



Flash Memory Summit

# New Standard for Remote Monitoring and Management of NVMe SSDs: NVMe-MI

John Wiedemeier, Teledyne-LeCroy  
and David Woolf, UNH-IOL



Flash Memory Summit

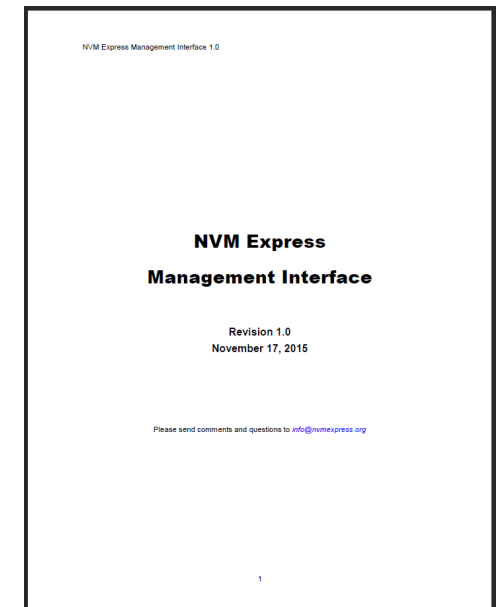
# Agenda

1. What is NVMe-MI
2. NVMe-MI Conformance Testing
3. Observations from the Plugfest
4. Future Work



# NVMe Management Interface (NVMe-MI)

- What is the NVMe Management Interface?
- A programming interface that allows out-of-band management of an NVMe Subsystem

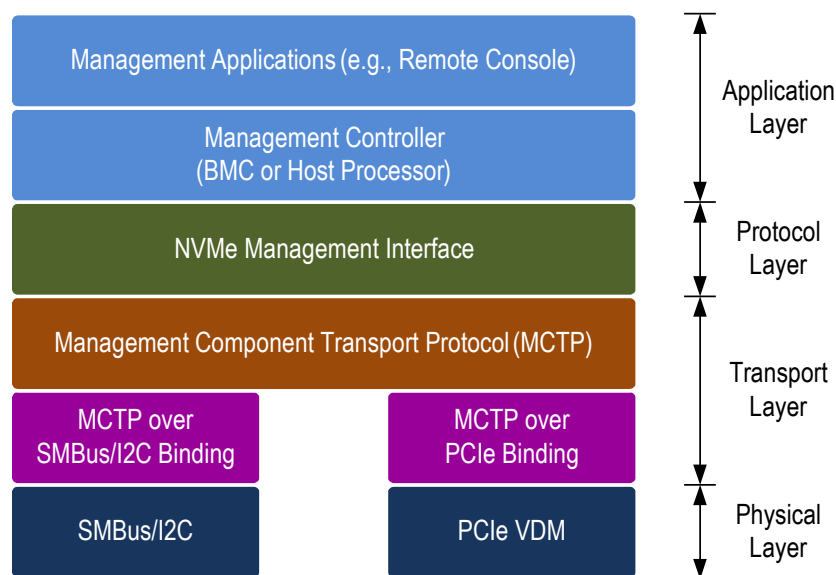




Flash Memory Summit

# NVMe Management Interface (NVMe-MI)

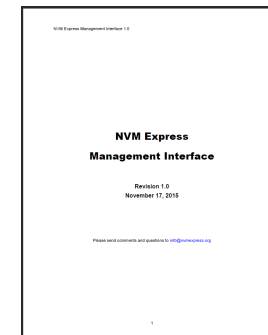
- NVMe-MI has the following key capabilities:
  - Discover devices that are present and learn capabilities of each device
  - Store data about the host environment; enabling a Management Controller to query the data later
  - Health and Temperature monitoring.
  - Multiple Command Slots to prevent a long latency command from blocking monitoring operations
  - Processor and operating system agnostic
  - A standard format for VPD and defined mechanisms to read/write VPD contents





# NVMe Management Interface (NVMe-MI)

- The NVMe Management Interface is used to
  - Send Command Messages which consist of standard NVMe Admin Commands that target a Controller within the NVM Subsystem
  - Send commands that provide access to the PCI Express configuration, I/O, and memory spaces of a Controller in the NVM Subsystem
  - Send Management Interface specific commands for inventorying, configuring and monitoring of the NVM Subsystem.
- Uses the concept of a “Management Endpoint”
- Management Endpoints may support the same commands regardless of the Physical medium used.





