STORAGE DEVELOPER CONFERENCE

SDC 23

BY Developers FOR Developers

What 10 Years of Drive Stats Data Can Tell Us

Andy Klein | Backblaze Inc.

Agenda

Andy Klein
8 years - Drive Stats Guy
25 years - Marketing
6 years - Sys Admin
8 years – Developer

- The data we've collected
- Hard drive failure rates
- SSDs: Failure rates and SMART stats
- A roundup on hard drive stats
- AI/ML and drive stats

SDC 23

The Hard Drive Environment

As of 6/30/2023

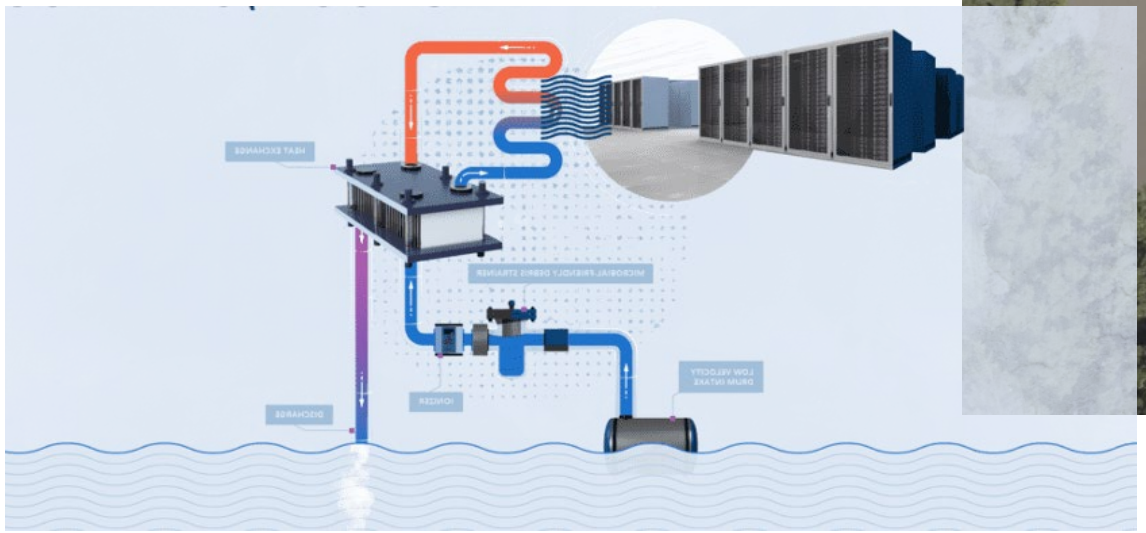
240,940
Active data drives
(Psst: 256,310 as of 9/1/2023)

3 Exabytes
Active data storage capacity

5 Data Centers
- Rancho Cordova, CA
- Phoenix, AZ
- Amsterdam, Netherlands
- Stockton, CA (Nautilus) *
- Reston, VA

The Nautilus Data Center

- Water cooled – No A/C
- Closed loop cooling
- 30% savings in electric
- No adverse effect on wildlife
- Located in a DHS protected site (Port of Stockton)



The Drive Stats Data

- Collected and stored since April 2013
  - Use smartmontools package to collect data
    - https://www.smartmontools.org/
  - All drives in operation each day to create a CSV file for each day
  - About 3,600 files, 400 million records, 17GB of ZIP data, 113GB of raw data
- The data is open source

Backblaze

www.backblaze.com/drivestats

SNIA IOTTA Repository

iotta.snia.org/traces/reliability

SDC 23

Drive Data Collected Each Day

Schema as of 3/31/2023

date	serial_number	model	capacity_bytes	failure	Smart_1_ normalized	Smart_1_ raw
3/15/23	Z305B2QN	ST4000DM000	4000787030016	0	98	2766
3/15/23	PL1331LAHG1S4H	HGST HMS5C4040ALE640	4000787030016	0	100	0
3/15/23	ZACH007	ST8000NM0055	8001563222016	1	81	139015
3/15/23	ZA130TTW	ST8000DM002	8001563222016	0	83	100901
3/15/23	ZA18CEBF	ST8000NM0055	8001563222016	0	81	140551
3/15/23	PL2331LAH3WYAJ	HGST HMS5C4040BLE640	4000787030016	0	100	0

More SMART stats ››››

SMART Stats:

- Collected using Smartmontools.
- There are 255 pairs of values per drive.

SMART Stat Attributes, e.g.

- Smart_1: Read Error Rate
- Smart_5: Reallocated Sector Count
- Smart_9: Power On Hours

en.wikipedia.org/wiki/S.M.A.R.T.

More Drives ⌄⌄⌄⌄

Drive Day: The data collected for one drive for one day.

SDC 23

More information is on the way

- New data fields in Q2: Vault ID and Pod ID
- Coming Soon: Data Center & Server Type
- Things we can learn
  - Failure Rates by Storage Server Type
  - Failure Rates by Data Center
  - Temperature by Storage Server Type
  - Temperature by Data Center

What's a Drive Failure

Reactive Failure

- The drive will not spin up or connect to the OS.
- The drive will not sync or stay synced in a storage array.

Proactive Failure

- Triggered by ATA, FSCK, etc.
- Combined with SMART data
- Reviewed by Backblaze before action is taken

Data Center:	Sac0
Pod:	pod-000-1113-01
Drive:	drive_0057
Tasks:	Replace Data Drive
Action:	Proactive
Reason:	High Offline Uncorrectable (SMART)
Brand:	HGST
Model:	HGST HUH721212ALN604
Serial:	8AJK007BH
Size:	12TB Drive
Notes:	5 Reallocated_Sector_Ct - 82 197 Current_Pending_Sector - 276 198 Offline_Uncorrectable - 266 199 UDMA_CRC_Error_Count - 0 9 Power_On_Hours - 23422 Found ATA error that is 2 hours old - CONSIDER REPLACING THIS DRIVE

Computing Annualized Failure Rate (AFR)

1. Define AFR cohort and period:
   a. Cohort = Drive Model: All models active as of 6/30/2023
   b. Period = Q2 2023
2. Obtain Drive Days and Drive Failures for the cohort and period.
   a. Drive Days = 21,408,175
   b. Drive Failures = 1,339
   c. Drive Count = 240,940
3. Apply Formula: AFR = (Drive Failures / (Drive Days / 365)) * 100

AFR = (1,339 / (21,408,175 / 365)) * 100 = **2.28%**

This method accounts for drives with different drive days within the period.

