STORAGE DEVELOPER CONFERENCE
SDC 21
BY Developers FOR Developers
Virtual Conference
September 28-29, 2021

# Drive Health Monitor (DHM) for HDD drives On-Prem (or core data center) and Cloud

DHM

Presented by

Mahmoud K. Jibbe, PhD , Technical Director NetApp

Charles Binford, Software Architect NetApp

# Agenda

- Motivation of DHM
- Goal of DHM
- DHM structure
- DHM Flow chart
- DHM Two sets of data, One Rolled Up Score
- Proposed Overall Threshold Values
- Days on List Charts
  - User Failed
  - Still Alive
- Drive Scoring Tables
- Conclusions and Takeaways

# Motivation of DHM

- Motivation
  - Customers experienced a data loss event and saw high drive failure rate >5% AFR
  - The high utilization and the age of the drives put system at risk of increasing the drive fallout rate. Leading to:
    - Outages: Loss of access
    - Possible data loss (Many recent examples of data risk/loss)
  - While most existing HDD systems are 5+ years old, future platforms are being designed with HDDs

# Goal of DHM

- Goal:  Find ways to minimize impact on drive fallout/failure
  - Predict failures and schedule proactive removal of bad drives to reduce risk of impacting customer data or operation
    - Outages: Loss of access
    - Performance impact
    - Possible data loss

# DHM structure

## Drive Monitor Approaches

1. **Drive Stats**: Drive's internal error handling (Log Sense)
2. **Errors observed by controller:** Reported / Detected Errors
3. Adjust some existing **SPFA thresholds** in Controller NVSRAM
   - SPFA = Synthetic Predictive Failure

## Drive Monitor Parameters

- **From "Read Error Counters" log page**
  - Recovered and unrecovered Read errors
- **From Read Defect Data**
  - Grown defect list from drive reassigned blocks
- **Computed metric**
  - Corrected Read Errors / TB_Read

## Drive Monitor Timeframes Considered

- **Lifetime values**
  - i.e. raw value read from drive
- **Long duration**
  - delta values between current and ~42 days prior (1000 hrs)
- **Short duration**
  - delta values over last 24 hours
- Apply **weight factor** to each drive on watch list to compute a score