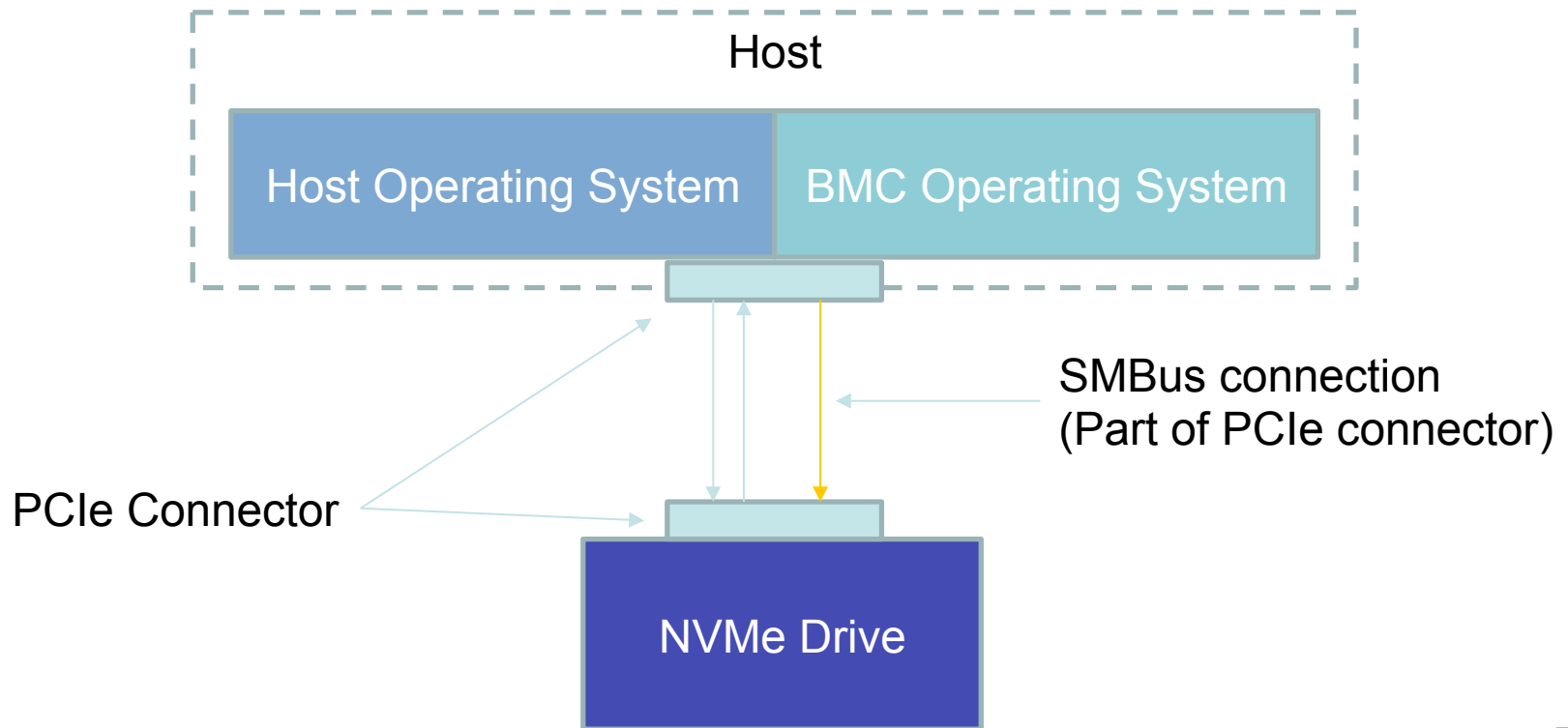
Flash Memory Summit

## Example: Temperature Monitoring

- Fan Control can be based on temperature reported by a specific drive
- BMC can monitor temperature out-of-band from data transfers
- Servers may contain many dozens of drives
  - Management Interface can be used to identify individual drives
  - Failures can be detected and reported remotely through networked infrastructure of BMCs
- Management Interface simplifies greatly the identification and fault finding within a large system