SDC 23

Drive Stats Reports

- Quarterly Reports
  - HDDs – report each quarter
  - SSDs – report twice a year
- Results for
  - Most recent quarter
  - Annual
  - Lifetime
- Find reports at
  - www.backblaze.com/blog
  - Search for drive stats

Backblaze Hard Drives Annualized Failure Rates for Q2 2023

Reporting period: 4/1/2023 thru 6/30/2023 for drives models active as of 6/30/2023

MFG	Model	Drive Size	Drive Count	Avg. Age (months)	Drive Days	Drive Failures	AFR
HGST	HMS5C4040ALE640	4TB	3,621	83.2	326,504	4	0.45%
HGST	HMS5C4040BLE640	4TB	11,934	80.1	1,083,231	22	0.74%
HGST	HUH728080ALE600	8TB	1,115	62	99,279	9	3.31%
HGST	**HUH728080ALE604**	**8TB**	**90**	**71.1**	**8,094**	**3**	**13.53%**
HGST	HUH721212ALE600	12TB	2,606	44.8	232,974	-	0.00%
HGST	HUH721212ALE604	12TB	13,203	27	1,181,748	42	1.30%
HGST	HUH721212ALN604	12TB	10,537	50.7	941,603	164	6.36%
Seagate	ST4000DM000	4TB	17,899	91.9	1,607,828	167	3.79%
Seagate	**ST6000DX000**	**6TB**	**883**	**98.3**	**80,411**	**3**	**1.36%**
Seagate	ST8000DM002	8TB	9,354	80.6	842,239	114	4.94%
Seagate	ST8000NM000A	8TB	153	11.6	12,088	-	0.00%
Seagate	ST8000NM0055	8TB	14,118	68.8	1,270,271	215	6.18%

... more models (31 this quarter)

Backblaze Hard Drives Lifetime Annualized Failure Rates

Reporting period: 4/20/2013 thru 6/30/2023 for drives models active as of 6/30/2023

MFG	Model	Drive Size	Drive Count	Drive Days	Drive Failures	AFR	Confidence Interval	
HGST	HMS5C4040BLE640	4TB	11,934	35,700,373	413	0.42%	0.40%	0.50%
HGST	**HUH728080ALE604**	**8TB**	**90**	**44,529**	**8**	**6.56%**	**2.80%**	**12.90%**
HGST	**HUH721212ALE604**	**12TB**	**13,203**	**10,730,974**	**172**	**0.59%**	**0.50%**	**0.70%**
Seagate	ST4000DM000	4TB	17,899	76,254,477	5,425	2.60%	2.60%	2.70%
Seagate	ST6000DX000	6TB	883	4,025,712	99	0.90%	0.70%	1.10%
Seagate	ST8000DM002	8TB	9,354	24,112,992	930	1.41%	1.30%	1.50%
Seagate	**ST8000NM000A**	**8TB**	**153**	**53,743**	**-**	**0.00%**	**0.00%**	**3.80%**
Seagate	ST12000NM001G	12TB	13,029	11,876,363	270	0.83%	0.80%	1.00%
Seagate	ST14000NM001G	14TB	10,790	9,351,776	312	1.22%	1.10%	1.40%
Seagate	ST16000NM001G	16TB	27,255	11,827,250	276	0.85%	0.80%	1.00%
Toshiba	MG07ACA14TA	14TB	38,101	37,163,688	987	0.97%	0.90%	1.00%
Toshiba	MG08ACA16TEY	16TB	5,289	2,910,620	36	0.45%	0.40%	0.80%
WDC	WUH721414ALE6L4	14TB	8,432	7,763,187	71	0.33%	0.30%	0.40%
WDC	WUH721816ALE6L4	16TB	14,099	3,699,632	32	0.32%	0.20%	0.50%

SDC 23

Lifetime Gold Medal Winners as of Q2 2023

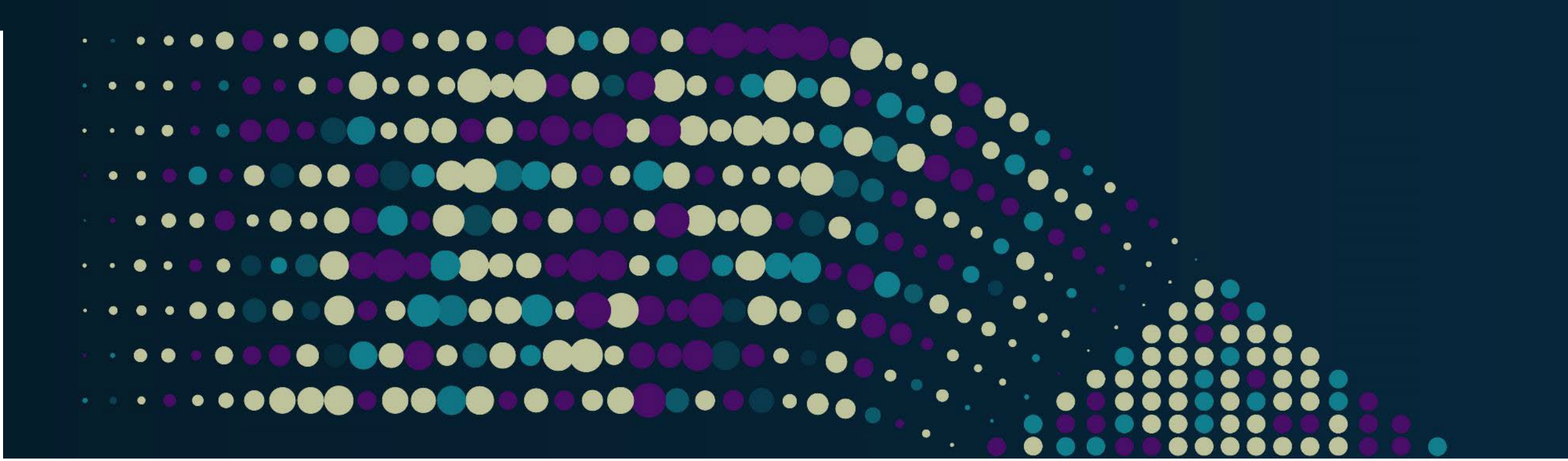
MFG	Model	Drive Size	AFR	Confidence Interval	
WDC	WUH721816ALE6L0 *(1)*	16TB	0.13%	0.10%	0.30%
WDC	WUH721816ALE6L4	16TB	0.32%	0.20%	0.50%
Toshiba	MG08ACA16TEY	16TB	0.45%	0.40%	0.80%
WDC	WUH721414ALE6L4	14TB	0.33%	0.30%	0.40%
Toshiba	MG07ACA14TA	14TB	0.97%	0.90%	1.00%
HGST	HUH721212ALE600 *(2)*	12TB	0.33%	0.20%	0.50%
HGST	HUH721212ALE604 *(2)*	12TB	0.46%	0.40%	0.60%
Seagate	ST12000NM001G	12TB	0.83%	0.80%	1.00%
HGST	HUH728080ALE600 *(3)*	8TB	0.94%	0.70%	1.20%
Seagate	ST6000DX000 *(3)*	6TB	0.90%	0.70%	1.10%
HGST	HMS5C4040BLE640 *(3)*	4TB	0.42%	0.40%	0.50%

(1) - Not available for retail sale in the US and Canada
(2) – Rebranded as Western Digital
(3) – Available as rehab drives only

