



Storage over PCIe protocol analysis and traffic generation techniques

Isaac Livny

Field Applications Engineer
Teledyne LeCroy Corporation

Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation

Layered protocols analysis over PCIe as transport mechanism

Split Tra	0	R→	2.5	Cfg	CfgRd0	RequesterID	CompleterID	Tag	TC	VC ID	DeviceID	Register	Status
	0		x8		00:00100	000:03:0	003:00:0	0	0	0	003:00:0	0x000	SC
Device ID	Vendor ID	Metrics	# LinkTras	Resp. time	Latency	Thrpt MB/s	Pld. Bytes	Time Stamp					
0xF015	0x10DF		2	556.000 ns	208.000 ns	6.861	4	0014 . 335 122 612 s					
Link Tra	16	R→	2.5	TLP	Cfg	CfgRd0	Length	RequesterID	Tag	DeviceID	Register	1st BE	VC ID
			x8	114		00:00100	1	000:03:0	0	003:00:0	0x000	1111	0
ExplicitACK	Packet #48668	Metrics	# Packets	Time Stamp									
			2	0014 . 335 122 612 s									
Packet	48665	R→	2.5	TLP	Cfg	CfgRd0	Length	RequesterID	Tag	DeviceID	Register	1st BE	
			x8	114		00:00100	1	000:03:0	0	003:00:0	0x000	1111	
LCRC	Time Delta	Time Stamp											
0xA275BADB	208.000 ns	0014 . 335 122 612 s											
Packet	48668	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Time Delta	Time Stamp				
			x8			114	0xF613	12.000 ns	0014 . 335 122 820 s				
Link Tra	17	R→	2.5	TLP	Cpl	CplID	Length	RequesterID	Tag	CompleterID	Status	BCM	Byte Cnt
			x8	4044		10:01010	1	000:03:0	0	003:00:0	SC	0	4
LwrAddr	Device ID	Vendor ID	VC ID	ExplicitACK	Metrics	# Packets	Resp. time	Pld. Bytes	Thrpt MB/s				
0x00	0xF015	0x10DF	0	Packet #48670		2	336.000 ns	4	11.353				
Time Stamp	0014 . 335 122 832 s												
Packet	48669	R→	2.5	TLP	Cpl	CplID	Length	RequesterID	Tag	CompleterID	Status	BCM	
			x8	4044		10:01010	1	000:03:0	0	003:00:0	SC	0	
Byte Cnt	LwrAddr	Device ID	Vendor ID	LCRC	Time Delta	Time Stamp							
4	0x00	0xF015	0x10DF	0xD977B772	332.000 ns	0014 . 335 122 832 s							
Packet	48670	R→	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Time Delta	Time Stamp				
			x8			4044	0xC5AA	-504.000 ns	0014 . 335 123 164 s				
Packet	48666	R→	2.5	SKIP	COM	SKIP Symbols	Idle	Time Stamp					
			x8		K28.5	K28.0 K28.0 K28.0	140.000 ns	0014 . 335 122 660 s					
Packet	48667	R→	2.5	DLLP	UpdateFC-NP	VC ID	HdrFC	DataFC	CRC 16	Time Delta	Time Stamp		
			x8			0	209	134	0xFF36	556.000 ns	0014 . 335 122 816 s		

Layered Protocols

Support in Analysis Tools



- Hierarchical view display capability with multi layer expansion into sub-layers
- Multi-view capabilities
- Processing capability of upper layer through scripting to adopt to specification changes
- Tooltip feature to highlight specification details
- Performance and statistical analysis per instruction, by segment and overall trace
- Compacting of repetitive traffic
- Compacting of multiple 32 bit transactions into 64 bit upper layer commands

Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation



SATA Express



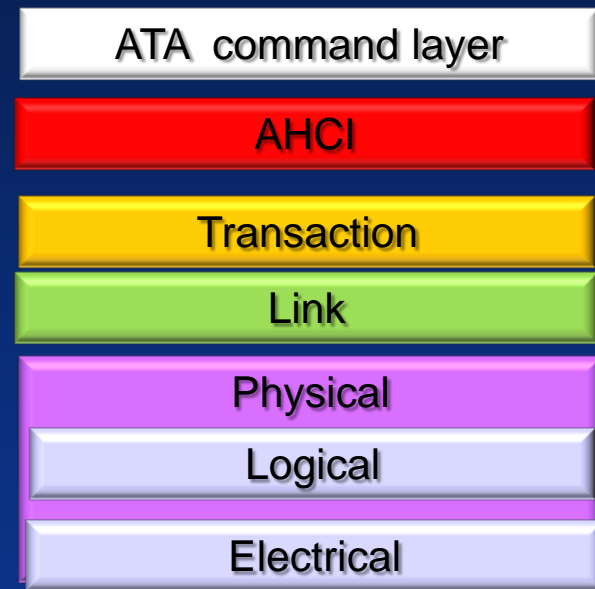
SATA Express

ATA command over Serial ATA Read DMA

ID	ATA Cmd.	1.5 G	Command	Input (H)	Normal Output (H)	ATA Command Data, 2560 Bytes	PM Port (H)	Protocol	Status	Metrics			
5.002.537.866 (s)	2		0xC8 : Read DMA	C8000500000040	00000400004050	00 01 02 03 04 05 06 07 08 09 0A 0B >>	0	0x07 : DMA	0x01 : Normal Output				
ID	Transport	1.5 G	FIS Type	PM Port (H)	C (H)	Command (H)	Features (H)	LBA Low (H)	LBA Mid (H)	LBA High (H)	Device (H)	LBA Low (exp) (H)	LBA Mid (exp) (H)
5.002.537.866 (s)	3		0x27 : Register Host to Device	0	1	C8	00	00	00	00	40	00	00
LBA High (exp) (H)	Features (exp) (H)	Sector Count (H)	Sector Count (exp) (H)	Control (H)	CRC (H)	Duration							
00	00	05	00	00	E0EA7CCB	4.426 (us)							
T2	Transport	1.5 G	FIS Type	PM Port (H)	Data, 2560 Bytes	CRC (H)	Duration						
5.015.078.960 (s)	4		0x46 : Data	0	00 01 02 03 04 05 06 07 08 09 0A 0B >>	89E129F5	32.666 (us)						
T2	Transport	1.5 G	FIS Type	PM Port (H)	I (H)	Status (H)	Error (H)	LBA Low (H)	LBA Mid (H)	LBA High (H)	Device (H)	LBA Low (exp) (H)	LBA Mid (exp) (H)
5.015.111.626 (s)	6		0x34 : Register Device to Host	0	1	50	00	04	00	00	40	00	00
LBA High (exp) (H)	Sector Count (H)	Sector Count (exp) (H)	CRC (H)	Duration									
00	00	00	47CA0EE1	40.373 (us)									

SATA Express

- The Serial ATA International Organization (SATA-IO) developed the specification
- This protocol combines the SATA AHCI software specification with the PCIe host interface
- SATA Express enables new devices to be developed that utilize the faster PCIe interface and maintain compatibility with a broad base of existing SATA applications
- Data Rate Support
 - PCIe 2.x at x2 link for 8GT/s data rate
 - PCIe 3.0 at x2 link for a 16GT/s data rate



AHCI HBA Registers

File Setup Record Generate Report Search View Tools Window Help

AHCI	RequesterID	CompleterID	CAP	S64A	SNCQ	SSNTF	SMPS	SSS	SALP	SAL	SCLO	ISS	SAM	SPM	FBSS	PMD	SSC	PSC	NCS	CCCS	EMS	SXS	NP	Time Delta	Time Stamp
0	x1	000:00:0	004:00:0	1	1	0	0	0	0	0	0	6.0 Gbps	0	0	0	1	0	0	31	0	0	0	7	22.976 µs	0008 . 574 012 896 s
1	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
2	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
3	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
4	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
5	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
6	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
7	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
8	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
9	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
10	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
11	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
12	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
13	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
14	x1	000:00:0	004:00:0	GHC	AE	Reserved	MRS	MRSM	IE	HR															
15	x1	000:00:0	004:00:0	IS	Interrupt Pending Status b00000000000000000000000000000000																				
16	x1	000:00:0	004:00:0	IS	Interrupt Pending Status b00000000000000000000000000000000																				

HBA Memory Registers

1. Port Control
2. Generic Host Control(GHC)

AHCI PCIe Configuration Space Registers

CAP: Host Capabilities
 GHC: Global Host Control
 IS: Interrupt Status
 PI: Ports Implemented

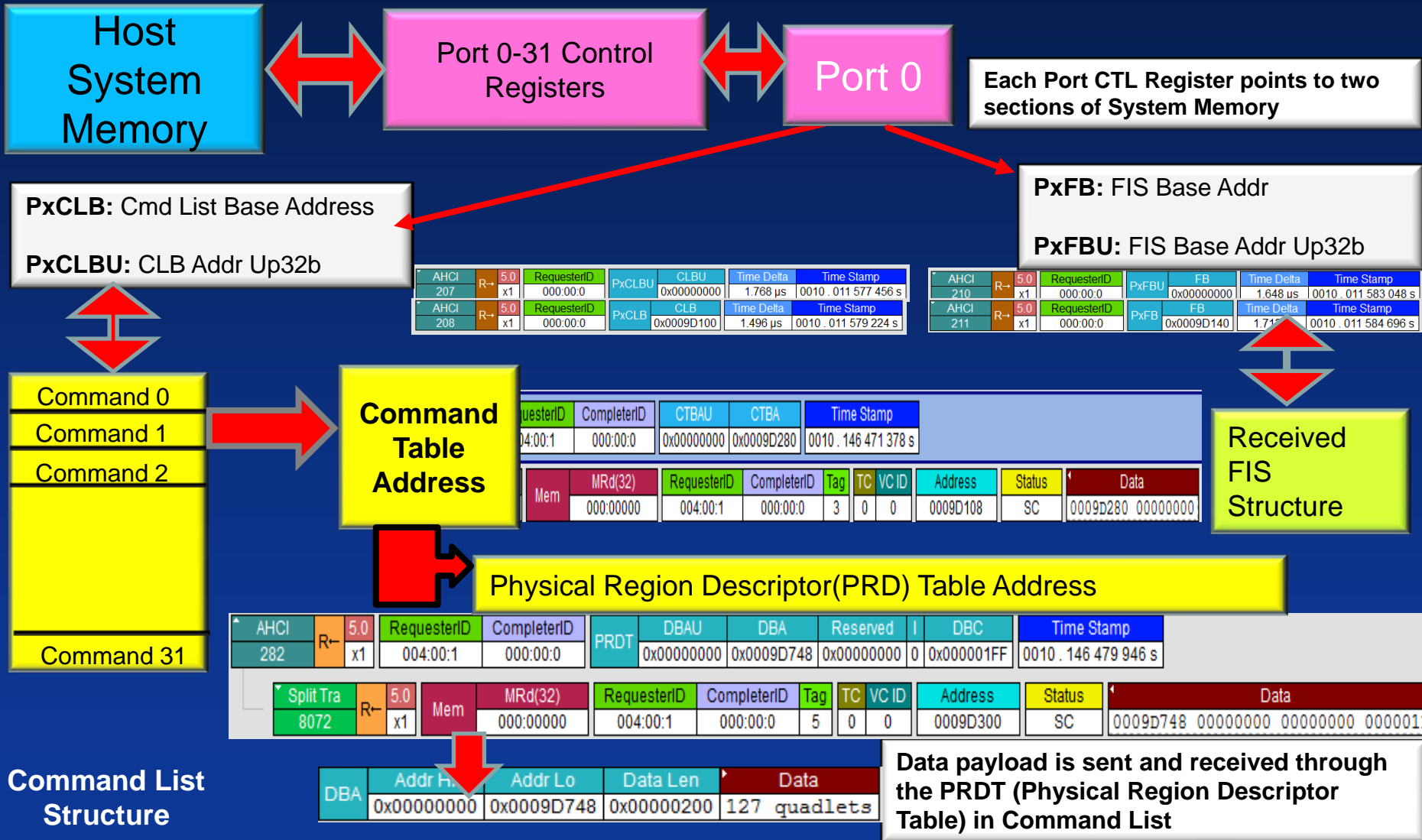
SATA Express Port Control Setup

AHCI	R→	5.0	RequesterID	PxCMD	ICC	ASP	ALPE	DLAE	ATAPI	APSTE	FBSCP	ESP	CPD	MPSP	HPCP	PMA	CPS	CR	FR	MPSS	CCS	FRE	CLO	POD	SUD	ST	Time Delta	Time Stamp				
203	R→	5.0	000:00:0	PxCMD	No-Op / Idle	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1.520 μs	0010 . 002 648 088 s				
204	R→	5.0	000:00:0	CompleterID	004:00:0	PxCMD	No-Op / Idle	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	812.776 μs	0010 . 002 649 608 s	
205	R→	5.0	000:00:0	PxCMD	No-Op / Idle	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1.688 μs	0010 . 003 462 384 s			
206	R→	5.0	000:00:0	CompleterID	004:00:0	PxCMD	No-Op / Idle	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	8.113 ms	0010 . 003 464 072 s
207	R→	5.0	000:00:0	PxCLBU	CLBU	0x00000000																										
208	R→	5.0	000:00:0	PxCLB	CLB	0x0009D100																										
209	R→	5.0	000:00:0	CompleterID	004:00:0	PxCLB	CLB	0x0009D100																								
210	R→	5.0	000:00:0	PxFBU	FB	0x00000000																										
211	R→	5.0	000:00:0	PxFB	FB	0x0009D140																										
212	R→	5.0	000:00:0	CompleterID	004:00:0	PxFB	FB	0x0009D140																								

Port control address of Command list setup

Port control address of Received FIS setup

AHCI System Memory





ATA command: Set Features

ATA	H	Port	Slot	Protocol	Command	Input	Features	Count	LBA	LBAExt	Device	7	6	5	4	0-3	Output	BSY	DRDY	DF	DSC	DRQ	CORR	IDC	ERR
18		0	0	ND	SET FEATURES	0x0003	0x0046	0x000000	0x000000		1	1	1	0	0x0		0	1	0	1	0	0	0	0	0

