



Flash Memory Summit

Using a PCIe Analyzer with NVMe-Based Products

Isaac Livny
Teledyne Corporation



Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - Analysis of NVMe Events

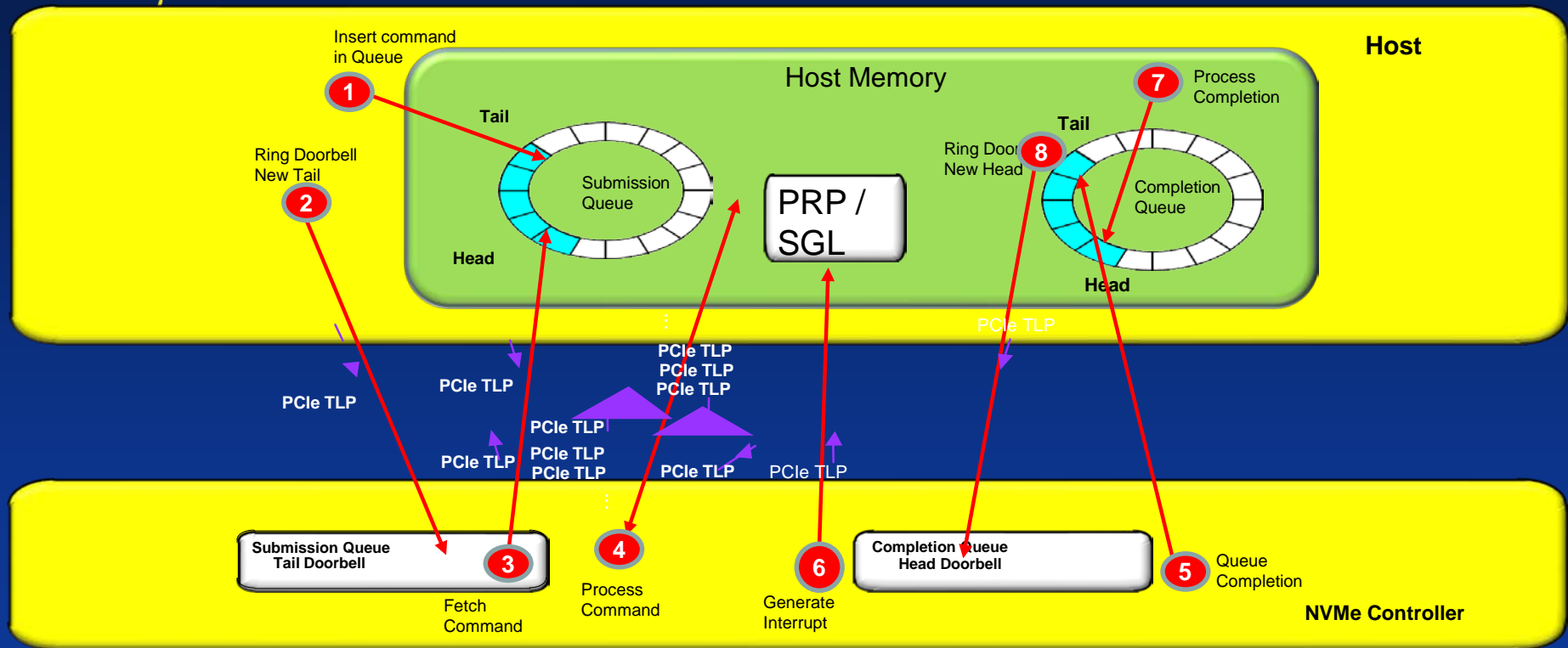


Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - NVMe Event capture and Analysis



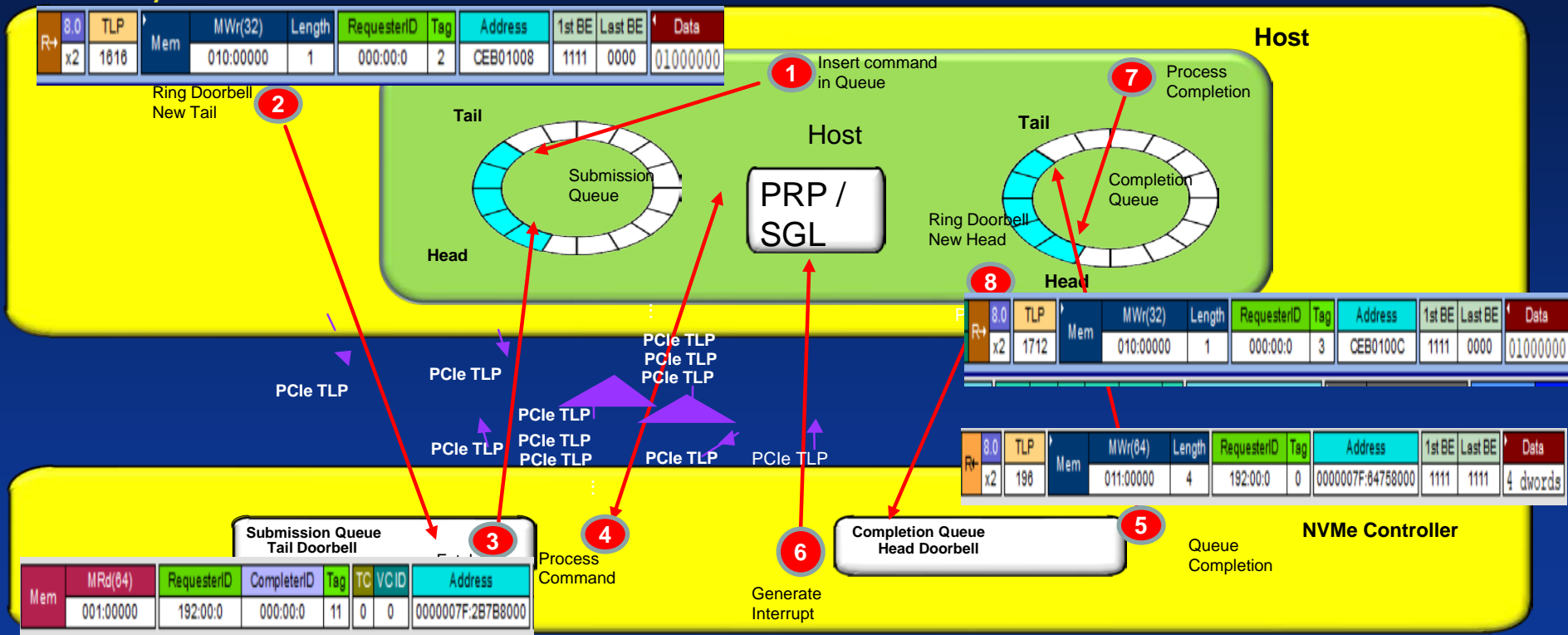
NVMe command execution





Normal mode Command execution

Flash Memory Summit





NVMe Command execution

Flash Memory Summit

NVMe 13	H	Device ID	ASQ	ASQB Address Low	ASQB Address High	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp							
		034:00:0		0x40BF4000	0x00000003	INTEL SSDPD2MD800G4		2	292.000 ns	0020 . 201 010 254 000 s							
NVMe 14	H	Device ID	ACQ	ACQB Address Low	ACQB Address High	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp							
		034:00:0		0x40963000	0x00000003	INTEL SSDPD2MD800G4		2	292.000 ns	0020 . 201 010 546 000 s							
NVMe 15	H	Device ID	OC	EN	CSS	MPS	AMS	SHN	IOSQES	IOQCES	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp		
		034:00:0		1	000	0	000	00	6	4	INTEL SSDPD2MD800G4		1	128.000 ns	0020 . 201 010 838 000 s		
NVMe 16	D	Device ID	CSTS	RDY	CFS	SHST	NSSRO	PP	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp				
		034:00:0		0	0	00	0	0	INTEL SSDPD2MD800G4		1	77.470 ms	0020 . 201 010 966 000 s				
Link Tra 425	R→	5.0	TLP	Msg	MsgD	Msg Routing	Length	RequesterID	Tag	Message Code	VID	Vendor Specific	Data	VCID	ExplicitACK	Metrics	# Pack
		x2	406		011:10011	Broadcast	1	000:00:0	0	Vendor_Defined_Type1	0x8086	0x00000080	1 dword	0	Packet #2990		2
NVMe 17	D	Device ID	CSTS	RDY	CFS	SHST	NSSRO	PP	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp				
		034:00:0		1	0	00	0	0	INTEL SSDPD2MD800G4		1	3.500 us	0020 . 305 489 954 000 s				
NVMe 18	H	Device ID	Table Update	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp								
		034:00:0		4 dwords	INTEL SSDPD2MD800G4		4	560.000 ns	0020 . 305 493 454 000 s								
Split Tra 208	R→	5.0	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VCID	Address	Status	Data	Metrics	# LinkTrans	Time Delta	Time Stamp	
		x2		000:000000	000:07:0	034:00:0	0	0	0	BFFFE000	SC	1 dword		2	3.508 us	0020 . 305 494 014 000 s	
NVMe 19	H	Device ID	Table Update	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp								
		034:00:0		3 dwords	INTEL SSDPD2MD800G4		3	189.608 us	0020 . 305 497 522 000 s								
NVMe 20	D	Device ID	VS	MNR	MJR	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp							
		034:00:0		0x02	0x0001	INTEL SSDPD2MD800G4		1	1.696 us	0020 . 305 687 130 000 s							
NVMe 21	D	Device ID	CAP	MCES	CCR	AMS	TC	DSTRD	NSSRS	CSS	MPSMIN	MPSMAX	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp
		034:00:0		2047	1	01	50	0	1	00000001	0	0	INTEL SSDPD2MD800G4		1	11.488 us	0020 . 305 688 826 000 s
NVMe Cmd 0	D	Device ID	OPC	1 warning(s)	SQID	CQID	CID	Data	PRP1	PRP2	CNS	CNTID	Controller C				
			Identify Controller	NVMe Cmd Warnings (2)	0x0000	0x0000	0x0000	1024 dwords	0x00000003:41F38000	0x00000000:00000000	Identify Controller	0x0000	Controller C				
NVMe 22	H	Device ID	QID	SQID	Admin SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp							
		034:00:0	0x0000	0x0001	0x0001	INTEL SSDPD2MD800G4		1	6.676 us	0020 . 305 700 314 000 s							
NVMe 23	H	Device ID	QID	CID	Address	ASQ	OPC	FUSE	PSDT	CID	NSID	MPTP	Address	PRP1	Address		
		034:00:0	0x0000	0x0000	00000003:40BF4000		Identify	Normal operation	PRP	0x0000	0x00000000		0x00000000:00000000	PRP1	0x00000003:41F38000		
NVMe 24	D	Device ID	QID	CID	Address	Identify Controller	Controller Capabilities	VID	SSVID	SN	MN						
		034:00:0	0x0000	0x0000	00000003:41F38000			0x8086	0x8086	CVEK625508Q1P0BGN	INTEL SSDPD2MD800G4						
NVMe 25	D	Device ID	QID	CID	Address	ACQ	SQID	QID	CID	P	DW0	Command Specific	ST	SCT	SC	M	DNR
		034:00:0	0x0000	0x0000	00000003:40963000		0x0001	0x0000	0x0000	1		0x00000000		Generic Command Status	Successful Completion	0	0
NVMe 26	D	Device ID	QID	CID	Address	Interrupt	Type	Vector	Message	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
		034:00:0	0x0000	0x0000	00000000:FEE20000		MSI-X	0	0x00004072	INTEL SSDPD2MD800G4		1	39.252 us	0020 . 305 819 900 s			