Criteria

> 2M Drive Days
<= 1.0% AFR
>= 0.5 Confidence Interval

SDC 23



Let’s Talk About SSDs

SSDs versus HDDs

The Tale of the Tape

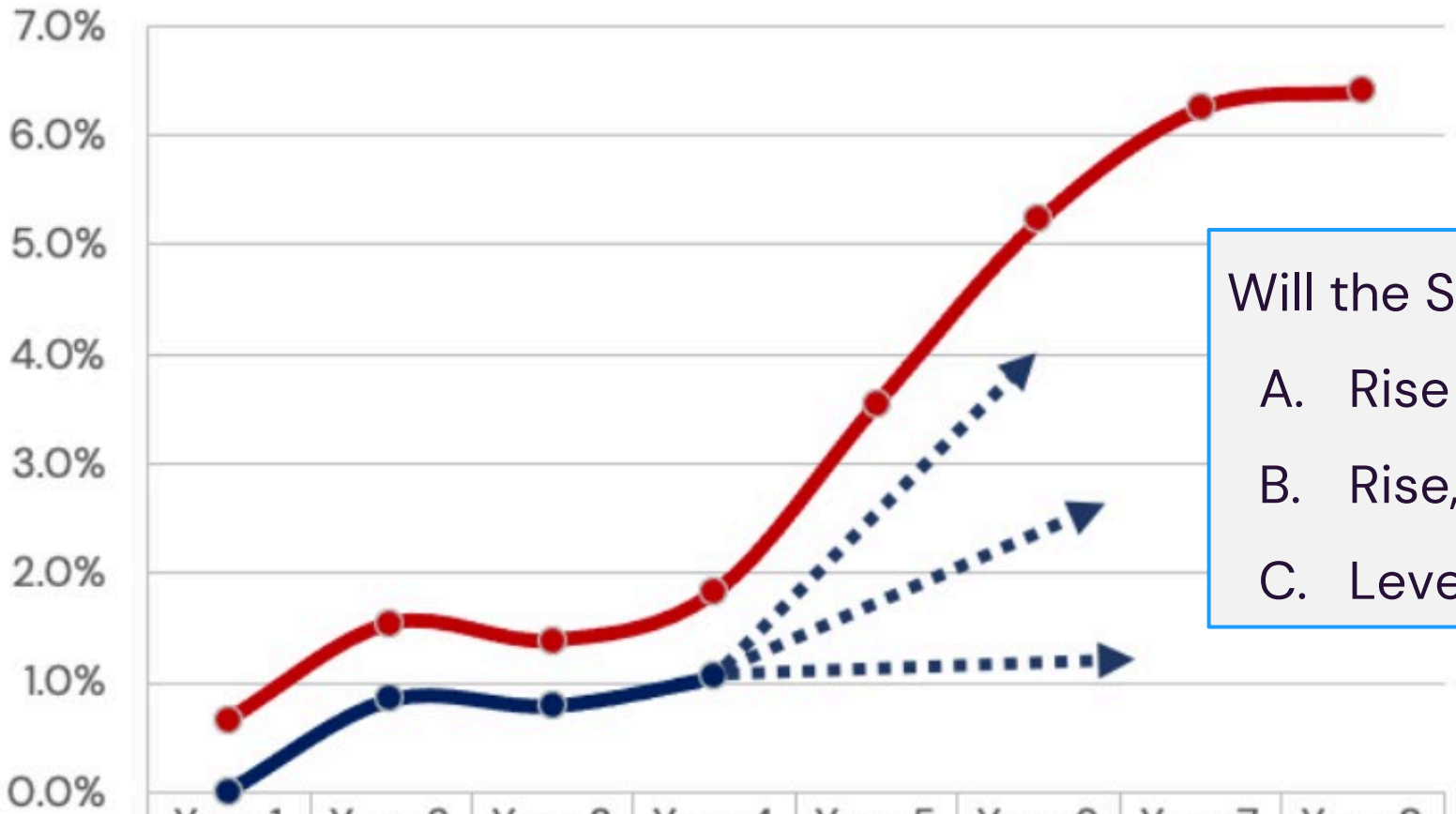
- Speed – SSD
- Electricity – SSD
- Cost/GB – HDD
- **Reliability – ???**

SSD versus HDD Reliability

Boot Drives

- Boot the server
- Regular daily activity to read/write/delete log files for system access and diagnostics.

Life SSD versus HDD Annualized Failure Rates thru 2021



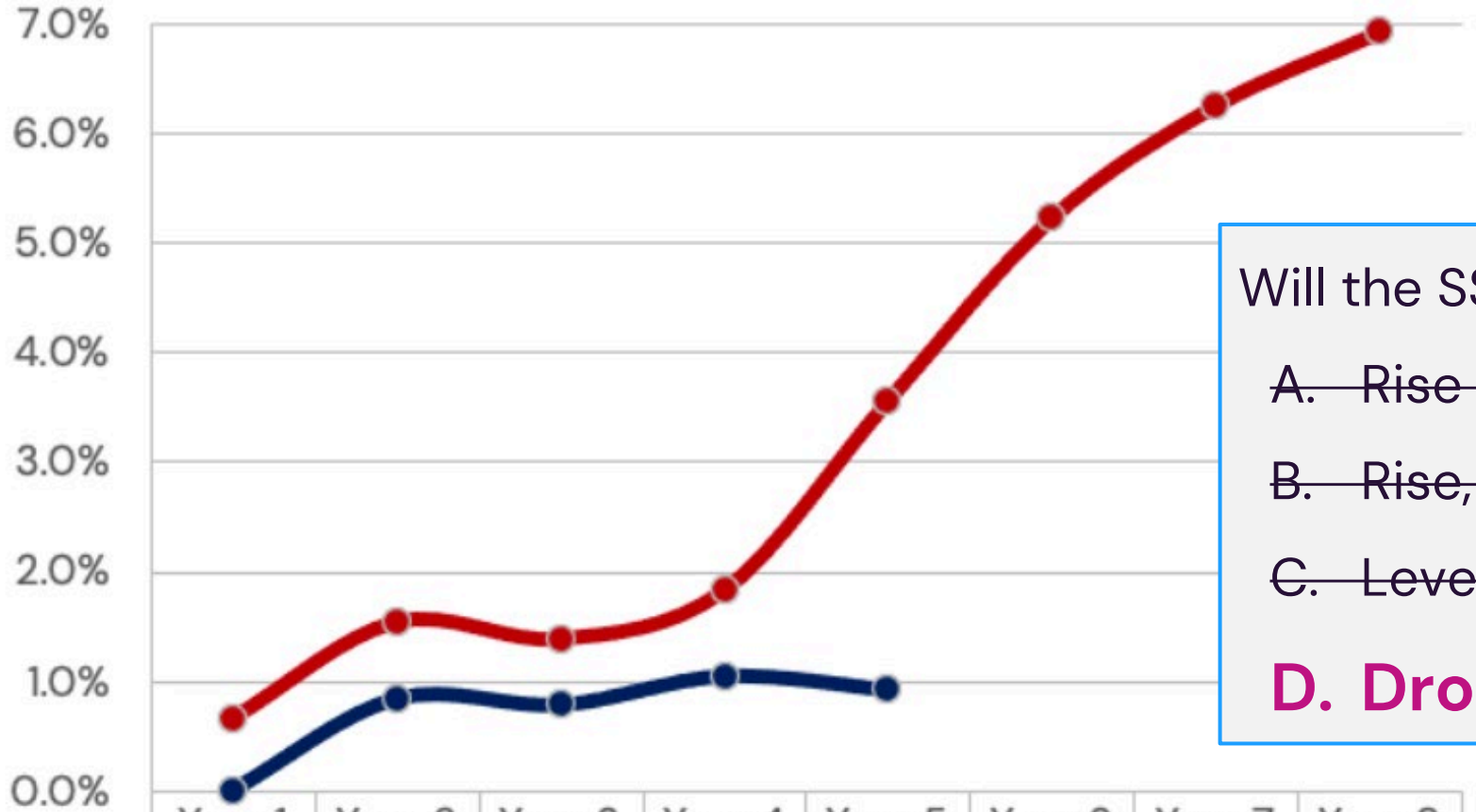
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
HDD Lifetime AFR	0.66%	1.54%	1.38%	1.83%	3.55%	5.23%	6.26%	6.41%
SSD Lifetime AFR	0.00%	0.84%	0.79%	1.05%				

HDD Lifetime AFR SSD Lifetime AFR

Will the SSD Failure Rate?