# Topology Overview

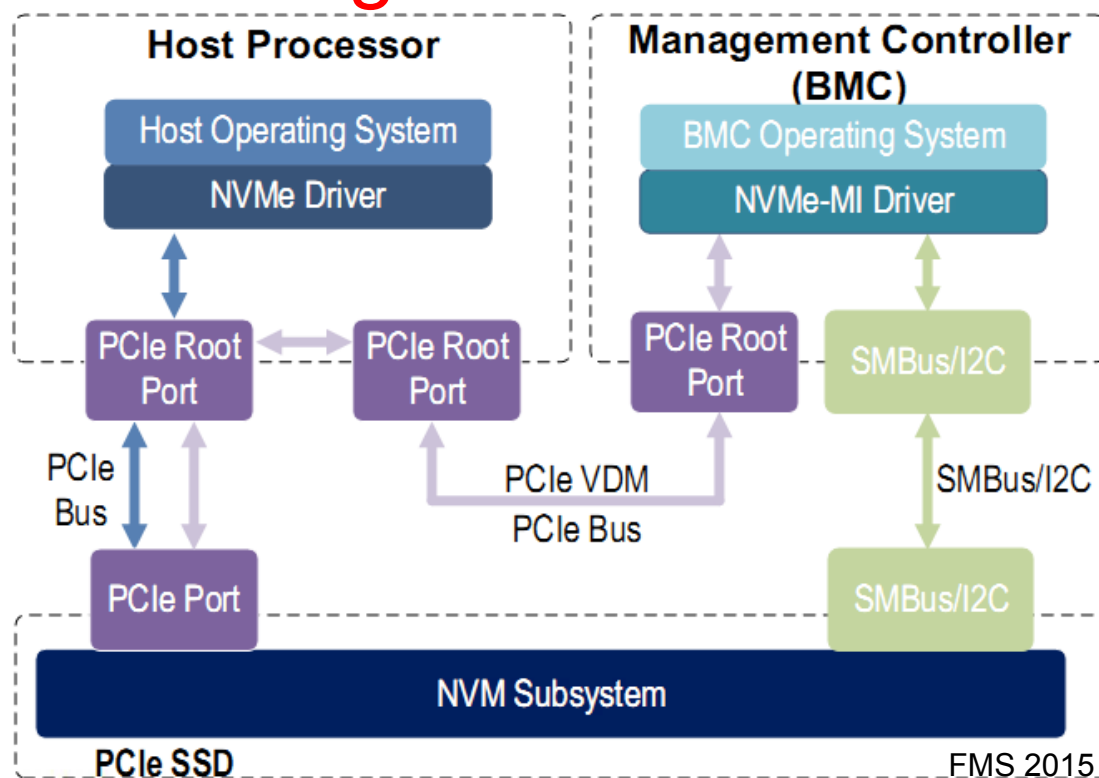




## Flash Memory Summit

- NVMe driver communicates to NVMe controllers over PCIe per NVMe Spec
- MC runs on its own OS on its own processor independent from host OS and driver
- Two OOB paths: PCIe VDM and SMBus
- PCIe VDMs are completely separate from in-band PCIe traffic though they share the same physical connection

# In-Band vs Out-of-band Management





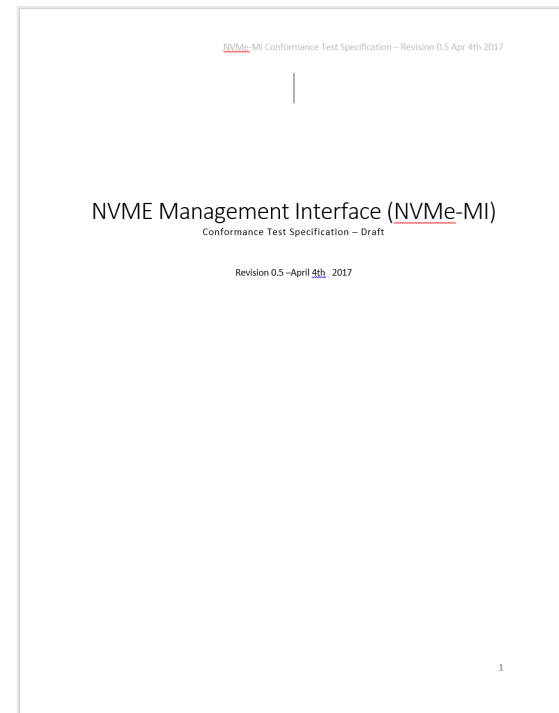


# NVMe-MI Conformance Testing



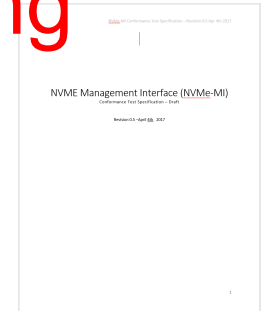
# NVMe-MI Conformance Test Specification

<https://www.iol.unh.edu/testing/storage/nvme/test-suites>



# NVMe-MI Conformance Testing

- Tests are split into groups:
  1. MCTP Base tests
  2. MCTP Control Message Tests
  3. MCTP Commands
  4. NVMe Error Handling
  5. NVMe Management Interface Tests
  6. NVMe-MI Message Processing Tests
  7. Control Primitives Tests
  8. Management Interface Commands
  9. NVMe Admin Command Set Tests
  10. Management Enhancement Tests
  11. Vital Product Data Tests
  12. Management Endpoint Reset Tests