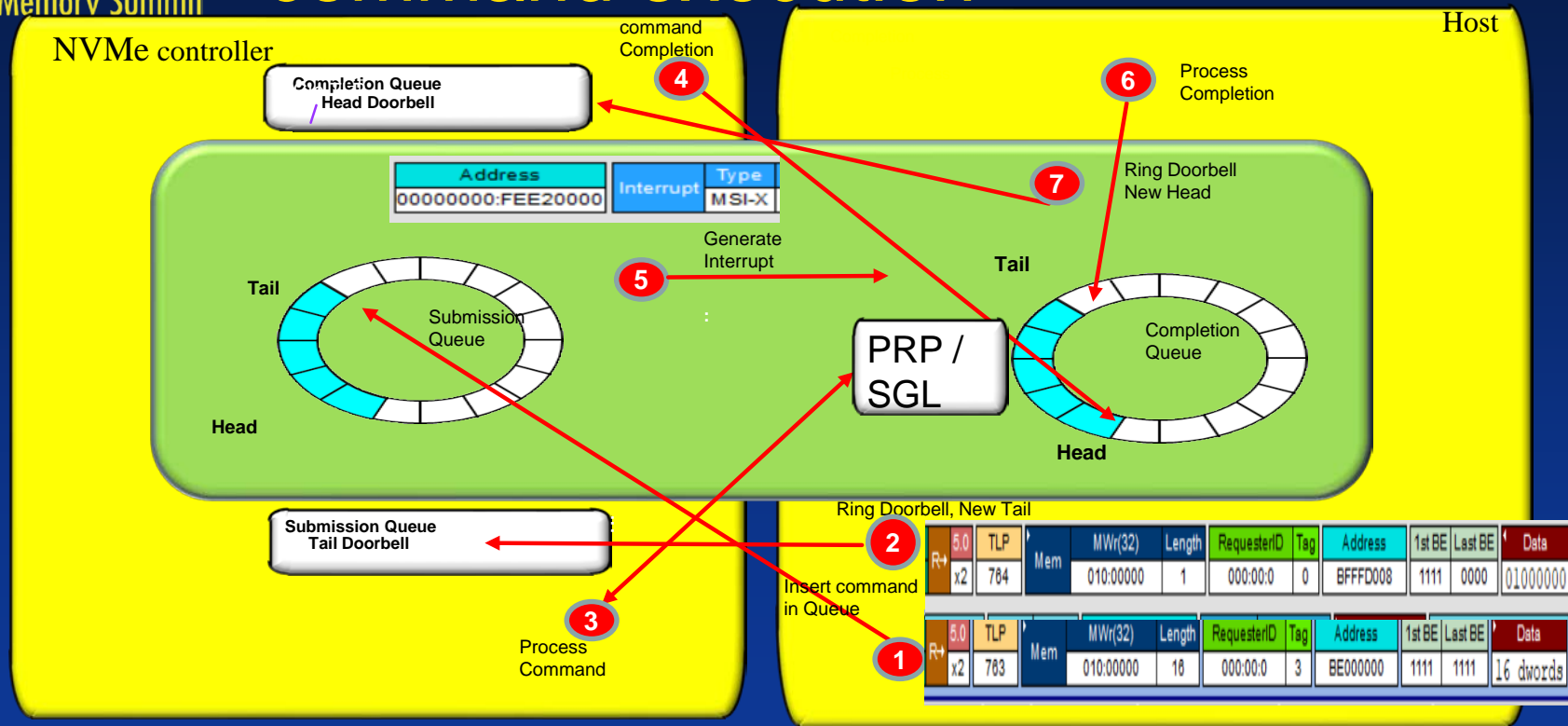


Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - Analysis of NVMe Events



Controller Memory Buffer command execution





Controller Memory Buffer

Flash Memory Summit

CMBSZ	SQS	CQS	LISTS	RDS	WDS	SZU	SZ
	1	0	0	0	0	4 KB	4096
CMBLOC	BIR	OFST	MN				
	0x4	0x00000	INTEL SSDPD2MD800G4				

NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB
42		Read	0x0001	0x0001	0x0000	1024 dwords	0x00000000:00000000	0x00000003:40A87000	0x00000000:00000000	0x00000000:00000000	0x0007

NVMe	H	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	PRP1	Address
305		034:00:0	0x0001	0x0000	00000000:BE000080		Read	Normal operation	PRP	0x0000	0x00000001		0x00000000:00000000	PRP1	0x00000003:40A8

Link Tra	R→	5.0	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VCID	ImplicitACK	Metrics	# Packets	Time
1110		x2	788		010:00000	16	000:00:0	2	BE000080	1111	1111	16 dwords	0	Packet #4097		1	140

NVMe	H	Device ID	QID	CID	SQyTDBL	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp
306		034:00:0	0x0001	0x0001		0x0003	INTEL SSDPD2MD800G4		1	42.964 us	0020.311028630000 s

Link Tra	R→	5.0	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VCID	ExplicitACK	Metrics	# Packets	Time D
1111		x2	787		010:00000	1	000:00:0	3	BFFFD008	1111	0000	1 dword	0	Packet #4097		2	42.964

NVMe	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp
307		034:00:0	0x0001	0x0000	00000003:40A87000		0x00000400	1024 dwords	INTEL SSDPD2MD800G4		16	4.512 us	0020.311028630000 s

NVMe	D	Device ID	QID	CID	Address	IOCC	SQHD	SQID	CID	P	DW0	Command Specific	ST	SCT	SC	M
308		034:00:0	0x0001	0x0000	00000003:40480020		0x0003	0x0001	0x0000	1		0x00000000		Generic Command Status	Successful Completion	0

Link Tra	R→	5.0	TLP	Mem	MWr(64)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VCID	ImplicitACK	Metrics	# Packets	Time D
1128		x2	714		011:00000	4	034:00:0	0	00000003:40480020	1111	1111	4 dwords	0	Packet #4123		1	

NVMe	D	Device ID	QID	CID	Address	Interrupt	Type	Vector	Message	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp
309		034:00:0	0x0001	0x0000	00000000:FEE20000		MSI-X	0	0x000040A2	INTEL SSDPD2MD800G4		1	5.248 us	0020.311028630000 s

NVMe	H	Device ID	QID	CID	CQyHDBL	IO CQH	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp
310		034:00:0	0x0001	0x0001		0x0003	INTEL SSDPD2MD800G4		1	7.736 us	0020.311081394000 s

Link Tra	R→	5.0	TLP	Mem	MWr(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VCID	ExplicitACK	Metrics	# Packets	Time D
1130		x2	788		010:00000	1	000:00:0	0	BFFFD00C	1111	0000	1 dword	0	Packet #4125		2	7.736



CMB command execution

NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	No frequ		
39	D	Read	0x0001	0x0001	0x0000	1024 dwords	0x00000000:00000000	0x00000003:40B59000	0x00000000:00000000	0x00000000:00000000	0x0007	0x0	0	0				
NVMe	H	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	PRP1	Address	PRP2	Address	
283	H	034:00:0	0x0001	0x0000	00000000:BE000000	Read	Normal operation	PRP	0x0000	0x00000001		0x00000000:00000000	0x00000003:40B59000		0x00000000:00000000			
NVMe	H	Device ID	QID	SQyTDBL	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp								
284	H	034:00:0	0x0001	0x0001	INTEL SSDPD2MD800G4		1	139.372 us	0020.310 007 674 000 s									
NVMe	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp					
285	D	034:00:0	0x0001	0x0000	00000003:40B59000	0x00000400	1024 dwords	INTEL SSDPD2MD800G4		16	4.488 us	0020.310 147 048 000 s						
NVMe	D	Device ID	QID	CID	Address	IOCC	SQHD	SQID	CID	P	DW0	Command Specific	ST	SCT	SC	M	DNR	MN
286	D	034:00:0	0x0001	0x0000	00000003:40480000	0x0001	0x0001	0x0000	1		0x00000000		Generic Command Status	Successful Completion	0	0		INTEL SSDPD2MD800G4
NVMe	D	Device ID	QID	CID	Address	Interrupt	Type	Vector	Message	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp				
287	D	034:00:0	0x0001	0x0000	00000000:FEE20000	MSI-X	0	0x000040A2	INTEL SSDPD2MD800G4		1	6.604 us	0020.310 151 574 000 s					
NVMe	H	Device ID	QID	CQyHDBL	IO CQH	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp								
288	H	034:00:0	0x0001	0x0001	INTEL SSDPD2MD800G4		1	634.988 us	0020.310 158 178 000 s									



Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - Analysis of NVMe Events



PRP List: points to a PRP data line

Flash Memory Summit

NVMe Cmd	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF	
56	D	Read	0x0001	0x0001	0x0033	32768 dwords	00000000:00000000	00000000:82403818	00000000:87A1E000	00000000:000E3980	0x00FF	0x0	0	1	No frequency information provided	
NVMe 414	H	Device ID	QID	CID	Address	IO SQT	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp				
414	H	001:00:0	0x0001	0x0033	00000000:87A20C00	0x0034	0x0034	SAMSUNG MZVPV256HDGL-00000	1	443.750 ns	0028 . 187 249 604 s					
NVMe 416	H	Device ID	QID	CID	Address	IO SQT	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp				
416	H	001:00:0	0x0001	0x0033	00000000:87A20C00	0x0034	0x0034	SAMSUNG MZVPV256HDGL-00000	1	443.750 ns	0028 . 187 249 604 s					
NVMe 417	H	Device ID	QID	CID	Address	IO SQT	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp				
417	H	001:00:0	0x0001	0x0033	00000000:87A1E000	0x0034	0x0034	SAMSUNG MZVPV256HDGL-00000	5	678.500 ns	0028 . 187 250 868 s					
Split Tra	R	8.0	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	Metrics	# LinkTrans	Time Delta	Time Stamp
1451	R	x4	Mem	MRd(32)	001:00:0	000:00:0	3	0	0	87A1E000	SC	0: 82404000 00000000 82405000 00000000 82406000 00000000 82407000 00000000 8: 82408000 00000000 82409000 00000000 8240A000 00000000 8240B000 00000000	2	688.500 ns	0028 . 187 251 546 s	
Split Tra	R	8.0	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	Metrics	# LinkTrans	Time Delta	Time Stamp
1452	R	x4	Mem	MRd(32)	001:00:0	000:00:0	4	0	0	87A1E040	SC	16 dwords	2	688.500 ns	0028 . 187 251 546 s	
Split Tra	R	8.0	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	Metrics	# LinkTrans	Time Delta	Time Stamp
1453	R	x4	Mem	MRd(32)	001:00:0	000:00:0	5	0	0	87A1E080	SC	16 dwords	2	686.500 ns	0028 . 187 252 234 s	
Split Tra	R	8.0	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	Metrics	# LinkTrans	Time Delta	Time Stamp
1454	R	x4	Mem	MRd(32)	001:00:0	000:00:0	6	0	0	87A1E0C0	SC	12 dwords	2	110.309 us	0028 . 187 252 920 s	
Split Tra	R	8.0	Mem	MRd(32)	RequesterID	CompleterID	Tag	TC	VC ID	Address	Status	Data	Metrics	# LinkTrans	Time Delta	Time Stamp
1455	R	x4	Mem	MRd(32)	001:00:0	000:00:0	7	0	0	87A1E0F0	SC	4 dwords	2	-27.309 us	0028 . 187 363 230 s	
NVMe 418	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
418	D	001:00:0	0x0001	0x0033	00000000:82403818	0x00000000	306 dwords	SAMSUNG MZVPV256HDGL-00000	19	782.000 ns	0028 . 187 335 920 s					
NVMe 419	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
419	D	001:00:0	0x0001	0x0033	00000000:82404000	0x00000400	1024 dwords	SAMSUNG MZVPV256HDGL-00000	40	1.649 us	0028 . 187 336 702 s					
NVMe 420	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
420	D	001:00:0	0x0001	0x0033	00000000:82405000	0x00000400	1024 dwords	SAMSUNG MZVPV256HDGL-00000	40	1.647 us	0028 . 187 338 352 s					
NVMe 421	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
421	D	001:00:0	0x0001	0x0033	00000000:82406000	0x00000400	1024 dwords	SAMSUNG MZVPV256HDGL-00000	40	1.653 us	0028 . 187 340 000 s					
NVMe 422	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
422	D	001:00:0	0x0001	0x0033	00000000:82407000	0x00000400	1024 dwords	SAMSUNG MZVPV256HDGL-00000	40	1.659 us	0028 . 187 341 652 s					
NVMe 423	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
423	D	001:00:0	0x0001	0x0033	00000000:82408000	0x00000400	1024 dwords	SAMSUNG MZVPV256HDGL-00000	40	1.649 us	0028 . 187 343 312 s					
NVMe 424	D	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
424	D	001:00:0	0x0001	0x0033	00000000:82409000	0x00000400	1024 dwords	SAMSUNG MZVPV256HDGL-00000	40	1.651 us	0028 . 187 344 962 s					

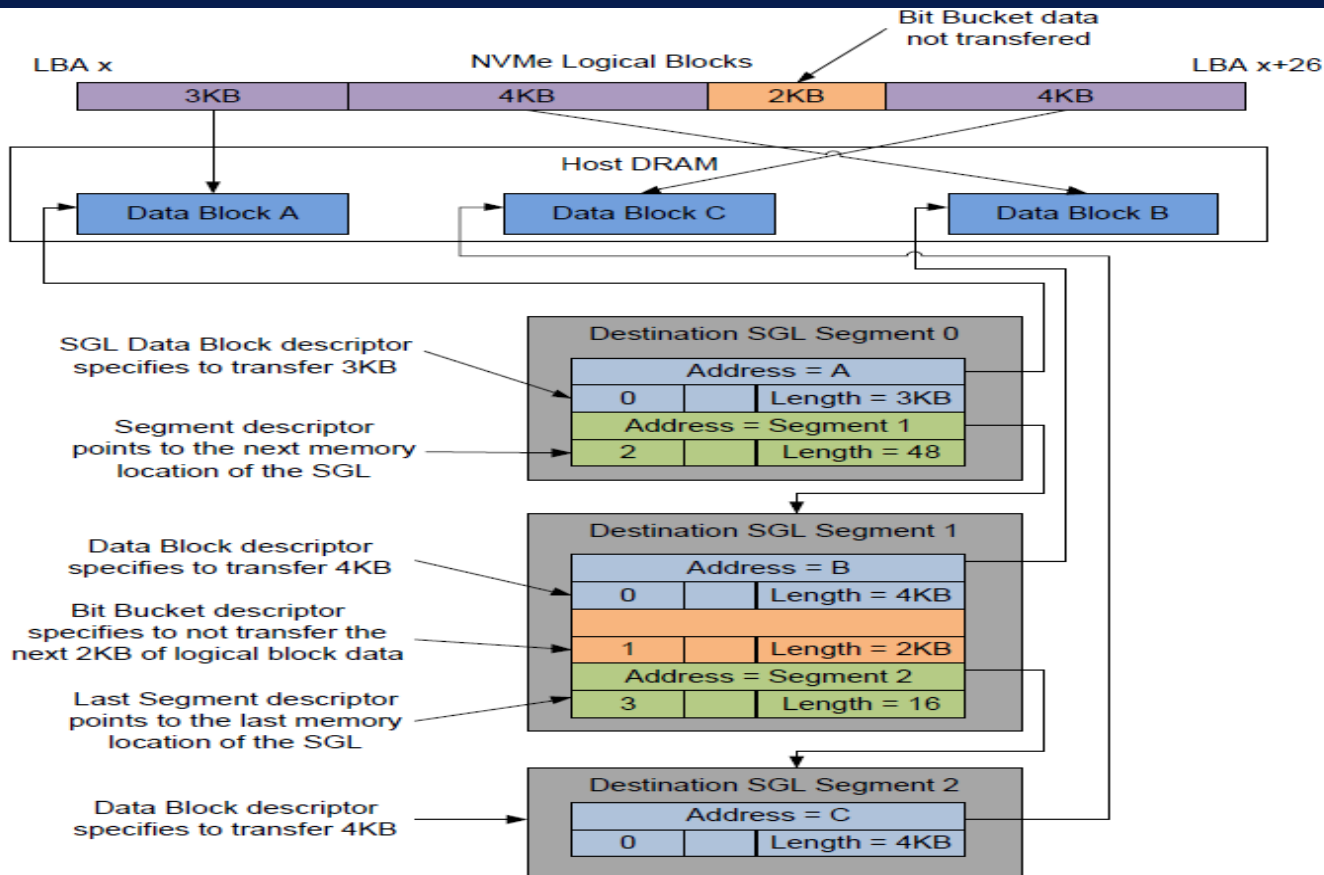


Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - Analysis of NVMe Events



SGL segments: SGL descriptor types





SGL descriptor rules

- Data Block Descriptor points to data payload
- Segment Descriptor points to next segment
- Last Segment Descriptor points to last segment
- Last segment cannot have a Segment or Last Segment Descriptor
- Bit Bucket descriptor has length but no pointer

