# Optimal Methods for Handling Data at Scale

## The way you access your storage has a profound impact on the vitality of the data

**Amit Chattopadhyay**

**Technologist/Director**
Western Digital
Advanced Reliability Engineering
amit.chattopadhyay@wdc.com

# Outline

1 **State of affairs in HDD Field Reliability**

- a HDD complexity leads to a myriad of possible failure modes
- b Macroscopic objects at molecular dimensions: **Head-disk Interface**
- c Critical Parameter: **Workload**

2 **HDD reliability modeling**

- a Historical model: derivation and limitations
- b Importance of workload and temperature

3 **Data Center Fleet Management: In-field Health Monitor**

# Core problem in HDD reliability

1. **Avoid an un-manageable catastrophe midway (or beyond) through a product's warranty life**
   *Referred to as 'wear-out' mode in reliability → unbounded Failure Rate at large times*

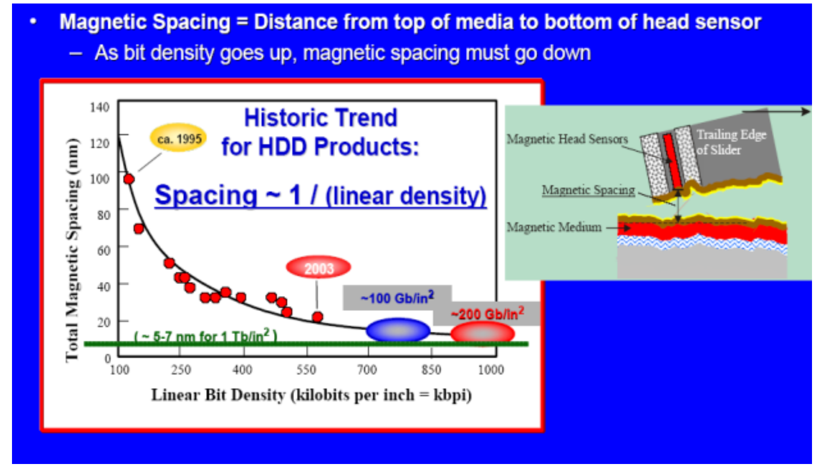
2. For the drives out in the field, develop a predictive analysis for failures and provide a real-time dashboard of drive health.

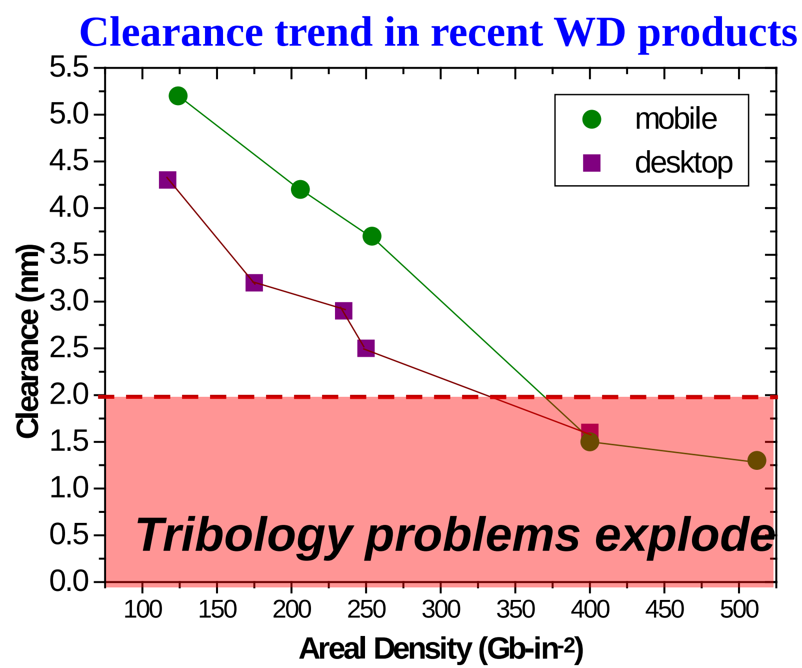
## *QUESTIONS*

- ❑ How do we define a product's "lifetime"?
- ❑ What are the key factors that determine the life of a drive?
- ❑ Can we help our customers perform effective Fleet Management in real time?