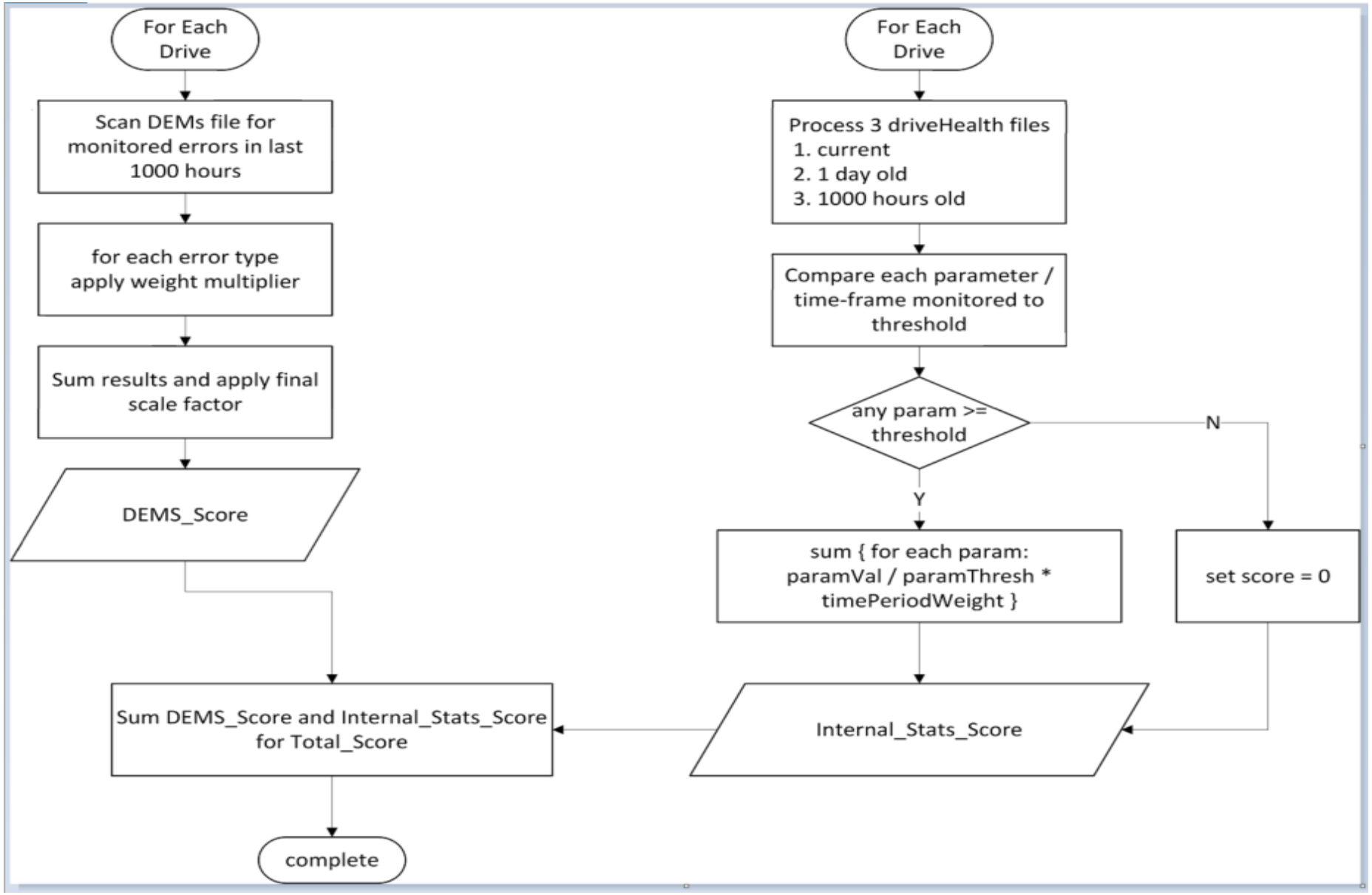
# DHM Data

- The monitor scripts look at two distinct sets of data.
  1. Data from the drive's log sense pages – *Internal Stats*
  2. Data from errors observed by the controller – *External Stats* (a.k.a. DEMS data )
- Each set of data considers multiple parameters that are combined using a weighted system to determine a score. Example:
  1. Aborted command counts higher than a Recovered Error.
- The two scores from the different data sets are added for a final score.

Note: DEMS (Drive Error Monitor System) – A file in our controller logs that contains history of each drive's error events.

# DHM Flow Chart



# Scoring:

| | Parameter | Count Multiplier Drive Type | Time Period | | Scoring Pseudo Code |
|---|---|---|---|---|---|
| **DEMS Score** | DRP | a | 1000 hours | | `For each parameter:`<br>`  demsScore += occurCnt * countMultiplier`<br><br>`demsScore = demsScore / 50` |
| | MED | b | | | |
| | ABT | c | | | |
| | REC | d | | | |
| | NCF | e | | | |
| | FAILED | f | | | |
| **Internal Stats Score** | **Parameter** | **Threshold** | **Weight** | **Time Period** | **Scoring Pseudo Code** |
| | GLL | t | 0.25 | lifetime | `for each parameter:`<br>`  if param > paramThresh:`<br>`  willScore = true`<br>`  break`<br>`intStatScore = 0`<br>`if willScore:`<br>`  for each parameter:`<br>`  pScore = parmVal / parmThresh * weight`<br>`  internalStatsScore += pScore` |
| | RUCORL | u | | | |
| | RCORL | v | | | |
| | GLD | w | 1.0 | 1000 hours | |
| | RUCORD | x | | | |
| | RCORD | y | | | |
| | Rcor/TB | z | 0.75 | 1-2 days | |

Table 1: show the individual stats that comprise the total-score and how each is weighted.

