

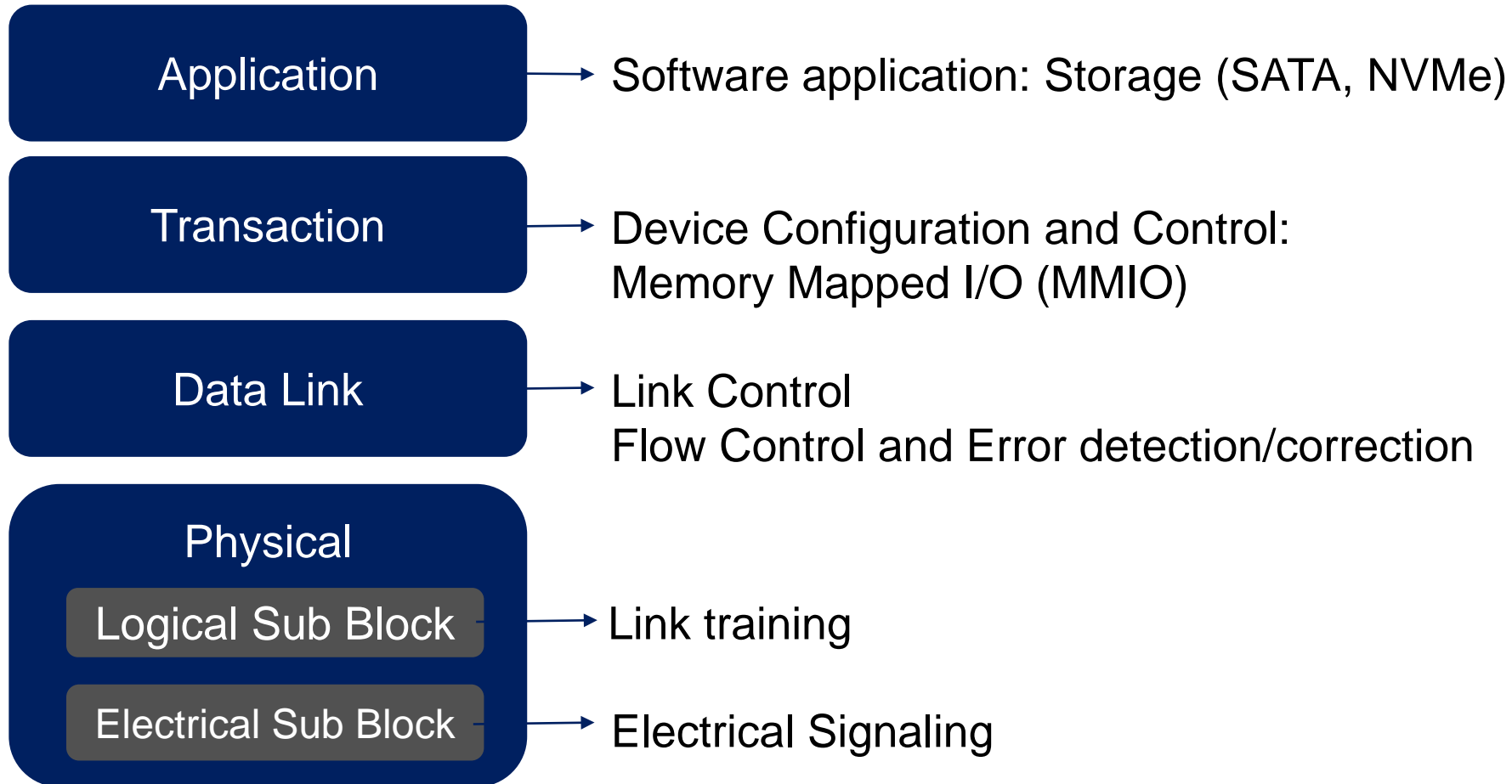
Session 101-A: SSD Testing Challenges, Part 1 (Testing Track)

