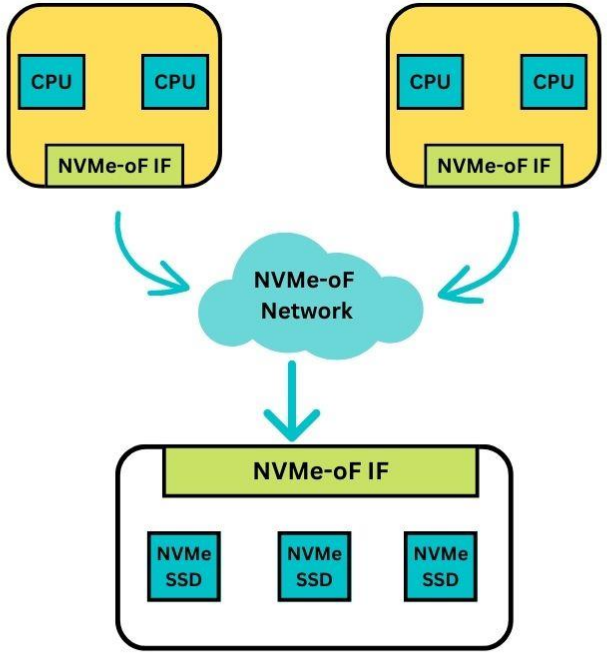

Fig 1: NVMe-oF usecase

Contd…

- ❑Capsules based transmission to boost real time data transmission
- ❑Low latency
- ❑Data protection and security
- ❑Support to various Fast Interconnect fabrics
- ❑Well defined Discovery mechanism to work with multiple types of transports
- ❑Reliable and Credit based flow control
- ❑Multi-Host support
- ❑Multi-channel support

NoF Variants

- NVMe-oF over RDMA
  - RoCE
  - Infiniband
  - iWARP
- NVMe-oF over Fiber Channel
- NVMe-oF over TCP/IP

RDMA

- Used for transferring data from address space of one memory to another with CPU involvement.
- Reduced latency: queue pair mechanism similar to NVMe
- Recommended for High speed computation applications: AI, Real time Data analytics, cloud computing.

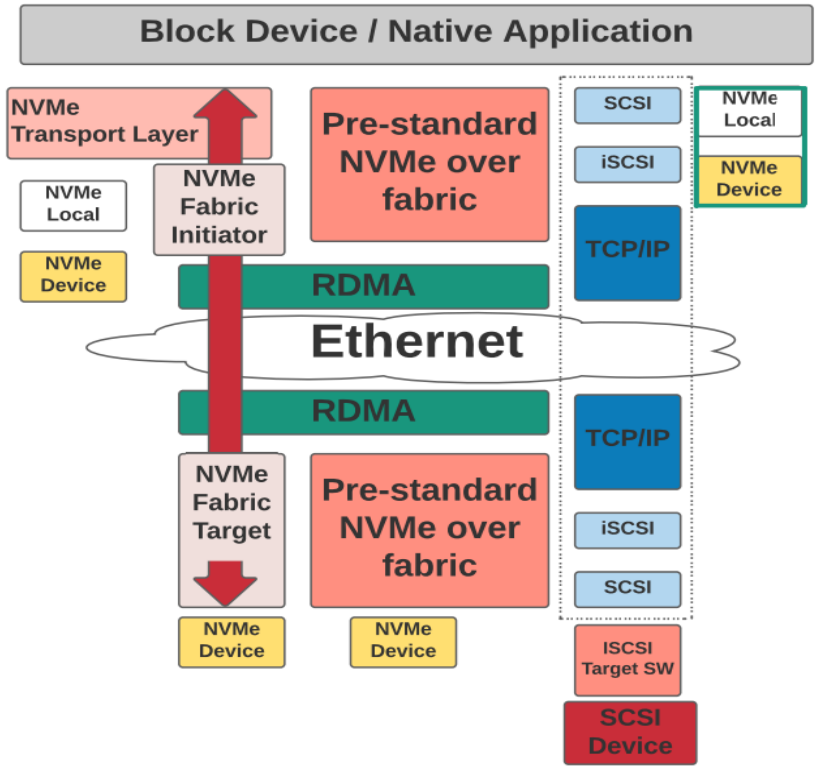

Fig 3: NVMe over RDMA, TCP/IP and Base NVMe

Fibre Channel

- Minimum implementation cost: utilizes the existing Fibre channel
- High bandwidth and respectable throughput despite reduced performance due to SCSI to NVMe translations as compared to other supported Transport layers.