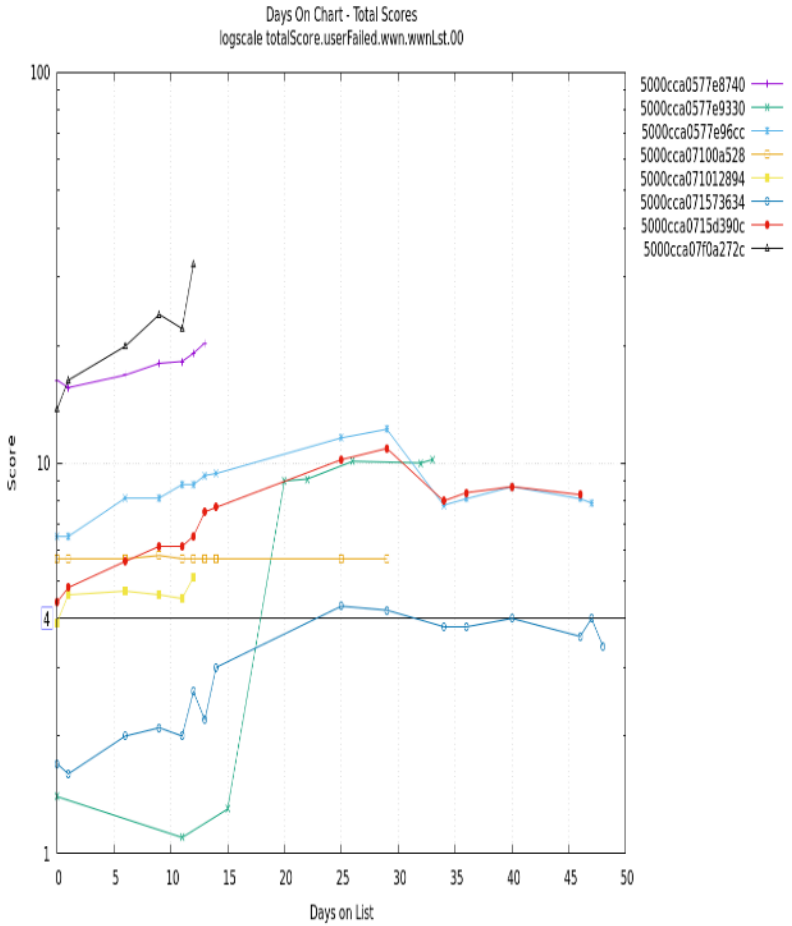
 

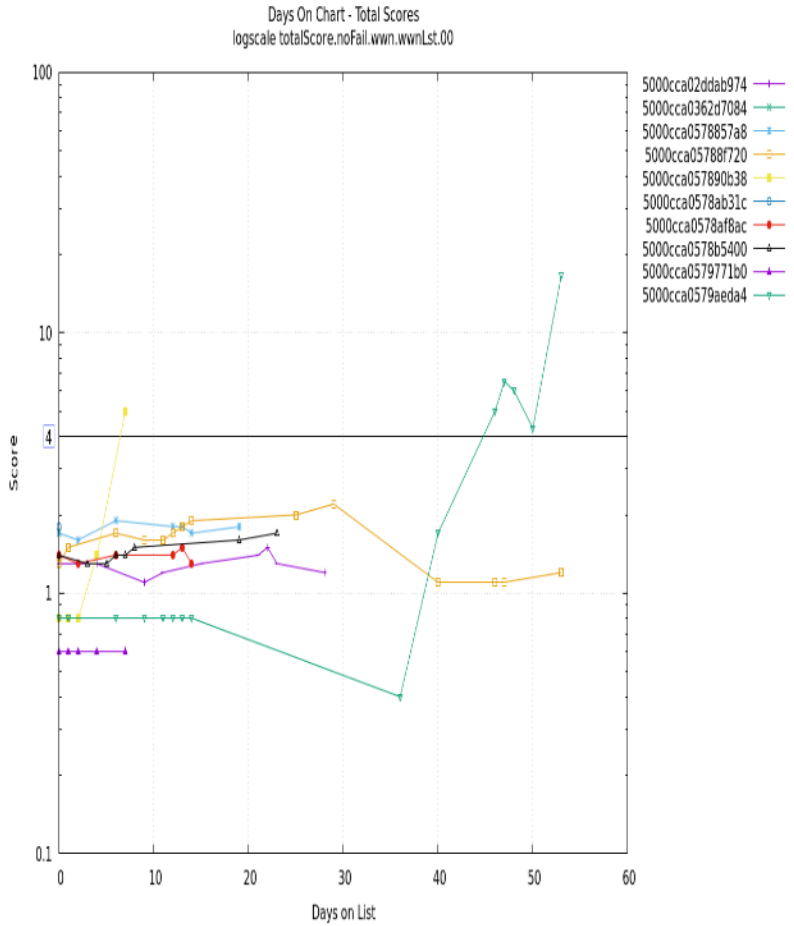
# Scoring Recap

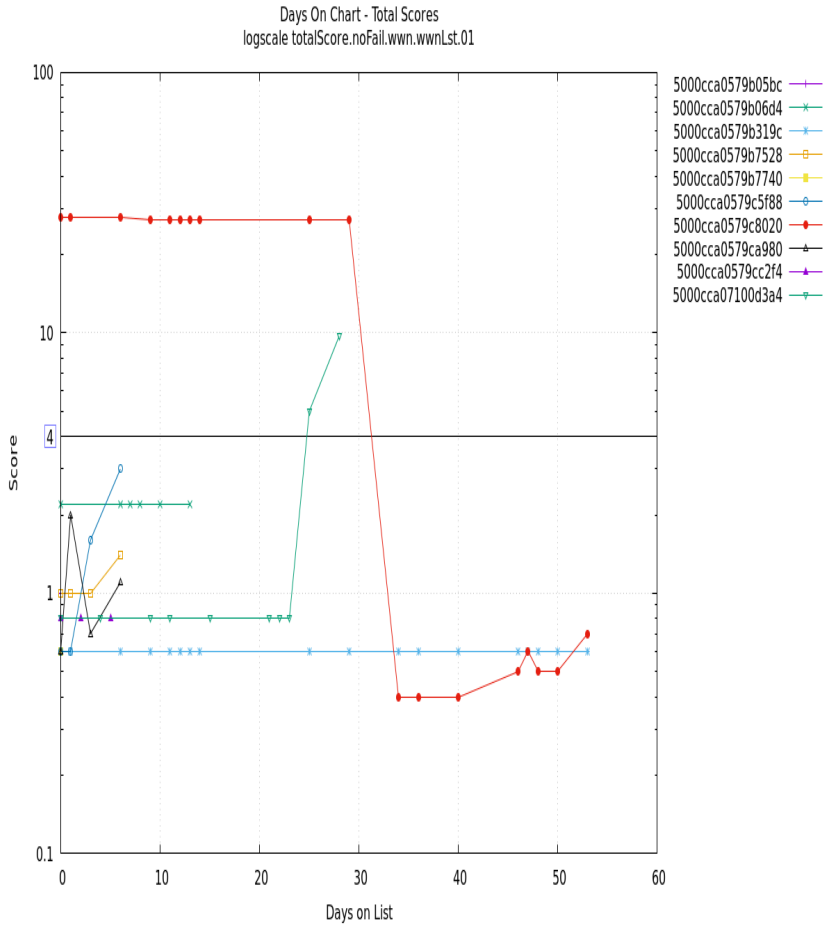
- The total score is the sum of the DEMS Score and the Internal Stats Score.
- The analysis of this document landed on a total score of X as the threshold for recommending removal (drive Model and type).
  - This value was chosen because it caught almost all of the drives failed by the controller and/or
  - Pulled by the user based previous analysis of the drive's history
  - while at the same time, Not flagging every drive that showed any sign of error.
- The goal here is to strike a balance between
  - Letting the drive stay in the system to use its error correction and internal retries to continue serve data; and
  - Pulling the drive before it gets to the point of risking data (dual drive failure) OR Causing performance issues (too many internal retries).

