



SDC 

STORAGE DEVELOPER CONFERENCE

SNIA  SANTA CLARA, 2017

Repurposing Depopulation: Extending storage device service lifetimes

Repurposing Depopulation

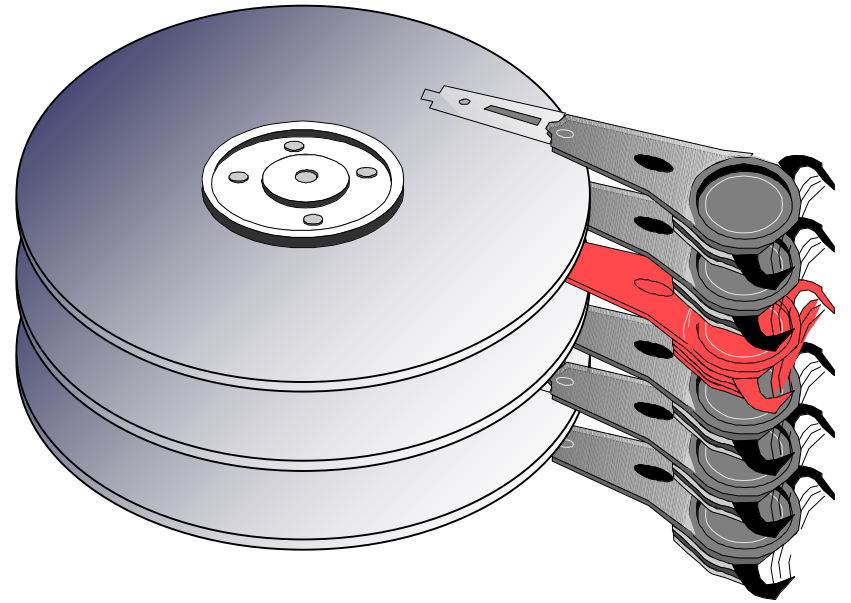
Extending storage device service lifetimes

Joe Breher
Technologist
Western Digital



Agenda

- ❑ Problem to be solved
- ❑ Solution building blocks
- ❑ Protocol flows
- ❑ Industry standard implementations
- ❑ Host use of feature
- ❑ Future developments



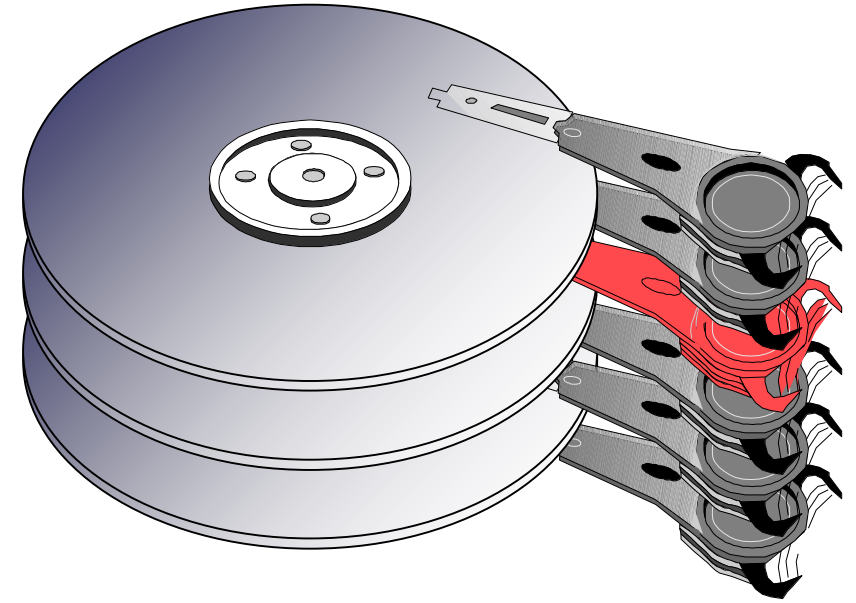
Foreword

- ❑ Broad participation
 - ❑ Standards Committees
 - ❑ T10 (SCSI)
 - ❑ T13 (ATA)
 - ❑ Co-proposers
 - ❑ Seagate
 - ❑ Toshiba
 - ❑ Western Digital
- ❑ Abstracted
 - ❑ Technology neutral
 - ❑ Speak here in terms of HDD
 - ❑ Applicable also to (e.g.) SSD
 - ❑ Standardized Implementation
 - ❑ Replacing earlier proprietary solution protocols



Problem to be solved

- ❑ Device reliability
 - ❑ Millions of hours of MTBF
 - ❑ Large yet limited
- ❑ Failure modes
 - ❑ Trending towards single element (head/disk) issues
 - ❑ Logical to physical mapping may not be knowable a priori (indirection)
- ❑ Central value proposition of depopulation
 - ❑ 'Offline' degraded element
 - ❑ *Other elements remain in service*