PCIe Protocol Analysis for SSDs

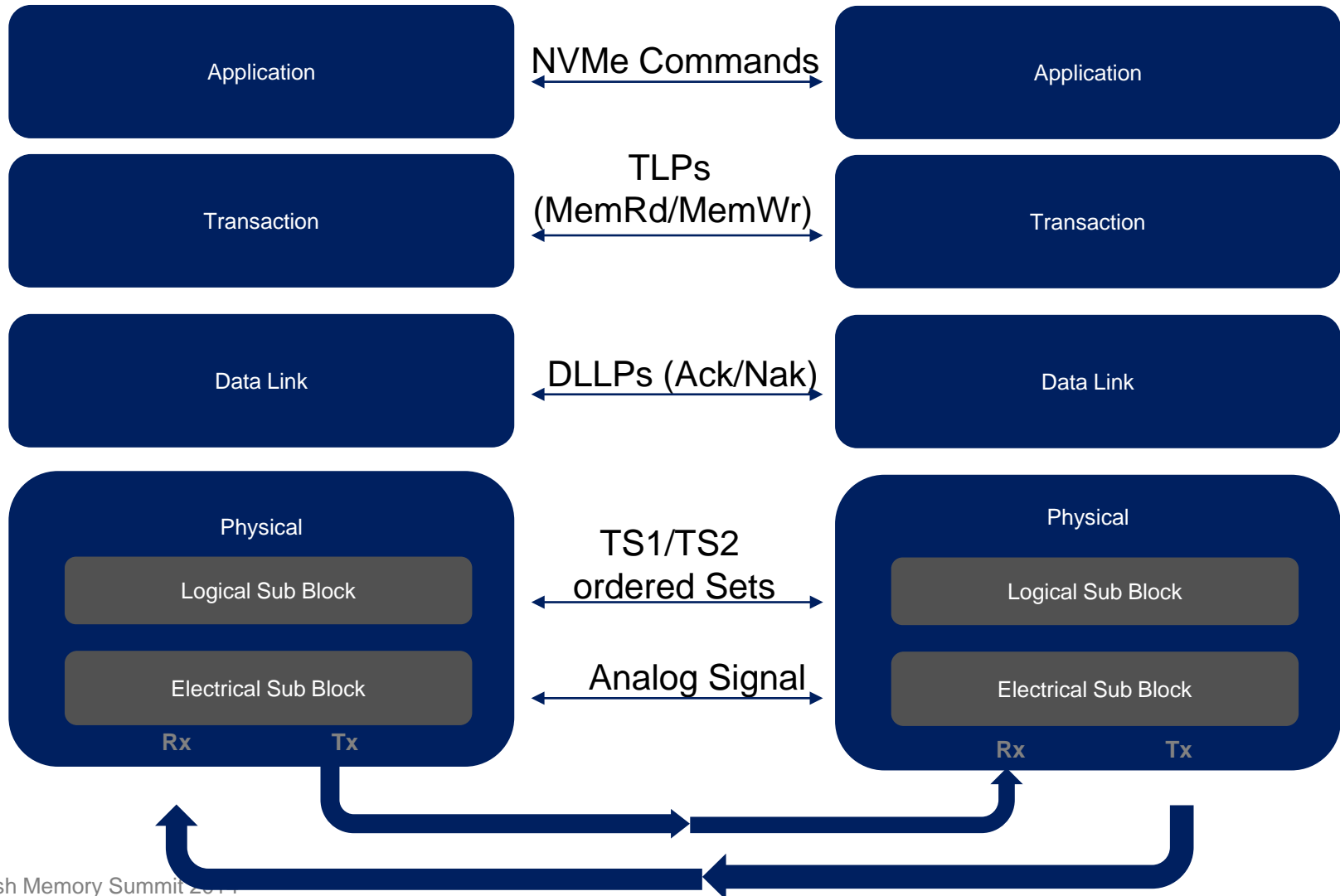
Don Schoenecker,
PCIe Product Manager

Keysight Technologies (Formerly Agilent Technologies)