# Current HDD Reliability Landscape

- Increases in areal density have largely come by decreasing the head-disk spacing
  - The Magnetics vs Tribology dilemma
- Current head-media spacings are at molecular dimensions: 1 – 2nm
  - Operating at these clearances can impact the HDD failure rates, and the failure pareto
  - As much as 70% of the failure pareto is attributable to low-clearance operation
- Operation at ever-decreasing spacings requires special attention be paid to:
  - Anamolous physical phenomenon in ultrathin films
  - Non-linear response functions, e.g. Adhesive and frictional forces
  - Mechanical tolerances
  - etc



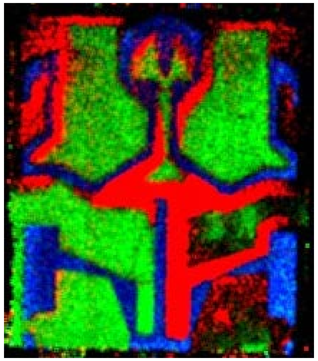



# Head-Disk Interface:

- HDD are complex devices that push the envelope of our physics understanding in a number of areas
- A monolayer of lubricant is used on disk surface to reduce adhesion / friction between slider and disk
- Lubricant pick-up by the slider is normal, **but lubricant pick-up in excess of one layer can pose a serious reliability risk**
  - HFW, OTW, modulation, and head degradation

**Optical image**

**Chemical image**



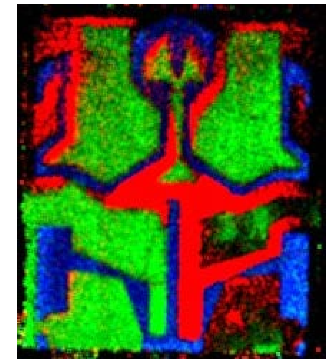
**Red = excess lube**

# Interface: at the edge of classical thermodynamics

- HDD are complex devices that push the envelope of our physics understanding in a number of areas
- A monolayer of lubricant is used on disk surface to reduce adhesion / friction between slider and disk
- Lubricant pick-up by the slider is normal, but lubricant pick-up in excess of one layer can pose a serious reliability risk
  - HFW, OTW, modulation, and head degradation