A. Rise like the HDDs
B. Rise, but slower
C. Level out

Life SSD versus HDD Annualized Failure Rates thru 2022



	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8
HDD Lifetime AFR	0.66%	1.54%	1.38%	1.83%	3.55%	5.23%	6.26%	6.93%
SSD Lifetime AFR	0.00%	0.84%	0.79%	1.05%	0.92%			

HDD Lifetime AFR SSD Lifetime AFR

Will the SSD Failure Rate?

~~A. Rise like the HDDs~~

~~B. Rise, but slower~~

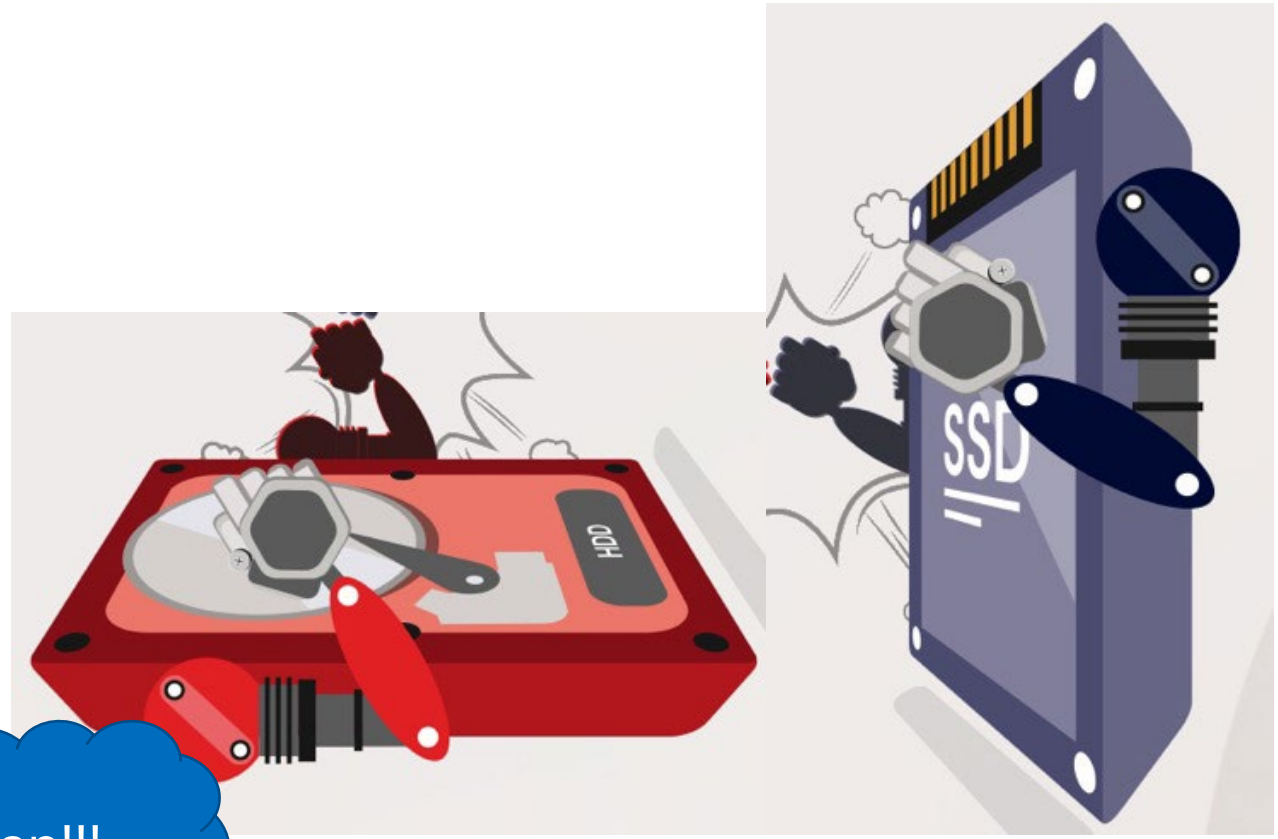
~~C. Level out~~

D. Drop

SSD versus HDD Reliability – Year 5

HDD Annualized Failure Rate (Lifetime)

3.55%

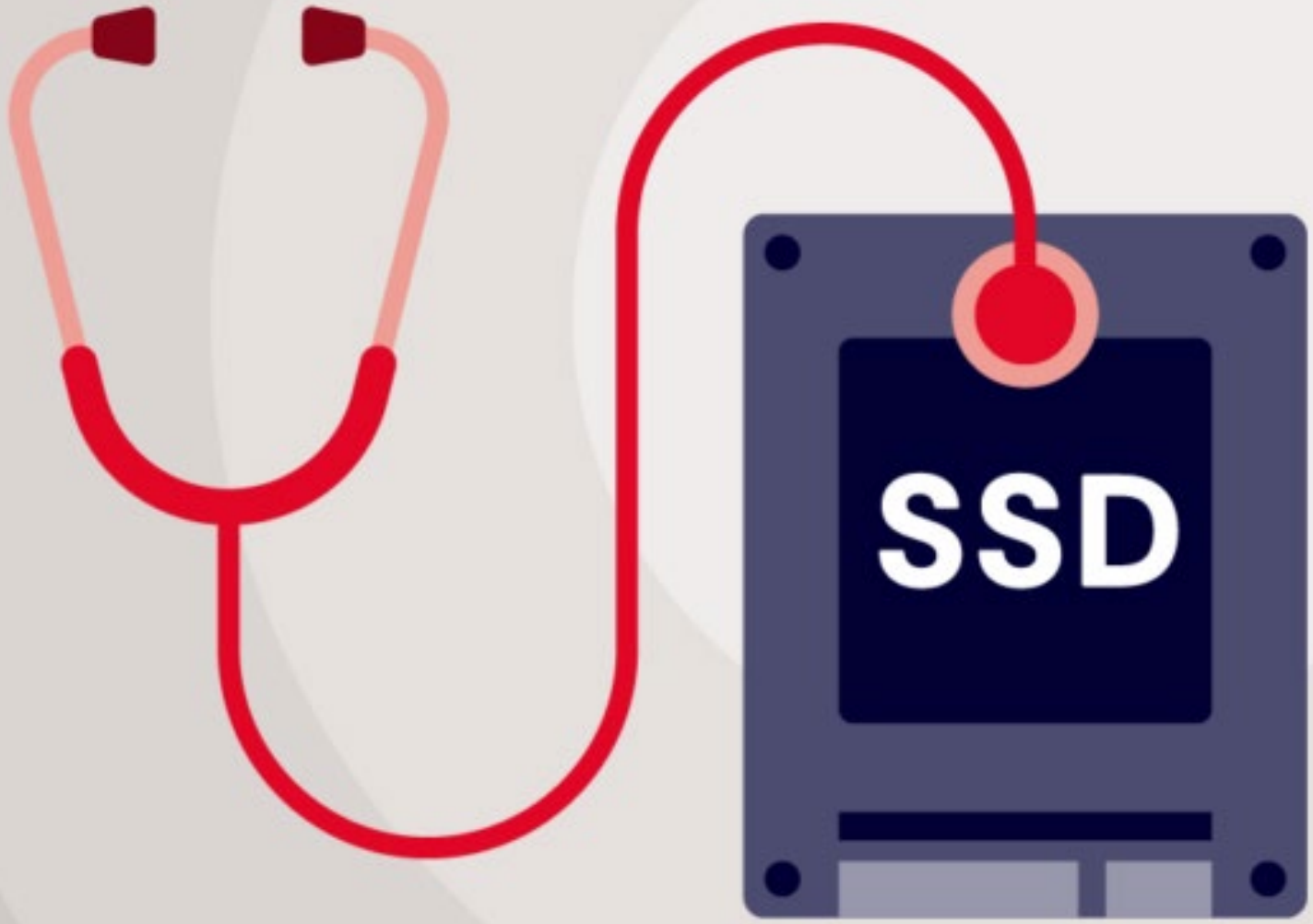


SSD Annualized Failure Rate (Lifetime)

0.92%

How SMART are SSDs?