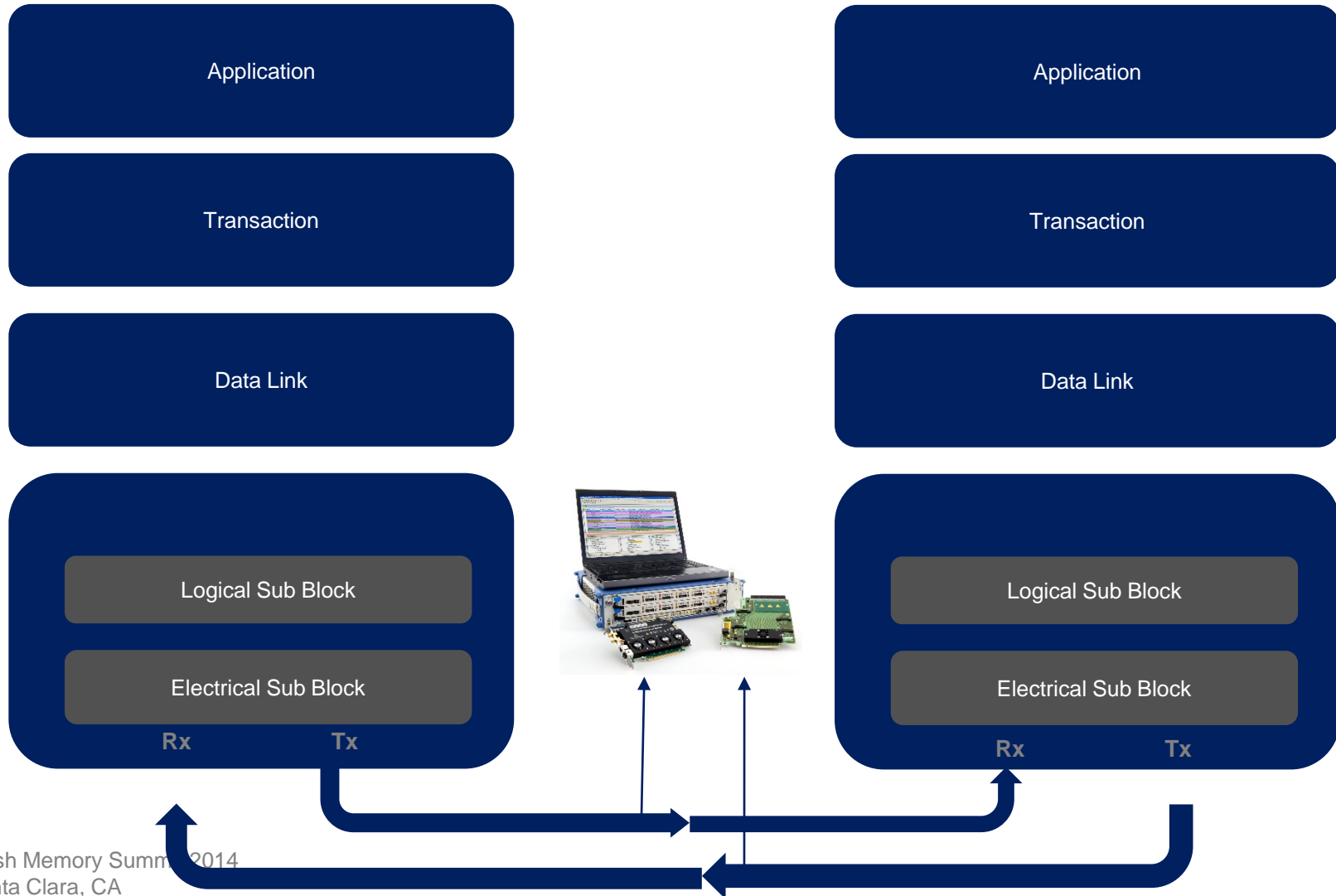
PCI Express : Layered Protocol



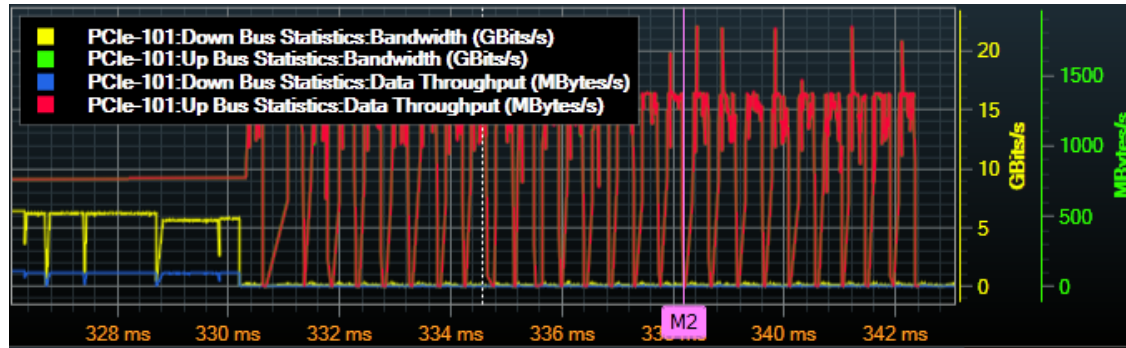
Communication of commands



Interposer Connection

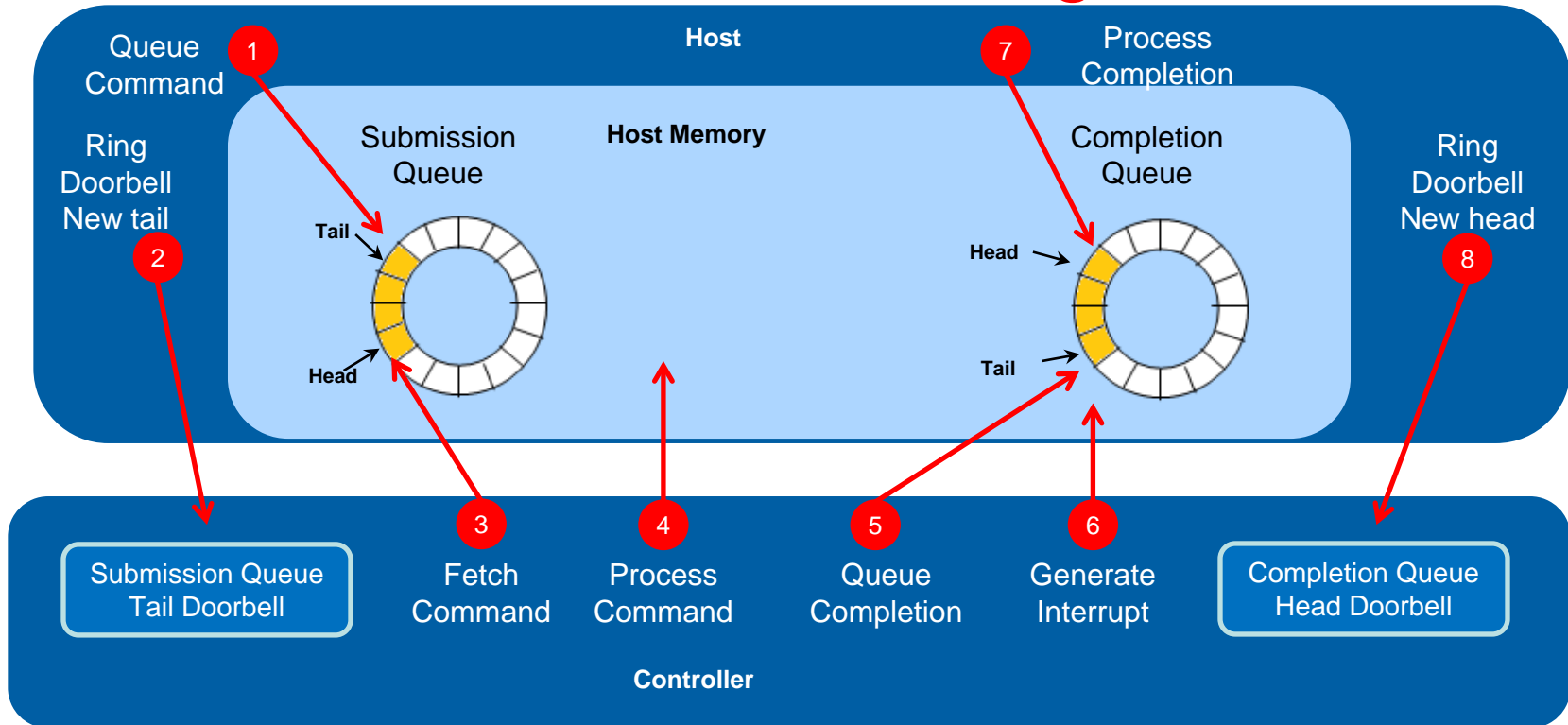


Why PCIe Transport?



- **PCIe is high performance**
 - Lowest latency (no HBA overhead)
 - Full duplex, multiple outstanding requests
 - Scalable port width (x1 to x16)
 - Scalable link speed (2.5 /5.0 /8.0 Gb/s/per lane)
- **PCIe is low cost**
 - High volume/commodity.
 - Eliminates Host Bus Adaptor (HBA) cost
- **PCIe power management capabilities**
 - Direct attach to CPU eliminates HBA power
 - Various low power levels (L0s, L1, L2) --- New* L1 substates
- **Maturity**