AHCI	H	Port	Slot	Protocol	Command Issue	Time Stamp
161029		0	0	PxCI	0000 0000 0000 0000 0000 0000 0000 0001	0111.666 032 030 s

Link Tra	R+	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	ImplicitACK	#Packets
325953		x4	1412		010:00000	1	000:00:0	0	FBB10138	1111	0000	01000000	0	Packet #650353	1

AHCI	H	Port	Slot	Address	Data Len	COMMAND HEADER	PRDTL	PMP	C	B	R	P	W	A	CFL	PRDBC	PRDBC	CTBA	CTBA	CTBAU	CTBAU	Time
161031		0	0	00000000:8D2FD000	4	0x0000	0x0	0	0	0	0	0	0	0	0x05	0	0	0x8D030000	0x8D030000	0x00000000	0x00000000	0111.666 032 030 s

Split Tra	R+	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data
162151		x4		000:00000	010:00:0	000:00:0	2	0	0	8D2FD000	SC	05000000 00000000 0000038D 00000000

AHCI	H	Port	Slot	Address	Data Len	FIS	FIS Type	PMP	C	Command	Features	LBA(7:0)	LBA(15:8)	LBA(23:16)	Device	LBA(31:24)	LBA(39:32)	LBA(47:40)	
161033		0	0	00000000:8D030000	5	0x05	Register FIS - H2D	0x0	1	0xEF	0x03	0x00	0x00	0x00	0xE0	0x00	0x00	0x00	0x00

Split Tra	R+	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data
162153		x4		000:00000	010:00:0	000:00:0	2	0	0	8D030000	SC	0: 2780EF03 000000E0 00000000 46000000 4: 00000000

AHCI	D	Port	Slot	Address	Data Len	RFIS	FIS Type	PMP	I	Status	Error	LBA(7:0)	LBA(15:8)	LBA(23:16)	Device	LBA(31:24)	LBA(39:32)	LBA(47:40)	Count(7:0)
161049		0	0	00000000:8D2E9040	5	0x05	Register FIS - D2H	0x0	1	0x50	0x00	0x00	0x00	0x00	0xE0	0x00	0x00	0x00	0x46

Link Tra	R+	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC
325992		x4	1458		010:00000	5	010:00:0	128	8D2E9040	1111	1111	0: 34405000 000000E0 00000000 46000000 4: 00000000	0



READ FPDMA QUEUED

ATA	Port	Slot	Protocol	Command	Input	Sector Count	Count	PRIO	NOQ Tag	RARC	LBA	LBA Ext	ICG	Hybrid Information	Device	FUA (7)	6	4	Output	BSY	DRDY	DF	DSC	DRQ	CORR	IDC	ERR	Address	Data	
4	D	0	DMQ	READ FPDMA QUEUED	Input	0x0010		00	28	0	0x000800	0x0000000	00000000	0x00		0	1	0	0	0	0	0	0	0	0	0	00000001:2BDFD000	2048 dwords		
AHCI	H	Port	PySACT Device Status				Time Delta		Time Stamp																					
776	0	0	0000 0100 0000 0000 0000 0000 0000 0000				48.000 ns		0032.189 641 282 s																					
AHCI	H	Port	PxIC Command Issue				Time Delta		Time Stamp																					
777	0	0	0000 0100 0000 0000 0000 0000 0000 0000				436.000 ns		0032.189 641 330 s																					
AHCI	H	Port	Slot	Address	Data Len	COMMAND HEADER		PRDTL	PMP	C	B	R	P	W	A	CFL	PRDBC	PRDBC	CTBA	CTBA	CTBAU	CTBAU	Time Delta		Time Stamp					
778	0	0	28	00000000:DEFF5340	4	0x0002		0x0	0	0	0	0	0	0	0x05		0		0xDCEB6400	0x00000000			1.320 us		0032.189 641 766 s					
AHCI	H	Port	Slot	Address	Data Len	CFIS	FIS Type	PMP	C	Command	Features	LBA(7:0)	LBA(15:8)	LBA(23:16)	Device	LBA(31:24)	LBA(39:32)	LBA(47:40)	Features(exp)	Count(7:0)	Count(15:8)	ICC	Control	Auxiliary	Auxiliary(exp)	T				
779	0	0	28	00000000:DCEB6400	5	Register FIS - H2D	0x0	1	0x60	0x10	0x00	0x00	0x08	0x00	0x40	0x00	0x00	0x00	0x00	0xD0	0x00	0x00	0x00	0x00	0x00	0x00	2			
AHCI	D	Port	Address	Data Len	RFIS	FIS Type	PMP	I	Status	Error	LBA(7:0)	LBA(15:8)	LBA(23:16)	Device	LBA(31:24)	LBA(39:32)	LBA(47:40)	Count(7:0)	Count(15:8)	Time Delta		Time Stamp								
780	0	0	00000000:DEFF5440	5	Register FIS - D2H	0x0	0	0x40	0x00	0x00	0x00	0x08	0x00	0x40	0x00	0x00	0x00	0x00	0xD0	0x00	8.062 ms		0032.189 672 166 s							
AHCI	D	Port	Address	Data Len	DSFIS	FIS Type	PMP	D	I	A	BI LOW	BI HIGH	Buffer Offset	Transfer Count	Time Delta		Time Stamp													
784	0	0	00000000:DEFF5400	7	DMA Setup FIS - Bi-directional	0x0	1	0	0	0x00000001A	0x00000000	0x00000000	0x00002000	128.000 ns		0032.177 734 054 s														
AHCI	H	Port	Slot	Address	Data Len	PRDT	DBA	DBAU	Reserved	I	DBC	PRDT	DBA	DBAU	Reserved	I	DBC	Time Delta		Time Stamp										
785	0	0	28	00000000:DCEB6480	8	0x2BDFD000	0x00000001	0x00000000	0	4095	0x0FAFE000	0x00000001	0x00000000	0	4095	1.464 us		0032.177 734 182 s												
AHCI	D	Port	Slot	Address	Data Len	DATA	Data	Time Delta		Time Stamp																				
786	0	0	28	00000001:2BDFD000	1024	1024 dwords	9.904 us		0032.177 735 646 s																					
AHCI	D	Port	Slot	Address	Data Len	DATA	Data	Time Delta		Time Stamp																				
787	0	0	28	00000001:CFAFE000	1024	1024 dwords	9.904 us		0032.177 745 550 s																					
AHCI	D	Port	Address	Data Len	SDBFIS	FIS Type	PMP	I	N	Status Lo	Status Hi	Error	Protocol specific				Time Delta		Time Stamp											
788	0	0	00000000:DEFF5468	2	Set Device Bits FIS - Device to Host	0x0	1	0	0x0	0x4	0x00	0000 0100 0000 0000 0000 0000 0000 0000	5.532 us		0032.177 755 454 s															
AHCI	D	Interrupt Pending Status				Time Delta		Time Stamp																						
789	0	0	0000 0000 0000 0000 0000 0000 0000 0001				2.312 us		0032.177 760 986 s																					
AHCI	D	Port	PxIS	CPDS	TFES	HBFS	HBDS	IFS	INFS	OFS	IPMS	PRCS	DMPS	PCS	DPS	UFS	SDBS	DSS	PSS	DHRS	Time Delta		Time Stamp							
790	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1.488 us		0032.177 763 298 s							
AHCI	D	Port	PxSERR	DIAG	X	F	T	S	H	C	D	B	W	I	N	ERR	E	P	C	T	M	I	Time Delta		Time Stamp					
791	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.520 us		0032.177 764 786 s						
AHCI	H	Port	PxIS	CPDS	TFES	HBFS	HBDS	IFS	INFS	OFS	IPMS	PRCS	DMPS	PCS	DPS	UFS	SDBS	DSS	PSS	DHRS	Time Delta		Time Stamp							
792	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	48.000 ns		0032.177 766 306 s							
AHCI	H	Interrupt Pending Status				Time Delta		Time Stamp																						
793	0	0	0000 0000 0000 0000 0000 0000 0000 0001				-7.003 ms		0032.177 766 354 s																					

ATA Command: READ BUFFER

ATA	Port	Slot	ATAError	ATAError	Protocol	Command	Input	Features	Count	LBA	LBAExt	Device	7	6	5	4	0-3	Address					
3	D	0	Incomplete	Logical error	PI	READ BUFFER		0x0000	0x0000	0x0000000	0x0000000	1	1	1	0	0x0		00000001:BC281020					
AHCI		H	Port	Slot	Command Issue			Time Stamp															
82776		H	0	0	PxCI	0000 0000 0000 0000 0000 0000 0000 0001			0111.539 022 878 s														
Link Tra		R	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	ImplicitACK	Metrics	# Packets						
165541		R	x4	3362		010:00000	1	000:00:0	0	FBB10138	1111	0000	01000000	0	Packet#331064		1						
AHCI		H	Port	Slot	Address	Data Len	COMMAND HEADER	PRDTL	PMP	C	B	R	P	W	A	CFL	PRDRC	PRDBC	CTBA	CTBAU	CTBAU	CTBAU	
82778		H	0	0	00000000:8D2FD000	4		0x0001	0x0	0	0	0	0	0	0	0x05	0	0	0x8D030000	0x00000000	0x00000000	0x00000000	0111
Split Tra		R	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data										
82761		R	x4		000:00000	010:00:0	000:00:0	2	0	0	8D2FD000	SC	05000100 00000000 0000038D 00000000										
AHCI		H	Port	Slot	Address	Data Len	Register FIS - H2D	PMP	C	Command	Features	LBA(7:0)	LBA(15:8)	LBA(23:16)	Device	LBA(31:24)	LBA(39:32)	LBA(47:40)					
82780		H	0	0	00000000:8D030000	4		0x0	1	0xE4	0x00	0x00	0x00	0x00	0x00	0xE0	0x00	0x00	0x00	0x00	0x00		
Split Tra		R	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data										
82763		R	x4		000:00000	010:00:0	000:00:0	2	0	0	8D030000	SC	0: 2780E400 000000E0 00000000 00000000 4: 00000000										
AHCI		H	Port	Slot	Address	Data Len	PRDT	DBA	DBAU	Reserved	I	DBC	Time Stamp										
82796		H	0	0	00000000:8D030080	4		0xBC281020	0x00000001	0x00000000	1	511	0111.539 048 110 s										
Split Tra		R	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data										
82779		R	x4		000:00000	010:00:0	000:00:0	0	0	0	8D030080	SC	201028BC 01000000 00000000 FF010080										
AHCI		D	Port	Slot	Address	Data Len	DATA	Data	Time Delta	Time Stamp													
82812		D	0	0	00000001:BC281020	128		128 dwords	6.648 us	0111.539 070 110 s													
AHCI		D	Port	Slot	Address	Data Len	PRDBC	PRDBC	Time Stamp														
82819		D	0	0	00000000:8D2FD004	1		512	0111.539 076 758 s														
Link Tra		R	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	ExplicitACK	Metrics	# Packets						
165627		R	x4	2962		010:00000	1	010:00:0	128	8D2FD004	1111	0000	00020000	0	Packet#331230		2						

Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation



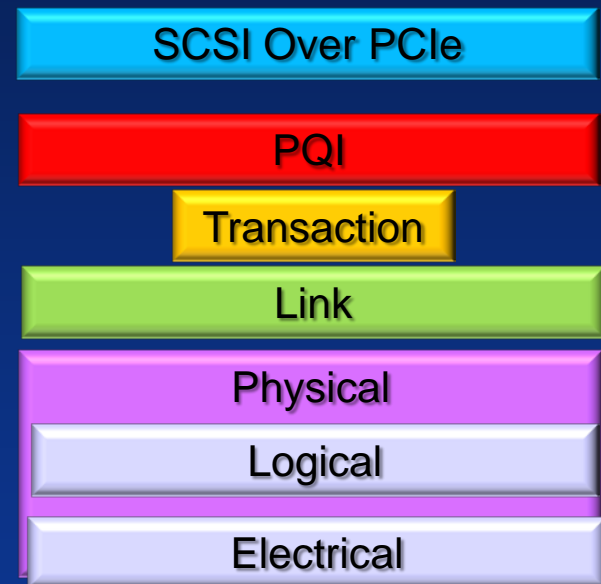
PCIe Architecture Queuing Interface (PQI)



SCSI Express

SCSI Over PCI Express(SOP/SOX)