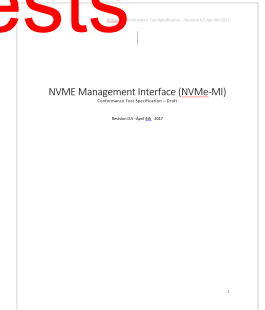




Flash Memory Summit

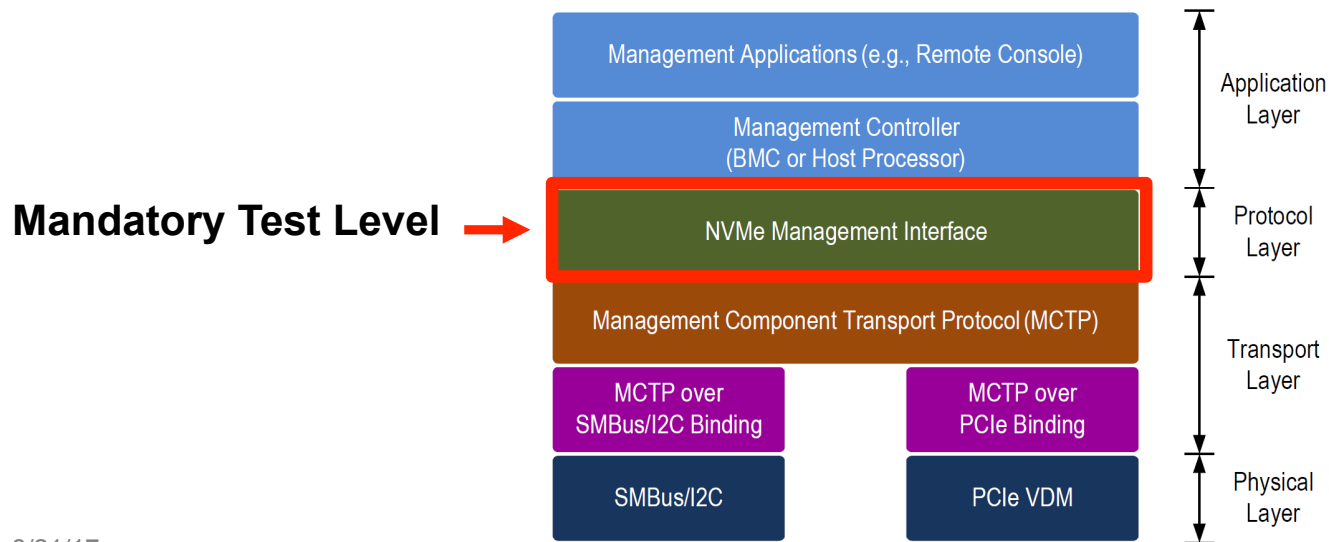
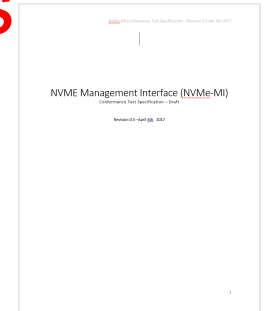
## Mandatory and Informational Tests

- Tests in the Test specification are marked as Mandatory or FYI
- Mandatory means the test is required for Integrators List testing
- FYI means the test will be run but the result will not count towards requirements for the integrators list
- FYI tests may become mandatory in the future



# Inband vs Out of Band Tests

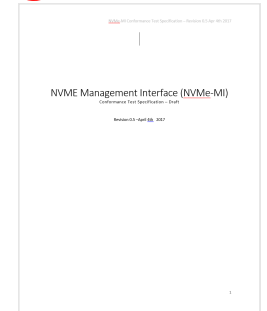
- NVMe-MI is independent of Physical Transport
- Tests apply to both In-band (PCIe VDM) or SMBus





# NVMe-MI Conformance Testing- Mandatory

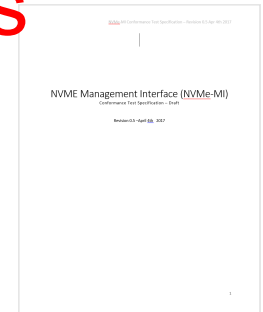
- Tests are split into groups:
  1. MCTP Base tests
  2. MCTP Control Message Tests
  3. MCTP Commands
  4. NVMe Error Handling
  5. NVMe Management Interface Tests
  6. NVMe-MI Message Processing Tests
  7. Control Primitives Tests
  8. Management Interface Commands
  9. NVMe Admin Command Set Tests
  10. Management Enhancement Tests
  11. Vital Product Data Tests
  12. Management Endpoint Reset Tests