# Proposed Overall Threshold Value
# User Failed, and Still Alive

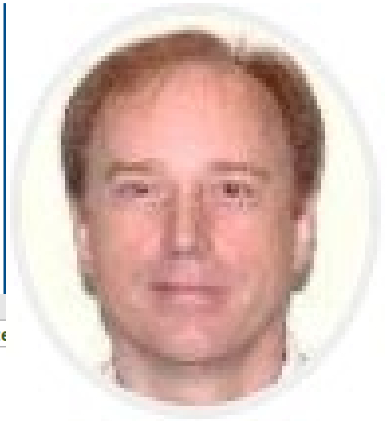


# Drive Scoring and Tracking tables



| WWN | TOTAL_SCORE | DEMS | FILTERED | TOTAL | FAILED | VG_Check |
|---|---|---|---|---|---|---|
| 5000cca057a18570 | TOTAL_SCORE | 12.5 | 58.7 | 71.2 | WRITE_OP | no check - missing drvVg data |
| 5000cca057893508 | TOTAL_SCORE | 20.7 | 35.0 | 55.7 | FastTo | no check - missing drvVg data |
| 5000cca0578929c0 | TOTAL_SCORE | 42.1 | 4.9 | 47.0 | WRITE_OP | no check - missing drvVg data |
| 5000cca07f0a272c | TOTAL_SCORE | 0.0 | 32.3 | 32.3 | USER_OP | no check - missing drvVg data |
| 5000cca0578b13b4 | TOTAL_SCORE | 13.4 | 7.0 | 20.4 | FastTo | no check - missing drvVg data |
| 5000cca0579aeda4 | TOTAL_SCORE | 12.7 | 3.9 | 16.6 | - | no check - missing drvVg data |
| 5000cca0710085d4 | TOTAL_SCORE | 14.3 | 0.0 | 14.3 | FastTo | no check - missing drvVg data |
| 5000cca07100d3a4 | TOTAL_SCORE | 3.2 | 10.0 | 13.2 | - | no check - missing drvVg data |
| 5000cca0579b0658 | TOTAL_SCORE | 5.9 | 6.4 | 12.3 | WRITE_OP | no check - missing drvVg data |
| 5000cca0577e96cc | TOTAL_SCORE | 0.0 | 12.2 | 12.2 | USER_OP | no check - missing drvVg data |
| 5000cca0715d390c | TOTAL_SCORE | 0.0 | 10.9 | 10.9 | USER_OP | no check - missing drvVg data |
| 5000cca07100a744 | TOTAL_SCORE | 0.0 | 10.4 | 10.4 | FastTo | no check - missing drvVg data |
| 5000cca0577e9330 | TOTAL_SCORE | 0.0 | 10.2 | 10.2 | USER_OP | no check - missing drvVg data |
| 5000cca0579cab1c | TOTAL_SCORE | 0.0 | 8.5 | 8.5 | WRITE_OP | no check - missing drvVg data |
| 5000cca0577e8740 | TOTAL_SCORE | 0.0 | 8.1 | 8.1 | USER_OP | no check - missing drvVg data |
| 5000cca071012894 | TOTAL_SCORE | 0.0 | 7.7 | 7.7 | USER_OP | no check - missing drvVg data |
| 5000cca057892d70 | TOTAL_SCORE | 5.2 | 1.7 | 6.9 | WRITE_OP | no check - missing drvVg data |
| 5000cca07f09097c | TOTAL_SCORE | 0.0 | 6.0 | 6.0 | - | no check - missing drvVg data |
| 5000cca07100a528 | TOTAL_SCORE | 0.0 | 5.8 | 5.8 | USER_OP | no check - missing drvVg data |