 - PCIe has been shipping for years and is a mature technology.

NVMe Queue Interface for Command Processing



Command Submission

1. Host writes command to submission queue
2. Host writes updated submission queue tail pointer to doorbell

Command Processing

3. Controller fetches command
4. Controller processes command

Command Completion

5. Controller writes completion to completion queue
6. Controller generates MSI-X interrupt
7. Host processes completion
8. Host writes updated completion queue head pointer to doorbell

Example of a NVMe Read operation

241544: NVMe Read

Address	Type	# PCI-Express TLP:	Size
NVMe Read	NVMe	1	512
0x000000007F101008	SQDB	1	
0x0000000078F7D000	SQ	1	
0x0000000078EDA2A0	PRP 1	4	512
0x0000000078F8D000	CQ	1	
0x000000007F10100C	CQDB	1	

Complete Payload: DWord, 4 Columns, Little Endian

```

00000000: D08EC033 8E7C00BC BED88EC0 00BF7C00
00000010: 0200B906 50A4F3FC CB061C68 0004B9FB
00000020: 8007BEBD 7C00007E 0E850F0B 10C58301
00000030: 18CDF1E2 55005688 051146C6 001046C6
00000040: AAB841B4 5D13CD55 FB810F72 0975AA55
00000050: 0001C1F7 46FE0374 80606610 7400107E
00000060: 00686626 66000000 680876FF 00680000
00000070: 0001687C B4001068 00568A42 13CDF48B
00000080: 10C4839F B814EB9E 00BB0201 00568A7C
00000090: 8A01768A 6E8A024E 6613CD03 FE1C7361
000000A0: 0C75114E 80007E80 008A840F 84EB80B2
000000B0: 8AE43255 13CD0056 819EEB5D 557DFE3E
    
```