Solution building blocks

- ❑ Device self-monitoring
 - ❑ Background monitoring of each element's performance
 - ❑ Compared against threshold
- ❑ Semi-asynchronous signal to host
 - ❑ Upon threshold cross – IE on SCSI, DSN on ATA
- ❑ Host query of device – GET PHYSICAL ELEMENT STATUS
 - ❑ Device returns PHYSICAL ELEMENT HEALTH for each physical element
- ❑ Host commands device to remove element
 - ❑ REMOVE ELEMENT AND TRUNCATE command



Upon Depopulation command

- ❑ Specified physical element is removed (i.e. becomes ineligible for storing data)
 - ❑ Reduction of physical capacity
- ❑ Device logical capacity truncated to reflect loss of physical capacity
 - ❑ Last LBA reduced (e.g. READ CAPACITY returns smaller value)
- ❑ Device engages optional initialization
 - ❑ E.g. Some implementations may require reformat of logical sectors across physical sectors

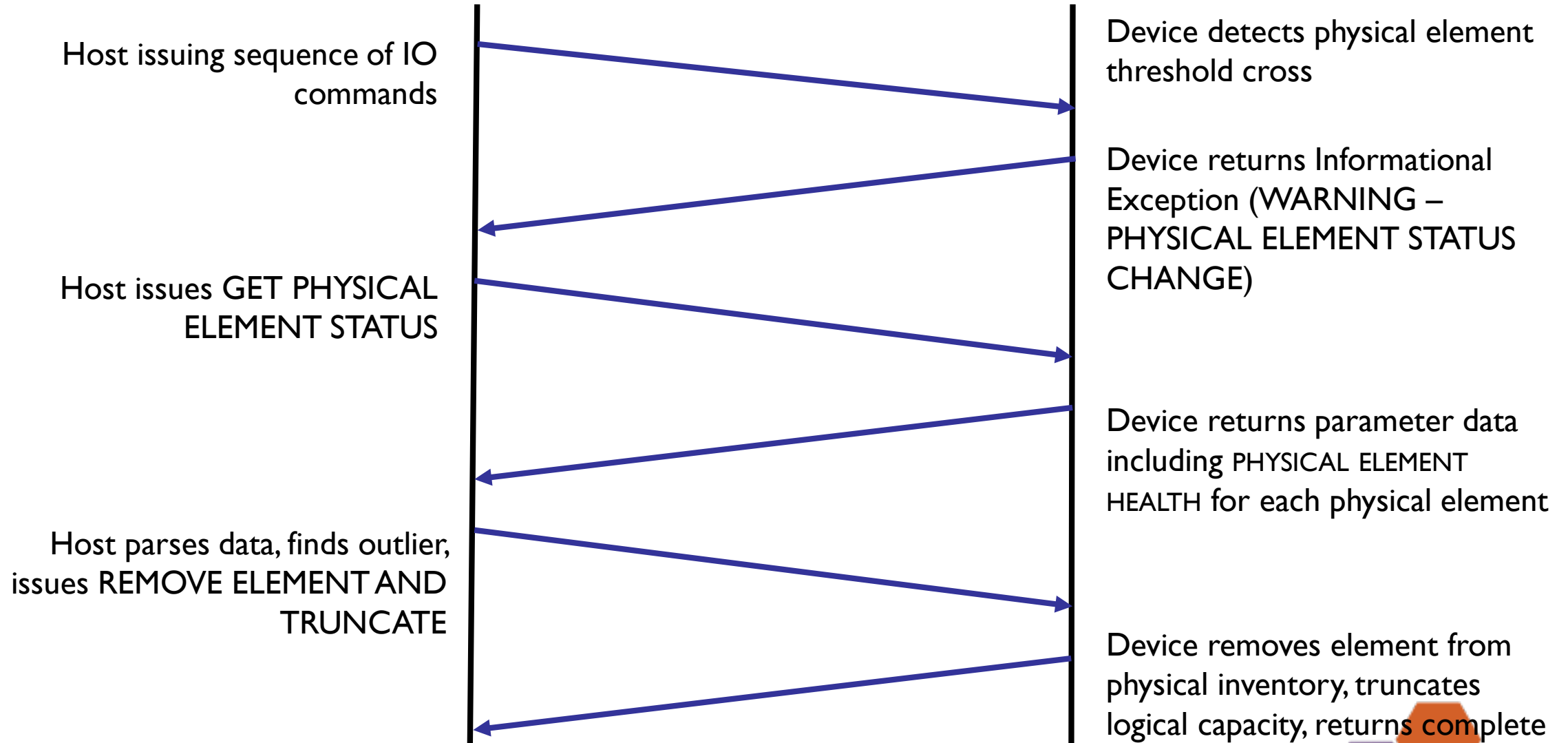


Result of Depopulation

- ❑ Device 'returns' as 'new', smaller, good device
- ❑ Reads return indeterminate data (until written)
 - ❑ Logical to physical mapping may be changed
 - ❑ No assumptions about any user data
 - ❑ If security consideration, host must 'clean'



