Fabric accelerators for NVMe-oF and NVMe/TCP storage arrays with MRAM

Pankaj Bishnoi
Director of Business Development
Everspin Technologies
Drivers of Modern Compute Architectures

Focus for Fabric Accelerator

NVMe Ecosystem
(SSID, NVMe over Fabrics)

Compute & Storage Accelerators
(GPGPU, FPGA etc.)

High Performance Fabric Accelerators
(< 1µS latency, 100G+ BW)

Persistent Memory Technologies
(STT-MRAM, Optane, NVDIMM-N & NVDIMM-P)

Next Gen Auto
(ADAS, Connected Car, Infotainment etc.)
High Performance Network / Fabric Drivers

- **Increasing Interface speeds**
  - 40 Gbps
  - 50 Gbps
  - 100Gbps+

- **Leading network technologies**
  - RoCE (RDMA over Converged Ethernet)
  - InfiniBand (RDMA based)
  - Low latency TCPDirect

- **New generation of acceleration** (CPU & Storage Offloads)
  - NVMeOF target offload
  - NVMe/TCP offload
  - RoCE protocol offload
  - TCP/UDP/IP stateless offload
  - Block level encryption
Fabric Accelerator Purpose
Higher Performance & Agility

- Provide sub-μSec latency from wire to application data persistence
  - Kernel bypass
  - Host CPU bypass
  - Host memory bypass
  - Peer-to-Peer data transfers
  - RDMA termination

- Offload CPU computation cycles

- Customer configurable offload engines
  - ARM CPU code or FPGA code

- Provide higher write/read data throughput

- Enable simpler, lower power and lower cost appliance designs
  - Without need for x86 Server CPUs i.e. target ARM
Data flow for NVMe-oF or NVMe/TCP

Completion Options:
1. Wait till data written to persistent media (ex: NVMe SSD)
2. Use NVDIMM as system memory

Latency driven by completion path to persistent device
What is STT-MRAM Persistent Memory?

**PERSISTENCE**
Maintains memory contents without requiring power

**PERFORMANCE**
SRAM & DRAM-like performance with low latency

**ENDURANCE**
Superior durability supports memory workloads without sophisticated management

**RELIABILITY**
Best-in-class robustness designed and tested for extreme conditions
1Gb STT-MRAM Device Characteristics

- **Persistent Memory**: 1Gb Capacity
- **Fast Performance**: 2.66GBps @ 16-bit Read or Write
- **Data Retention**: Up to 10Yr @ 85C Reliability
- **Data Endurance**: 10 Billion cycles Reads or Writes
- **Read/Write Size**: Byte Level Access
- **Operating Temps**: 0° C to +85° C
- **Design Flexibility**: DDR4 (x8/x16) Interfaces
- **System Flexibility**: DDR Compatible BGA Device Packages
Completion architecture options with STT-MRAM:

1. On-board the NVMe Fabric storage offload device
2. Sister acceleration card next to offload device
STT-MRAM Value Proposition & Benefits

- **Higher Performance with Bypass Assist**
  - Acts as power loss protected write burst data buffer on the fabric/network controller card for offload engines
  - Providing at point persistent write data completion
  - Eliminates the multi-microseconds latency path (host CPU – Kernel – host memory – application – storage stack – persistent device) before data can be committed to a persistent device
  - Act as a persistent RDMA burst buffer. Must investigate opportunity related to SNIA Proposed extensions for RDMA commit aka “RDMA Flush”

- **Provide bigger working persistent memory region on offload device**

- **No Batteries** – Natively power loss protected persistent memory

- **NVMe-oF & NVMe/TCP Competitive Differentiation**
  - Eliminates need of a NVDIMM based server w/ software and chipset complexity. Can be deployed to millions of existing installed base of servers world wide

- **Flexible and profitable selling**
  - Get higher operating margins
  - Monolithic or Upgrade Option (Low risk)

STT-MRAM reduces latency of NVMe-OF or NVMe/TCP by providing at point power loss protected persistent write data completion buffer
Backup
STT-MRAM: Usage Models and Application Fit

Application Requirements

- Write Performance
- QoS (Low Latency)
- Reliability
- Persistence
- Endurance
STT-MRAM: Low Latency Write Burst Buffer

Get higher overall system application performance by using STT-MRAM as write buffer

Incoming Data
- Variable Rate
- Bursts
- Latency sensitive

Application Requirements
- Power Loss Protection
- Persistent Data
- Low Latency & High Performance

Written to in big block sizes