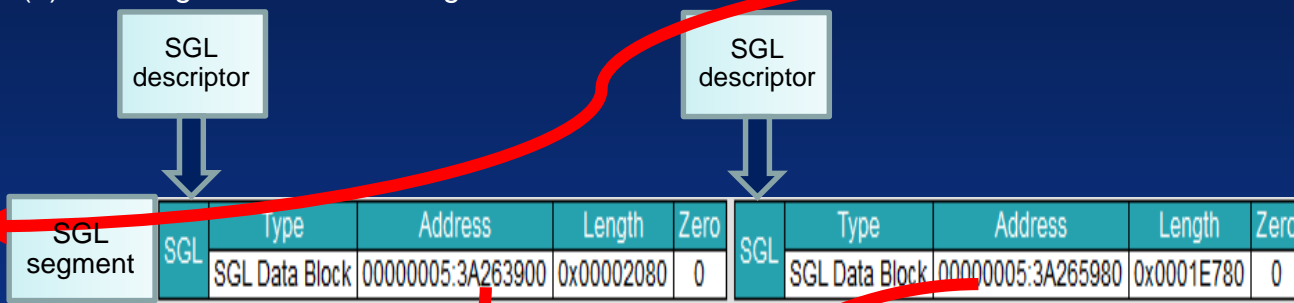


SGL decode transaction layer view

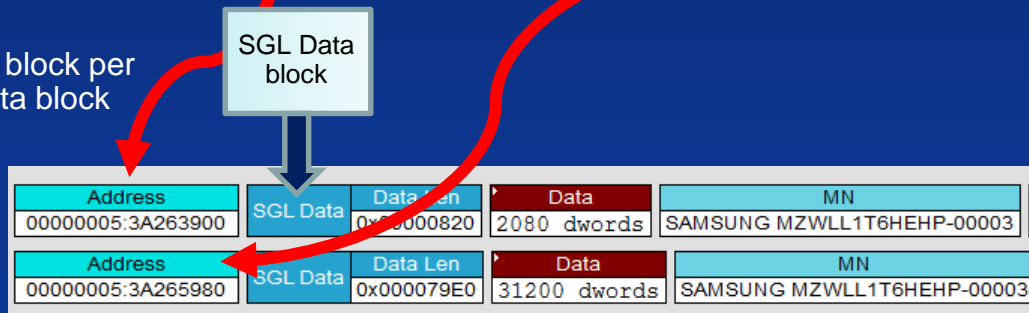
(1) command line, indicating the first SGL segment for the command and decoding its fields.

IOSQ	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	SGL	Type	Address	Length	Zero
	Read	Normal operation	SGL	0x0006	0x00000001		00000000:00000000		SGL Last Segment	00000000:84245880	0x00000020	0

(2) SGL segment line decoding its fields



(3) SGL data block per each SGL data block descriptor



NVMe	H	Device ID	QID	CID	Address	IOSeq	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	Length	Zero	SLBA	PRCHK	PRACT	FUA	LR	AF	AL	SR	I			
350		192.00.0	0x0001	0x0008	0000007F.2B7B8200		Read	Normal operation	SGL	0x0008	0x00000001		0x00000000.00000000		0	0x00000000.61CC8018	0x0007	PRINFO	100	0	0	0	0	No frequency info	None	0	0

EILBRT	ELBAT	ELBATM	MN	# Link & Split Trans	Time Delta	Time Stamp
0x61CC8018	0xBEEF	0xFFFF	INTEL SSDPD21K015TAM	1	9.135 us	0165. 058 943 850 250 s

NVMe	H	Device ID	QID	CID	Address	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
352		192.00.0	0x0001	0x0008	0000007F.72600000	SGL Data Block	0x0000007F.7DC60000	0x00000053	0		SGL Data Block	0x0000007F.7DC60053	0x00000053	0		SGL Data Block	0x0000007F.7DC600A6	0x00000053	0		SGL Data Block	0x0000007F.7DC600F9	0x00000053	0

SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0	SGL Data Block	0x0000007F.7DC6014C	0x00000053	0	0	SGL Data Block	0x0000007F.7DC6019F	0x00000053	0	0	SGL Data Block	0x0000007F.7DC601F2	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60245	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60298	0x00000053	0

Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0	0	SGL Data Block	0x0000007F.7DC602EB	0x00000053	0	0	SGL Data Block	0x0000007F.7DC6033E	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60391	0x00000053	0	0	SGL Data Block	0x0000007F.7DC603E4	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60437	0x00000053	0

Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0x00000053	0	0	SGL Data Block	0x0000007F.7DC6048A	0x00000053	0	0	SGL Data Block	0x0000007F.7DC604DD	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60530	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60583	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60647	0x00000053	0

Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0x0000007F.7DC605D8	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60629	0x00000053	0	0	SGL Data Block	0x0000007F.7DC6067C	0x00000053	0	0	SGL Data Block	0x0000007F.7DC606CF	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60722	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60775	0x00000053	0

Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
SGL Data Block	0x0000007F.7DC60775	0x00000053	0	0	SGL Data Block	0x0000007F.7DC607C8	0x00000053	0	0	SGL Data Block	0x0000007F.7DC6081B	0x00000053	0	0	SGL Data Block	0x0000007F.7DC6086E	0x00000053	0	0	SGL Last Segment	0x0000007F.726001C0	0x00000170	0

MN	# Link & Split Trans	Time Delta	Time Stamp
INTEL SSDPD21K015TAM	2	9.833 us	0165. 058 952 985 250 s

NVMe	H	Device ID	QID	CID	Address	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
353		192.00.0	0x0001	0x0008	0000007F.726001C0	SGL Data Block	0x0000007F.7DC608C1	0x00000053	0		SGL Data Block	0x0000007F.7DC60914	0x00000053	0		SGL Data Block	0x0000007F.7DC60967	0x00000053	0		SGL Data Block	0x0000007F.7DC609BA	0x00000053	0

SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0	SGL Data Block	0x0000007F.7DC60A0D	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60A60	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60AB3	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60B06	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60B59	0x00000053	0

Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0	0	SGL Data Block	0x0000007F.7DC60BAC	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60BF8	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60C52	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60CA5	0x00000053	0	0	SGL Data Block	0x0000007F.7DC60CF8	0x00000053	0

Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0x00000054	0	0	SGL Data Block	0x0000007F.7DC60D4C	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60DA0	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60DF4	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60E48	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60F1B	0x00000054	0

Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero	SGL	Type	Address	Length	Zero
0x0000007F.7DC60E9C	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60EF0	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60F44	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60F98	0x00000054	0	0	SGL Data Block	0x0000007F.7DC60FEC	0x00000054	0	0	SGL Data Block	0x0000007F.7DC61048	0x00000054	0

MN	# Link & Split Trans	Time Delta	Time Stamp
INTEL SSDPD21K015TAM	3	34.651 us	0165. 058 962 818 250 s

NVMe	D	Device ID	QID	CID	Address	SGL Data	Data Len	Data	MN	# Link & Split Trans	Time Delta	Time Stamp
354		192.00.0	0x0001	0x0008	0000007F.7DC60000	0x00000005	0:	FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF	INTEL SSDPD21K015TAM	2	203.000 ns	0165. 058 997 469 750 s
							4:	FFFFFFFF				

NVMe	D	Device ID	QID	CID	Address	SGL Data	Data Len	Data	MN	# Link & Split Trans	Time Delta	Time Stamp
355		192.00.0	0x0001	0x0008	0000007F.7DC600A4	0x00000005	0:	0000FFFF FFFFFFFF FFFFFFFF FFFFFFFF	INTEL SSDPD21K015TAM	2	198.750 ns	0165. 058 997 672 750 s
							4:	FFFFFFFF				

NVMe	D	Device ID	QID	CID	Address	SGL Data	Data Len	Data	MN	# Link & Split Trans	Time Delta	Time Stamp
356		192.00.0	0x0001	0x0008	0000007F.7DC6014C	0x00000005	0:	FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF	INTEL SSDPD21K015TAM	2	164.500 ns	0165. 058 997 871 500 s
							4:	FFFFFFFF				

Santa Clara, CA
August 2019



Bit bucket descriptor

NVMe	H	Device ID	QID	SQyTDBL	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp										
631		198:00:0	0x0001		0x0069	HUSMR7638BDP3Y1		1	265.000 ns	0010.178 690 834 250 s										
NVMe	H	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	SGL	Type	Length	SLBA	NLB		
632		198:00:0	0x0001	0x0080	00000000:87A48A00		Read	Normal operation	SGL	0x0080	0x00000001		0x00000000:00000000		SGL Bit Bucket	0x0001FFED	0x00000000:00081848	0x00FB		
NVMe	D	Device ID	QID	CID	Address	IOCC	SQHD	SQID	CID	P	DWO	Command Specific	ST	SCT	SC	M	DNR	MN	Metrics	# Link
633		198:00:0	0x0001	0x0080	00000000:87A47680		0x0069	0x0001	0x0080	1		0x00000000		Generic Command Status	Successful Completion	0	0	HUSMR7638BDP3Y1		
NVMe	H	Device ID	QID	COyHDBL	IO CQH	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp										
635		198:00:0	0x0001		0x0069	HUSMR7638BDP3Y1		1	2.000 sec	0010.178 759 233 250 s										

- Last Segment has only a Bit Bucket Descriptor
- No segment descriptor in Last Segment



Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - Analysis of NVMe Events



NVMe Analyzer mode - Goals

- Real time analysis of NVMe products and
 - Identify specific NVMe events and scenarios
 - identify performance bottlenecks within a long recording
- PCIe analysis engine Needs to understand the NVMe command queuing mechanism
 - PCIe analysis engine self-learns queue and payload bindings



NVMe mode Long Recordings analysis

- Dropped idles, SKPs, EDSs, DLLPs and data payload
- Memory Utilization Model Assumptions
 - 16GB memory dedicated per direction
 - Capture Duration Doubles for Gen4 NVMe analyzer
 - Recording stops as soon as either side fills up
 - SSD rate for read is 2 GB / sec or 16 Gb/sec
 - This implies a Gen3 x4 link with 60% utilization
 - Assume 16 pages / command
 - Assume 2 doorbells
 - Assume no interrupt aggregation
- Each TLP to occupy 1.09 to 1.43 memory blocks on average



Recording duration using 32K DW NVMe Transfers

Best Case
Gen 1 x1 Lane Width

Worst Case
Gen 3 x4 Lane Width

TLP size (dw)	Link Utilization (%)				
	100	90	80	70	60
32	139.8164	155.3516	174.7705	199.7378	233.0274
64	128.817	143.13	161.0213	184.0243	214.695
128	123.3173	137.0192	154.1466	176.1676	205.5289
256	120.5675	133.9638	150.7093	172.2392	200.9458
512	119.1925	132.4361	148.9907	170.275	198.6542
1024	118.5051	131.6723	148.1313	169.293	197.5084
2048	118.1613	131.2904	147.7017	168.8019	196.9356
4096	117.9895	131.0994	147.4868	168.5564	196.6491