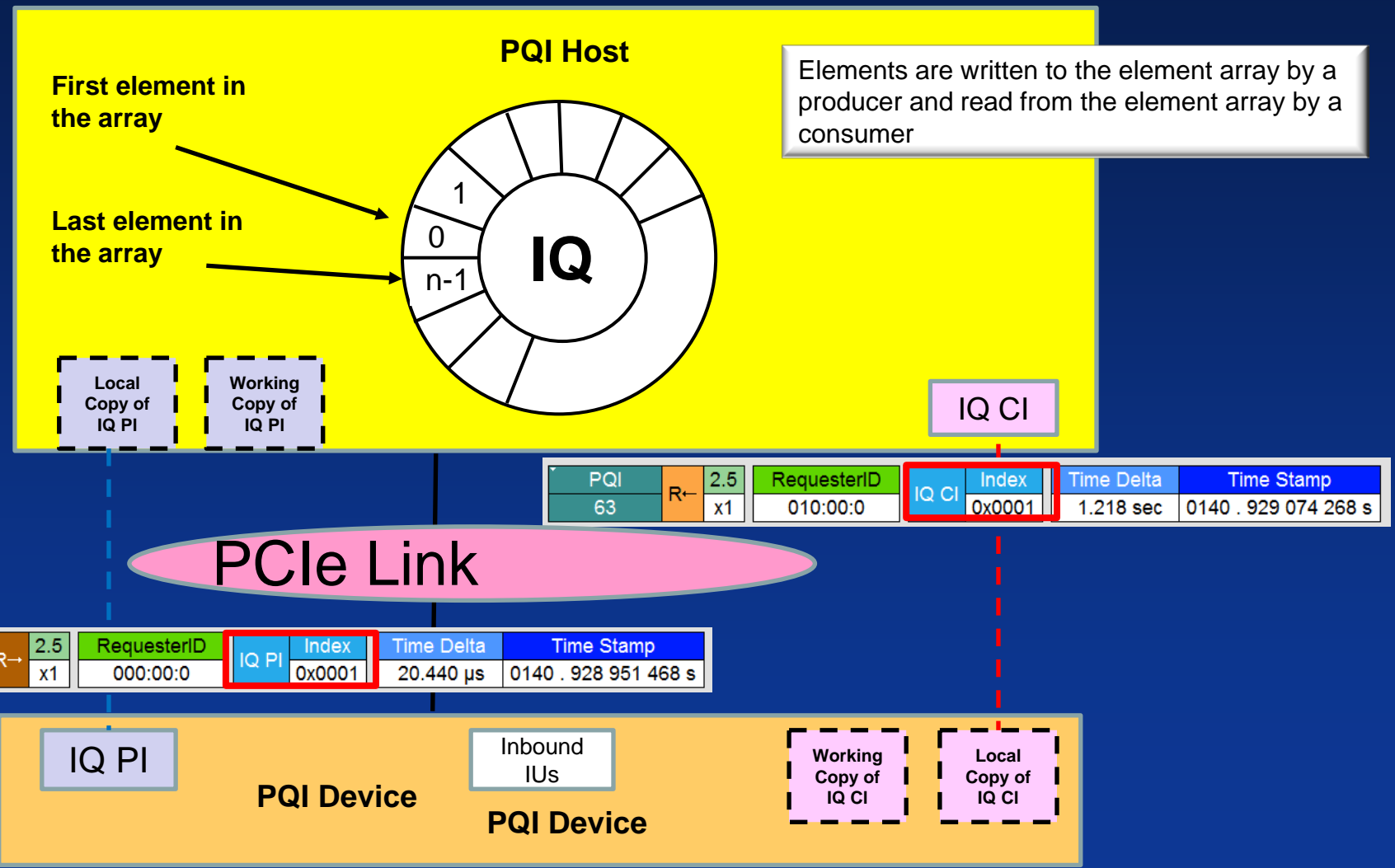
- Developed by T10 Committee
- Compliance with SCSI Architectural Model
- Proposed support for SOP target ports interfacing to flash devices, RAID controllers, and other SCSI peripheral device types
- Targeting PCIe 3.0



Preliminary
Protocol Stack

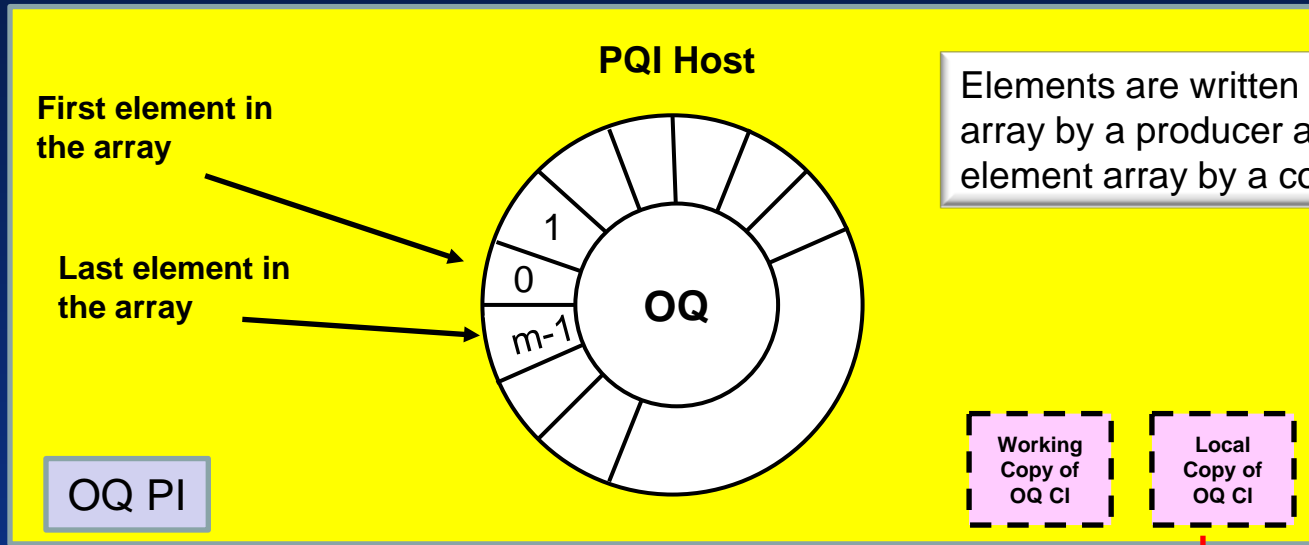


SCSI Express SOP/PQI IQ(Inbound Queue)



SCSI Express SOP/PQI

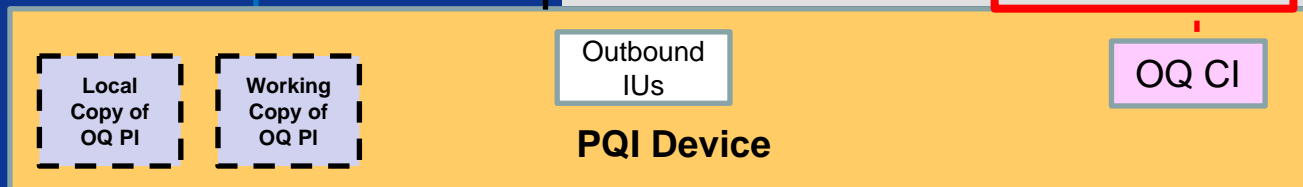
OQ(Outbound Queue)



PQI	R←	2.5	RequesterID	OQ PI	Index	Time Delta	Time Stamp
65		x1	010:00:0	0x0001		10.616 μs	0142 . 779 495 972 s



PQI	R→	2.5	RequesterID	OQ CI	Index	RI	Time Delta	Time Stamp
60		x1	000:00:0	0x00000001	0		9.702 ms	0140 . 919 249 844 s

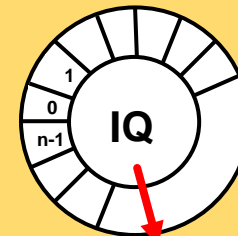
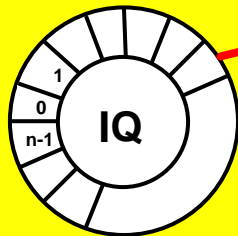


Creating Administrative and Operational Queues

Creating an Operational Inbound Queue

PQI	R+	2.5	RequesterID	CompleterID	Admin IU	IT	CF	IU LEN	ROQID	WA	Req ID	Op Code	IQ ID	IQEAA Hi	IQEAA Lo	IQCIA Hi
38		x1	010:00:0	000:00:0		0x60	0x00	0x003C	0x0000	0x0000	0x0004	Create Operational IQ	0x0000	0x00000000	0xBF70D000	0x00000000
IQCIA Lo	Num of Elements	Element Len	IMN	WFR	CC	MIN CT	MAX CT	Protocol	Time Delta	Time Stamp						
0xBF7089C8	0x00000040	0x0040	0x0000	0	0x0000	0x0000	0x0000	SOP	79.776 μs	0140 . 040 995 716 s						

Response

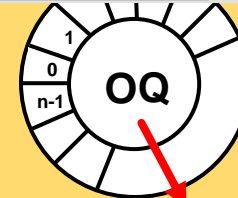
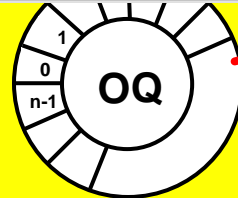


PQI	R+	2.5	RequesterID	Admin OU	IT	CF	IU LEN	Reserved	WA	Req ID	OP Code	Status	QPI Offset Hi	IQPI offset Lo	Time Delta
39		x1	010:00:0		0xE0	0x00	0x003C	0x0000	0x0000	0x0004	Create Operational IQ	Success	0x00000000	0x00001010	22.792 μs

Creating an Operational Outbound Queue

PQI	R+	2.5	RequesterID	CompleterID	Admin IU	IT	CF	IU LEN	ROQID	WA	Req ID	Op Code	OQ ID	OQEAA Hi	OQEAA Lo	OQPIA Hi
32		x1	010:00:0	000:00:0		0x60	0x00	0x003C	0x0000	0x0000	0x0000	Create Operational OQ	0x0000	0x00000000	0xBF712000	0x00000000
OQPIA Lo	Num of Elements	Element Len	IMN	WFR	CC	MIN CT	MAX CT	Protocol	Time Delta	Time Stamp						
0xBF708BD8	0x00000040	0x0040	0x0003	0	0x0000	0x0000	0x0000	SOP	101.840 μs	0140 . 038 995 284 s						

Response



PQI	R+	2.5	RequesterID	Admin OU	IT	CF	IU LEN	Reserved	WA	Req ID	OP Code	Status	OQCI Offset Hi	OQCI offset Lo	Time Delta
33		x1	010:00:0		0xE0	0x00	0x003C	0x0000	0x0000	0x0003	Create Operational OQ	Success	0x00000000	0x00001038	22.816 μs

Admin Queue

Operational Queue

SCSI Express Initialization

PQI	RequesterID	CompleterID	PAF	STATUS	Time Delta	Time Stamp																				
5	x1	000:01:0	010:00:0	Idle	1.551 ms	0140 . 032 247 892 s																				
6	x1	000:01:0	010:00:0	ADMIN Queue Parameter	ADMIN IQ ELEMENTS	ADMIN OQ ELEMENTS	INTERRUPT MESSAGE	Time Delta	Time Stamp																	
7	x1	000:01:0	010:00:0	CAPABILITY	MAX ADMIN IQ	MAX ADMIN OQ	AIQE LEN	AOQE LEN	Time Delta	Time Stamp																
8	x1	000:01:0	010:00:0	CAPABILITY	MAX ADMIN IQ	MAX ADMIN OQ	AIQE LEN	AOQE LEN	Time Delta	Time Stamp																
9	x1	000:00:0	010:00:0	ADMIN Queue Parameter	ADMIN IQ ELEMENTS	ADMIN OQ ELEMENTS	INTERRUPT MESSAGE	Time Delta	Time Stamp																	
10	x1	000:00:0	010:00:0	ADMIN IQ	Offset Hi	Offset Lo	Time Delta	Time Stamp																		
11	x1	000:00:0	010:00:0	ADMIN IQ CI	Offset Hi	Offset Lo	Time Delta	Time Stamp																		
12	x1	000:00:0	010:00:0	ADMIN OQ	Offset Hi	Offset Lo	Time Delta	Time Stamp																		
13	x1	000:00:0	010:00:0	ADMIN OQ PI	Offset Hi	Offset Lo	Time Delta	Time Stamp																		
14	x1	000:01:0	010:00:0	PAF	STATUS	Time Delta	Time Stamp																			
15	x1	000:00:0	010:00:0	PAF	STATUS	Time Delta	Time Stamp																			
16	x1	000:01:0	010:00:0	PAF	STATUS	Time Delta	Time Stamp																			
17	x1	000:01:0	010:00:0	ADMIN IQ PI	Offset Hi	Offset Lo	Time Delta	Time Stamp																		
18	x1	000:01:0	010:00:0	ADMIN OQ CI	Offset Hi	Offset Lo	Time Delta	Time Stamp																		
19	x1	000:00:0	010:00:0	RequesterID	Index	Time Delta	Time Stamp																			
20	x1	010:00:0	000:00:0	Admin IU	IT	CF	IU LEN	ROQID	WA	Req	Op Code	IQ ID	OQEAA Hi	OQEAA Lo	OQPIA Hi	OQPIA Lo	Num of Elements	Element Len	IMN	WFR	CC	MIN CT	MAX CT	Protocol	Time Delta	
											Create Operational OQ	0x0000	0x00000000	0xBF70E000	0x00000000	0xBF708BC3	0x00000040	0x0040	0x0001	0	0x0000	0x0000	0x0000	SOP	100.680 μs	
21	x1	010:00:0	000:00:0	Admin OQ	IT	CF	IU LEN	Reserved	WA	Req ID	OP Code	Status	OQCI Offset Hi	OQCI offset Lo	Time Delta	Time Stamp										
											Create Operational OQ	Success	0x00000000	0x00001018	22.792 μs	0140 . 035 085 540 s										

Creating Administrative IQ and OQ Queues

Creating Operational OQ Queue



SCSI Express SOP Transfer Packet

Advancing Producer in Inbound Queue

PQI	Req	2.5	RequesterID	IQ PI	Index	Time Delta	Time Stamp
55	R	x1	000:00:0	0x0001	21.304 µs	0140 . 113 039 492 s	

PQI	Req	2.5	RequesterID	CompleterID	SOP IU	IT	CF	IU LEN	ROQID	WA	Req ID	Data Dir	PARTIAL	DATA LEN	SCSI CDB	OP Code	SA	CDB Info	LBA	CDB Info	LEN	CDB Info	CONTROL
56	R	x1	010:00:0	000:00:0	0x10	0x00	0x003C	0x0003	0x0000	0x0000	0x0007	3	0	0x00000010	0xA0	0x0	0x0	0x00000000	0x00001000	0x00000000	0x00	0x00000000	

Advancing Consumer in Inbound Queue

PQI	Req	2.5	RequesterID	IQ CI	Index	Time Delta	Time Stamp
57	R	x1	010:00:0	0x0001	190.879 ms	0140 . 113 162 956 s	

Link Tra	Req	2.5	TLP	Msg	MsgD	Msg Routing	Length	RequesterID	Tag	Message Code	VID	Data	VC ID	Explicit ACK	Metrics	# Packets	Time Delta	Time Stamp
9610	R	x1	851	011:10011	Broadcast	1	000:00:0	0	Vendor_Defined_Type1	0x8086	1 dword	0	Packet #18600		2	135.073 ms	0140 . 604 041 748 s	