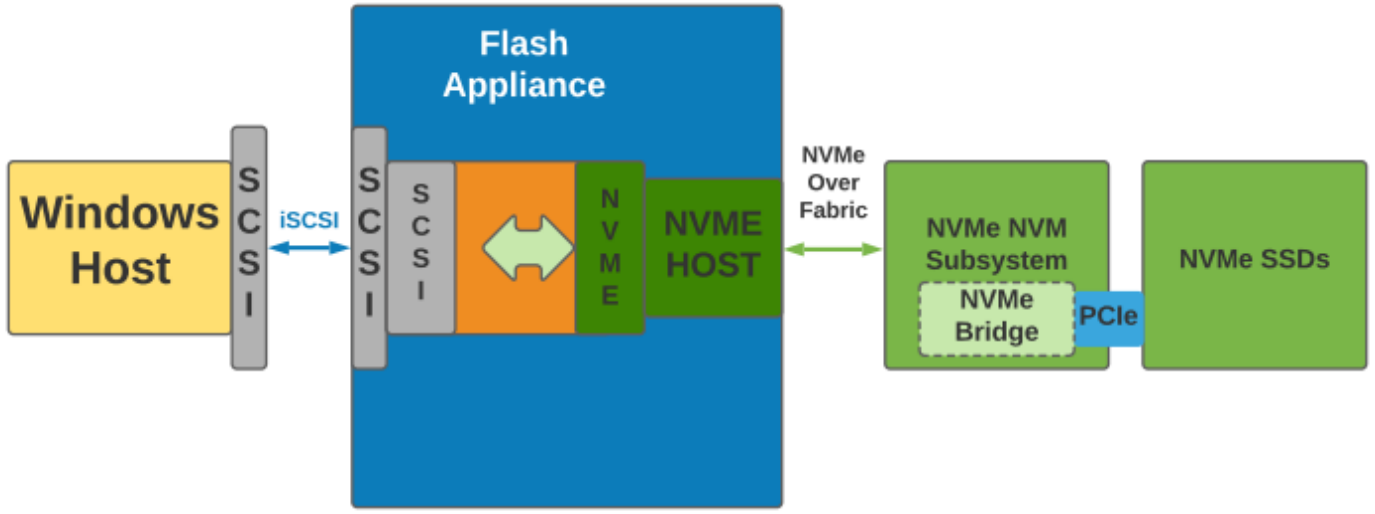

Fig 2: NVMe-oF Using Fibre Channel showing SCSI to NVMe Transalations

TCP/IP

- Flexibility to work over any Ethernet network
- TCP/IP commonly accepted protocol for communication over Network
- Compatibility with the existing infrastructure to communicate over the network
- Number of layers overshadows the advantages NoF offers

Performance Metrices

Underlying Transport	Latency	Performance	Cost
Fibre Channel	Lower	High	Medium
RoCE (RDMA over Converged Ethernet)	Lowest	High	High
TCP/IP	High	Medium	Low

Fig 4: Performance Vectors and Compatible Transport Mechanisms

NoF QVIP Solution

Basic characteristics

- Coverage driven verification
  - ✓Require an exhaustive test plan based on specification
  - ✓Achieving 100% coverage ensures complete functional testing based on test plan
  - ✓Particularly helpful in finding bugs for complex design code
- Assertion based verification
  - ✓Highlights any protocol violation
  - ✓Ensures bug detection exactly when and where they occur
  - ✓Significantly reduces complex debug process

NoF QVIP Features

- ❑Support for NoF specification 1.1
- ❑Acknowledgement Error Handling mechanism
- ❑NoF Command Capsules
- ❑SGL support for in-capsule Data Transfer
- ❑NoF with RDMA : Alignment with RDMA which also used queuing mechanism for transport

Contd…

- Agility across storage Command Set
  - Highly configurable command structure
  - Wide pool of APIs to set/get command fields
  - Simultaneous execution of commands belonging to different command sets
  - Wide range of command sets
    - NoF Command Set
    - NVM Command Set
    - ZNS Command Set
    - KV Command Set
  - Variety of configurations related to features and customer ease of use

Contd….