TLP size (dw)	Link Utilization (%)				
	100	90	80	70	60
32	8.875066	9.861185	11.09383	12.67867	14.79178
64	8.176861	9.085402	10.22108	11.68123	13.6281
128	7.827759	8.69751	9.784699	11.18251	13.04627
256	7.653208	8.503564	9.56651	10.93315	12.75535
512	7.565932	8.406591	9.457415	10.80847	12.60989
1024	7.522294	8.358105	9.402868	10.74613	12.53716
2048	7.500475	8.333862	9.375594	10.71496	12.50079
4096	7.489566	8.32174	9.361958	10.69938	12.48261

**Note: Capture Duration Doubles when using Gen4 analyzer

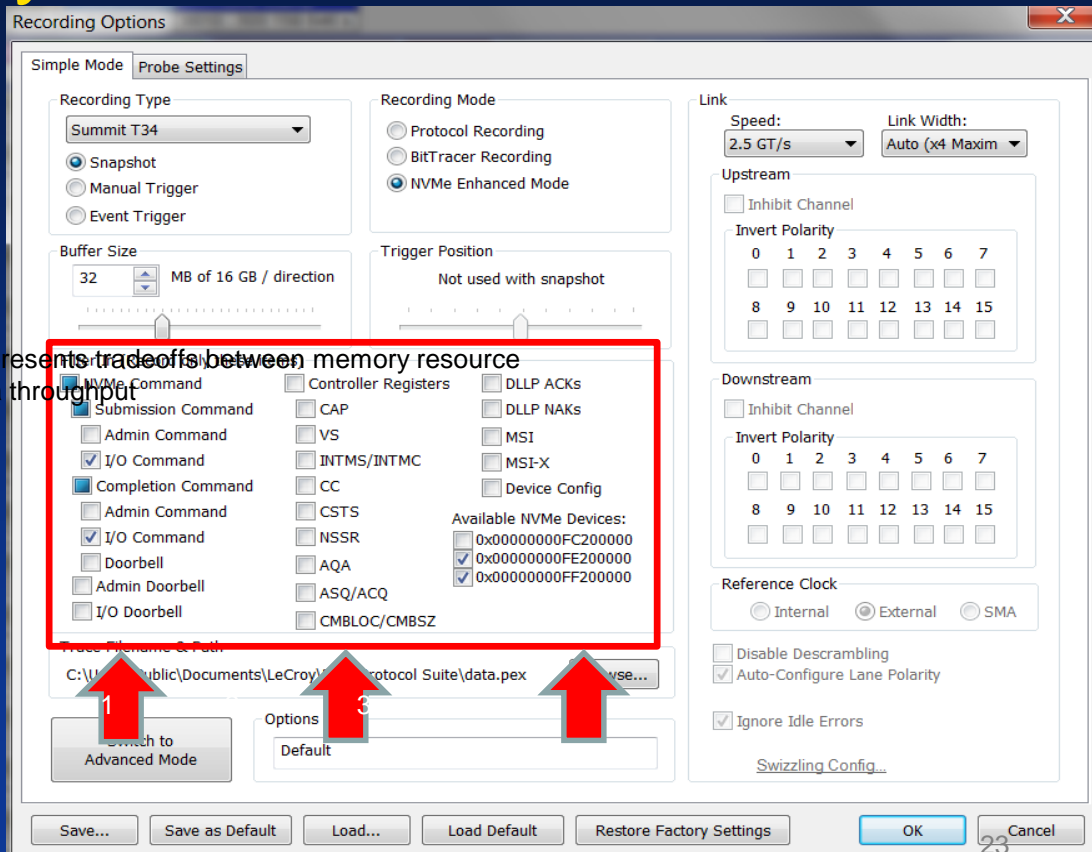


Flash Memory Summit

NVMe analyzer mode

- Filter in items
 1. Entities that form NVMe Commands
 2. NVMe Control Registers
 3. PCIe entities related to NVMe traffic
 4. ACK/NAK DLLPs
 5. Interrupts
 6. PCIe configuration map

Queue allocation presents tradeoffs between memory resource allocation and data throughput





NVMe Analyzer mode

Event Properties

NVMe Filters

Label:
NVMe Filters

NVMe Filters | **Actions**

Filter In (Record only these items)

<input checked="" type="checkbox"/> NVMe Command	<input checked="" type="checkbox"/> Controller Registers	<input type="checkbox"/> DLLP ACKs
<input checked="" type="checkbox"/> Submission Command	<input type="checkbox"/> CAP	<input type="checkbox"/> DLLP NAKs
<input type="checkbox"/> Admin Command	<input type="checkbox"/> VS	<input type="checkbox"/> MSI
<input checked="" type="checkbox"/> I/O Command	<input type="checkbox"/> INTMS/INTMC	<input type="checkbox"/> MSI-X
<input checked="" type="checkbox"/> Completion Command	<input checked="" type="checkbox"/> CC	<input checked="" type="checkbox"/> Device Config
<input type="checkbox"/> Admin Command	<input type="checkbox"/> CSTS	<input type="button" value="Set Namespace IDs"/>
<input checked="" type="checkbox"/> I/O Command	<input type="checkbox"/> NSSR	Available NVMe Devices:
<input type="checkbox"/> Doorbell	<input checked="" type="checkbox"/> AQA	<input checked="" type="checkbox"/> 0x00000000EF510000
<input type="checkbox"/> Admin Doorbell	<input checked="" type="checkbox"/> ASQ/ACQ	
<input type="checkbox"/> I/O Doorbell	<input type="checkbox"/> CMBLOC/CMBSZ	
	<input type="checkbox"/> BPINFO	
	<input type="checkbox"/> BPRSEL	
	<input type="checkbox"/> BPMBL	



Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - Analysis of NVMe Events



Flash Memory Summit

NVMe Analyzer mode

NVMe Cmd	H	OPC	SQID	CQID	CID	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF	ACCL	SEQR	INCOM	ILBF		
3	H	Write	0x0005	0x0005	0x0146	0x00000000:00000000	0x00000000:9C7A1000	0x00000000:00000000	0x00000000:0060808	0x0007	0x0	0	0	DSM	No frequency information provided	None	0	0	0x0000		
NVMe	H	Device ID	QID	SQyTDBL	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp											
15	H	005:00:0	0x0005		0x0147	Samsung SSD 950 PRO 256GB		1	725.000 ns	0023.903117589500 s											
NVMe	H	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	PRP1	Address	PRP2	Address	SLBA	NLB		
16	H	005:00:0	0x0005	0x0146	00000001:10425180	Write	Normal operation	PRP		0x0146	0x00000001		0x00000000:00000000	0x00000000:9C7A1000		0x00000000:00000000		0x00000000:0060808	0x0007		
NVMe	D	Device ID	QID	CID	Address	IOCC	SQHD	SQID	CID	P	DW0	Command Specific	ST	SCT	SC	M	DNR	MN	Metrics	# Link & Split Trans	Time
17	D	005:00:0	0x0005	0x0146	00000001:08BF9460		0x0147	0x0005	0x0146	0		0x00000000		Generic Command Status	Successful Completion	0	0	Samsung SSD 950 PRO 256GB		1	166
NVMe	D	Device ID	QID	CID	Address	Interrupt	Type	Vector	Message	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp							
18	D	005:00:0	0x0005	0x0146	00000000:FEE220C		MSI-X	5	0x00004980	Samsung SSD 950 PRO 256GB		1	42.732 us	0023.903422094750 s							
NVMe	H	Device ID	QID	CQyHDBL	IO CQH	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp											
19	H	005:00:0	0x0005		0x0147	Samsung SSD 950 PRO 256GB		1	159.344 us	0023.903464826750 s											
NVMe Cmd	H	OPC	SQID	CQID	CID	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF	ACCL	SEQR	INCOM	EILBF		
0	H	Read	0x0008	0x0008	0x0000	0x00000000:00000000	0x00000004:01684000	0x00000000:00000000	0x00000000:00127660	0x0007	0x0	0	0	DSM	No frequency information provided	None	0	0	0x0000		
NVMe	H	Device ID	QID	SQyTDBL	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp											
0	H	001:00:0	0x0008		0x0001	SAMSUNG MZVLB1T0HALR-00000		1	704.000 ns	0002.279277773500 s											
NVMe	H	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	PRP1	Address	PRP2	Address	SLBA	NLB		
1	H	001:00:0	0x0008	0x0000	00000004:044B0000	Read	Normal operation	PRP		0x0000	0x00000001		0x00000000:00000000	0x00000004:01684000		0x00000000:00000000		0x00000000:00127660	0x0007		
NVMe	D	Device ID	QID	CID	Address	IOCC	SQHD	SQID	CID	P	DW0	Command Specific	ST	SCT	SC	M	DNR	MN	Metrics	# Link & Split Trans	Time
28	D	001:00:0	0x0008	0x0000	00000004:044DC000		0x000E	0x0008	0x0000	1		0x00000000		Generic Command Status	Successful Completion	0	0	SAMSUNG MZVLB1T0HALR-00000		1	
NVMe	H	Device ID	QID	CQyHDBL	IO CQH	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp											
29	H	001:00:0	0x0008		0x0001	SAMSUNG MZVLB1T0HALR-00000		1	-92.427 us	0002.279382738000 s											



NVMe Analyzer mode

- NVMe Read Command Packet Level associated by queue address
 - Completion TLP with “mysterious” address field