Protocol flows



Command processing

- ❑ May execute as background processes
 - ❑ Depopulate
 - ❑ Truncate
 - ❑ Initialize (optional)
- ❑ Very limited set of commands allowable during processing
 - ❑ Most queries (INQUIRY, GET PHYSICAL ELEMENT STATUS, ...)
- ❑ Host can poll for completion
 - ❑ GET PHYSICAL ELEMENT STATUS header
 - ❑ Device advertises DEPOPULATION TIME (VPD, IDENTIFY DEVICE data log)



Industry standard implementations

- ❑ T10 has incorporated into SBC-4
 - ❑ SCSI Block Commands - 4
 - ❑ 2016 Nov – current is [sbc4r14](#)
- ❑ T13 has incorporated into ACS-4
 - ❑ ATA Command Set - 4
 - ❑ 2017 Feb – current is [di529r18](#)



SCSI Standard SBC-4

- ❑ Model: 4.38 Repurposing depopulation
 - ❑ Describes mechanism, notification, operations, effects on other commands
- ❑ 5.7 GET PHYSICAL ELEMENT STATUS
 - ❑ Query command description
- ❑ 5.25 REMOVE ELEMENT AND TRUNCATE
 - ❑ Operational command
- ❑ Annex J Using repurposing depopulation



GET PHYSICAL ELEMENT STATUS

Table 56 — GET PHYSICAL ELEMENT STATUS command

Byte	Bit	7	6	5	4	3	2	1	0
0		OPERATION CODE (9Eh)							
1		Reserved				SERVICE ACTION (17h)			
2		Reserved							
...									
5									
6	(MSB)	STARTING ELEMENT							
...									
9									
10	(MSB)	ALLOCATION LENGTH							
...									
13									
14		FILTER	Reserved			REPORT TYPE			
15		CONTROL							

STARTING ELEMENT and ALLOCATION LENGTH allow host to use sequence of commands to iterate across all elements

FILTER specifies returned list contain: all elements; or only interesting ones

REPORT TYPE specifies list of: all elements; or only storage elements



Physical element descriptor data

Table 59 — GET PHYSICAL ELEMENT STATUS parameter data

Byte	Bit	7	6	5	4	3	2	1	0
0	(MSB)	NUMBER OF DESCRIPTORS							
...									
3									
4	(MSB)	NUMBER OF DESCRIPTORS RETURNED							
...									
7									
8	(MSB)	IDENTIFIER OF ELEMENT BEING DEPOPULATED							
...									
11									
12		Reserved							
...									
31									
physical element status descriptor list									
32		physical element status descriptor [first] (see 5.7.2.2)							
...									
63									
		:							
n - 32		physical element status descriptor [last] (see 5.7.2.2)							
...									
n									

Table 60 — Physical Element Status descriptor format

Byte	Bit	7	6	5	4	3	2	1	0
0		Reserved							
...									
3									
4	(MSB)	ELEMENT IDENTIFIER							
...									
7									
8		Reserved							
...									
13									
14		PHYSICAL ELEMENT TYPE							
15									
16	(MSB)	ASSOCIATED CAPACITY							
...									
23									
24		Reserved							
...									
31									

01h to 63h: within manufacturer's spec limit
 64h: at manufacturer's spec limit
 65h to CFh: outside manufacturer's spec limit
 FDh: depopulation error
 FEh: depopulation in progress
 FFh: successfully depopulated



REMOVE ELEMENT AND TRUNCATE

ELEMENT IDENTIFIER
specifies element to be
removed

REQUESTED CAPACITY
specifies logical capacity
upon completion.
Zero specifies device
shall choose.
If unable to meet, device
aborts command.

Table 104 — REMOVE ELEMENT AND TRUNCATE command

Byte	Bit	7	6	5	4	3	2	1	0
0		OPERATION CODE (9Eh)							
1		Reserved				SERVICE ACTION (18h)			
2	(MSB)	REQUESTED CAPACITY							
...									
9									
10	(MSB)	ELEMENT IDENTIFIER							
...									
13									
14		Reserved							
15		CONTROL							