PCI-Express Packets: Nested (Time Ordered) | Show Completions | Acks

ID	Packet	Timestamp	Delta Time
SQDB	Memory Write 32b (Gen1,2)	0 s	0 s
SQ	Memory Read 32b (Gen1,2)	256 ns	256 ns
SQ	Completion with Data (Gen1,2)	405 ns	149 ns
PRP 1.1	Memory Write 32b (Gen1,2)	64.025 us	63.620 us
PRP 1.2	Memory Write 32b (Gen1,2)	64.061 us	36 ns
PRP 1.3	Memory Write 32b (Gen1,2)	64.099 us	38 ns
PRP 1.4	Memory Write 32b (Gen1,2)	64.135 us	36 ns
CQ	Memory Write 32b (Gen1,2)	65.455 us	1.320 us
CQDB	Memory Write 32b (Gen1,2)	68.763 us	3.308 us

Example of a NVMe Read operation

241544: NVMe Read

Address	Type	# PCI-Express TLPs	Size
NVMe Read			
0x000000007F101008	SQDB		
0x0000000078F7D000	SQ		
0x0000000078EDA2A0	PRP 1		512
0x0000000078F8D000	CQ	1	
0x000000007F10100C	CQDB	1	

Complete Payload: DWord, 4 Columns, Little Endian

```

00000000: D08EC033 8E7C00BC BED88EC0 00BF7C00
00000010: 0200B906 50A4F3FC CB061C68 0004B9FB
00000020: 8007BEBD 7C00007E 0E850F0B 10C58301
00000030: 18CDF1E2 55005688 051146C6 001046C6
00000040: AAB841B4 5D13CD55 FB810F72 0975AA55
00000050: 0001C1F7 46FE0374 80606610 740010
00000060: 00686626 66000000 680876FF 006
00000070: 0001687C B4001068 00568A42
00000080: 10C4839F B814EB9E 00BB0201
00000090: 8A01768A 6E8A024E 6613CD00
000000A0: 0C75114E 80007E80 008A840F 84EB80B2
000000B0: 8AE43255 13CD0056 819EEB5D 557DFE3E
    
```

PCI-Express Packets: Nested (Time Ordered), Show Completions, Acks

ID	Packet	Timestamp	Delta Time
SQDB	Memory Write 32b (Gen1,2)	0 s	0 s
SQ	Memory Read 32b (Gen1,2)	256 ns	256 ns
SQ	Completion with Data (Gen1,2)	405 ns	
PRP 1.1	Memory Write 32b (Gen1,2)		
PRP 1.2	Memory Write 32b (Gen1,2)		
PRP 1.3	Memory Write 32b (Gen1,2)		38 ns
PRP 1.4	Memory Write 32b (Gen1,2)	64.135 us	36 ns
CQ	Memory Write 32b (Gen1,2)	65.455 us	1.320 us
CQDB	Memory Write 32b (Gen1,2)	68.763 us	3.308 us

NVMe commands

PCIe TLPs

Application Data



Key PCIe tests

- **Equalization testing:** Signal quality is critical for success operation at Gen3 8GT/s
- **LTSSM analysis:** State transitions to handle link up, recovery, and power management
- **Packet capture and decode:** PCIe packets, responses, configuration and device enumeration
- **Performance Analysis:** Response times and process delays can greatly impact data throughput
- **Flow Control analysis:** Included in the performance analysis, clearly identify and track credit starvation issues.

Sampl	Time	Downstr	Ups	Lin	Seq	T	Le	Address, Register	H	Payload
245039	33.748789316	s	CfgRd_0	Gen2	D36	00	001	Register Number=32		
238808	33.748791840	s		CplD	Gen2	4DE	00	001		018AE005
245040	33.748792416	s	CfgWr_0	Gen2	D37	00	001	Register Number=32		018A00CA
238809	33.748795151	s		Cpl	Gen2	4DF	00	000		
245041	33.748796444	s	MemWr_32	Gen2	D38	00	001	Address=7F10 2000		FEE00000
245042	33.748796518	s	MemWr_32	Gen2	D38	00	001	Address=7F10 2004		00000000

Protocol Analysis of NVMe

