NVMe® Management Interface (NVMe-MI™) 2.0 Specification Updates

Sponsored by NVM Express organization, the owner of NVMe® specifications

Speaker

Myron Loewen

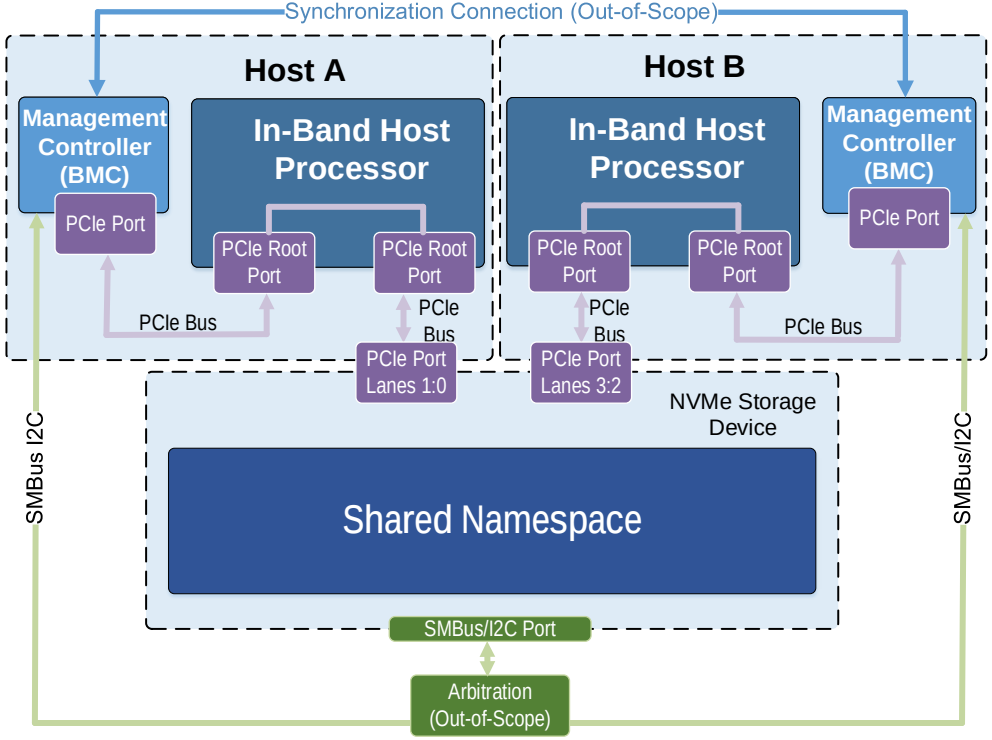
SOLIDIGM

NVMe® Board of Directors

What's changed for NVMe-MI™ 2.0?

- High Availability Out-of-Band Management (TP6034a)
- Out-of-Band Asynchronous Events (TP6035a)
- I3C Upgrade to SMBus Communications (TP6037)
- Cleanup unspecified cases around timings, reset, state machines, etc. to improve interop between a BMC and the NVMe® device (TP6027b, TP6032, TP6033a, and TP6038).
- Updated specification references to match latest versions (TP6036)

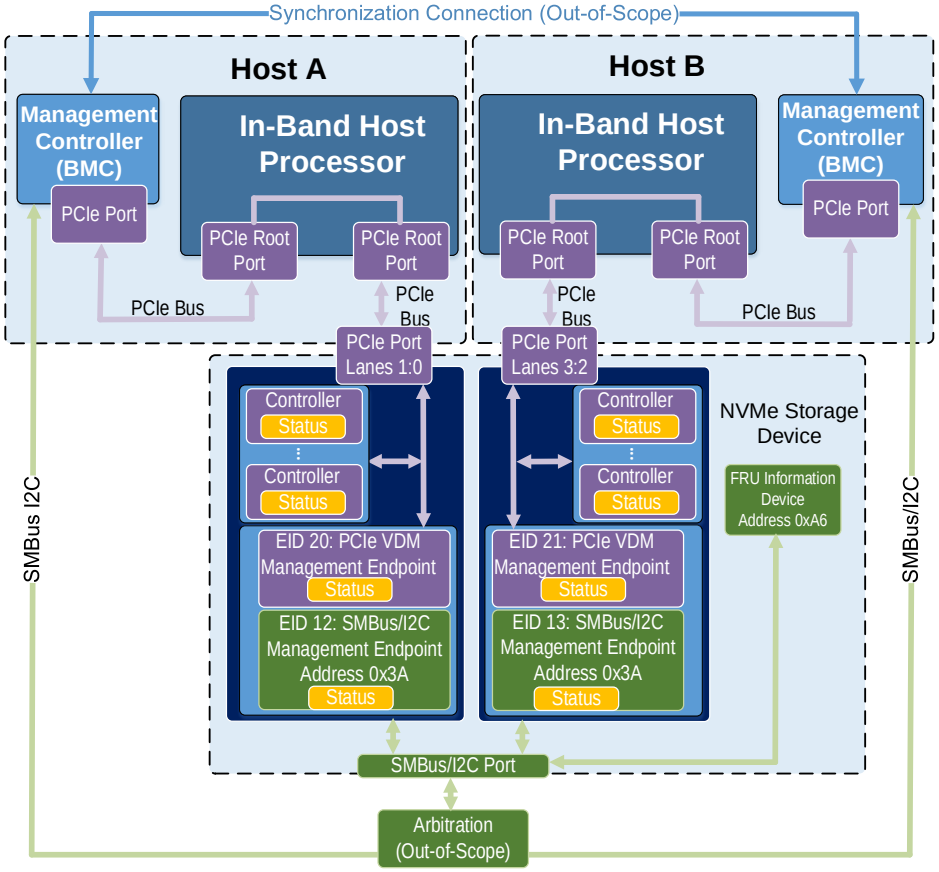
NVMe® Management Interface (NVMe-MI™) High Availability Out-of-Band Management