Packet 0	8.0 x4	TLP 424	Mem	MWr(32)	Length 1	RequesterID 000:00:0	Tag 0	Address F7101040	1st BE 1111	Last BE 0000	Data 1 dword	LCRC 0xA493CEA7	Idle 697.750 ns	Time Sta 0002.279277
Packet 1	8.0 x4	TLP 425	Cpl	CpID	Length 16	RequesterID 001:00:0	Tag 0	CompleterID 000:00:0	Status SC	BCM 0	Byte Cnt 64	Lwr Addr 0x00	Address 00000004:044B0000	Data 16 dwords
Packet 2	8.0 x4	TLP 426	Mem	MWr(32)	Length 1	RequesterID 000:00:0	Tag 0	Address F7101040	1st BE 1111	Last BE 0000	Data 1 dword	LCRC 0x8DAC4761	Idle 680.500 ns	Time Sta 0002.279290
Packet 3	8.0 x4	TLP 427	Cpl	CpID	Length 16	RequesterID 001:00:0	Tag 0	CompleterID 000:00:0	Status SC	BCM 0	Byte Cnt 64	Lwr Addr 0x40	Address 00000004:044B0040	Data 16 dwords
Packet 4	8.0 x4	TLP 428	Mem	MWr(32)	Length 1	RequesterID 000:00:0	Tag 0	Address F7101040	1st BE 1111	Last BE 0000	Data 1 dword	LCRC 0xE07CCF7E	Idle 710.000 ns	Time Sta 0002.279297
Packet 5	8.0 x4	TLP 429	Cpl	CpID	Length 16	RequesterID 001:00:0	Tag 0	CompleterID 000:00:0	Status SC	BCM 0	Byte Cnt 64	Lwr Addr 0x00	Address 00000004:044B0080	Data 16 dwords
Packet 6	8.0 x4	TLP 430	Mem	MWr(32)	Length 1	RequesterID 000:00:0	Tag 0	Address F7101040	1st BE 1111	Last BE 0000	Data 1 dword	LCRC 0x9ED42437	Idle 689.750 ns	Time Sta 0002.279305
Packet 7	8.0 x4	TLP 431	Cpl	CpID	Length 16	RequesterID 001:00:0	Tag 0	CompleterID 000:00:0	Status SC	BCM 0	Byte Cnt 64	Lwr Addr 0x40	Address 00000004:044B00C0	Data 16 dwords



Agenda

- NVMe Command Execution
 - Host Controller Interface
 - Controller Memory Buffer
- NVMe Data Structures
 - Physical Region Page
 - Scattered Gathered List
- NVMe Analysis Engine
 - Recording NVMe transactions
 - Analysis of NVMe Events



NVMe performance analysis

- As NVMe technology matures leading to a need to maximize performance
- What's special about NVMe performance vs general PCIe performance
- differences in performance analysis techniques between SSD drives and traditional magnetic drives.



NVMe latency categories

- Doorbell to command submission
- Command submission to Data transfer
- Command sub mission to command completion
- Command completion to interrupt



NVMe Performance management

- PCI express flow control credit starvation.
- Tradeoffs between memory resource allocation and data throughput
- Queue allocation presents tradeoffs between command latency and throughput
- Virtual channels manage traffic in fabric
 - Allocate traffic classes to Virtual channels



NVMe performance criteria

- Response time
 - Transmission of the complete transfer from the beginning of the PCIe packet to the end of the last PCIe packet of this NVMe command
- Latency time
 - Time from the last PCIe packet of the NVMe command submission to the first PCIe packet of the NVMe command completion
- Throughput
 - NVMe command payload of all simultaneous coincident commands during the processing of current command divided by response time



Instantaneous Performance Metrics

Flash Memory Summit

NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF
21		Read	0x0001	0x0001	0x0003	128 dwords	00000000:00000000	00000002:2DD9B000	00000000:00000000	00000000:00000000	0x0000	0x0	0	0		No frequency information provided
ACCL	SEQR	INCOM	EILBRT	ELBAT	ELBATM	ST	SCT	SC	Device ID	MN	Explicit SQyTDBL	Explicit IOSQ	Explicit IOCQ	Explicit CQyHDBL	NSID	Metrics
None	0	0	0x00000000	0x0000	0x0000	ST	Generic Command Status	Successful Completion	006:00:0	NVMeLeCroy000000	NVMe #120	NVMe #121	NVMe #123	NVMe #125	0x00000001	
# NVMe Trans	Resp. time	Latency	Pld. Bytes	Thrpt MB/s	IOPS	SDBl - CDBl	SDBl - CCmd	SCmd - CCmd	Time Delta	Time Stamp						
6	683.652 us	464.192 us	512	0.714	1462.733	682.240 us	631.476 us	466.192 us	692.512 us	0039.509104202 s						
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF
22		Read	0x0001	0x0001	0x0004	128 dwords	00000000:00000000	00000002:2DFC3CC0	00000000:00000000	00000000:00000000	0x0000	0x0	0	0		No frequency information provided
ACCL	SEQR	INCOM	EILBRT	ELBAT	ELBATM	ST	SCT	SC	Device ID	MN	Explicit SQyTDBL	Explicit IOSQ	Explicit IOCQ	Explicit CQyHDBL	NSID	Metrics
None	0	0	0x00000000	0x0000	0x0000	ST	Generic Command Status	Successful Completion	006:00:0	NVMeLeCroy000000	NVMe #126	NVMe #127	NVMe #129	NVMe #131	0x00000001	
# NVMe Trans	Resp. time	Latency	Pld. Bytes	Thrpt MB/s	IOPS	SDBl - CDBl	SDBl - CCmd	SCmd - CCmd	Time Delta	Time Stamp						
6	812.300 us	509.776 us	512	0.601	1231.072	810.888 us	760.884 us	511.840 us	13.514 ms	0039.509796714 s						
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF
23		Read	0x0001	0x0001	0x0005	128 dwords	00000000:00000000	00000002:2DFCD980	00000000:00000000	00000000:00000000	0x0000	0x0	0	0		No frequency information provided
ACCL	SEQR	INCOM	EILBRT	ELBAT	ELBATM	ST	SCT	SC	Device ID	MN	Explicit SQyTDBL	Explicit IOSQ	Explicit IOCQ	Explicit CQyHDBL	NSID	Metrics
None	0	0	0x00000000	0x0000	0x0000	ST	Generic Command Status	Successful Completion	006:00:0	NVMeLeCroy000000	NVMe #132	NVMe #133	NVMe #135	NVMe #137	0x00000001	
# NVMe Trans	Resp. time	Latency	Pld. Bytes	Thrpt MB/s	IOPS	SDBl - CDBl	SDBl - CCmd	SCmd - CCmd	Time Delta	Time Stamp						
6	681.364 us	470.128 us	512	0.717	1467.644	679.984 us	626.980 us	472.144 us	696.584 us	0039.523310538 s						
NVMe Cmd	D	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF
24		Read	0x0001	0x0001	0x0006	128 dwords	00000000:00000000	00000002:2DD9D000	00000000:00000000	00000000:00000000	0x0000	0x0	0	0		No frequency information provided
ACCL	SEQR	INCOM	EILBRT	ELBAT	ELBATM	ST	SCT	SC	Device ID	MN	Explicit SQyTDBL	Explicit IOSQ	Explicit IOCQ	Explicit CQyHDBL	NSID	Metrics
None	0	0	0x00000000	0x0000	0x0000	ST	Generic Command Status	Successful Completion	006:00:0	NVMeLeCroy000000	NVMe #138	NVMe #139	NVMe #141	NVMe #143	0x00000001	
# NVMe Trans	Resp. time	Latency	Pld. Bytes	Thrpt MB/s	IOPS	SDBl - CDBl	SDBl - CCmd	SCmd - CCmd	Time Delta	Time Stamp						
6	780.748 us	475.760 us	512	0.625	1280.823	779.112 us	726.268 us	477.808 us	804.880 us	0039.524007122 s						