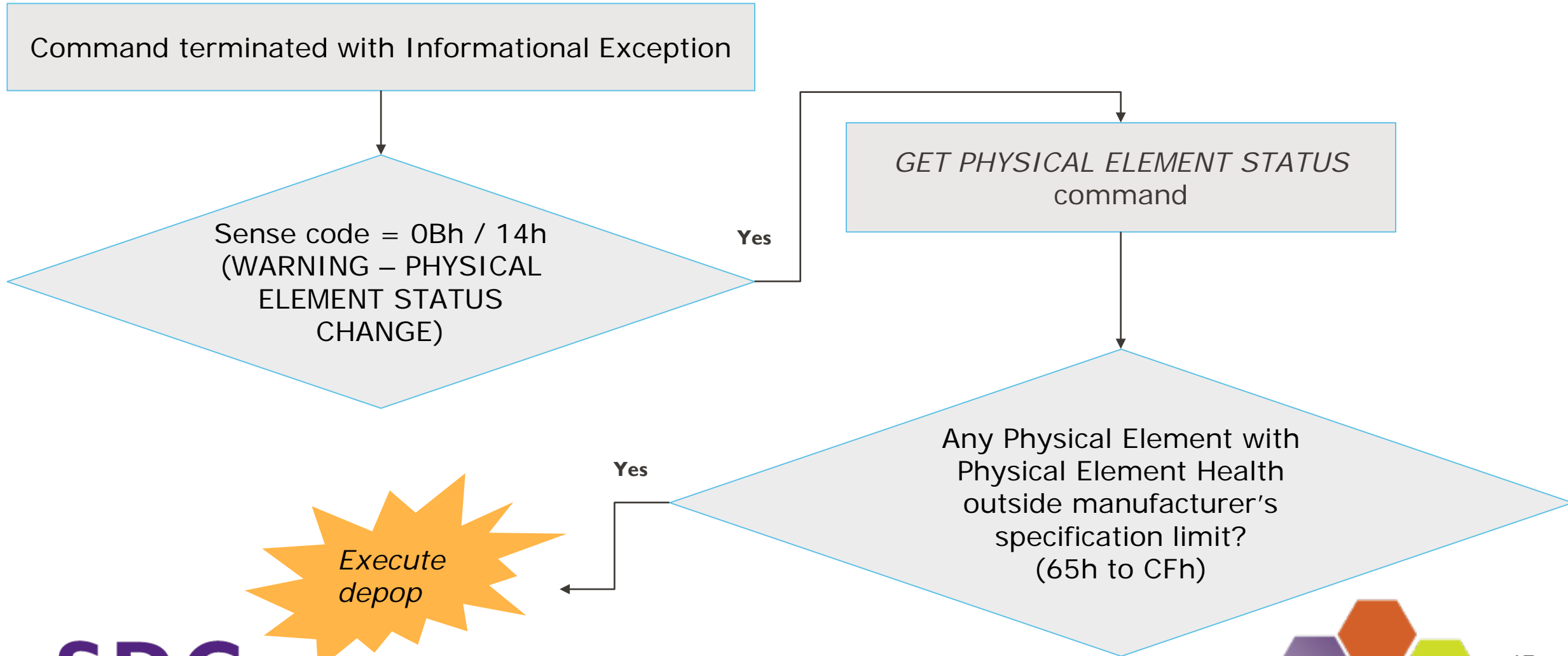


Host implementation 'howto' sketch

- ❑ Monitor for additional sense code of WARNING– PHYSICAL ELEMENT STATUS CHANGE
- ❑ Upon that ASC/Q, issue GET PHYSICAL ELEMENT STATUS
 - ❑ Parse for physical element health outside manufacturer's spec limit (65h to CFh)
 - ❑ Use case dependent: may save data not part of redundancy set
- ❑ Issue REMOVE ELEMENT AND TRUNCATE
- ❑ Upon completion, treat as new, somewhat smaller device



SCSI Host - Degraded head identification sequence



Follow on designs

- ❑ WD has also prototyped Data Preserving Depopulation
- ❑ Allows retention of user data on non-removed elements
 - ❑ Split Depopulation, Truncation into separate commands
 - ❑ Add Amputation, Regeneration, additional queries to manage LBA holes
- ❑ Proposals to T10 & T13 withdrawn
 - ❑ Complexity viewed as not warranted at this time
 - ❑ Will reintroduce if customer pull



Calls to action

- ❑ Implement in your host storage management stack
- ❑ Provide feedback on utility and design
 - ❑ T10.org
 - ❑ T13.org
 - ❑ joe.breher@wdc.com
- ❑ Desire Data Preserving variant? Let us know!



Backup material

□ Thank You!

Western Digital®



Learning Objectives

- ❑ Understand the need for Repurposing Depopulation and its main benefits
- ❑ Learn real-life applications of Repurposing Depopulation
- ❑ Understand the architecture and design of the Repurposing Depopulation feature
- ❑ Understand protocol flows as it as implemented in T10 (SCSI) and T13 (ATA)
- ❑ Learn how to apply the benefits of Repurposing Depopulation in your products