A look at SSD SMART Stats



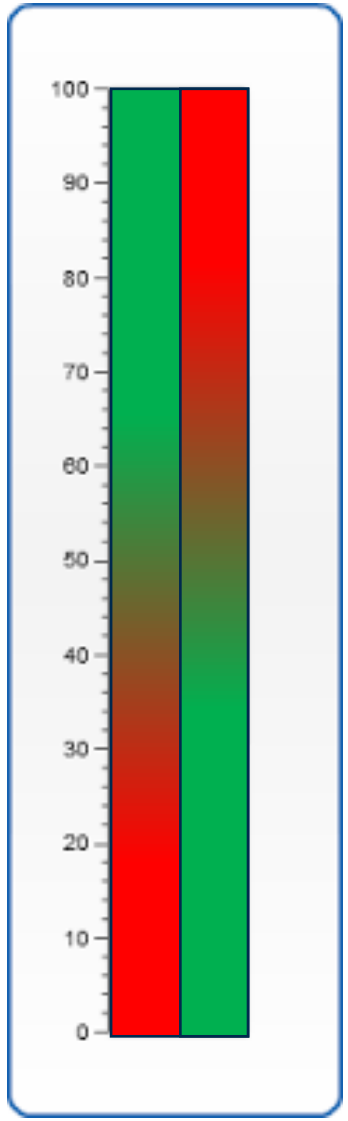
Making Sense of SSD SMART Stats
https://www.backblaze.com/blog/making-sense-of-ssd-smart-stats/

Let's Get SMARTed

- Selected three SSDs from our collection – all are 250GB models
  - Seagate BarraCuda 120 SSD ZA250CM10003 – 20 SMART attribute pairs
  - WDC WDS250G2B0A – 25 SMART attribute pairs
  - Crucial CT250MX500SSD1 – 23 SMART attribute pair
- Common Attributes for all three models
  - SMART 9: Power-On Hours
  - SMART 12: Power Cycle Count
  - SMART 194: Temperature
  - SMART 173: SSD Wear Leveling
  - SMART 174: Unexpected Power Loss Count

Lifetime Percentage

- SMART 169: Remaining Lifetime Percentage (WDC)
  - The approximate life left from a combination of program-erase cycles and available reserve blocks of the device. A new SSD starts at “100” for the Normalized value and decrease to “0”.
- SMART 231: Life Left (Seagate)
  - The approximate life left from a combination of program-erase cycles and available reserve blocks of the device. A new SSD starts at “100” for the Normalized value and decreases with a threshold value of “10” for replacement. A “0” value means the drive is in read-only mode.
- SMART 202: Percentage of Lifetime Used (Crucial)
  - How much of the drive’s projected lifetime has been used at any point in time. A new SSD will report “0” for the Normalized value and rise to “100,” as it being used.



Interesting SSD Attributes

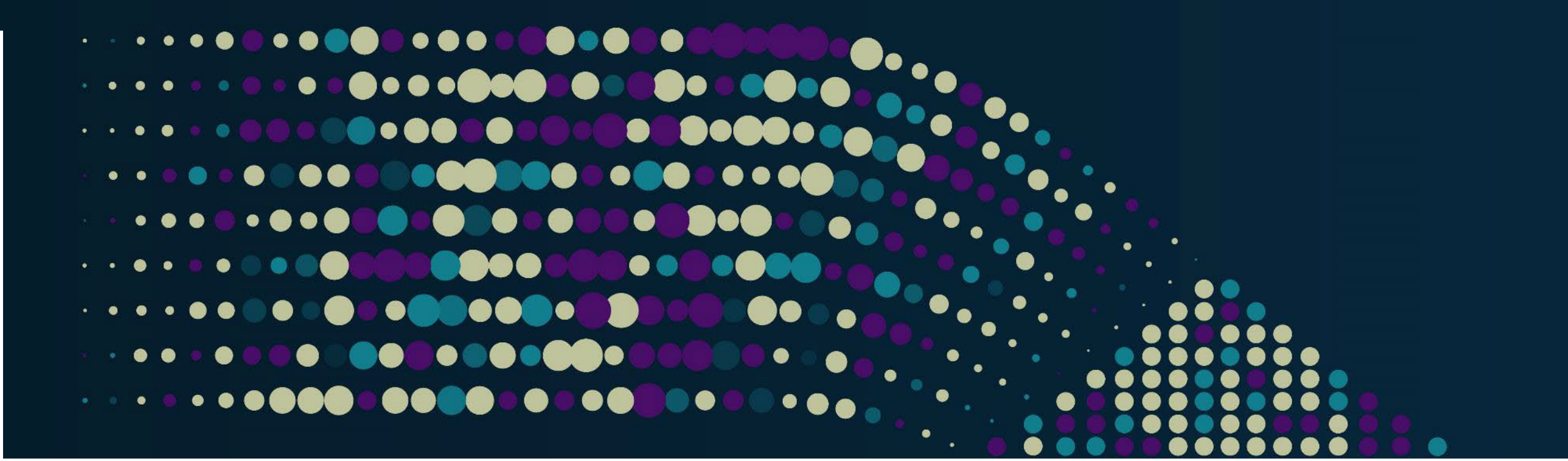
- SMART 230: Drive Life Protection Status (WDC)
  - A comparison of the SSD's usage trajectory versus the expected life curve
    - Yes, there is a usage trajectory calculation and value (SMART 169)
    - Yes, there is a defined expected life curve
- SMART 232: Endurance Remaining (Seagate and WDC)
  - Percentage of physical erase cycles completed versus the maximum number the drive is designed to endure
  - Does not consider available reserved blocks but can be useful given that erasing SSD blocks at an accelerated rate often leads to having to utilize available reserved blocks later
- SMART 210: RAIN Successful Recovery Page Count (Crucial)
  - Redundant Array of Independent NAND (RAIN) is similar data redundancy using RAID
  - RAIN redundancy is accomplished on the SSD itself and is user transparent

How Long Can You Store an SSD Without Power?