| 5000cca0710102f8 | TOTAL_SCORE | 2.9 | 2.5 | 5.4 | - | no check - missing drvVg data |
| 5000cca057890b38 | TOTAL_SCORE | 5.0 | 0.0 | 5.0 | - | no check - missing drvVg data |
| 5000cca071573634 | TOTAL_SCORE | 0.5 | 3.7 | 4.2 | USER_OP | no check - missing drvVg data |
| 5000cca0579c5f88 | TOTAL_SCORE | 3.0 | 0.0 | 3.0 | - | no check - missing drvVg data |
| 5000cca0579771b0 | TOTAL_SCORE | 0.6 | 2.0 | 2.6 | - | no check - missing drvVg data |
| 5000cca071571950 | TOTAL_SCORE | 0.0 | 2.4 | 2.4 | - | no check - missing drvVg data |
| 5000cca0579ca980 | TOTAL_SCORE | 1.1 | 1.3 | 2.4 | - | no check - missing drvVg data |
| 5000cca0579b06d4 | TOTAL_SCORE | 2.2 | 0.0 | 2.2 | - | no check - missing drvVg data |
| 5000cca05788f720 | TOTAL_SCORE | 0.0 | 2.2 | 2.2 | - | no check - missing drvVg data |
| 5000cca0578857a8 | TOTAL_SCORE | 0.0 | 2.1 | 2.1 | - | no check - missing drvVg data |
| 5000cca07f099b48 | TOTAL_SCORE | 0.0 | 1.8 | 1.8 | - | no check - missing drvVg data |
| 5000cca0578ab31c | TOTAL_SCORE | 1.8 | 0.0 | 1.8 | - | no check - missing drvVg data |
| 5000cca0362d7084 | TOTAL_SCORE | 0.0 | 1.7 | 1.7 | - | no check - missing drvVg data |
| 5000cca0578b5400 | TOTAL_SCORE | 0.0 | 1.7 | 1.7 | - | no check - missing drvVg data |
| 5000cca02ddab974 | TOTAL_SCORE | 0.0 | 1.5 | 1.5 | - | no check - missing drvVg data |
| 5000cca0578af8ac | TOTAL_SCORE | 0.0 | 1.5 | 1.5 | - | no check - missing drvVg data |
| 5000cca0579b7528 | TOTAL_SCORE | 1.4 | 0.0 | 1.4 | - | no check - missing drvVg data |
| 5000cca07f09ee70 | TOTAL_SCORE | 0.0 | 1.4 | 1.4 | - | no check - missing drvVg data |
| 5000cca0579c8020 | TOTAL_SCORE | 0.0 | 1.3 | 1.3 | - | no check - missing drvVg data |
| 5000cca07f0943a8 | TOTAL_SCORE | 0.0 | 1.2 | 1.2 | - | no check - missing drvVg data |
| 5000cca0579b05bc | TOTAL_SCORE | 1.0 | 0.0 | 1.0 | - | no check - missing drvVg data |
| 5000cca0579cc2f4 | TOTAL_SCORE | 0.8 | 0.0 | 0.8 | - | no check - missing drvVg data |
| 5000cca0579b7740 | TOTAL_SCORE | 0.6 | 0.0 | 0.6 | - | no check - missing drvVg data |