Link Tra	Req	2.5	TLP	Mem	MWrr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	Explicit ACK	Metrics	# Packets	Time Delta	Time Stamp
9611	R	x1	568	010:00000	4	010:00:0	0	14281120	1111	1111	4 dwords	0	Packet #18602		2	180.095 ms	0140 . 739 115 076 s	

Advancing Producer in Outbound Queue

PQI	Req	2.5	RequesterID	OQ PI	Index	Time Delta	Time Stamp			
58	R	x1	010:00:0	0x90	0x00	0x000C	0x0003	0x0000	0x0007	0x0007
59	R	x1	010:00:0	0x0001	10.664 µs	0140 . 919 233 540 s				

Link Tra	Req	2.5	TLP	Mem	MWrr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	Explicit ACK	Metrics	# Packets	Time Delta	Time Stamp
9614	R	x1	571	010:00000	1	010:00:0	0	FEE22000	1111	0000	1 dword	0	Packet #18608		2	5.640 µs	0140 . 919 244 204 s	

PQI	Req	2.5	RequesterID	OQ CI	Index	Time Delta	Time Stamp
60	R	x1	000:00:0	0x00000001	9.702 ms	0140 . 919 249 844 s	

Advancing Consumer in Outbound Queue

Courtesy SanDisk 2012



PQI initialization

▼ PQI 6	H	Admin Queue Parameter	ADMIN IQ ELEMENTS	ADMIN OQ ELEMENTS	INTERRUPT MESSAGE
			64	64	0
▼ PQI 7	H	Admin IQ Mem Addr	Address		
			00000000:BF709000		
▼ PQI 8	H	Admin IQ CI Mem Addr	Address		
			00000000:BF708190		
▼ PQI 9	H	Admin OQ Mem Addr	Address		
			00000000:BF70A000		
▼ PQI 10	H	Admin OQ PI Mem Addr	Address		
			00000000:BF7081B0		
▼ PQI 11	D	AQ Config Function	Function and status		
			IDLE		
▼ PQI 12	H	AQ Config Function	Function and status		
			Create AQ Pair		
▼ PQI 13	D	AQ Config Function	Function and status		
			IDLE		
▼ PQI 14	D	Admin IQ PI Offset	Offset		
			00000000:00001000		
▼ PQI 15	D	Admin OQ CI Offset	Offset		
			00000000:00001008		



PQI: Create Operational Queues

PQI	H	AIQ	Admin IQ PI	Index										
16				0x0001										
PQI	H	AIQ	Elem	Adm IQ IU	Type	Rsv	Cmp	Len	WA	RQ ID	OP	Queue ID	Queue Element Array Address	
17			0x0000		Adm RQ	0x0	0x0	0x003C	0x0000	1	Create OP OQ	0x0001	00000000:BF70E000	
OQ PI Address		OQ Number of Elements		OQ Element Length		IMN	WFR	CC	Min CT	Max CT	Op Queue Protocol			
00000000:BF708BC8		64		64 (1024 bytes)		1	0	0	0	0	SOP			
PQI	D	A OQ	Elem	Adm OQ IU	Type	Rsv	Cmp	Len	WA	RQ ID	OP	Status	OQ CI Offset	
18			0x0000		Adm RSP	0x0	0x0	0x003C	0x0000	1	Create OP OQ	Good	00000000:00001018	
PQI	D	A OQ	Admin OQ PI	Index										
19				0x0001										
Link Tra	R←	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	
9563		x1	539		010:00000	1	010:00:0	0	FEE00000	1111	0000	B0400000	0	
Explicit ACK		Metrics		# Packets										
Packet #18511				2										
PQI	D	AIQ	Admin IQ CI	Index										
20				0x0001										
PQI	H	A OQ	Admin OQ CI	Index	RI									
21				0x0001	0									
PQI	H	AIQ	Admin IQ PI	Index										
22				0x0002										
PQI	H	AIQ	Elem	Adm IQ IU	Type	Rsv	Cmp	Len	WA	RQ ID	OP	Queue ID	Queue Element Array Address	
23			0x0001		Adm RQ	0x0	0x0	0x003C	0x0000	2	Create OP OQ	0x0002	00000000:BF710000	
OQ PI Address		OQ Number of Elements		OQ Element Length		IMN	WFR	CC	Min CT	Max CT	Op Queue Protocol			
00000000:BF708BD0		64		64 (1024 bytes)		2	0	0	0	0	SOP			
PQI	D	A OQ	Elem	Adm OQ IU	Type	Rsv	Cmp	Len	WA	RQ ID	OP	Status	OQ CI Offset	
24			0x0001		Adm RSP	0x0	0x0	0x003C	0x0000	2	Create OP OQ	Good	00000000:00001028	
PQI	D	A OQ	Admin OQ PI	Index										
25				0x0002										
Link Tra	R←	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	
9571		x1	544		010:00000	1	010:00:0	0	FEE00000	1111	0000	B0400000	0	



PQI: Create Op Queue detailed

PQI	H	A IQ	Admin IQ PI	Index																
16				0x0001																
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID							
9558		x1	831		010:00000	1	000:00:0	7	F8FF1000	1111	0000	01000000	0							
Explicit ACK													Metrics	# Packets	Resp. time	Pld. Bytes	Thrpt MB/s			
Packet #18501														2	2.120 us	4	1.799			
PQI	H	A IQ	Elem	Adm IQ IU	Type	Rsv	Cmp	Len	WA	RQ ID	OP	Queue ID	Queue Element Array Address							
17			0x0000		Adm RQ	0x0	0x0	0x003C	0x0000	1	Create OP OQ	0x0001	00000000:BF70E000							
OQ PI Address													OQ Number of Elements	OQ Element Length	IMN	WFR	CC	Min CT	Max CT	Op Queue Protocol
00000000:BF708BC8													64	64 (1024 bytes)	1	0	0	0	0	SOP
Split Tra	R→	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status									
4632		x1		000:00000	010:00:0	000:00:0	0	0	0	BF709000	SC									
Data												Metrics	# LinkTras							
0: 60003C00 00000000 01001101 00000000 00E070BF 00000000 C88B70BF 00000000													2							
8: 40000000 40000100 00000000 00000000 00000000 00000000 00000000 00000000																				
PQI	D	A OQ	Elem	Adm OQ IU	Type	Rsv	Cmp	Len	WA	RQ ID	OP	Status	OQ CI Offset							
18			0x0000		Adm RSP	0x0	0x0	0x003C	0x0000	1	Create OP OQ	Good	00000000:00001018							
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE									
9561		x1	537		010:00000	16	010:00:0	0	BF70A000	1111	1111									
Data												VC ID	Explicit ACK	Metrics						
0: E0003C00 00000000 01001100 00000000 18100000 00000000 00000000 00000000												0	Packet #18507							
8: 00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000																				
# Packets													Resp. time	Pld. Bytes	Thrpt MB/s					
2													2.000 us	64	30.518					
PQI	D	A OQ	Admin OQ PI	Index																
19				0x0001																
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	Explicit ACK						
9563		x1	539		010:00000	1	010:00:0	0	FEE00000	1111	0000	B0400000	0	Packet #18511						
Metrics													# Packets	Resp. time	Pld. Bytes	Thrpt MB/s				
2													1.760 us	4	2.167					
PQI	D	A IQ	Admin IQ CI	Index																
20				0x0001																
PQI	H	A OQ	Admin OQ CI	Index	RI															
21				0x0001	0															



SCSI command: WRITE (10)

SCSI	SCSI Error	SCSI CDB	OpCode	WRPROTECT	DPO	FUA	FUA_NV	Obsolete	Logical Block Addr	GROUP NUMBER	Transfer Len	CONTROL
5	Error in SOP Sub-Tra	WRITE(10)	0x2A	0x0	Not set	Not set	Not set	Not set	0x000007B0	0x00	8	0x00

Data													Status	LUN
0:	5FC70167	A383C49B	56BF8AE4	D53BC84A	BA9D2D50	752FA850	4CF46AFA	B336393F	B4685E8D	39D8398C	9CBA37A1	15905BF9	GOOD	0x0
12:	BD5CE0AA	785454A1	25115B61	C6AB18A0	154382B3	8D121353	5EF26D40	68ACEED5	AB3F388D	54CE4363	D7226551	51023715		
24:	9FB2D796	C99AA193	A2A920D1	0693C359	5E6B1F7F	DD683329	9A45448E	ABB3A5C3	61E136CB	3D25496D	CB06BE92	FB4DB47B		
36:	BA5DA111	A9D34ABA	42338E0B	60251EE2	26D23FB0	711F011A	2F49399C	9B00F80E	8DD8F2B5	CCE3C1D6	BF092162	EC82A733		
48:	D89BEE65	B873C6B5	325643B1	7E7F8DDC	9060E7DE	1F36201C	2E6608B2	845A8019	AE4E9E28	6D5EDB62	DADC095C	024F5C0E		
60:	C161CB8A	72358C10	253817CD	4A94F037	C5917CF7	ACD60339	DA9CDC14	43FD009C	9A15F979	0B43EE8A	68146E27	9097CE92		
72:	1F1E39D8	1A657C39	8D18EC3C	AC169C3B	E17E6EAB	F27F3C17	17BFED93	BCD79552	EB980A6A	8CD499A9	4EDA3633	2E96B20D		
84:	D65FBD76	6B261EC6	8CE33D10	4516CA0C	3C4FC1B5	0224089E	9A07187A	EE004BD5	691F88B3	D0EAAAF5	1656BC3B	09AF3E6D		
96:	E05ACA0F	5488DE83	359DE82A	2375F906	06035814	9B25574D	5280030A	DC39738B	1E4D2D25	5EF67E7E	4E20BB61	083F6D29		
108:	20B35970	1B92A508	CBBCF789	AA63E9E4	7798ACB4	1EECE65A	37A17066	BD3B13AD	385DEB4F	79C549EA	CE2BB1F8	F3D18355		
120:	38010398	5E7DF680	A6669660	CC5474BF	8F124B54	5A068508	F5A763AD	AB198A9F	...					

Metrics	# SOP IUs
	3

SOP	H	Type	Cmp	Rstr	Len	Rsp QID	WA	RQ ID	Data Dir	Partial	Fence	Size	CDB	SGL	Address	Length
15		Lim Cmd	0x0	0x0	0x003C	0x0001	0x0000	0x002A	Data-Out	0	0	0x00001000	16 bytes		00000000:447CD000	0x00001000

Zero	Type	Address	Length	Zero	Type
0x0	Data Block	00000000:00000000	0x00000000	0x0	Data Block

PQI	H	IQ ID	Elem	Op IQ IU	Rstr	Rstr	Cmp	Len	Data
331		0x0001	0x0000		0x10	0x0	0x0	0x003C	60 bytes

SOP	H	Tra	Address	Length	Data
16		15	00000000:447CD000	0x00001000	1024 dwords

SOP	D	1 error(s)	Rsvd not zero	Type	Cmp	Rstr	Len	WA	RQ ID	Nexus ID
17			Field: Rsvd	Success	0x0	0x0	0x000C	0x0000	0x002A	0x0000



SCSI command: Report LUNS

SCSI	SCSI Error	SCSI CDB	OpCode	SELECT REPORT	AllocLen	CONTROL	Data	Status	LUN	Metrics	# SOP IUs
0	Error in SOP Sub-Tra	REPORT LUNS	0xA0	0x0000	16	0x00	4 dwords	GOOD	0x0		3

SOP	H	Type	Cmp	Rstr	Len	Rsp QID	WA	RQ ID	Data Dir	Partial	Fence	Size	CDB	SGL	Address	Length	Zero
0		Lim Cmd	0x0	0x0	0x003C	0x0003	0x0000	0x0007	Data-Out	0	0	0x00000010	16 bytes		00000000:14281120	0x00000010	0x0

Type	Address	Length	Zero	Type
Data Block	00000000:00000000	0x00000000	0x0	Data Block

PQI	H	IQ ID	Elem	Op IQ IU	Rstr	Rstr	Cmp	Len	Data
53		0x0001	0x0010		0x10	0x0	0x0	0x003C	60 bytes

Split Tra	R	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	Metrics	# LinkTras
4638	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	2	0	0	BF711000	SC	16 dwords	2
4643	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	3	0	0	BF711040	SC	16 dwords	2
4919	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	3	0	0	BF711080	SC	16 dwords	2
6950	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	1	0	0	BF7110C0	SC	16 dwords	2
6951	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	2	0	0	BF711100	SC	16 dwords	2
6952	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	3	0	0	BF711140	SC	16 dwords	2
6953	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	0	0	0	BF711180	SC	16 dwords	2
6956	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	2	0	0	BF7111C0	SC	16 dwords	2
6963	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	3	0	0	BF711200	SC	16 dwords	2
6966	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	1	0	0	BF711240	SC	16 dwords	2
6973	R	x1	Mem	MRd(32)	000:00000	010:00:0	000:00:0	2	0	0	BF711280	SC	16 dwords	2

SOP	D	Tra	Address	Length	Data
1		0	00000000:14281120	0x00000010	4 dwords

SOP	D	1 error(s)	Rsvd not zero	Type	Cmp	Rstr	Len	WA	RQ ID	Nexus ID
2		Rsvd not zero	Field: Rsvd	Success	0x0	0x0	0x000C	0x0000	0x0007	0x0000

Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation

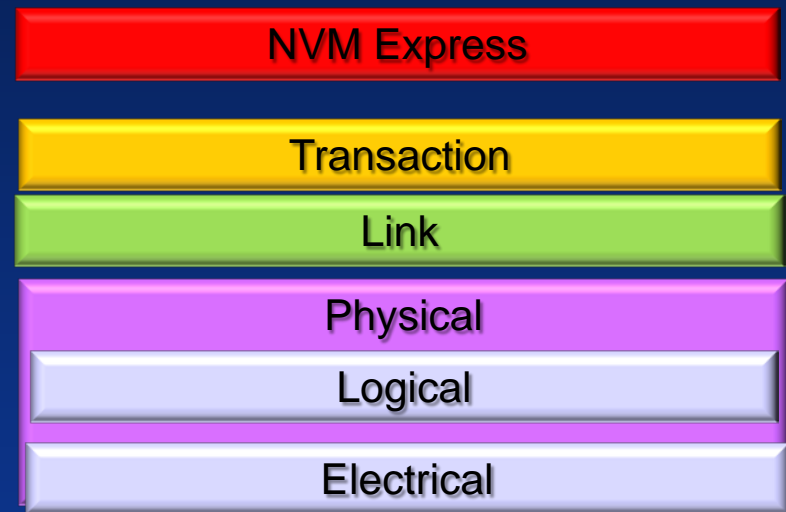


NVM Express





- The NVMHCI Workgroup released the NVM Express 1.0 specification on March 1, 2011 and is available at www.nvmexpress.org
- NVMe is a standardized high performance queuing interface and command set optimized for PCIe SSDs
- NVMe is scalable from client to enterprise applications



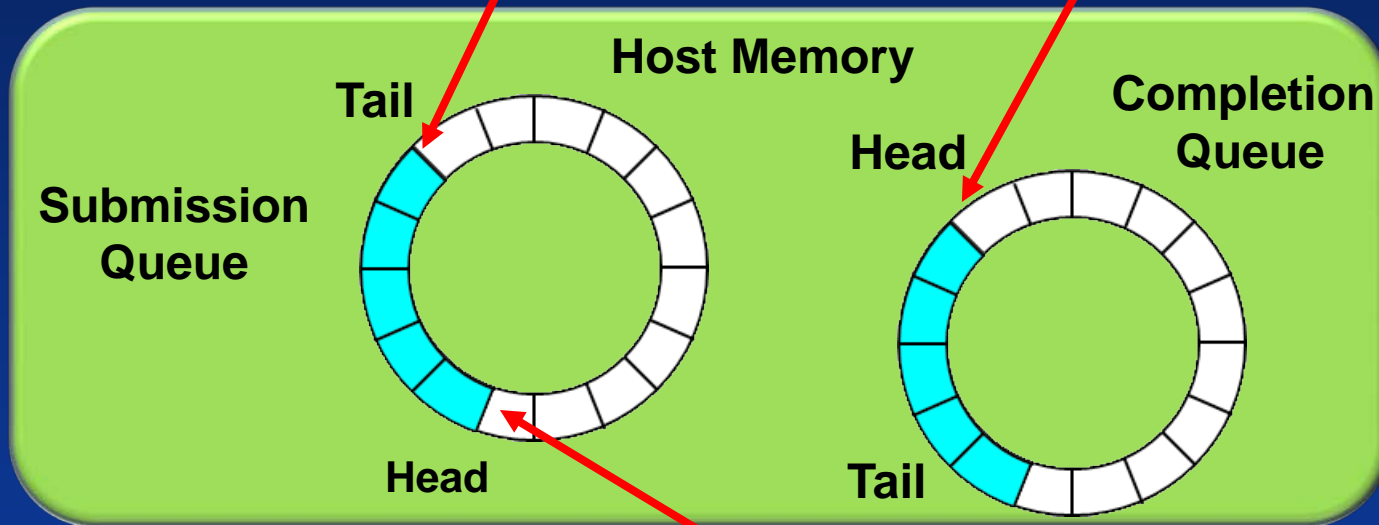
Submission and Completion Queues

NVM	R→	2.5	RequesterID	Admin SQT QID = 0	Time Delta	Time Stamp
5		x4	000:00:0	0x0001	352.000 ns	0048 . 780 381 524 s

Submission Queue Tail

Completion Queue Head Pointer

NVM	R→	2.5	RequesterID	Admin CQH QID = 0	Time Delta	Time Stamp
22		x4	000:00:0	0x0003	271.860 μs	0078 . 827 032 124 s

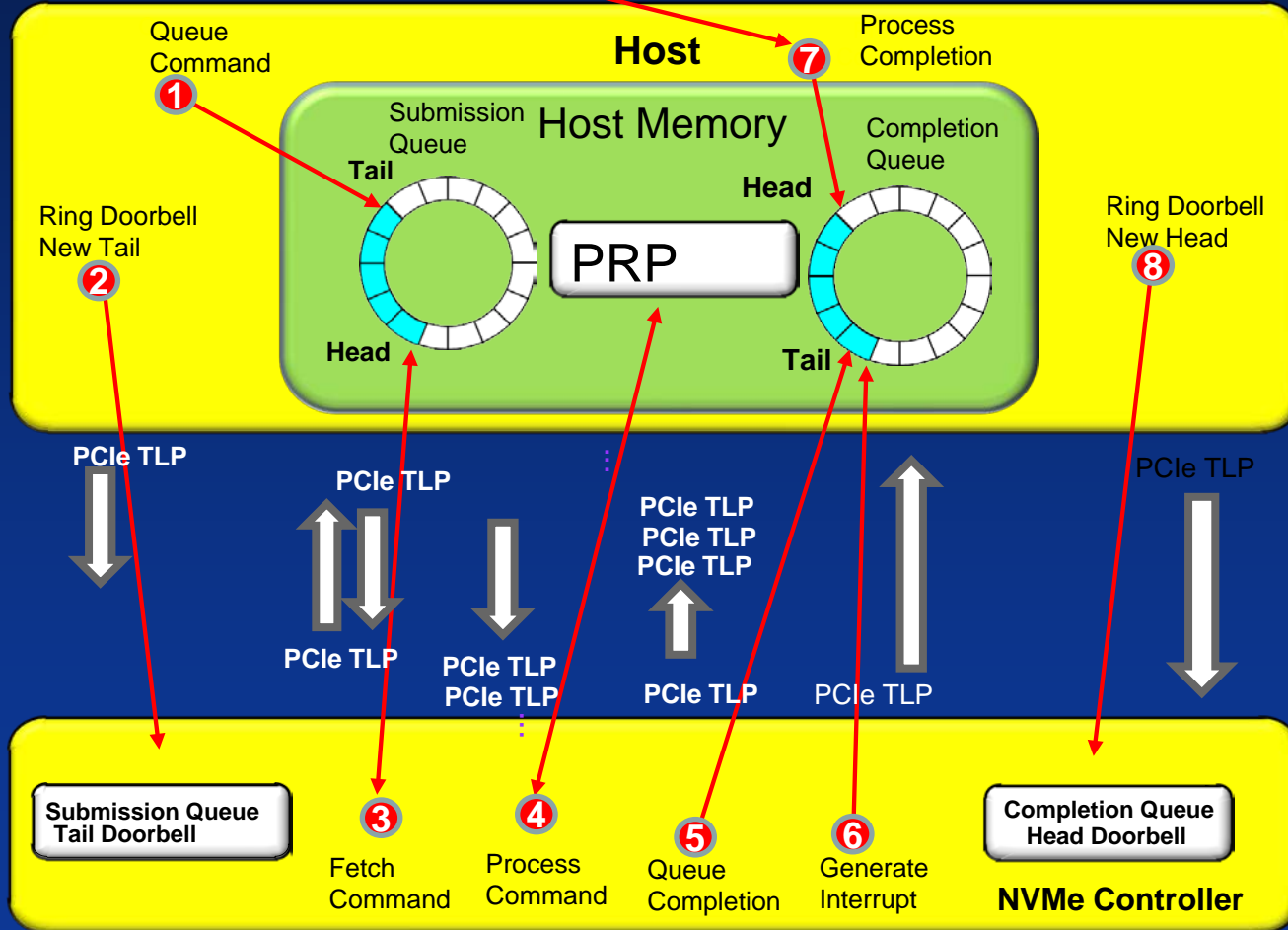


NVM	R←	2.5	RequesterID	Command Completion	SQHD	SQID	CID	P	ST	SC	SCT	M	DNR	Time Delta	Time Stamp
111		x4	007:00:0	0x00000000	0x0023	0x0000	0x0022	1		0x0000	0x00	0	0	26.752 μs	0030 . 998 377 752 s

Submission Queue Head Pointer

Circular Queuing Interface

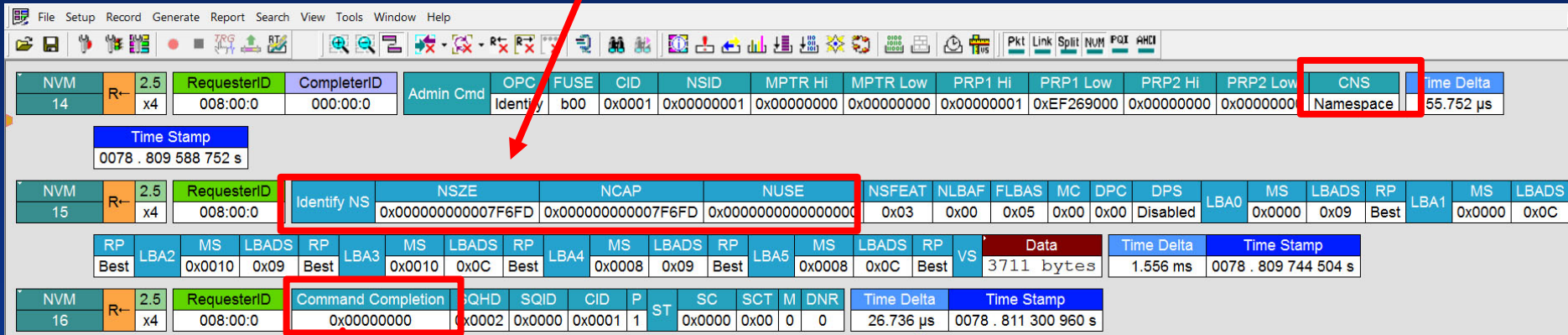
NVM	R←	2.5	RequesterID	Command Completion	SQHD	SQID	CID	P	ST	SC	SCT	M	DNR	Time Delta	Time Stamp
111		x4	007:00:0	0x00000000	0x0023	0x0000	0x0022	1		0x0000	0x00	0	0	26.752 μs	0030 . 998 377 752 s



Viewing the NVM Express 1.0c

Identify Namespace Capabilities command

Namespace capabilities



NVM	RequesterID	CompleterID	Admin Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	CNS	Time Delta					
14	008:00:0	000:00:0	Identify	b00	0x0001	0x00000001	0x00000000	0x00000000	0x00000000	0x00000001	0xEF269000	0x00000000	0x00000000	Namespace	55.752 μs					
Time Stamp 0078 . 809 588 752 s																				
NVM	RequesterID	Identify NS	NSZE	NCAP	NUSE	NSFEAT	NLBAF	FLBAS	MC	DPC	DPS	LBA0	MS	LBADS	RP	LBA1	MS	LBADS		
15	008:00:0	0x000000000007F6FD	0x000000000007F6FD	0x0000000000000000	0x03	0x00	0x05	0x00	0x00	Disabled		0x0000	0x09	Best		0x0000	0x0C			
RP	LBA2	MS	LBADS	RP	LBA3	MS	LBADS	RP	LBA4	MS	LBADS	RP	LBA5	MS	LBADS	RP	VS	Data	Time Delta	Time Stamp
Best	0x0010	0x09	Best	0x0010	0x0C	Best	0x0008	0x09	Best	0x0008	0x0C	Best						3711 bytes	1.556 ms	0078 . 809 744 504 s
NVM	RequesterID	Command Completion	QID	SQID	CID	P	ST	SC	SCT	M	DNR	Time Delta	Time Stamp							
16	008:00:0	0x00000000	0x0002	0x0000	0x0001	1		0x0000	0x00	0	0	26.736 μs	0078 . 811 300 960 s							

Command Completion



Viewing the NVM Express 1.0c

Creating an I/O queue and Read command example

Create IO Submission Queue

NVM	RequesterID	CompleterID	Admin Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	QSIZE	QID	CQID	QPRIQ	PC
150	008:00:0	000:00:0	Create I/O SQ	b00	0x0024	0x00000000	0x00000000	0x00000000	0x00000000	0xEF31A000	0x00000000	0x00000000	0x00000000	1023	3	3	Medium	1
NVM	RequesterID	CompleterID	Command Completion	SQHD	SQID	CID	P	ST	SC	SCT	M	DNR	Time Delta	Time Stamp				
151	008:00:0	000:00:0	0x00000000	0x0025	0x0000	0x0024	1		0x0000	0x00	0	0	26.736 us	0048 . 897 517 192 s				

Courtesy SanDisk 2012

Data Transfer

Command Completion

Physical Region Page pointer

NVM	RequesterID	CompleterID	IO Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	SLBA	LR	FUA	PRINFO	
195	008:00:0	000:00:0	Read	b00	0x0000	0x00000001	0x00000000	0x00000000	0x00000000	0x00000001	0xEED31000	0x00000000	0x00000000	0x00000000:00000000	0	0	0x0	
NLB	DSM	Incompressible	SR	AL	AF	EILBRT	ELBAT	ELBATM	Time Delta	Time Stamp								
0x0000		0	0	None	None	0x00000000	0x0000	0x0000	163.968 μs	0079 . 553 228 928 s								
NVM	RequesterID	CompleterID	CMD PRP	Addr Hi	Addr Lo	Data Len	Data	Time Delta	Time Stamp									
196	008:00:0	000:00:0	0x00000001	0xEED31000	0x00000200	127 quadlets		313.872 μs	0079 . 553 392 896 s									
NVM	RequesterID	CompleterID	Command Completion	SQHD	SQID	CID	P	ST	SC	SCT	M	DNR	Time Delta	Time Stamp				
197	008:00:0	000:00:0	0x00000000	0x0001	0x0010	0x0000	1		0x0000	0x00	0	0	26.720 μs	0079 . 553 706 768 s				

Courtesy SanDisk 2012

Data



Viewing the NVM Express 1.0c NVMe Multiple Pointer Based Transactions

Physical Region Page (PRP)