- Current thinking
  - About a year for new drives, decreasing from there as drives age
- Using SMART Attributes to compute storage time

Crucial (a) SMART 202 Normalized	Seagate (b) SMART 231 Normalized	WDC (b) SMART 169 Normalized	Maximum Storage Period
0	100	100	1 year
50	50	50	6 months
100	10	0	1 month

(a) – Documented by Crucial
(b) – No documentation found



What About Hard Drives

Some random bits about what we know about hard drives

SDC 23

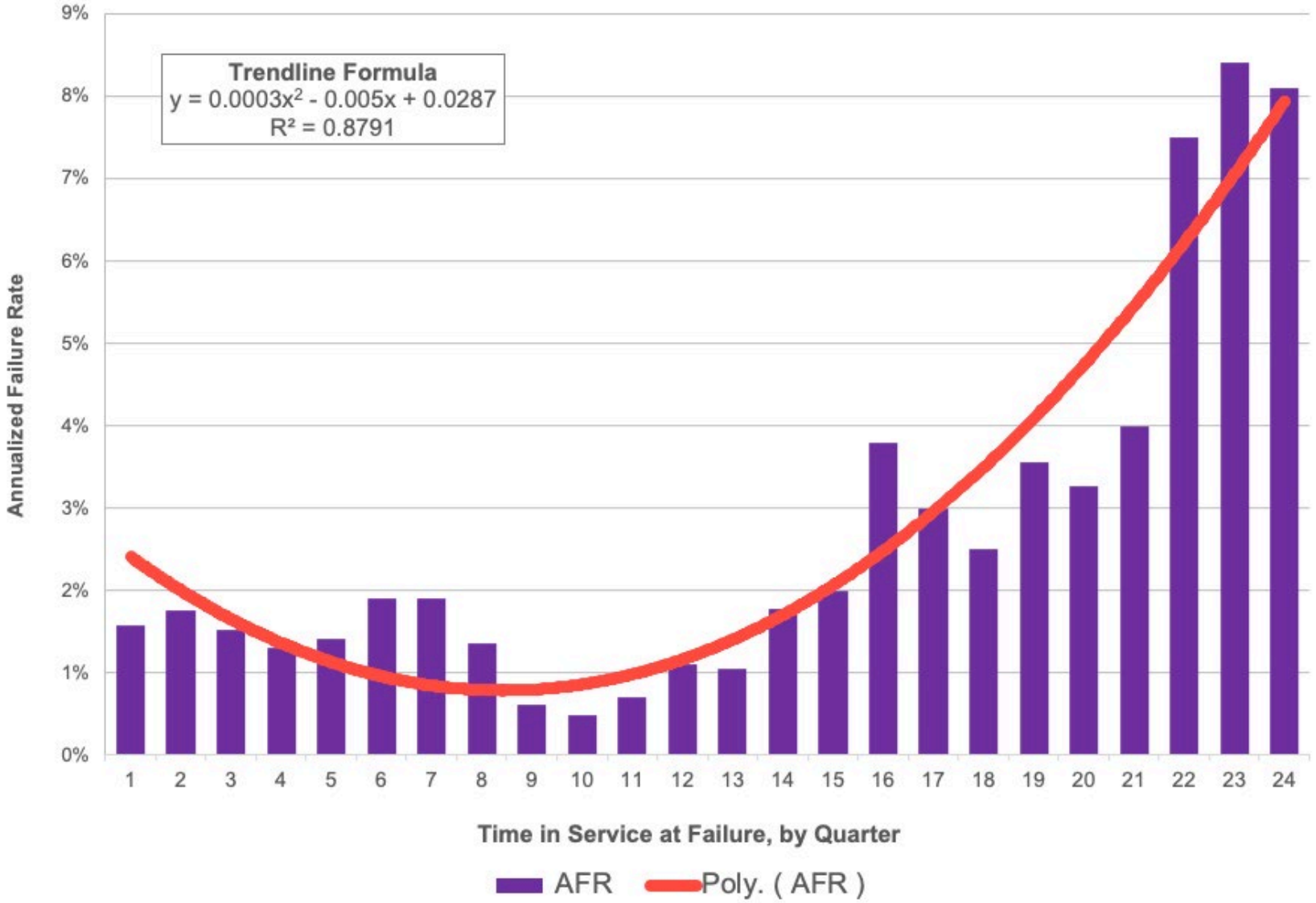
Average Age of Drive Failure

SOURCE	FAILED DRIVE COUNT	AVERAGE FAILED AGE
Secure Data Recovery	2,007	2 years 10 months
Backblaze	17,155 (all models)	2 years 6 months
Backblaze	3,379 (models no longer in production)	2 years 7 months

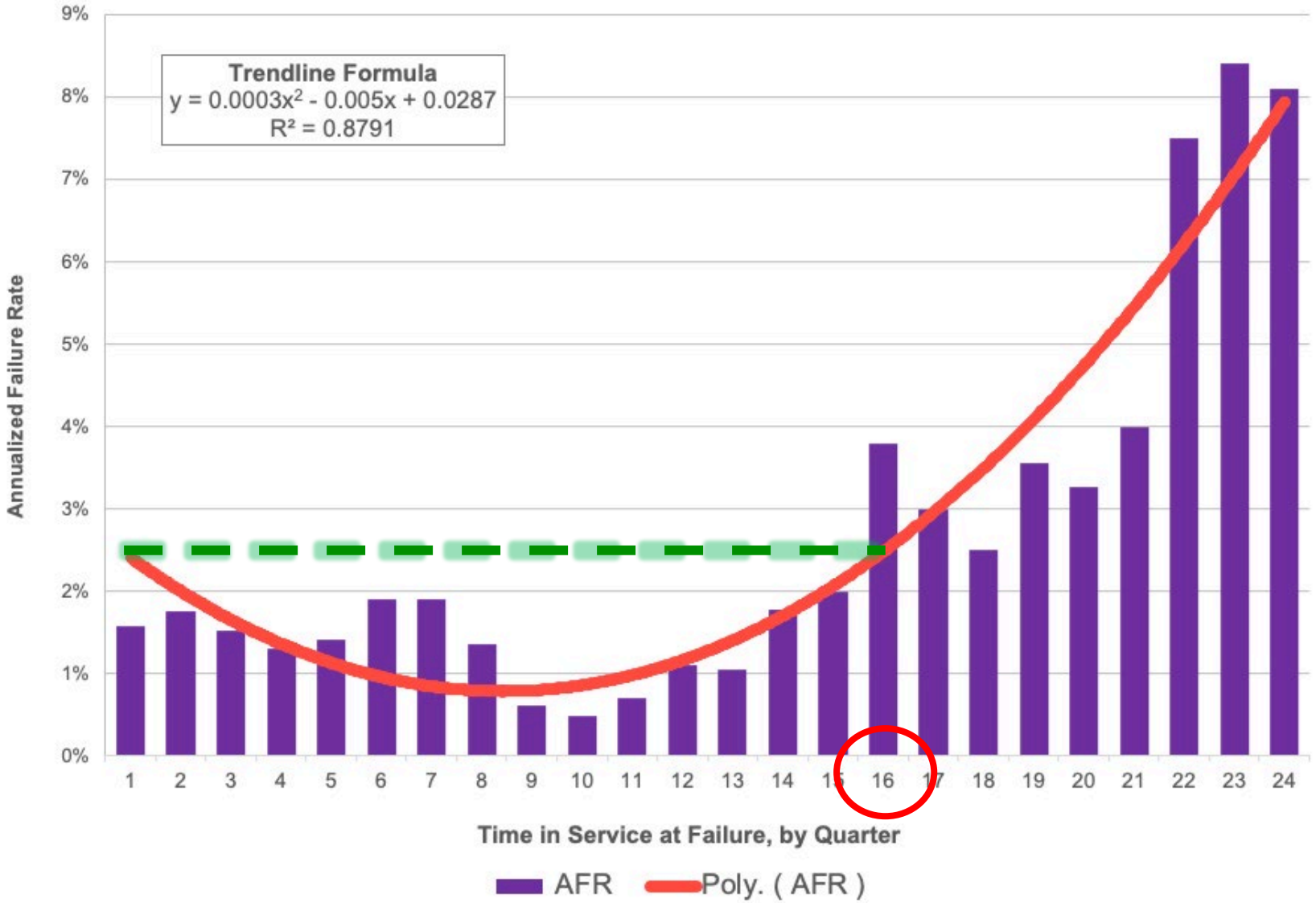
Average Age of Drive Failure by Drive Size