Problems:

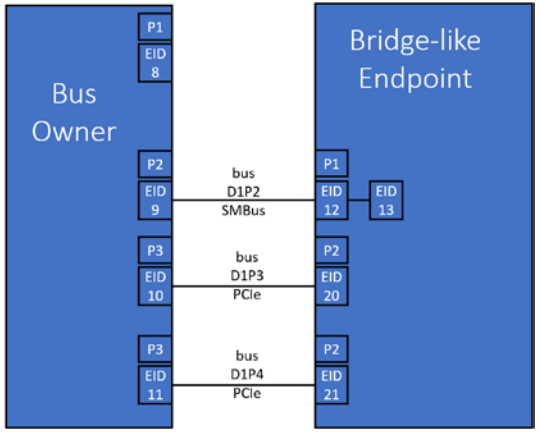
- Limitation of one Management Component Transport Protocol (MCTP) endpoint per physical address for non-bridge devices
- NVMe drives have a single System Management Bus (SMBus) port with a single physical SMBus address
- Difficult to manage a single NVMe drive between redundant host processors and Baseboard Management Consoles (BMCs)
- Would like second host to take over if the first fails and avoid downtime
- Difficult to share a single MCTP endpoint between two BMCs due to conflicts with shared status/state

High Availability Solution – MCTP Bridging for Endpoints

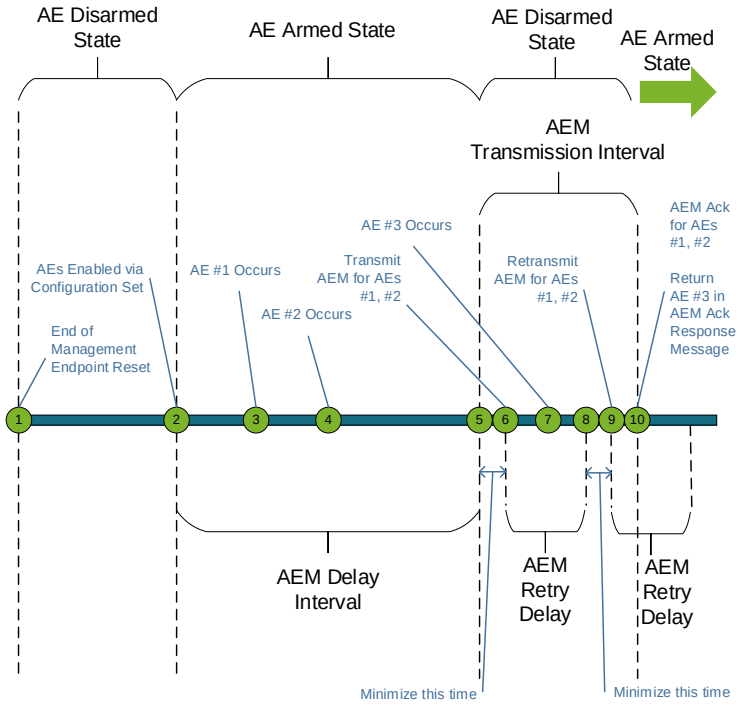


Solution:

- MCTP bridging on endpoints to allow multiple MCTP endpoints per physical address
- Unique instance of status/state per MCTP endpoint



NVMe® Management Interface (NVMe-MI™) Out-of-Band Asynchronous Events



Problems:

- Polling required for muxed SMBus topologies
- Polling penalty increases with drive count

Solution:

- PCIe® VDM and I3C hubs allow async events
- Added NVMe-MI async events
- Events for changes in health, temperature, inventory, security state, etc.
- Easily extensible to new events
- Support for event coalescing and adjustable rate limiting
- BMC can subscribe to select events
- Automatic retries for reliability

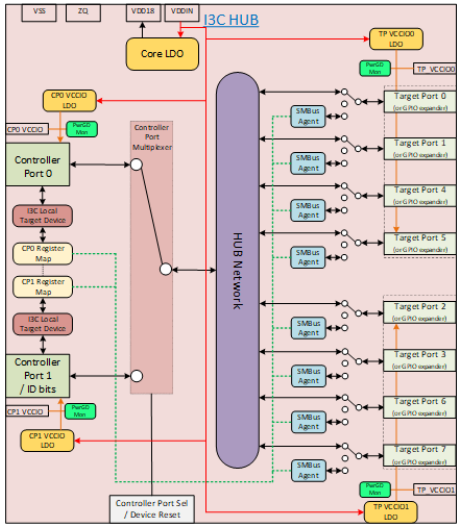
NVMe-MI™ over MIPI I3C Basic℠ on 2-Wire Bus

Problems with SMBus:

- Large data transfers like firmware updates or log page reads can take minutes to complete
- Typical SMBus architectures only connect one SSD at a time to the BMC, serialization of SSD accesses multiplies latency

Solution:

- Leverage PCIe defined methodology to enable I3C while providing legacy SMBus
- MIPI I3C Basic is at least an order of magnitude faster than SMBus
- I3C enables the BMC to control the clock signal to prevent security compromises
- I3C hubs allow BMC to interleave traffic to all SSDs, any SSD can respond when bus is idle



Questions?


nvm
EXPRESS®
Architected for Performance