Response time

NVMe Cmd	OPC	SQID	CQID	CID	Data	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACC	ACCF	ACCL	SEQR	INCOM	EILBRT	ELBAT	ELBATM	ST
21	D	Read	0x0001	0x0001	0x0003	128 dwords	00000000:00000000	00000002:2D9B000	00000000:00000000	00000000:00000000	0x0000	0x0	0	0	DSM	None	0	0	0	0x00000000	0x0000	0x0000	ST
Generic Command Status		Successful Completion		Device ID	MN	Explicit SQyTDBL	Explicit IOSQ	Explicit IOCC	Explicit CQyHDBL	NSID	Metrics	# NVMe Tr	Resp. time	Latency	Bytes	Thrpt MB/s	IOPS	SDbl - CDBl	SDbl - CCmd	SCmd - CCmd			
				006:00:0	NVMeLeCroy000000	NVMe #120	NVMe #121	NVMe #123	NVMe #125	0x00000001		6	683.652 us	464.192 us	512	0.714	1462.733	682.240 us	631.476 us	466.192 us			
Time Delta		Time Stamp																					
1.428 us		0039 : 509 104 202 s																					
NVMe	Device ID	QID	SQyTDBL	IO SQT	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp														
120	H	006:00:0	0x0001	0x0004	NVMeLeCroy000000		1	1.428 us	0039 : 509 104 202 s														
Link Tra	R	2.5	TLP	Mem	MW(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	Explicit ACK	Metrics	# Packets	Time Delta	Time Stamp					
1927	R	x1	929	Mem	010:00000	1	000:00:0	0	FE201008	1111	0000	04000000	0	Packet #377117		2	1.428 us	0039 : 509 104 202 s					
Packet	R	2.5	TLP	Mem	MW(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	LCRC	Time Delta	Time Stamp								
377116	R	x1	929	Mem	010:00000	1	000:00:0	0	FE201008	1111	0000	1 dwords	0x551D6F4D	1.428 us	0039 : 509 104 202 s								
Packet	R	2.5	DLLP	ACK	AckNak_Seq_Num	CRC 16	Time Delta	Time Stamp															
377117	R	x1			929	0xD93D	163.856 us	0039 : 509 105 630 s															
NVMe	Device ID	QID	CID	Address	IOSQ	OPC	FUSE	PSDT	CID	NSID	MPTR	Address	PRP1	PRP2	Address	SLBA	NLB	PRINFO	PRCHK	PRACT	FUA	LR	
121	H	006:00:0	0x0001	0x0003	00000002:2E0830C0	Read	Normal operation	PRP	0x0003	0x00000001		00000000:00000000	00000002:2D9B000	00000000:00000000	00000000:00000000	0x0000	0x0000	000	0	0	0	0	
AF	AL	SR	EILBRT	ELBAT	ELBATM	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp													
No frequency info	None	0	0	0x00000000	0x0000	0x0000	NVMeLeCroy000000	1	358.206 us	0039 : 509 269 486 s													
NVMe	Device ID	QID	CID	Address	PRP Data	Data Len	Data	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp											
122	D	006:00:0	0x0001	0x0003	00000002:2D9B000	128 dwords	NVMeLeCroy000000	NVMeLeCroy000000		4	107.984 us	0039 : 509 627 694 s											
NVMe	Device ID	QID	CID	Address	IOCC	SOHD	SQID	CID	P	DW0	Reserved	SCT	SC	M	DNR	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp			
123	D	006:00:0	0x0001	0x0003	00000002:2E093030	0x0004	0x0001	0x0003	1	0x00000000	0x00000000	Generic Command Status	Successful Completion	0	0	NVMeLeCroy000000		1	121.648 us	0039 : 509 735 678 s			
NVMe	Device ID	QID	CID	Address	Interrupt	Type	Vector	Message	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp										
124	D	006:00:0	0x0001	0x0003	00000000:FEE3F00C	MSI	1	0x000049A9	NVMeLeCroy000000		1	29.116 us	0039 : 509 757 326 s										
NVMe	Device ID	QID	COyHDBL	IO CQH	MN	Metrics	# Link & Split Trans	Time Delta	Time Stamp														
125	H	006:00:0	0x0001	0x0004	NVMeLeCroy000000		1	1.380 us	0039 : 509 786 442 s														
Link Tra	R	2.5	TLP	Mem	MW(32)	Length	RequesterID	Tag	Address	1st BE	Last BE	Data	VC ID	Explicit ACK	Metrics	# Packets	Time Delta	Time Stamp					
1936	R	x1	931	Mem	010:00000	1	000:00:0	0	FE20100C	1111	0000	04000000	0	Packet #377135		2	1.380 us	0039 : 509 786 442 s					

Command level metrics

Latency time

Response time



Flash Memory Summit

Performance Analysis via Trace Expert

- Drill down to  get more reports

Performance Analysis

- [Link Transaction Performance](#)
- [Split Transactions Performance](#)
- [NVMe Performance](#)

Link Transaction Performance

Performance

Transaction Type	Total	# Packets (Min)	# Packets (Avg)	# Packets (Max)	Resp. time (Min)	Resp. time (Avg)	Resp. time (Max)	Pld. Bytes (Min)	Pld. Bytes (Avg)	Pld. Bytes (Max)
MsgD	1	2	2.00	2	1.412 us	1.412 us	1.412 us	4	4.00	4
CfgRd0	1072	2	2.00	2	1.236 us	1.375 us	1.620 us	0	0.00	0
CpID	565132	2	2.00	2	84.000 ns	1.636 us	2.396 us	1	63.87	64
CfgWr0	86	2	2.00	2	1.316 us	1.401 us	1.604 us	1	2.47	4
Cpl	163	2	2.00	2	68.000 ns	188.340 ns	308.000 ns	0	0.00	0
MWr(32)	2939	2	2.00	2	76.000 ns	1.011 us	1.692 us	4	4.00	4
MRd(32)	3	2	2.00	2	1.404 us	1.625 us	1.740 us	0	0.00	0
MRd(64)	71725	2	2.00	2	76.000 ns	202.820 ns	468.000 ns	0	0.00	0
MWr(64)	29700	2	2.00	2	148.000 ns	696.740 ns	988.000 ns	16	123.37	128
	670821									

Memory Writes Performance

Requester, TC	Total	Resp. time (Min)	Resp. time (Avg)	Resp. time (Max)	Pld. Bytes (Min)	Pld. Bytes (Avg)	Pld. Bytes (Max)	Thrpt MB/s (Min)	Thrpt MB/s (Avg)	Thrpt MB/s (Max)
000:00:0, TCO	1973	1.268 us	1.407 us	1.692 us	4	4.00	4	2.255	2.715	3.008
006:00:0, TCO	30667	76.000 ns	681.170 ns	988.000 ns	4	119.60	128	10.039	162.247	207.603
	32640									

Split Transactions Performance

Overall Performance

Requester -> Completer	Total	# Packets (Min)	# Packets (Avg)	# Packets (Max)	Resp. time (Min)	Resp. time (Avg)	Resp. time (Max)
000:00:0 -> 000:00:0	6	2	2	2	102.880 us	123.771 us	179.312 us
000:00:0 -> 006:00:0	1155	2	2	2	99.648 us	118.388 us	647.712 us
006:00:0 -> 000:00:0	71718	2	2	9	1.848 us	6.673 ms	9.859 ms
	72879						

More Reports





Categorized Performance analysis

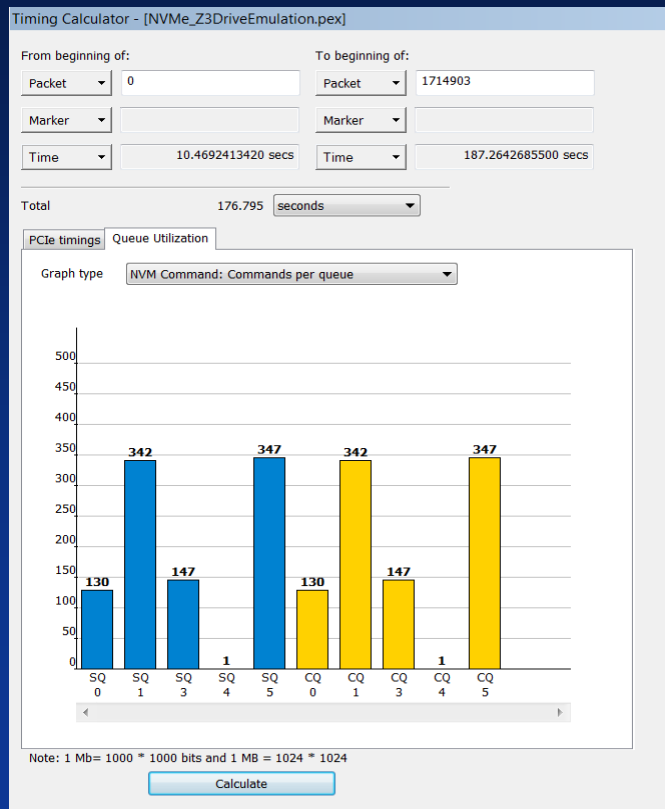
Read Requests Performance

Requester -> Completer	Total	Thrpt MB/s (Min)	Thrpt MB/s (Avg)	Thrpt MB/s (Max)	Resp. time (Min)	Resp. time (Avg)	Resp. time (Max)	Latency (Min)	Latency (Avg)	Latency (Max)
000:03:0 -> 000:00:0, Cfg TCO	416	0.995	21.175	1502.353	378.000 ns	1.172 us	3.834 us	204.000 ns	1.009 us	3.662 us
000:03:0 -> 002:00:0, Cfg TCO	5314	7.914	6189.918	11387.661	338.000 ns	376.760 ns	482.000 ns	174.000 ns	210.050 ns	304.000 ns
000:03:0 -> 003:04:0, Cfg TCO	5198	7.981	6833.574	11387.661	338.000 ns	382.470 ns	478.000 ns	174.000 ns	216.120 ns	304.000 ns
000:03:0 -> 004:00:0, Cfg TCO	1044	0.245	22.834	1526.181	2.522 us	3.152 us	15.562 us	2.360 us	2.986 us	15.396 us
000:03:0 -> 003:05:0, Cfg TCO	5194	8.048	6841.318	11387.661	338.000 ns	382.450 ns	474.000 ns	174.000 ns	215.810 ns	312.000 ns
000:03:0 -> 005:00:0, Cfg TCO	1024	0.093	23.177	1507.059	2.554 us	3.202 us	41.170 us	2.390 us	3.035 us	40.994 us
000:03:0 -> 003:06:0, Cfg TCO	5190	8.048	6844.803	11387.661	338.000 ns	382.440 ns	474.000 ns	174.000 ns	216.210 ns	312.000 ns
000:03:0 -> 006:00:0, Cfg TCO	1024	0.082	23.386	1523.765	2.526 us	3.198 us	46.490 us	2.370 us	3.032 us	46.322 us
000:03:0 -> 003:07:0, Cfg TCO	5190	7.981	6845.320	11387.661	338.000 ns	382.410 ns	478.000 ns	174.000 ns	216.160 ns	304.000 ns
000:03:0 -> 007:00:0, Cfg TCO	1024	0.081	23.116	1521.356	2.530 us	3.212 us	46.918 us	2.386 us	3.045 us	46.774 us
000:03:0 -> 004:00:0, Mem TCO	2089	0.005	0.889	16.124	3.954 us	6.335 us	807.882 us	3.796 us	6.168 us	807.708 us
004:00:0 -> 000:00:0, Mem TCO	277	155.702	186.383	468.472	274.000 ns	332.680 ns	456.000 ns	178.000 ns	230.490 ns	300.000 ns
000:03:0 -> 005:00:0, Mem TCO	2083	0.005	0.892	15.949	3.962 us	6.252 us	804.430 us	3.790 us	6.084 us	804.254 us
005:00:0 -> 000:00:0, Mem TCO	190	155.249	184.474	214.913	284.000 ns	329.720 ns	378.000 ns	168.000 ns	227.480 ns	274.000 ns
000:03:0 -> 006:00:0, Mem TCO	2078	0.005	0.873	16.188	3.962 us	6.371 us	805.082 us	3.786 us	6.203 us	804.902 us
006:00:0 -> 000:00:0, Mem TCO	541	61.527	1042.631	2765.319	288.000 ns	425.760 ns	526.000 ns	174.000 ns	248.020 ns	354.000 ns
000:03:0 -> 007:00:0, Mem TCO	2071	0.005	0.859	1.904	3.958 us	6.363 us	813.442 us	3.794 us	6.195 us	813.274 us
007:00:0 -> 000:00:0, Mem TCO	535	58.388	1051.490	2770.340	288.000 ns	428.030 ns	516.000 ns	178.000 ns	248.270 ns	338.000 ns
004:00:0 -> 000:31:6, Mem TCO	429	78.250	148.991	200.774	304.000 ns	416.230 ns	780.000 ns	192.000 ns	312.830 ns	676.000 ns
005:00:0 -> 000:31:6, Mem TCO	393	109.382	153.033	204.816	298.000 ns	402.730 ns	558.000 ns	194.000 ns	300.040 ns	454.000 ns
006:00:0 -> 000:31:6, Mem TCO	637	62.196	271.715	2601.922	302.000 ns	399.540 ns	614.000 ns	190.000 ns	283.240 ns	510.000 ns
007:00:0 -> 000:31:6, Mem TCO	872	44.703	285.667	2726.789	302.000 ns	395.230 ns	578.000 ns	198.000 ns	277.180 ns	474.000 ns
42813										