Drive Size	Failure Count	Average Age of Failure		Drives Remaining	Avg Age of Remaining	
		Years	Months		Years	Months
1 TB	76	4	7	0		
1.5 TB	704	3	9	0		
2 TB	309	3	6	0		
3 TB	2,168	2	1	0		
4 TB	6,098	3	2	34,614	7	1
5 TB	2	0	8	0		
6 TB	168	2	1	884	8	1
8 TB	2,160	3	5	25,163	5	11
10 TB	120	3	10	1,159	5	4
12 TB	3,579	1	8	60,240	2	11
14 TB	1,368	1	7	59,402	2	4
16 TB	399	0	10	55,431	1	2

Drive Failure Over Time: 2022



Drive Failure Over Time: 2022



Predicting Hard Drive Failure

2016

Paper: Predicting Disk Replacement towards Reliable Data Centers

Authors: Botezatu, Mirela & Giurgiu, Ioana & Bogojeska, Jasmina & Wiesmann, Dorothea. (2016).

Location: *https://dl.acm.org/doi/10.1145/2939672.2939699*

...several others...

2021

Paper: Interpretable Predictive Maintenance for Hard Drives

Authors: Maxime Amram, Jack Dunn, Jeremy J. Toledano, Ying Daisy Zhuo

Location: *https://www.sciencedirect.com/science/article/pii/S2666827021000219*

Using Machine Learning to Predict Hard Drive Failure

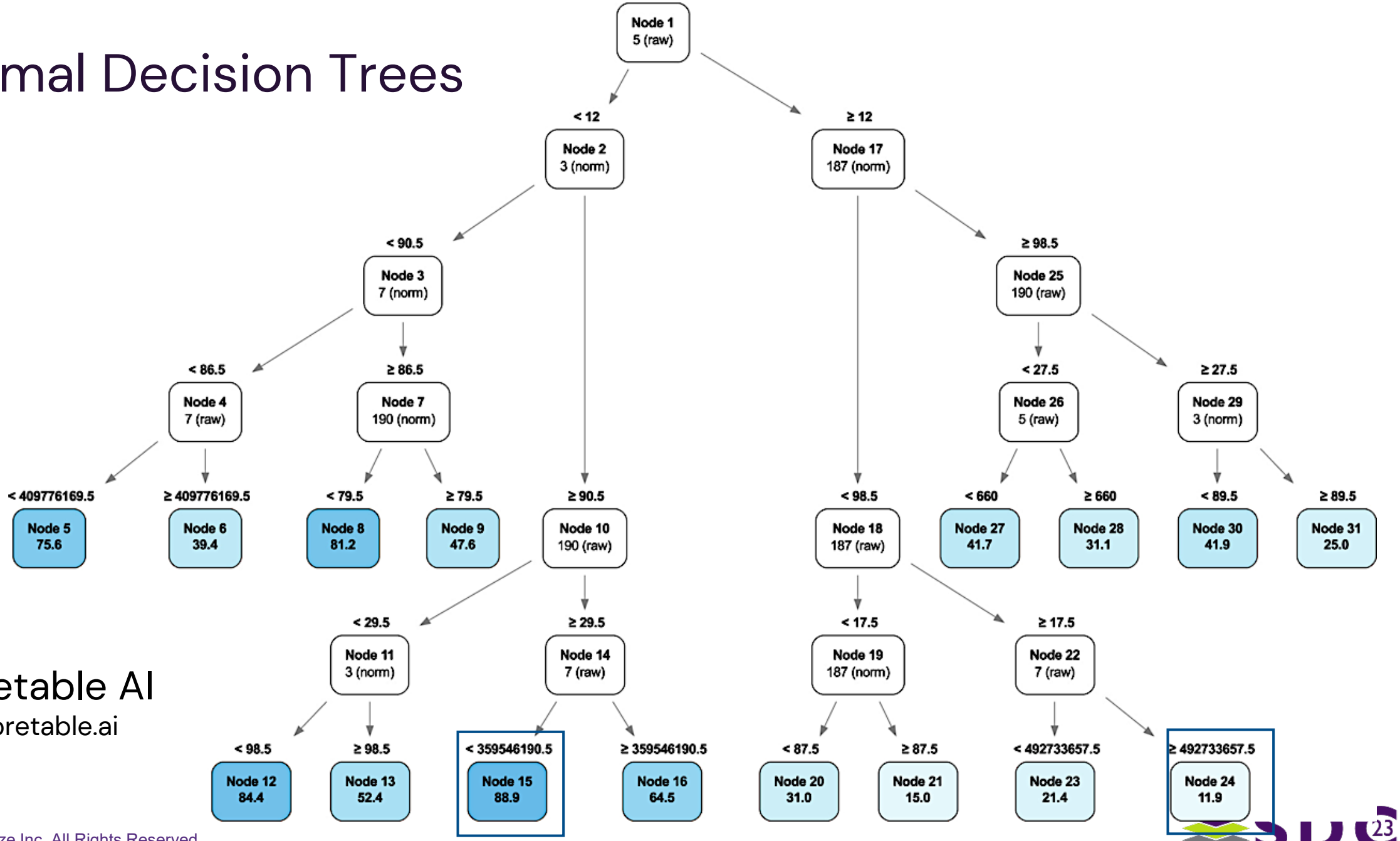
Backblaze Drive Stats Data
+ Optimal Decision Trees
+ Survival Curves