# Mandatory Tests

- Test 4.1 - NVMe-MI Invalid Opcode – Mandatory Test
- Test 4.2 - NVMe-MI Reserved Identifier – Mandatory Test
- Test 4.3 - NVMe-MI Health Status Change – Mandatory Test
- Test 4.4 - NVMe-MI Reserved Configuration Identifier – Mandatory Test
- Test 4.5 - NVMe-MI MAXRENT Error – Mandatory Test
- Test 4.6 - NVMe-MI Reserved Data Structure Type – Mandatory Test
- Test 4.8 - NVMe-MI Invalid VPD Write Status – Mandatory test
- Test 4.9 - NVMe-MI Invalid Parameter Status – Mandatory Test
- Test 5.1 – NVMe-MI Message Type – Mandatory Test
- Test 5.2 – NVMe-MI Message IC – Mandatory Test
- Test 5.4 – NVMe-MI Command Slot – Mandatory Test
- Test 5.6 – NVMe-MI MCTP packet padding – Mandatory Test
- Test 5.7 – NVMe-MI Message Integrity Check – Mandatory Test
- Test 6.1 – NVMe-MI Reserved Fields – Mandatory Test
- Test 6.2 – NVMe-MI Error Response Code – Mandatory Test
- Test 7.1 – NVMe-MI Response Tag – Mandatory Test
- Test 7.2 – NVMe-MI Response Message – Mandatory Test
- Test 7.4 – NVMe-MI Response Message Replay – Mandatory Test
- Test 7.5 – NVMe-MI Response Replay Offset (RRO) – Mandatory Test
- Test 8.1 – NVMe-MI Response Header – Mandatory Test
- Test 8.2 – NVMe-MI Configuration Set – Mandatory Test
- Test 8.3 – NVMe-MI Config Get Response – Mandatory Test
- Test 8.4 – NVMe-MI Health Status Poll – Mandatory Test
- Test 8.5 – NVMe-MI Controller Health Status Poll – Mandatory Test
- Test 8.6 – NVMe-MI Read Data Structure – Mandatory Test
- Test 8.7.1 – NVMe-MI Verify NVMSSI Data Length – Mandatory Test
- Test 8.7.2 – NVMe-MI PortInfo Data Length – Mandatory Test
- Test 8.7.3 – NVMe-MI CtrlrList Data Length – Mandatory Test
- Test 8.7.4 – NVMe-MI CtrlrInfo Data Length – Mandatory Test
- Test 8.7.5 – NVMe-MI OptCmds Data length – Mandatory Test

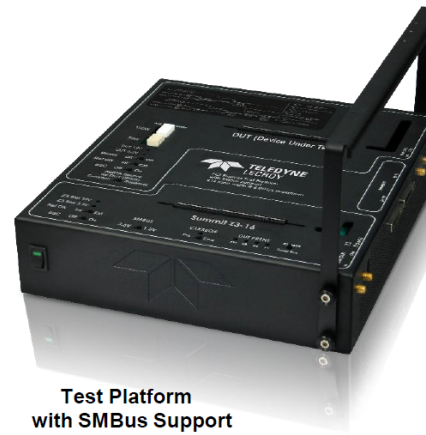


# MCTP/NVMe-MI Conformance Process

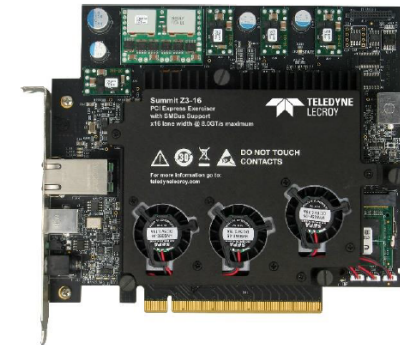
- Set up equipment
  - Z3-16 Exerciser
  - T34 Analyzer
  - Test platform
  - NVMe-MI Capable DUT
- Select Test Cases
- Power on DUT
- Run Test Cases
  - Log File Created
  - Trace File Created



Summit T34 Analyzer



Test Platform  
with SMBus Support



Summit Z3-16 Exerciser  
with SMBus Support





# Analyzing MCTP Trace Files

- Set view to MCTP Message Level/Command Level
- Navigate to SMBus packets
- What you should see is...



Packet	400k	SMBus	Addr	Wr	Cmd	B Cnt	S Addr	*	MCTP Header	Dst EID	Src EID	MCTP Message	Data
10	SMBus	MCTP	1100001	0	0x0F	13	1000000	1		0x00	0x00	Middle pkt	8 bytes

- What this is doing/saying is...
  - There is SMBus traffic that is MCTP
- Maybe what happens/should happen next
  - May have SMBus traffic that is not MCTP
  - May have MCTP traffic that is not NVMe-MI



# MCTP Packet Creation in Z3-16 Trainer Script

Teledyne LeCroy PCIe Protocol Suite - ALPHA - [C:\Users\Gordon.Getty\Desktop\PETrainer Script 1.peg]

File Setup Record Generate Report Search View Tools Window Help

Packet View

Packet	100k	SMBus	Addr	Wr	Cmd	B Cnt	S Addr	MCTP Header	Dst EID	Src EID	MCTP Message	Data	PEC
0	SMBus	MCTP	1110 000	0	0x0F	13	0000 000	1	0x00	0x00	Middle pkt	8 bytes	0xFA