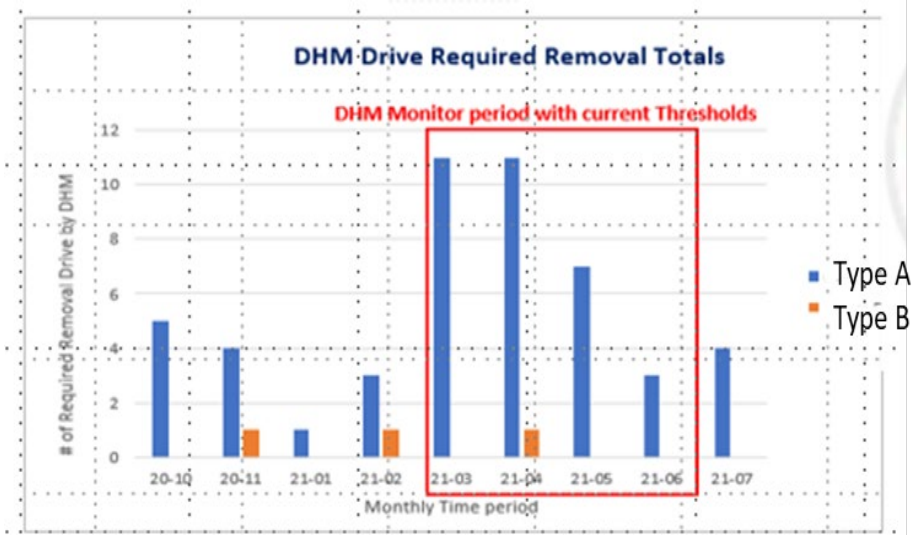
| | A | B | C | D | M | N | O | P | Q | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | System | Tray/Slo | S/N | WWN | Feb 26 Update | | Mar 5 Update | | Mar 9 Update | |
| 20 | sys_3.6 | t0,s19 | KXJEG76X | 5000cca057893508 | Recent DEMS: 20<br>**StagIO jump 19** since 2/11<br>Failed FastIO 2/26<br>**Lftm RCORL: 9K**<br>**GLL: 260**<br>**GLD (60d):257**<br>**RCORD(60d):9K**<br>**RCpTB(3.7d): 324** | F | | | | |
| 21 | sys_1.20 | t3,s3 | KXJVUSLX | 5000cca057a18570 | Since 2/22:<br>**+4 Fast**<br>**+49 StagIO**<br>**+30 Medium**<br>**+88 Rec**<br>Recent DEMS: 167<br>**Lftm RCOR: 12k**<br>**58 day RCORD 12K**<br>**GLL: 93**<br>**GLD (58d): 75**<br>**RCORD(58d): 60**<br>**RCpTB(3.6d): 19K** | RR | new filter score: 38 (**top** of non-failed)<br>**RCORD (30 day) 11k vs. RCORL of 12k**<br>**GLD of -46**<br>**RCor/TB: 1972**<br>**Recent DEMS 342**<br>RUCORD: -43 | RR | didn't look | Failed: Mar 11<br>10K jump in RCOR between 2/22 and 2/26<br>Recovered, medium, StatIO and FastTO all jumped in same timeframe (130 DEMS over those 4 days),<br>Recent DEMS: 320<br>GLL: 140<br>Filter score 38.9<br>**Appeared on Filtered List** | F |
| 22 | sys_1.6 | t3,s23 | KXJS803W | 5000cca0579b0658 | Since 2/11: +10 Med, +7 Rec<br>Recent Dems: 94<br>Lftm RCORL: 2K<br>10 day RCpTB: 99 | M | new filter score: 4.3 (6th from top of non-failed)<br>recent **DEMS: 154** (aborts, MED, REC)<br>**RCORD (27 day): 1911 (vs. RCORL 2159)**<br>RCor/TB: 99 | RR | didn't look | Failed: Mar 13<br>Recent DEMS: 194<br>Filter score: 4.9<br>RCORD 2.3K<br>**Appeared on Filtered List with score of 4.6 on 2/22** - | F |
| 23 | sys_4.2 | t3,s6 | KXJEEGXX | 5000cca0578929c0 | Recent DEMS score jumped from 21 to 202 in 4 days.<br>Gwth jump to 6 entries since 2-11.<br>Corrected errors jump 432 since 2-11<br>Daily Corr/TB jumped to 80 on 2/26 | Monitor | new filter score: 3.1 (8th from top of non-failed)<br>**RCORD (38 day) 1281 (vs. RCORL 1282)**<br>RCor/TB: 85<br>Recent DEMS: 808 (151 aborts, plus Med, Rec) | RR | didn't look | Failed: Mar 10<br>**Appeared on filtered list with score of 3.2 on 3/3.**<br>**> week notice.** | F |