= Predictions on long-term & short-term drive health

Drive Stats Data: https://www.backblaze.com/b2/hard-drive-test-data.html

SDC 23

Optimal Decision Trees



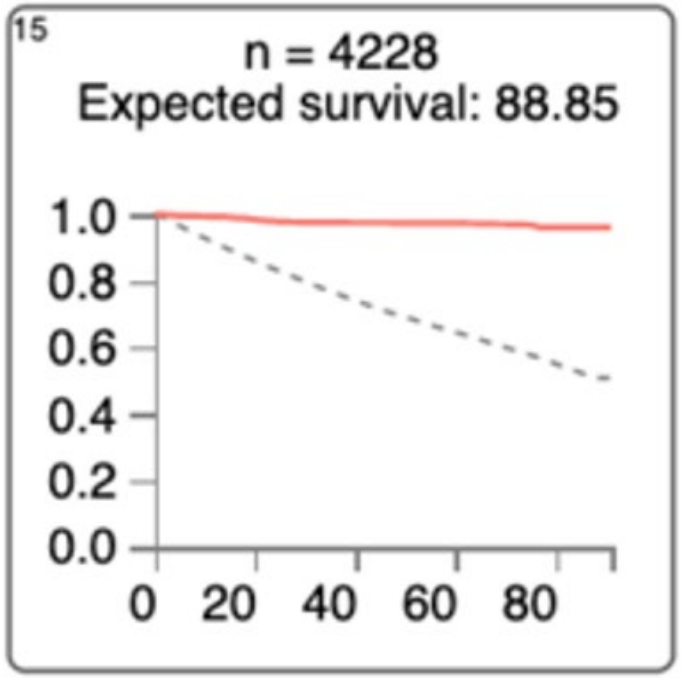
Interpretable AI
www.interpretable.ai

Predicting Short-term Drive Health

Example survival curves for selected cohorts

Healthy Drives

Node 15



5 (raw) < 12

3 (norm) < 90.5

10 (raw) >= 29.5

7 (raw) >= 359546190.5

Interpretable AI

www.interpretable.ai

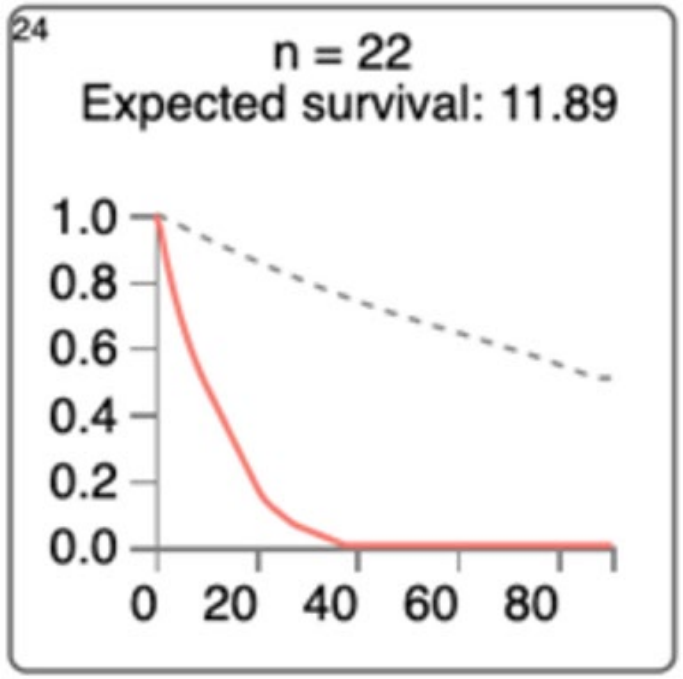
SDC 23

Predicting Short-term Drive Health

Example survival curves for selected cohorts

Unhealthy Drives

Node 24



5 (raw) >= 12

187 (norm) < 98.5

187 (raw) >= 17.5

7 (raw) >= 492733657.5

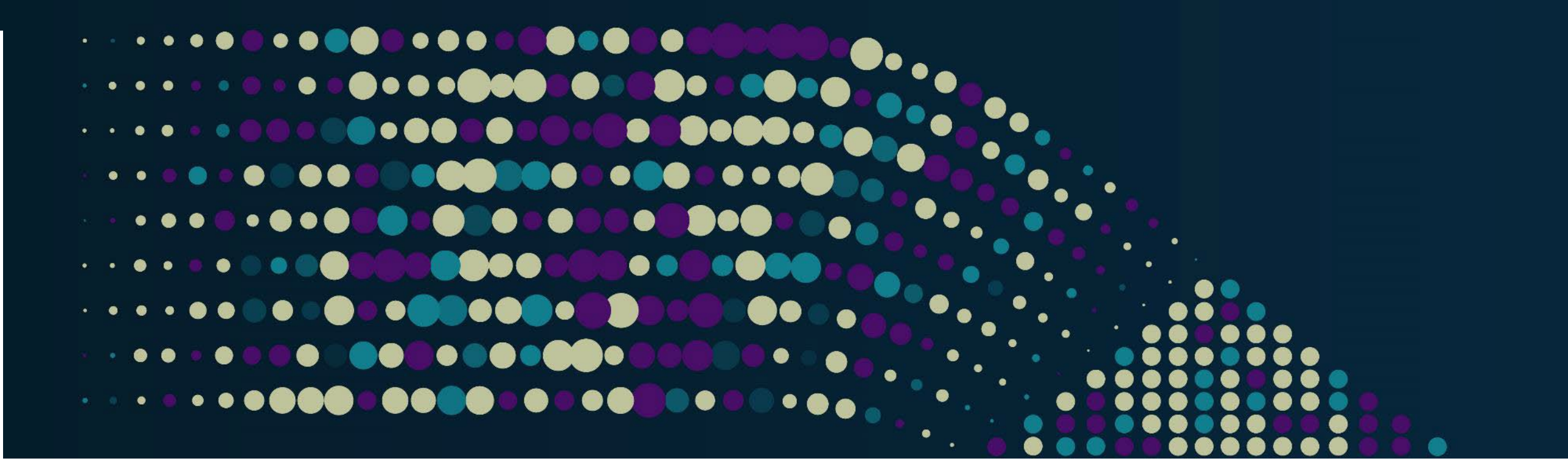
Interpretable AI

www.interpretable.ai

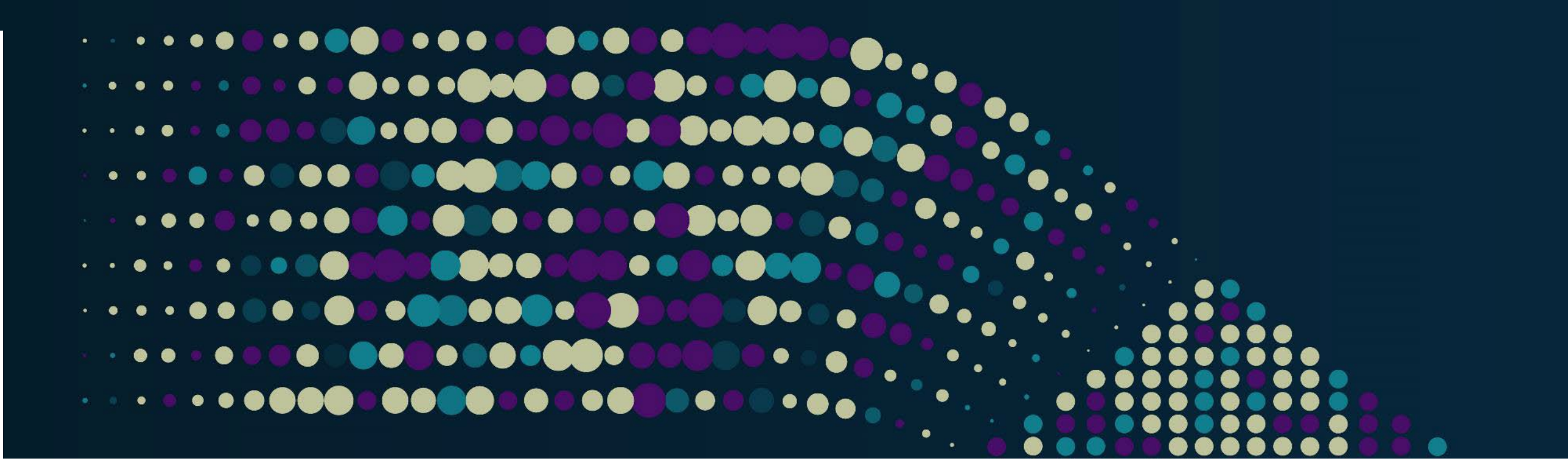
Summary

- The data we've collected
- Hard drive failure rates
- SSDs: Failure rates and SMART stats
- A roundup on hard drive stats
- AI/ML and drive stats

SDC 23



Questions



Please take a moment to rate this session.

Your feedback is important to us.