NVM	RequesterID	IO SQT QID	Time Delta	Time Stamp
89	x8	000:00:0	452.000 ns	0009.438 505 794 s
90	x8	129:00:0	000:31:1	

NVM	RequesterID	CompleterID	IO Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	SLBA	LR	FUA	PRINFO	NLB	DSM	Incompressib
90	x8	000:31:1	Read	b00	0x000A	0x00000001	0x00000000	0x00000000	0x00000000	0x00000008	0x8FF41000	0x00000004	0x2FBFFB00	0x00000000:00000800	0	0	0x0	0x001F	DSM	0

NVM	RequesterID	CompleterID	PRP LIST PTR	Addr Hi	Addr Lo	Data Len	Data	Time Delta	Time Stamp
91	x8	000:31:1		0x00000008	0x2FBFFB00	0x00000000	0: 8FF42000 00000008 8FF43000 00000000 4: 8FF44000 00000008 00000000 00000000.	168.000 ns	0009.438 507 150 s

NVM	RequesterID	CompleterID	CMD PRP	Addr Hi	Addr Lo	Data Len	Data	Time Delta	Time Stamp
92	x8	129:00:0		0x00000008	0x8FF41000	0x00001000	1023 quadlets	1.816 µs	0009.438 516 318 s
93	x8	129:00:0		0x00000008	0x8FF42000	0x00001000	1023 quadlets	2.124 µs	0009.438 509 166 s
94	x8	129:00:0		0x00000008	0x8FF43000	0x00001000	1023 quadlets	1.284 µs	0009.432 511 290 s
95	x8	129:00:0		0x00000008	0x8FF44000	0x00001000	0: 00010000 00090008 00530052 00550054 00570056 00590058 005B005A 005D005C 00090008 00530052 00550054 00570056 12: 00190018 00210020 004B004A 004D004C 004F004E 00510050 00530052 00550054 00210020 004B004A 004D004C 004F004E 24: 00310030 00390038 00450044 00370036 00330032 003B003A 003D003C 003F003E 00390038 00450044 00370036 00330032 36: 00490048 00510050 005D005C 004F004E 004B004A 00530052 00550054 00570056 00510050 005D005C 004F004E 004B004A 48: 00410060 00490048 00550054 00470046 00430042 004B004A 004D004C 004F004E 00490048 00550054 00470046 00430042 60: 00590058 00810080 008D008C 007F007E 007B005A 00830082 00850084 00870086 00810080 008D008C 007F007E 007B005A 72: 00910090 00C000CC 00C000CE 00D100D0 00D300D2 00D500D4 009D009C 009F009E 00C000CC 00C000CE 00D100D0 00D300D2 84: 00A900A8 00C500C4 00C700C6 00C900C8 00CB00CA 00CD00CC 00B500B4 00B700B6 00C500C4 00C700C6 00C900C8 00CB00CA 96: 00C100C0 00C900C8 00D500D4 00C700C6 00C300C2 00CB00CA 00CD00CC 00CF00CE 00D100D0 00D300D2 00D500D4 00D700D6 108: 00D900D8 00C100C0 00CD00CC 00DF00DE 005D005C 00110010 00130012 00150014 00C900C8 00CB00CA 00CD00CC 00CF00CE 120: 00D900D8 00D500D4 00F700D6 00550054 00290028 002B002A 002D002C ...		

NVM	RequesterID	Command Completion	SQID	SQID	CID	P	SC	SCI	M	DNK	Time Delta	Time Stamp
96	x8	0x00000000	0x000B	0x0001	0x000A	1	ST	0x0000	0x00	0	48.000 ns	0009.438 514 894 s

PRP List of pointers to memory addresses

Data

Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation



NVMe Device script Submission Queue Setup

NVM 0	R→	2.5 x4	RequesterID 000:00:0	AQA	ASQS 383	ACQS 383	Time Stamp 0029.315 690 784 s							
Link Tra 11783	R→	2.5 x4	TLP 2237	Mem	MWr(32) 010:00000	Length 1	RequesterID 000:00:0	Tag 10	Address B1A00024	1st BE 1111	Last BE 0000	Data 017F017F	VC ID 0	Exp Pack
NVM 1	R→	2.5 x4	RequesterID 000:00:0	ASQ	ASQB AddressHi 0x00000000	ASQB AddressLow 0x7F55A000	Time Stamp 0029.315 690 956 s							
Link Tra 11784	R→	2.5 x4	TLP 2238	Mem	MWr(32) 010:00000	Length 2	RequesterID 000:00:0	Tag 11	Address B1A00028	1st BE 1111	Last BE 1111	Data 7F55A000 00000000		
NVM 2	R→	2.5 x4	RequesterID 000:00:0	ACQ	ACQB AddressHi 0x00000000	ACQB AddressLow 0x7F561000	Time Stamp 0029.315 691 120 s							
Link Tra 11785	R→	2.5 x4	TLP 2239	Mem	MWr(32) 010:00000	Length 2	RequesterID 000:00:0	Tag 8	Address B1A00030	1st BE 1111	Last BE 1111	Data 7F561000 00000000		

- The host creates a command for execution within the appropriate Submission Queue
- Admin submission queue base register 0x7F55A000 is written to controller register, later used by controller to fetch the read command from host memory

```
# NVM 0
Wait=TLP {
    TLPTType=MWr32
    Length = 1 }
#NVM 1
Wait=TLP {
    TLPTType=MWr32
    Length = 2 }
```



NVMe Device script Command fetch

NVM	2.5	RequesterID	CompleterID	Admin Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	CNS	Time Stamp
8	R-	x4	007:00:0	000:00:0	Identify	b00	0x0000	0x00000000	0x00000000	0x00000000	0x00000000	0x7F5FA000	0x00000000	0x00000000	Controller	0030 . 972 808 984 s

Split Tra	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data
5845	R-	x4	000:00000	007:00:0	000:00:0	0	0	0	7F55A000	SC	0: 00000006 00000000 00000000 00000000 00000000 00000000 7F5FA000 00000000 8: 00000000 00000000 00000001 00000000 00000000 00000000 00000000 00000000

NVM	2.5	RequesterID	Identify CDS	VID	SSVID	SN	MN	FR	RAB	IEEE	MIC	MDTS	OACS	ACL	AERL	FRMW	LPA	ELPE	NPSS	AVSOC	SQES	CQES	NN	ONCS	FUSES	FNA	VWC	AWUN	AWUPF	N
9	R-	x4	007:00:0	0x111D	0x0000				0	0x0	0	0	0x0000	255	0	4	0	0	0	0	102	68	0x00000004	0x0000	0x0000	0	0	0x00FF	0x00FF	N

EXLAT	RRT	RRL	RWT	RWL	VS	Data	Time Stamp
0x00000000	0x00	0x00	0x00	0x00		1023 bytes	0030 . 972 981 232 s

Link Tra	2.5	TLP	Mem	MW(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	ExplicitACK	Metrics	# Packets	Time Delta	Time Stamp
11828	R-	x4	2128	010:00000	64	007:00:0	0	7F5FA000	1111	1111	64 dwords	0	Packet#23593		2	79.632 us	0030 . 972 981 232 s

- The controller fetches the command(s) in the Submission Queue from host memory
- The Admin submission queue base address register 0x7F55A000 is read from implemented memory resource through field substitution

```

packet = TLP {
  TLPType = MRd64
  Length = 0x10
  RequesterId = (8:0:0)
  Tag = 0x0
  LastDwBe = 0xF
  FirstDwBe = 0xF
  AddressHi = ( FROM_MEM32_A, 0x28 )
              # 0x1
  AddressLo = ( FROM_MEM32_A, 0x2C )
              # 0xEF224000 }
  
```


Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation

Testing NVMe Express

- NVMe Registers emulation
 - Setup Admin Queue
 - Doorbell Registers
- Admin Commands
 - Delete I/O Submission Queue
 - Create I/O Submission Queue
 - Delete I/O Completion Queue
 - Create I/O Completion Queue
 - Get Log Page
 - Identify
 - Abort
 - Set Features
 - Get Features
 - Format NVM
 - Extensible for Vendor Specific Commands



Summit Z3-16 Protocol Exerciser

Pre-Silicon Device emulation

FlashMemory
SUMMIT



RTL Design

RTL Test Vectors

SimPass

CATC Trace

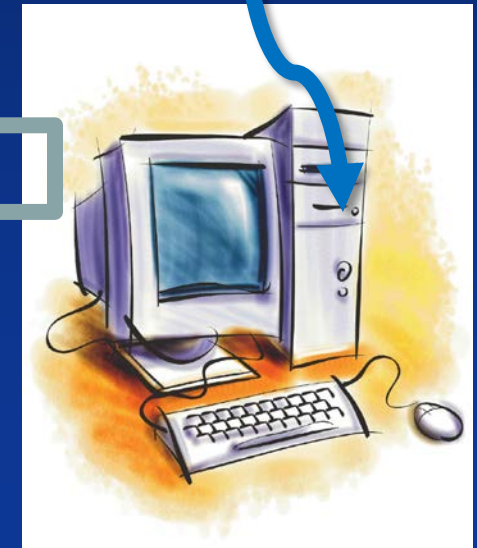
PE Tracer
Export to
Generation
script

Z3 Exerciser
Generation Script

Summit Z3



Summit T3



Loading Config Space and Implementing Memory Space

Write Address Space

	File Path:	Offset (bytes):	Size (bytes):
<input checked="" type="checkbox"/> Cfg :	cripts\sys_device_cfg_space_w_msi_x_1_bar		
<input type="checkbox"/> Mem64 :		0x00000000	0x20000000
<input type="checkbox"/> Mem32 A :		0x00000000	0x08000000
<input checked="" type="checkbox"/> Mem32 B :	F:\Data_All\NVMe_Emulation\identify1	0x00000000	0x00002000
<input type="checkbox"/> IO A :		0x00000000	0x00000100
<input type="checkbox"/> IO B :		0x00000000	0x00000100

Buttons: Clear, Write, Cancel

Summit Z3-16 SN:63055

Role	Speed	Link State
Device	2.5 GT/s	???

Summit T3-16 SN:62102

DS	x16	2.5	Speed	Link State
DS	x16	2.5	●●●●●●●●●●	●●●●●●●●●●
US	x16	2.5	●●●●●●●●●●	●●●●●●●●●●

Ready



SSD Drive Emulation

NVM	R→	2.5	RequesterID	ASQS	ACQS	Time Delta	Time Stamp
4	x4	x4	000:00:0	127	127	1.464 µs	0199 . 362 722 052 s

NVM	R→	2.5	RequesterID	ASQB AddressHi	ASQB AddressLow	Time Delta	Time Stamp
5	x4	x4	000:00:0	0x00000002	0x3D61B000	2.872 µs	0199 . 362 723 516 s

NVM	R→	2.5	RequesterID	ACQB AddressHi	ACQB AddressLow	Time Delta	Time Stamp
6	x4	x4	000:00:0	0x00000002	0x3D61D000	2.904 µs	0199 . 362 726 388 s

NVM	R→	2.5	RequesterID	EN	CSS	MPS	AMS	SHN	IOSQES	IOCQES	Time Delta	Time Stamp
7	x4	x4	000:00:0	1	NVM command set	0	b00	b00	0	0	34.200 µs	0199 . 362 729 292 s

NVM	R→	2.5	RequesterID	CompleterID	RDY	CFS	SHST	Time Delta	Time Stamp
8	x4	x4	000:00:0	008:00:0	1	0	b00	14.988 ms	0199 . 362 763 492 s

NVM	R→	2.5	RequesterID	Admin SQT QID = 0	Time Delta	Time Stamp
9	x4	x4	000:00:0	0x0001	192.276 µs	0199 . 377 751 104 s

NVM	R→	2.5	RequesterID	CompleterID	Admin Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	CNS
10	x4	x4	008:00:0	000:00:0	Identify	b00	0x0000	0x00000000	0x00000000	0x00000000	0x00000000	0x00000002	0x3D2DA160	0x00000002	0x3D2DB000	Controller

Time Delta	Time Stamp
1.179 ms	0199 . 377 943 380 s

NVM	R→	2.5	RequesterID	Identify CDS	VID	SSVID	SN	MN	FR	RAB	IEEE	MIC	MDTS	OACS	ACL	AERL	FRMW	LPA	ELPE	NPSS	AVSCC	SQES
11	x4	x4	008:00:0	0x1570	0x1570		NVMeLeCroy000000	0x00000000AD55A46B	0	0x0	0	0	0	0x0303	0	6	0	63	0	4	0	0

CQES	NN	ONCS	FUSES	FNA	VWC	AWUN	AWUPF	NVSCC	PSD0	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	PSD1	MP	ENLAT	EXLAT	RRT	RRL
0	0x44660000	0x1	0x0	0	2	0x0	0x0	0	0x0000	0x00000000	0x00000000	0x00	0x00	0x00	0x00	0x00	0x0000	0x00000000	0x00000000	0x00	0x00	

RWT	RWL	PSD2	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	PSD3	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	PSD4	MP	ENLAT	EXLAT	RRT	RRL
0x00	0x00	0x0000	0x00000000	0x00000000	0x00	0x00	0x00	0x00	0x00	0x0000	0x00000000	0x00000000	0x00	0x00	0x00	0x00	0x00	0x0000	0x00000000	0x00000000	0x00	0x00	

RWT	RWL	VS	Data	Time Delta	Time Stamp
0x00	0x00		1023 bytes	663.720 µs	0199 . 379 122 220 s

NVM	R→	2.5	RequesterID	Command Completion	SQHD	SQID	CID	P	ST	SC	SCT	M	DNR	Time Delta	Time Stamp
12	x4	x4	008:00:0	0x00000000	0x0001	0x0000	0x0000	1		0x0000	0x00	0	0	18.232 µs	0199 . 379 785 940 s

Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	ExplicitACK	Metrics	# Packets	Time Delta
1965	x4	x4	995		010:00000	1	008:00:0	0	FEE3F00C	1111	0000	1 dword	0	Packet #7599		2	10.420 µs