Timing Calculator Queue View

- View submission and completion Queues
- Compare Queues to see where overloading is occurring
- Verify if submission and completion queues are equal and that nothing was lost





NVMe Command IOPS Statistical Information

File Setup Record Generate Report Search View Tools Window Help

Trace View

Packet	R+	TLP	Mem	MW(r64)	Length	RequestedID	Tag	Address	1st BE	Last BE	Data	LCRC	Time Delta	Time Stamp
46763	x1	2575		011.00000	4	006.00.0	0	00000004.650D630	1111	1111	4 dwords	0x164A8F7C	13.849 us	0000.270.319.004 s

NVMe Cmd	OpC	SQID	CQID	CID	MPTR	PRP1	PRP2	SLBA	NLB	PRINFO	FUA	LR	DSM	ACCF	ACCL	SEQR	INCOM
15588	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006510	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0
15589	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006511	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0
15590	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006511	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0
15591	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006511	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0
15592	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006511	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0
15593	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006511	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0
15594	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006511	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0
15595	Read	0x000A	0x000A	0x0000	00000000.00000000	00000004.4188F000	00000000.00000000	00000000.00006511	0x0000	0x0	0	0	DSM	No frequency information provided	None	0	0

QuickTiming markers not set

Traffic Summary Report | Bus Utilization

NVMe Command IOPS

Ops per sec (Ops/s)

Time (us)

QuickTiming markers not set

Traffic Summary Report | Bus Utilization

NVMe Command IOPS

SPLIT: Pending Requests
 SPLIT: Response Time
 SPLIT: Latency Time
 SPLIT: Throughput Per Transaction
 SPLIT: Response Time & SPLIT: Latency Time
 Memory Writes: Response Time
 Memory Writes: Throughput
 Packet length
 Link Utilization
 Data Throughput
 Packet Count
 NVMe Command IOPS
 ATA ID: Pending Transactions
 Downstream Packet length
 Upstream Packet length
 FC: Upstream Posted Header & FC: Upstream Nonposted Header & FC: Upstream Completion Header
 FC: Upstream Posted Data & FC: Upstream Nonposted Data & FC: Upstream Completion Data
 FC: Downstream Posted Header & FC: Downstream Nonposted Header & FC: Downstream Completion Header
 FC: Downstream Posted Data & FC: Downstream Nonposted Data & FC: Downstream Completion Data

Ready

Search: Find



Conditional performance analysis using Verification scripting

- Extract metrics within a defined range
- LBA drive access pattern
- Queue access distribution
- Low power states entry / exist
- Multiple NVMe commands per TLP referenced by time stamps



VSE script example

```
OnStartScript()      {
  ReportText("OnStartScript called...");
  ReportText("\n\nRunning...\n");
  EventCount = 0;
  SendAllChannels();
  SendAllTraceEvents();
  #SendLevelOnly( _NVMC );
  #SendLevelOnly( _SPLIT );
  #ScriptForDisplayOnly();
  # Comment the line below - if you want to enable output from ReportText()-functions.
  DisableOutput();
  filePtr = OpenFile("C:\\Documents\\test.csv");
  WriteString(filePtr,"Start time, Response time, latencyTime, LBA, Length, QID, FUA"); }
```



VSE script example

Flash Memory Summit

```
ProcessEvent() {
    respTime= in.Metric_ResponseTime;
    latencyTime=in.Metric_LatencyTime;
    time=in.Time;
    throughput=in.Metric_Throughput;
    CMD = in.nvmcCommandOpCode;
    NLB = in.Read_NLB;
    SLBA0 = in.Read_SLBA_DW0;
    SLBA1 = in.Read_SLBA_DW1;
    SLBA = in.Read_SLBA;
    SQID = in.nvmcSubmissionQueueID;
    if( CMD!= 1 )    FUA = 0;           else    FUA = in.Write_FUA;
    if (SQID!= 0) { ReportText(FormatEx("%s,%s,%d,%s,%d,%d,%d",CSV_Val_TimeStamp_Seconds(in.Time ),
                                        CSV_Val_TimeStamp_Seconds(respTime), CMD, (SLBA), (NLB+1), SQID, FUA));
        WriteString(filePtr,FormatEx("%s,%s,%s,%s,%d,%d,%d",CSV_Val_TimeStamp_Seconds(in.Time ),
                                        CSV_Val_TimeStamp_Seconds(respTime),CSV_Val_TimeStamp_Seconds(latencyTime),
                                        (SLBA), (NLB+1), SQID, FUA)); }
    if( EventCount == MAX_NUMBER_OF_EVENTS ) ScriptPassed(); EventCount++; return Complete(); }
```

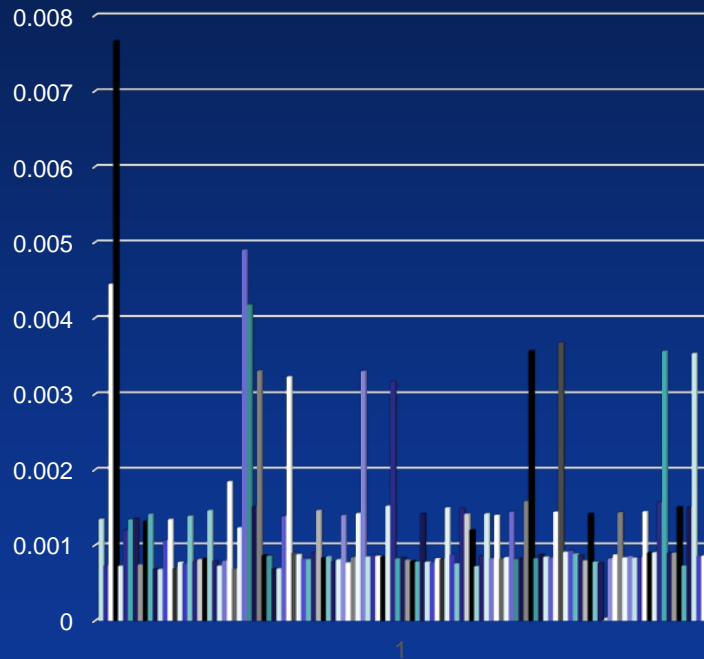
Santa Clara, CA
August 2019



SSD traffic statistics

Start time	Response time	Command	LBA	Length	QID	FUA
97.44338	0.000819348	2	0x8000000000000000	1	3	0
97.44427	0.00087518	2	0x8000000000000000	1	3	0
97.44522	0.001432268	2	0xC000000000000000	4	3	0
97.4476	0.0008339	2	0x8000000000000000	1	5	0
97.44904	0.00085046	2	0x8000000000000000	1	1	0
97.45029	0.000832564	2	0x8000000000000000	1	3	0
97.45125	0.000837508	2	0x8000000000000000	1	3	0
97.4522	0.00144462	2	0xC000000000000000	4	3	0
97.45464	0.000901324	2	0x8000000000000000	1	5	0
97.4576	0.000903012	2	0x8000000000000000	1	3	0
97.4586	0.001566428	2	0xC000000000000000	4	3	0
97.46027	0.003569364	2	0x8000000000000000	13	3	0
97.46389	0.00089634	2	0x7FE8070000000000	1	3	0
97.46483	0.00089858	2	0x80F4030000000000	1	3	0
97.46662	0.001507236	2	0x8000000000000000	1	5	0
97.47032	0.00072846	2	0x8000000000000000	1	1	0
97.47108	0.001501092	2	0xC000000000000000	4	3	0
97.47268	0.003540652	2	0x8000000000000000	13	3	0
97.47632	0.000855788	2	0x7FE8070000000000	1	3	0
97.47727	0.0008609	2	0x80F4030000000000	1	3	0
97.47907	0.000801716	2	0x8000000000000000	1	3	0
97.48059	0.000653244	2	0x8000000000000000	1	5	0
97.48137	0.000685132	2	0x8000000000000000	1	5	0
97.48209	0.001389612	2	0xC000000000000000	4	5	0
97.48421	0.00069338	2	0x8000000000000000	1	5	0
97.48543	0.000734724	2	0x8000000000000000	1	5	0

Response time





NVMe Analyzer mode summary

- Keep only NVMe transactions
 - Link layer and physical layer traffic truncated
 - PRP/SGL Data payload truncated
- NVMe transactions selectively included
- Long recordings made possible
- Enables utilization of advanced metrics
 - Trap long latency events