- Convenient stimulus generation
  - ❑Built in sequence library to generate stimulus
  - ❑Randomization of sequences to target all scenarios
  - ❑Error injection in command to verify erroneous behavior
  - ❑Various modes for command execution, e.g., blocking, non-blocking
  - ❑NoF Bring-up Sequence
    - ✓Backdoor
    - ✓Frontdoor

Easier debugging

- Support for strong built-in debug
  - ✓Debug ID associated with each command
  - ✓ Complete command parameters

DEBUG ID	BDF	SRC	RW	TYPE	QID	QPN	CMD	MISC	CID	SGL					NSID	STS
										TYPE	SUBT	KEY	LEN	ADDR		
0000000000000001	----	H	W	SQE	0	2	CONNECT	-----	0000	KEYED	0	888D329B	00000400	00002B37C4D46000	00000001	-------
nof_host_beat.log [+]																
000100010000001A	0001	H	W	SQE	1	4	WR	2	0000	KEYED	0	49C25196	00000630	0000841912DD7000	00000001	-------
000100010000001B	0001	H	W	SQE	1	4	WR	3	0001	KEYED	0	BAD94321	00000840	0000BAA117772000	00000001	-------
000100010000001C	0001	H	W	SQE	1	4	WR	4	0002	KEYED	0	B8E222A7	00000A50	0000F33D2F2AE000	00000001	-------
000100010000001D	0001	H	W	SQE	1	4	WR	5	0003	KEYED	0	A27FD4C8	00000C60	00000F14B2F6C000	00000001	-------
000100010000001E	0001	H	W	SQE	1	4	WR	6	0004	KEYED	0	D07FD110	00000E70	0000BF7EA4407000	00000001	-------
000100010000001A	0001	D	R	SGLD	1	4	----------	-----	----	--------	----	------	------	--------------	--------	-------
000100010000001A	0001	D	W	CQE	1	4	----------	00005	0000	--------	----	------	------	--------------	--------	GEN00
000100010000001B	0001	D	R	SGLD	1	4	----------	-----	----	--------	----	------	------	--------------	--------	-------
000100010000001B	0001	D	W	CQE	1	4	----------	00005	0001	--------	----	------	------	--------------	--------	GEN00
000100010000001C	0001	D	R	SGLD	1	4	----------	-----	----	--------	----	------	------	--------------	--------	-------

Fig 5: Unique debug ID to identify each command

Contd…

- Identify NSIDs out of a large pool
  - ✓Print frequently required characteristics of namespace

ENABLED COMMAND SETS ON BDF 0x100

COMMAND SET	STATUS
NVM	ENABLED
ZNS	ENABLED
KV	ENABLED

NAMESPACE DATA STRCT BDF 0x100

FORMAT NUM	IS EXT LBA	LBA DATA SIZE	METADATA SIZE	NO. OF ZONES	NS TYPE	NSID	PIF	ZONE SIZE
0	1	4096	16	NA	NVME_NVM_CS	0x1	NVME_32B_GUARD	NA
1	1	4096	16	NA	NVME_NVM_CS	0x2	NVME_64B_GUARD	NA
2	1	4096	16	NA	NVME_NVM_CS	0x3	NVME_16B_GUARD	NA
0	1	4096	16	4	NVME_ZNS_CS	0x4	NVME_32B_GUARD	262
1	1	4096	16	4	NVME_ZNS_CS	0x5	NVME_64B_GUARD	262
2	1	4096	16	4	NVME_ZNS_CS	0x6	NVME_16B_GUARD	262

FORMAT NUM	MAX KEY LEN	MAX NUM KEYS	MAX VAL LEN	NS TYPE	NSID
0	16	1000	8192	NVME_KV_CS	0x7

Fig 6: Useful characteristics of namespaces

Thank You !!!