Time Stamp
0199 . 379 804 172 s

NVM	R→	2.5	RequesterID	Admin CQH QID = 0	Time Delta	Time Stamp
13	x4	x4	000:00:0	0x0001	13.547 ms	0199 . 379 814 592 s



Setting up Controller Registers

Submission queue size

Submission queue BAR

Completion queue BAR

Enable doorbell execution

Submission queue tail doorbell

NVM	R→	2.5	RequesterID	ASQS	ACQS	Time Stamp							
4	x4		000:00:0	127	127	0199.362722052 s							
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	
1920	x4		961		010:00000	1	000:00:0	0	FE400024	1111	0000	1 dword	
NVM	R→	2.5	RequesterID	ASQB AddressHi	ASQB AddressLow	Time Stamp							
5	x4		000:00:0	0x00000002	0x3D61B000	0199.362723516 s							
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	
1921	x4		962		010:00000	1	000:00:0	0	FE400028	1111	0000	00B0613D	
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	
1922	x4		963		010:00000	1	000:00:0	0	FE40002C	1111	0000	02000000	
NVM	R→	2.5	RequesterID	ACQB AddressHi	ACQB AddressLow	Time Stamp							
6	x4		000:00:0	0x00000002	0x3D61D000	0199.362726388 s							
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	
1923	x4		964		010:00000	1	000:00:0	0	FE400030	1111	0000	1 dword	
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	
1924	x4		965		010:00000	1	000:00:0	0	FE400034	1111	0000	1 dword	
NVM	R→	2.5	RequesterID	EN	CSS	MPS	AMS	SHN	IOSQES	IOCQES	Time Stamp		
7	x4		000:00:0	1	NVM command set	0	b00	b00	0	0	0199.362729292 s		
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	
1925	x4		966		010:00000	1	000:00:0	0	FE400014	1111	0000	1 dword	
NVM	R→	2.5	RequesterID	CompleterID	RDY	CFS	SHST	Time Stamp					
8	x4		000:00:0	008:00:0	1	0	b00	0199.362763492 s					
Split Tra	R→	2.5	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	
959	x4			000:00000	000:00:0	008:00:0	1	0	0	FE40001C	SC	1 dwe	
NVM	R→	2.5	RequesterID	Admin SQT QID = 0	Time Stamp								
9	x4		000:00:0	0x0001	0199.377751104 s								
Link Tra	R→	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	
1928	x4		968		010:00000	1	000:00:0	0	FE401000	1111	0000	1 dword	

Identify Command Execution

NVM	2.5	RequesterID	CompleterID	Admin Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	CNS	Time Stamp
10	R- x4	008:00:0	000:00:0	Identify	b00	0x0000	0x00000000	0x00000000	0x00000000	0x00000000	0x00000002	0x3D2DA160	0x00000002	0x3D2DB000	Controller	0199.377943380 s

Split Tra	2.5	Mem	MRd(64)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status
960	R- x4	001:00000	008:00:0	000:00:0	0	0	0	0	00000002:3D61B000	SC

Data	Metrics	# LinkTras	Time Delta	Time Stamp
0: 06000000 00000000 00000000 00000000 00000000 00000000 60A12D3D 02000000		2	1.179 ms	0199.377943380 s
8: 00B02D3D 02000000 01000000 00000000 00000000 00000000 00000000 00000000				

NVM	2.5	RequesterID	VID	SSVID	SN	MN	FR	RAB	IEEE	MIC	MDTS	OACS	ACL	AERL	FRMW	LPA	ELPE	NPSS	AVSCC	SQES	CQES	NN	ONCS
11	R- x4	008:00:0	Identify CDS	0x1570	0x1570	NVMeLeCroy000000	0x000000008D0DB8E1	0	0x0	0	0	0x0303	0	6	0	63	0	4	0	0	0	0x44660000	0x1

FUSES	FNA	VWC	AWUN	AWUPF	NVSCC	PSD0	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	PSD1	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	PSD2	MP	ENLAT	EXLAT	RRT
0x0	0	2	0x0	0x0	0	0x0000	0x00000000	0x00000000	0x00	0x00	0x00	0x00	0x00	0x0000	0x00000000	0x00000000	0x00	0x00	0x00	0x00	0x00	0x0000	0x00000000	0x00000000	0x00	

RRL	RWT	RWL	PSD3	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	PSD4	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	VS	Data	Time Stamp
0x00	0x00	0x00	0x0000	0x00000000	0x00000000	0x00	0x00	0x00	0x00	0x00	0x0000	0x00000000	0x00000000	0x00	0x00	0x00	0x00	0x00	1023 bytes	0199.379122220 s	

Link Tra	2.5	TLP	Mem	MW(64)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	Explicit ACK	Metrics	# Packets	Time Delta	Time Stamp
1931	R- x4	961	011:00000	8	008:00:0	0	00000002:3D2DA160	1111	1111	8 dwords	0	Packet#7531	2	36.208 μs	0199.379122220 s		
1932	R- x4	962	011:00000	32	008:00:0	0	00000002:3D2DA180	1111	1111	32 dwords	0	Packet#7533	2	19.888 μs	0199.379158428 s		
1933	R- x4	963	011:00000	32	008:00:0	0	00000002:3D2DA200	1111	1111	32 dwords	0	Packet#7535	2	23.128 μs	0199.379178316 s		
1934	R- x4	964	011:00000	32	008:00:0	0	00000002:3D2DA280	1111	1111	32 dwords	0	Packet#7537	2	18.512 μs	0199.379201444 s		
1935	R- x4	965	011:00000	32	008:00:0	0	00000002:3D2DA300	1111	1111	32 dwords	0	Packet#7539	2	18.664 μs	0199.379219956 s		
1936	R- x4	966	011:00000	32	008:00:0	0	00000002:3D2DA380	1111	1111	32 dwords	0	Packet#7541	2	18.904 μs	0199.379238620 s		

Identify Command Execution

System Memory command

Address	Data
0x3D61B000	0x3D2DA160

NVMe Controller registers

Register	Address	Data
ASQB	0x28	3D61B000

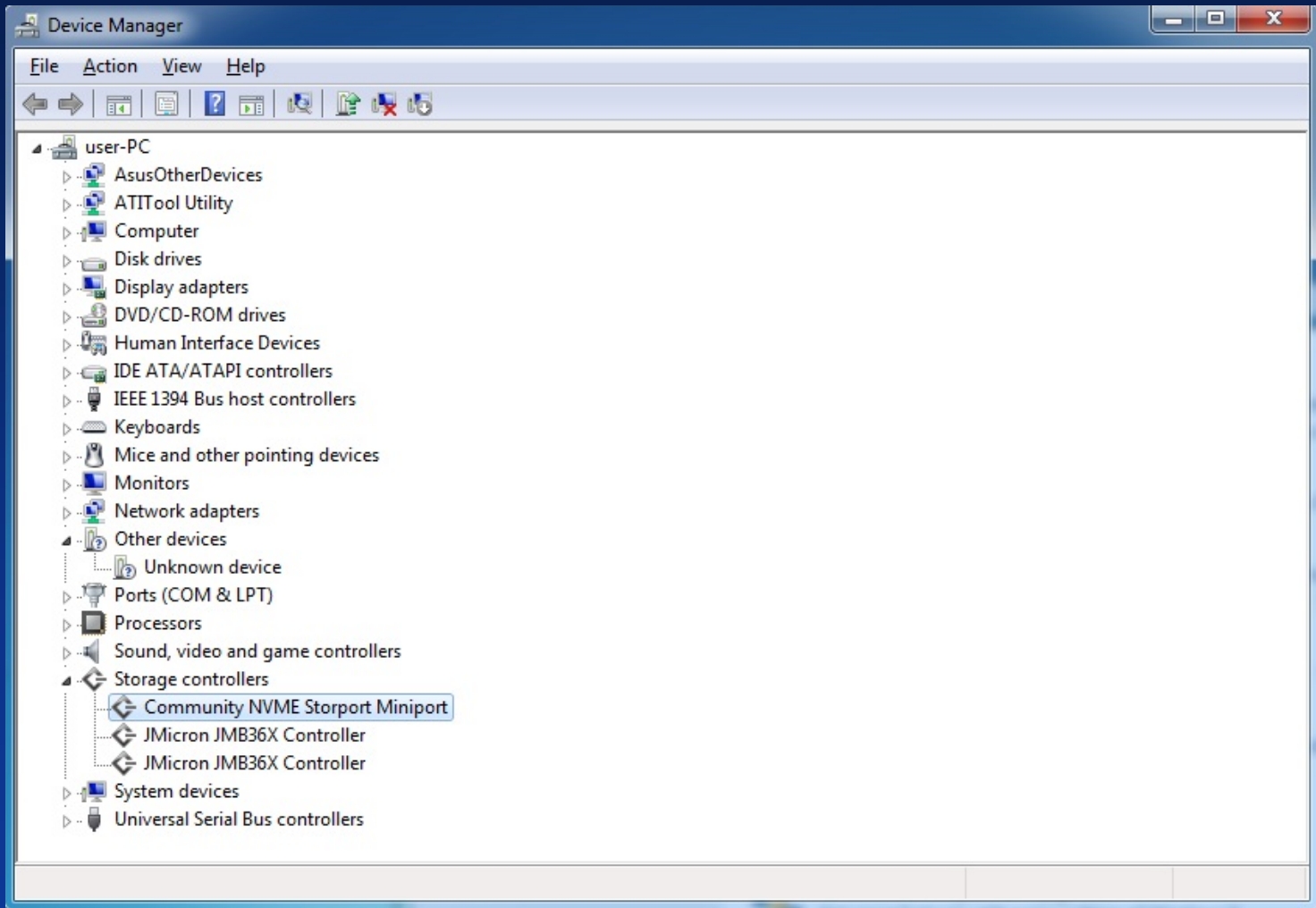
System Memory Data

Address	Data
0x3D2DA160	0x70157015

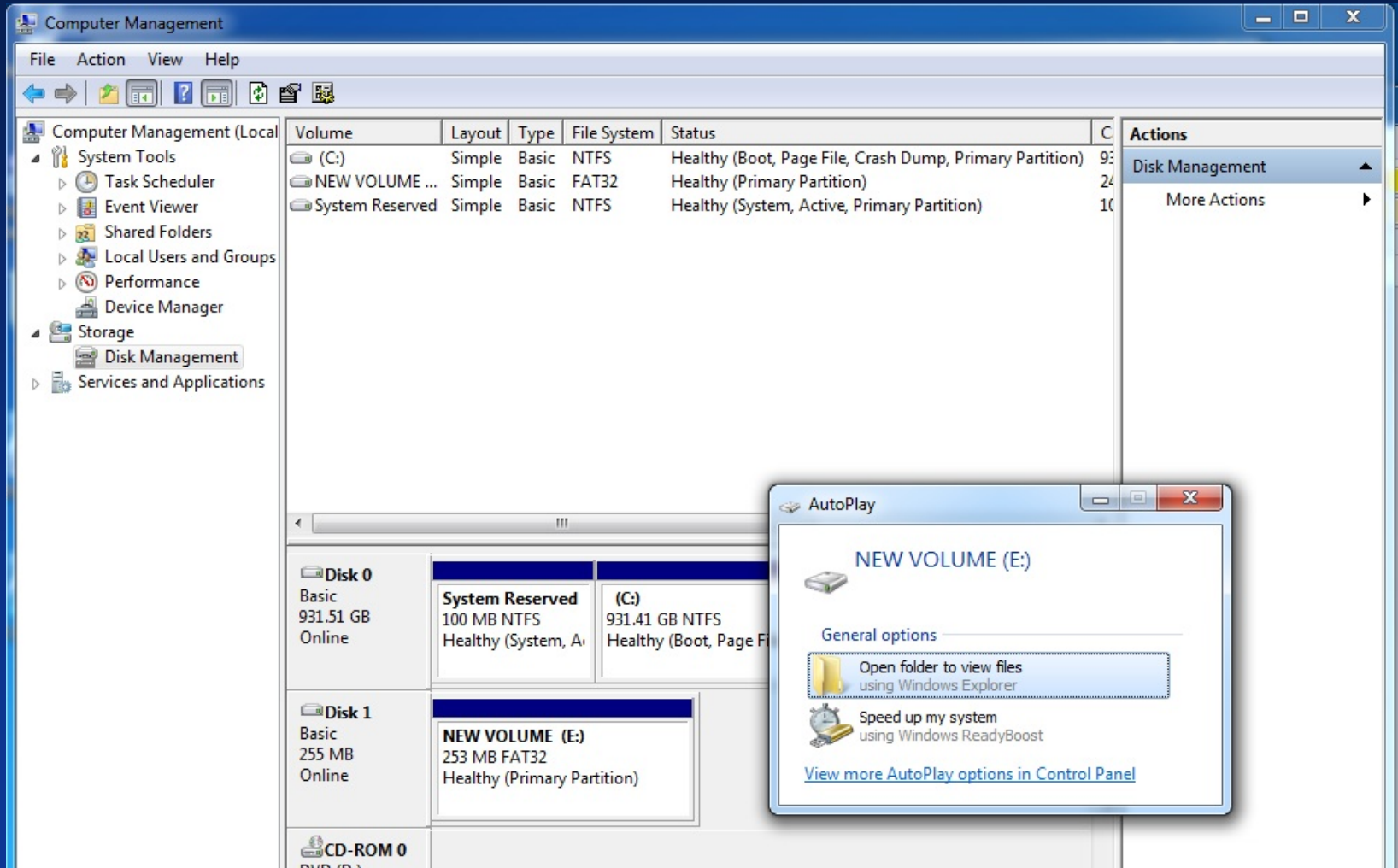
NVMe Device memory space

Address	Data
0x3D61B000	3D2DA160

Emulated NVMe Device Shown in Device Manager



Emulated Drive Shown in Disk Management



The screenshot shows the Windows Computer Management console. The left pane shows the navigation tree with 'Storage' > 'Disk Management' selected. The main pane displays a table of volumes:

Volume	Layout	Type	File System	Status	C
(C:)	Simple	Basic	NTFS	Healthy (Boot, Page File, Crash Dump, Primary Partition)	93
NEW VOLUME ...	Simple	Basic	FAT32	Healthy (Primary Partition)	24
System Reserved	Simple	Basic	NTFS	Healthy (System, Active, Primary Partition)	10

Below the table, the 'Disk 1' properties are shown:

- Disk 1**
- Basic
- 255 MB
- Online
- NEW VOLUME (E:)**
- 253 MB FAT32
- Healthy (Primary Partition)

An AutoPlay dialog box is open for the 'NEW VOLUME (E:)', showing options to 'Open folder to view files using Windows Explorer' and 'Speed up my system using Windows ReadyBoost'.

Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation

Testing NVM Express

nvm
EXPRESS



Summit T3-16
Analyzer

Summit T3-8
Analyzer

Summit T28
Analyzer

Summit Z3-16
Exerciser with
Test Platform

- NVM Commands
 - ✓ Write
 - ✓ Read
 - ✓ Compare
 - ✓ Extensible for Vendor Specific Commands
- Queue Management
- Come up in Device Manager
- Extensible Vendor Specific Features (for Get/Set Features)
- Complete commands via fused Commands (i.e. Compare & Write)



Initialization example

; This script performs basic initialization of an NVMe device

; Set up BARs;

**; Enable bus master, memory space,
set interrupt disable;**

; Set Max payload and read request in Device Control;

; Write MSI-X table and enable MSI-X.

```
include="nvme_definitions.peg"
```

```
packet="Temp_ConfigWrite0"
```

```
{
```

```
    Register = 0x10
```

```
    Payload = ( BAR0_ADDRESS_FLIPPED )
```

```
}
```

```
wait=TLP { TLPType = Cpl }
```

```
packet="Temp_ConfigWrite0"
```

```
{
```

```
    Register = 0x14
```

```
    Payload = ( 0 )
```

```
}
```

```
wait=TLP { TLPType = Cpl }
```

; This script performs NVM specific initialization by writing Controller registers on the device

```
include="nvme_definitions.peg"
```

; Set ACQS and ASQS in AQA – Admin Queue Attributes register

```
packet="Temp_OneDwordWrite"
```

```
{
```

```
    Address = ( CONTROLLER_REGISTERS_BASE + 0x24 )
```

```
        Payload = ( 7F007F00 )
```

```
}
```

; Set Admin submission Queue address base ASQB high and low . This address corresponds to the base address set for Mem_64 Host region in the generation options file "host_go.gen"

; ASQ – Admin Submission Queue Base Address low

```
packet="Temp_OneDwordWrite"
```

```
{
```

```
    Address =
```

```
    ( CONTROLLER_REGISTERS_BASE + 0x28 )
```

```
        Payload = ( 0080AA2F ) }
```



Exerciser Features for Storage over PCIe validation

- Read completion payload storage for later processing to implement command queuing
- Branch upon write payload and procedure activation to implement doorbell registers
- DMA descriptor implementation and the use of descriptor data through field substitution
- Creation of data structures in emulator memory
- Trace export to generation file with different timing options
- Extraction of configuration file from trace and import to device emulator

Agenda

- Storage over PCI Express[®] architecture
 - SATA Express / AHCI
 - SCSI Express / SOP – PQI
 - NVM Express
- Command queue generation example
- Emulating an SSD controller
- Emulating an SSD host
- Command Validation



Command Validation- NVMe - Generation script

```
include="nvme_start.peg"
```

```
; Pre-program command in the Host Memory Region. This is the Mem_64 Host region
```

```
; defined in the generation options file "nvme_host_gen_options.gen"
```

```
; Command data is copied from the trace taken a the last call
```

```
AddressSpace=Write
```

```
{
```

```
Location=Mem64
```

```
Offset = 0
```

```
Size = 128
```

```
LoadFrom =
```

```
(0x06 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00  
0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x90 0xB1 0x25 0x3F 0x08 0x00 0x00 0x00  
0x00 0xC0 0x25 0x3F 0x08 0x00 0x00 0x00 0x01 0x00 0x00 0x00 0x00 0x00 0x00 0x00  
0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00
```

```
0x06 0x00 0x01 0x00 0x01 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00  
0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0xE0 0x47 0x3E 0x02 0x00 0x00 0x00  
0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00  
0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 )
```

```
} # Identify Controller and Identify Namespace
```




Command Validation- NVMe - Generation script

; Write Admin Submission queueTail Doorbell Register

```
packet="Temp_OneDwordWrite"
```

```
{
```

```
    Address = ( CONTROLLER_REGISTERS_BASE + 0x1000 )
```

```
    Payload = ( 01000000 )
```

```
}
```

**; Wait for the Controller to process the command. This will include writing the Identify data,
; writing the Admin completion Queue, and sending the MSI-X interrupt at vector**

```
wait=TLP
```

```
{
```

```
    TLPType = MWr32
```

```
    Address = ADMIN_INT_VECTOR_ADDRESS
```

```
}
```

; Write Admin Completion Queue Head Doorbell Register

```
packet="Temp_OneDwordWrite"
```

```
{
```

```
    Address = ( CONTROLLER_REGISTERS_BASE + 0x1004 )
```

```
    Payload = ( 01000000 )
```

```
}
```



Command Validation- NVMe - Verification script

Constant definitions

```
# Test stage definitions, should be sequential
const STAGE_NVME_CONFIG      = 0;
const STAGE_ADMIN_DORBELL_1  = 1;
const STAGE_READ_CMD_1      = 2;
const STAGE_TRANSFER_DATA_1  = 3;
const STAGE_WRITE_CPL_1     = 4;
const STAGE_SEND_INTERRUPT_1 = 5;
const STAGE_ADMIN_DORBELL_2  = 6;
const STAGE_READ_CMD_2      = 7;
const STAGE_TRANSFER_DATA_2  = 8;
const STAGE_WRITE_CPL_2     = 9;
const STAGE_SEND_INTERRUPT_2 = 10;
```

Variable declarations

```
set Admin_SQB_Low = 0;
set Admin_SQB_High = 0;
set Admin_CQB_Low = 0;
set Admin_CQB_High = 0;
set PRP1_High = 0;
set PRP1_Low = 0;
set PRP2_High = 0;
set PRP2_Low = 0;
set Cmd_Dw10 = 0;
set CurrentIdentifyXferredLength = 0;
set TestStage = 0;
set CurrentChannel = 0;
```



Command Validation- NVMe - Verification script

Function: OnStartScript()

Description: The application calls this function at the beginning of the script execution.

OnStartScript()

```
{ ReportText( "Verifying Identify Command..." );  
  SendAllChannels();  
  SendLevelOnly( _LINK );  
  SendTraceEvent( _LINK_CONFIG );  
  SendTraceEvent( _LINK_COMPLETION );  
  SendTraceEvent( _LINK_MEMORY );  
  Admin_SQB_Low = 0; # initialize variables  
  Admin_SQB_High = 0;  
  Admin_CQB_Low = 0;  
  Admin_CQB_High = 0;  
  PRP1_High = 0;  
  PRP1_Low = 0;  
  PRP2_High = 0;  
  PRP2_Low = 0;  
  Cmd_Dw10 = 0;  
  CurrentIdentifyXferredLength = 0;  
  TestStage = STAGE_NVME_CONFIG;}
```



Command Validation- NVMe- Verification script

```
# Function: ProcessEvent()
# Description: Entry point of the script.
# The application calls this function every time it finds the relevant trace event.
ProcessEvent()
{
    CurrentChannel = in.Channel;
    event_type = in.TraceEvent;
    # transaction status checking
    if( in.TransactionStatus == LINK_TRA_STATUS_INCOMPLETE )
    {
        FailTest( "Transaction wasn't complete at the Link Layer" );
        return null;
    }
    select
    {
        event_type == _LINK_CONFIG      : ProcessCfgRequest();
        event_type == _LINK_COMPLETION  : ProcessCompletion();
        event_type == _LINK_MEMORY     : ProcessMemReadOrWrite();
    };
    return Complete();
}
```



NVMe Command Validation Resulting Trace

Split Tra	R	2.5	Cfg	CfgRd0	RequesterID	CompleterID	Tag	TC	VC ID	DeviceID	Register	Status	Class Code	Revision ID	Metrics	# LinkTras	Time Delta	Time Stamp						
9	x1	x1	000:00100	064:02:0	129:00:0	0	0	0	0	129:00:0	0x008	SC	0x010802	0x01	2	165.760 us	0002 . 746 951 600 s							
Split Tra	R	2.5	Cfg	CfgRd0	RequesterID	CompleterID	Tag	TC	VC ID	DeviceID	Register	Status	Base Address Register 0	Metrics	# LinkTras	Time Delta	Time Stamp							
10	x1	x1	000:00100	064:02:0	129:00:0	0	0	0	0	129:00:0	0x010	SC	0xEC030004	2	159.864 us	0002 . 747 117 360 s								
NVM	R	2.5	RequesterID	CC	EN	CSS	MPS	AMS	SHN	IOSQES	IOCQES	Time Delta	Time Stamp											
0	x1	x1	064:02:0	0	0	NVM command set	0	b00	b00	0	0	26.888 us	0002 . 747 277 224 s											
NVM	R	2.5	RequesterID	AQA	ASQS	ACQS	Time Delta	Time Stamp																
1	x1	x1	064:02:0	127	127	18.936 us	0002 . 747 304 112 s																	
NVM	R	2.5	RequesterID	ASQ	ASQB AddressHi	ASQB AddressLow	Time Delta	Time Stamp																
2	x1	x1	064:02:0	0x00000004	0x2FAA8000	40.600 us	0002 . 747 323 048 s																	
NVM	R	2.5	RequesterID	ACQ	ACQB AddressHi	ACQB AddressLow	Time Delta	Time Stamp																
3	x1	x1	064:02:0	0x00000004	0x2FAAA000	37.608 us	0002 . 747 363 648 s																	
NVM	R	2.5	RequesterID	CC	EN	CSS	MPS	AMS	SHN	IOSQES	IOCQES	Time Delta	Time Stamp											
4	x1	x1	064:02:0	1	0	NVM command set	0	b00	b00	0	0	775.660 ms	0002 . 747 401 256 s											
NVM	R	2.5	RequesterID	SQyTDBL	Admin SQT QID = 0	Time Delta	Time Stamp																	
5	x1	x1	064:02:0	0x0001	704.000 ns	0003 . 523 061 552 s																		
NVM	R	2.5	RequesterID	CompleterID	Admin Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	CNS	Time Delta	Time Stamp						
6	x1	x1	129:00:0	000:00:0	Identify	b00	0x0000	0x00000000	0x00000000	0x00000000	0x00000000	0x00000008	0x3F25B190	0x00000008	0x3F25C000	Controller	275.656 us	0003 . 523 062 256 s						
NVM	R	2.5	RequesterID	Identify CDS	VID	SSVID	SN	MN	FR	RAB	IEEE	MIC	MDTS	OACS	ACL	AERL	FRMW	LPA	ELPE	NPSS	AVSQC	SQES	CQES	NN
7	x1	x1	129:00:0	0x111D	0x1D11	IDT_SN_0000	IDT V.Board NVMe SSD	1.5.0	1	0x313233	1	0	0x0006	3	5	5	1	63	0	1	102	68	0x00000001	
			ONCS	FUSES	FNA	VWC	AWUN	AWUPF	NVSCC	PSD0	MP	ENLAT	EXLAT	RRT	RRL	RWT	RWL	vs	Data	Time Delta	Time Stamp			
			0x0006	0x0000	0	1	0xFFFF	0xFFFF	1	0x0BB8	0x00000064	0x00000064	0x00	0x00	0x00	0x00	1024 bytes	2.007 ms	0003 . 523 337 912 s					
NVM	R	2.5	RequesterID	Command Completion	SQHD	SQID	CID	P	ST	SC	SCT	M	DNR	Time Delta	Time Stamp									
8	x1	x1	129:00:0	0x00000000	0x0001	0x0000	0x0000	1	0x0000	0x00	0	0	76.352 us	0003 . 525 344 736 s										
Link Tra	R	2.5	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	ExplicitACK	Metrics	# Packets	Time Delta	Time Stamp						
66	x1	x1	1070	010:00000	1	129:00:0	0	FFABF00C	1111	0000	1 dword	0	Packet#133	2	925.880 us	0003 . 525 421 088 s								
NVM	R	2.5	RequesterID	CQyHDBL	Admin CQH QID = 0	Time Delta	Time Stamp																	
9	x1	x1	064:02:0	0x0001	161.856 us	0003 . 526 346 968 s																		
NVM	R	2.5	RequesterID	SQyTDBL	Admin SQT QID = 0	Time Delta	Time Stamp																	
10	x1	x1	064:02:0	0x0002	728.000 ns	0003 . 526 508 824 s																		
NVM	R	2.5	RequesterID	CompleterID	Admin Cmd	OPC	FUSE	CID	NSID	MPTR Hi	MPTR Low	PRP1 Hi	PRP1 Low	PRP2 Hi	PRP2 Low	CNS	Time Delta	Time Stamp						
11	x1	x1	129:00:0	000:00:0	Identify	b00	0x0001	0x00000001	0x00000000	0x00000000	0x00000000	0x00000002	0x3E47E000	0x00000000	0x00000000	Namespace	211.320 us	0003 . 526 509 552 s						

The logo for Flash Memory Summit features a stylized yellow sunburst with multiple rays. The text 'Flash Memory' is written in a blue, sans-serif font, with 'Flash' in a larger size than 'Memory'. Below this, the word 'SUMMIT' is written in white, uppercase letters on a blue rectangular background.

Flash Memory Summit Summary

- Full SCSI and ATA command level storage over PCIe decodes essential to SSDs traffic analysis
- Hierarchical expandable display
- NVM Express, SCSI Express, and SATA Host and Device emulation are now implemented using traditional PCIe exercisers
- Official UNH IOL and NVMe consortium conformance test setup
- Support for SFF8639 and M.2 form factors through dedicated interposers