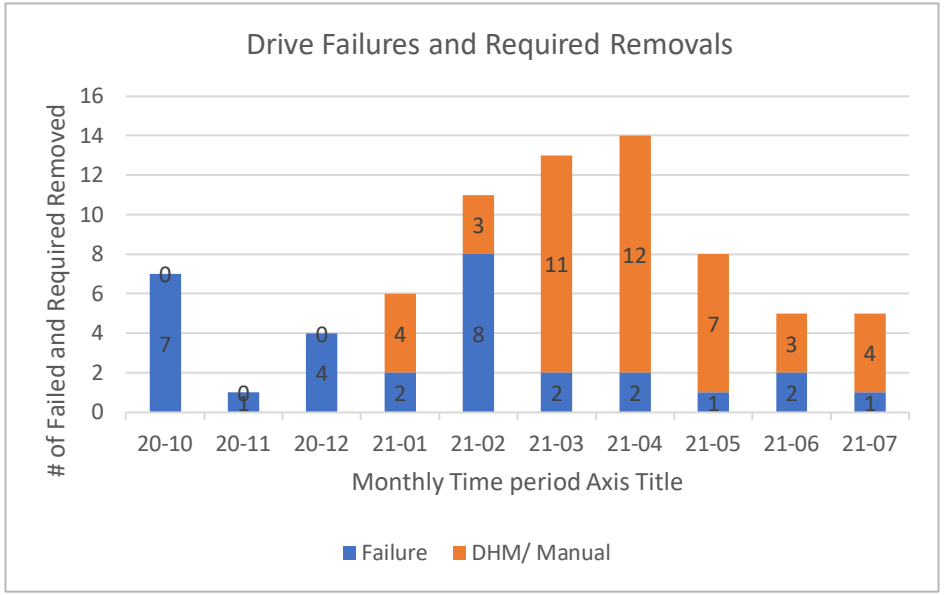
 

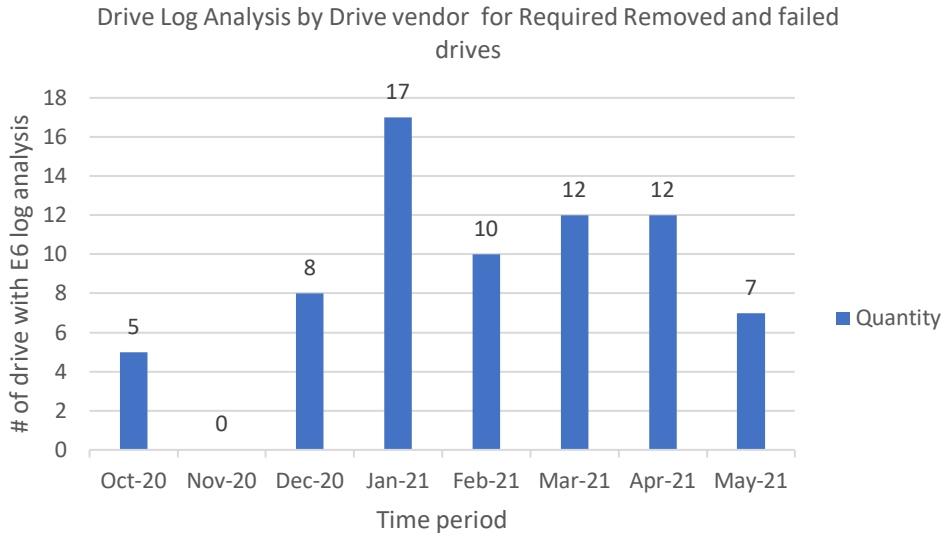
# Drive Health Monitor Statistics and results

With 1600 HDD drives: Reduction from >5% AFR in Jan of 2020 to < 2% AFR in July of 2021 in spite of drives' ages > 5 years

- Drive Required Removal (RR) increased in March thru May as DHM removed weak drives
- Drive Failure (DF) remained low since use of current DHM thresholds March 2021
- Significant Log analysis by Drive Vendors throughout deployment to review failures for trends
- DHM likely helped avoid multiple URS/Dead Volumes through proactive drive removal



# Q /A

# Thank You

# Please take a moment to rate this session.

Your feedback is important to us.