QuickTiming markers not set

Generation Script Editor

```

1 Packet=SMBus
2 {
3   DestSlaveAddr = 1100001
4   CommandCode = MCTP
5   ByteCntSMBus = 13
6   MCTP_HDR_RevId = 1
7   MCTP_HDR_DestEPID = 0x00
8   MCTP_HDR_SrcEPID = 0x00
9   MCTP_MSG_Type = 1
10  Payload = 0xffff
11 }
12

```

Parameter	Value
DestSlaveAddr	1100001
CommandCode	MCTP
ByteCntSMBus	13
SrcSlaveAddr	
MCTP_HDR_RevId	1
MCTP_HDR_Version	
MCTP_HDR_DestEPID	0x00
MCTP_HDR_SrcEPID	0x00
MCTP_HDR_SOM	
MCTP_HDR_EOM	
MCTP_HDR_Psn	
MCTP_HDR_TO	
MCTP_HDR_Tag	
MCTP_MSG_IC	
MCTP_MSG_Type	1
MCTP_MSG_RqBit	
MCTP_MSG_Dbit	
MCTP_MSG_InstID	
MCTP_MSG_CmdCode	
Payload	0xffff

Summit Z3-16 SN:1200    Role: Host    Speed: 2.5 GT/s x8    Link State: Detect Quiet    InitFC State: Complete

Summit T34 SN:1600    US:    DS:    ●●●●

Ready    Ln 3, Col 28    Search: Fwd

8/21/17

18

## Testing of Different Form Factors

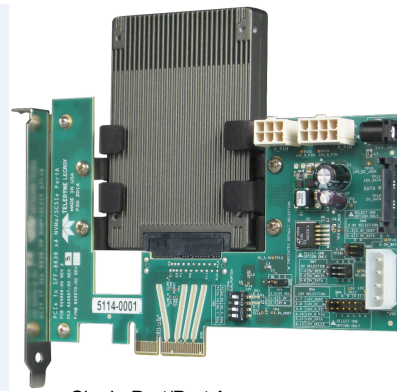
- Testing of U.2 and M.2 Form Factors done using adapters
- Protocol is basically the same regardless of form factor



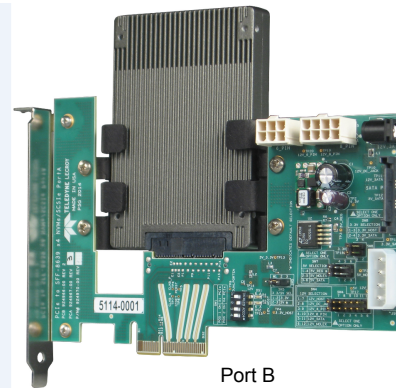


## Dual Port U.2

- Dual Port PCIe/NVMe devices exist now.
  - A Dual Port device has 2 separate independent PCIe links
  - It requires 2 analyzers and a separate Interposer to debug
- Compliance Testing requires each port to be tested separately



Single Port/Port A



Port B



# Running NVMe-MI Conformance Tests

- Select Desired Tests in the Link Expert Menu

The screenshot shows the LinkExpert 0.98 - BETA software interface. The main window displays a 'Diagnostics' menu on the left with categories like Link Establishment, Link Maintenance, Flow Control, Performance, and Power Management. A 'Select Tests' dropdown menu is open, showing a list of test categories including LinkExpert System Level Tests, PCIe 3.0 Compliance Package Tests, NVMe-MI/MCTP SMBus Tests, NVMe-MI/MCTP VDM Tests, Validation Tests (Endpoint), and Validation Tests (Root Complex). An arrow points from the 'NVMe-MI/MCTP SMBus Tests' option in this menu to a dialog box titled 'Add/Remove PCIe diagnostics to run'. This dialog box shows a list of tests under the 'NVMe-MI/MCTP SMBus Tests' group, with several tests selected (indicated by checked checkboxes). The tests listed are:

Test	Description
<input type="checkbox"/> Test 3-5 MCTP EndPoint Discovery	The intent of this test is to ensure that proper Completion messages are returned for MCTP Command EndPoint Discovery.
<input type="checkbox"/> Test 3-6 MCTP Command Get Endpoint ID	The intent of this test is to ensure that the response returns with successful completion and that all the reserved bits of the response are set to zero in the Get Endpoint ID Response.
<input checked="" type="checkbox"/> Test 4-1 NVMe-MI Invalid Opcode	The intent of this test is to ensure that a DUT returns the Invalid Command Opcode Status upon Management Interface Commands with undefined opcode.
<input checked="" type="checkbox"/> Test 4-10 NVMe-MI Invalid Command Size	he intent of this test is the ensure that Invalid Command size response shall be sent when Command message contains an Invalid Command Size.
<input checked="" type="checkbox"/> Test 4-2 NVMe-MI Reserved Identifier	The intent of this test is to ensure that the using the reserved bits in Configuration identifier ...
<input checked="" type="checkbox"/> Test 4-3 NVMe-MI Health Status Change	The intent of this test is to ensure that the using the Configuration Get - Health Status Change returns a succesful response and no response data.
<input checked="" type="checkbox"/> Test 4-4 NVMe-MI Reserved Identifier 2	The intent of this test is to ensure that the using a reserved Configuration Identifier returns an invalid parameter error response for NVMe-MI Configuration SET.
<input checked="" type="checkbox"/> Test 4-5 NVMe-MI Health Status Poll MA...	The intent of this test is to ensure that the Health Status Poll MAXRENT returns the proper error response.
<input type="checkbox"/> Test 4-6 NVMe-MI Read Data Structure - ...	he intent of this test is to ensure that the reading reserved data structure types returns the proper Invalid Parameter error response.
<input type="checkbox"/> Test 4-7 NVMe-MI VPD Read Invalid Par...	The intent of this test is to ensure that the performing VPD Read with Data Length plus Data Offset being greater than the size of the VPD returns the ...
<input type="checkbox"/> Test 4-8 NVMe-MI Invalid VPD Write Stat...	The intent of this test is the ensure that the VPD Write returns Invalid Parameter Error Response when the Data length and Data Offset Fields are greater than...
<input type="checkbox"/> Test 4-9 NVMe-MI Invalid Parameter Sta...	The intent of this test is the ensure that Invalid Parameter Status response shall be sent when Command message contains invalid parameter...



# Running NVMe-MI Conformance Tests

Select "Run All"

The screenshot shows the LinkExpert 0.98 - BETA software interface. The 'Diagnostics' tab is active, displaying a table of NVMe-MI/MCTP SMBus Tests. A red box highlights the 'Run All' button in the 'Action' column of the table. A red arrow points from the text 'Select "Run All"' to this button. The 'Analysis Results' section shows a 'Device Under Test' of 'Unknown Device' and a 'Diagnostics Summary' of '0 OUT OF 0 TESTS'. The 'Analyzer/Exerciser Devices' section shows an Analyzer (Summit T34 SN:17176) and an Exerciser (Z3-16 SN:17136).

Diagnostic	Status	Action
<b>NVMe-MI/MCTP SMBus Tests</b>		
Test 4-1 NVMe-MI Invalid Opcode	NOT STARTED	▶ ×
Test 4-10 NVMe-MI Invalid Command Size	NOT STARTED	▶ ×
Test 4-2 NVMe-MI Reserved Identifier	NOT STARTED	▶ ×
Test 4-3 NVMe-MI Health Status Change	NOT STARTED	▶ ×
Test 4-4 NVMe-MI Reserved Identifier 2	NOT STARTED	▶ ×
Test 4-5 NVMe-MI Health Status Poll MA...	NOT STARTED	▶ ×



Flash Memory Summit

# Results

Result

Log Files

The screenshot displays the LinkExpert 0.98 - BETA software interface. The top navigation bar includes 'File', 'Diagnostics', and 'Help'. The main window is divided into several sections:

- Diagnostics:** A table listing various tests under the category 'NVMe-MI/MCTP SMBus Tests'. The first test, 'Test 4-1 NVMe-MI Invalid Opcode', is marked as 'PASSED'. Other tests like 'Test 4-10 NVMe-MI Invalid Command Size' and 'Test 4-2 NVMe-MI Reserved Identifier' are marked as 'NOT STARTED'.
- Analysis Results:** A summary section showing 'Device Under Test: Unknown Device' and a '100%' completion rate. It also displays a 'Diagnostics Summary' with 1 out of 1 tests completed, and a 'Results Summary' with 0 warnings and 0 errors.
- Log Files:** A detailed view of the test log for 'Test 4-1 NVMe-MI Invalid Opcode'. The log shows the test description, the execution of the stimulus, and the successful completion of the verification script. The log text includes: 'Run test: Controller ID parser', 'Executing stimulus: C:\Users\Public\Documents\LeCroy\PCIe Protocol Suite\ScriptAutomationTestTool\TrainerScripts\Endpoint\MCTP\_NVMeMi\special\_ctrl\_id\_peg', 'Trace captured - starting Verification', 'C:\Users\Public\Documents\LeCroy\PCIe Protocol Suite\ScriptAutomationTestTool\VerificationScripts\Endpoint\special\_ctrl\_id\_parser.pevs', 'Running verification script...', 'Test Description: This verification script processes the trace containing the recording of Read NVMe-Data Structure - Controller ID from the DUT and produces the necessary definitions files used by Trainer Scripts in LeCroy's PCI Express NVMe-MI Compliance Testing package.', 'Controller ID parser', 'Verifying test stage 1 (Performing MCTP initialization).', 'Verifying test stage 2 (Processing Read NVMe-MI - Controller List Response and Storing Controller ID in Config File). The response was returned successfully!', 'The first controller ID is :35', 'C:\Users\Public\Documents\LeCroy\PCIe Protocol Suite\ScriptAutomationTestTool\TrainerScripts\Endpoint\ctrl\_id\_definition.peg', and '----- P A S S E D -----'.
- Analyzer/Exerciser Devices:** A section showing the connected devices. The analyzer is 'Summit T34 SN:17176' with a speed of '2.5 Gb/s x1'. The exerciser is 'Z3-16 SN:17136' with a speed of 'HbL 2.5 Gb/s N/A'. The link state is 'Detect/Quiet'.



Flash Memory Summit

# Agenda

1. What is NVMe-MI
2. NVMe-MI Conformance Testing
3. Observations from the Plugfest
4. Future Work





Flash Memory Summit

## May 2017 Plugfest

- First public testing of NVMe-MI
- New Track in NVMe Plugfest #7
- Goal to list first products on NVMe-MI Integrators List





Flash Memory Summit

# NVMe-MI Integrators List

6 Products as of  
8/7/17

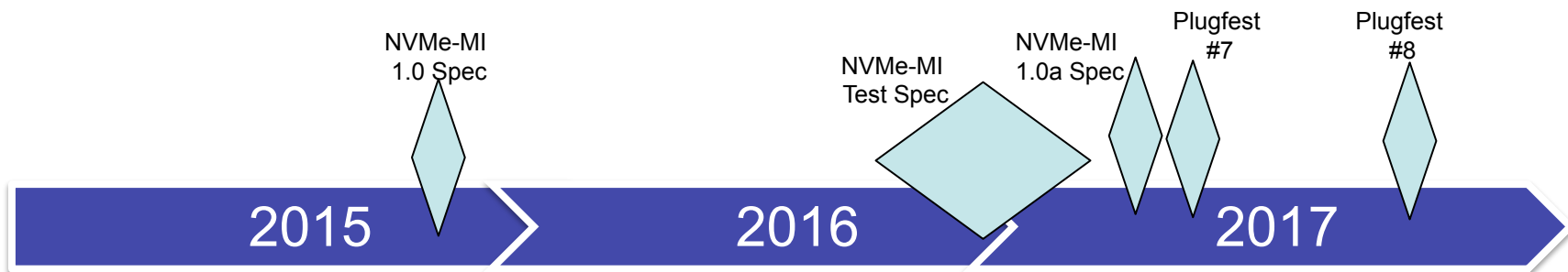
[www.iol.unh.edu/  
registry/nvme-mi](http://www.iol.unh.edu/registry/nvme-mi)

NVMe-MI Devices					
Product	Product Type	Firmware Version	Interop Program Revision	Date Listed	Further Info
Intel® Optane™ SSD DC P4800X Series	NVMe SSD	E0000320	v7.0	7/27/2017	<a href="#">Intel Product Page</a>
Intel® SSD DC P4500 Series	NVMe SSD	QDV10130-003	v7.0	7/27/2017	<a href="#">Intel Product Page</a>
Intel® SSD DC P4608 Series	NVMe SSD	QDV10130-003	v7.0	7/27/2017	<a href="#">Intel Product Page</a>
Intel® SSD DC P4600 Series	NVMe SSD	QDV10130-003	v7.0	7/27/2017	<a href="#">Intel Product Page</a>
Intel® SSD DC P4501 Series	NVMe SSD	QDV10130-003	v7.0	7/27/2017	<a href="#">Intel Product Page</a>
Samsung NVMe 172X Series	NVMe SSD	GPNAuB3Q	v7.0	6/27/2017	<a href="http://samsung.com">samsung.com</a> MiKyeong Kang [mkkang.kang@samsung.com]



# Plugfest Observations

- SMBus/I2C only, no In-band / Vendor Defined Messages (VDM) tested yet
- Some products are using 1.0 spec, some products using 1.0a spec.





Flash Memory Summit

## Future Work

- Support for ECNs
  - ECNs 1,2,3, captured in 1.0a spec release
  - Some changes have interop impact
  - Need to Test for this
- Verify In-band/VDM support



Flash Memory Summit

## Future Work

- Interop Tests
  - As MI-Compliant servers become available, add interop to the program.
  - Prove that interface is OS agnostic
    - Exercise MI via OS functions
- Plugfest #8 Oct 30-Nov 3 @ UNH-IOL
  - NVMe, NVMe-MI, NVMeoF Integrators Lists