Optical image

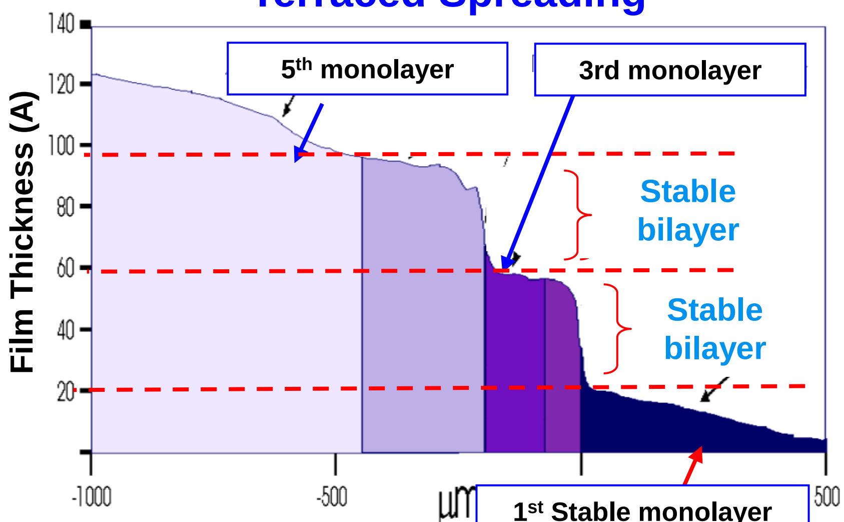
Chemical image



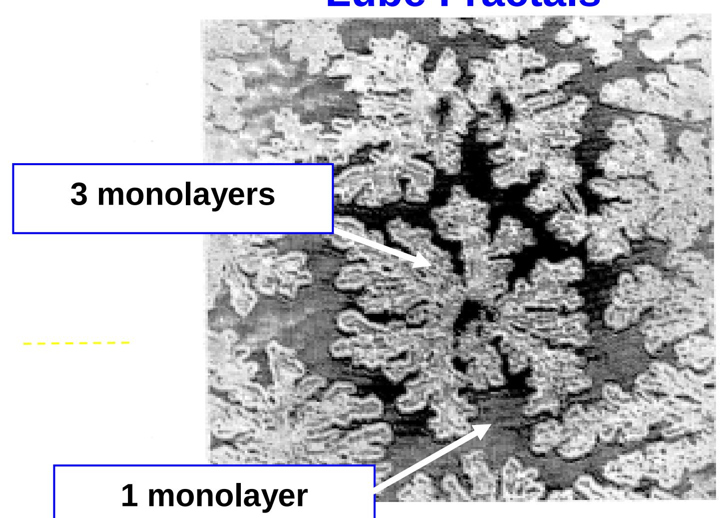
Red = excess lube

- Monomolecular films can show anamolous behavior
  - Non-classical thermodynamics: physical properties are dependent on the amount of material present

**Terraced Spreading**

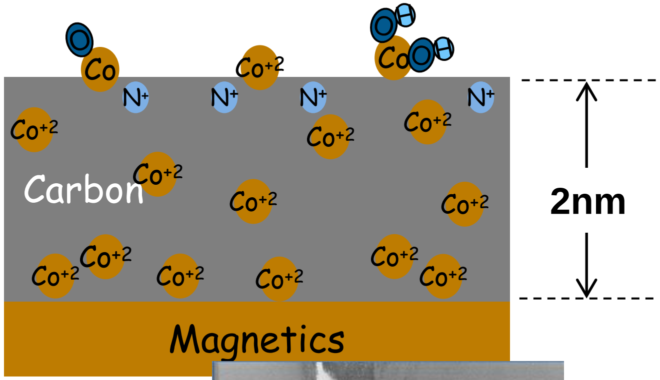



**Lube Fractals**

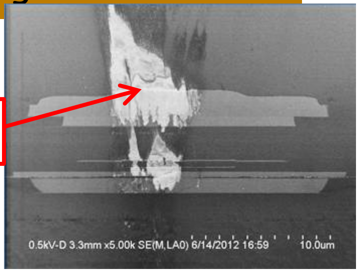


# …..with some Chemistry and Materials Science

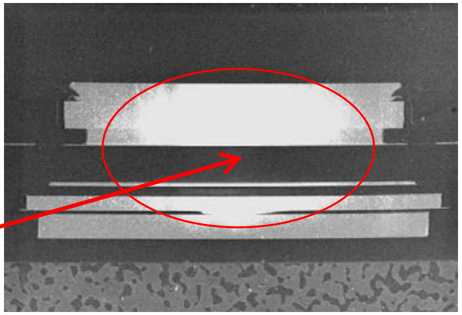
- **Protective diamond-like carbon (DLC) overcoats are used on both the head and disk surfaces**
  - Constant pressure to thin these overcoats for the purpose of decreasing the magnetic separation distance …..and thereby improving magnetic signal / areal density
  - Current DLC thicknesses on the order of 2nm
  - As these thicknesses are reduced, various reliability risks can arise




- **Disk example: migration of magnetic material to head-disk interface**
  - Corrective action required modification to carbon deposition energetics / process



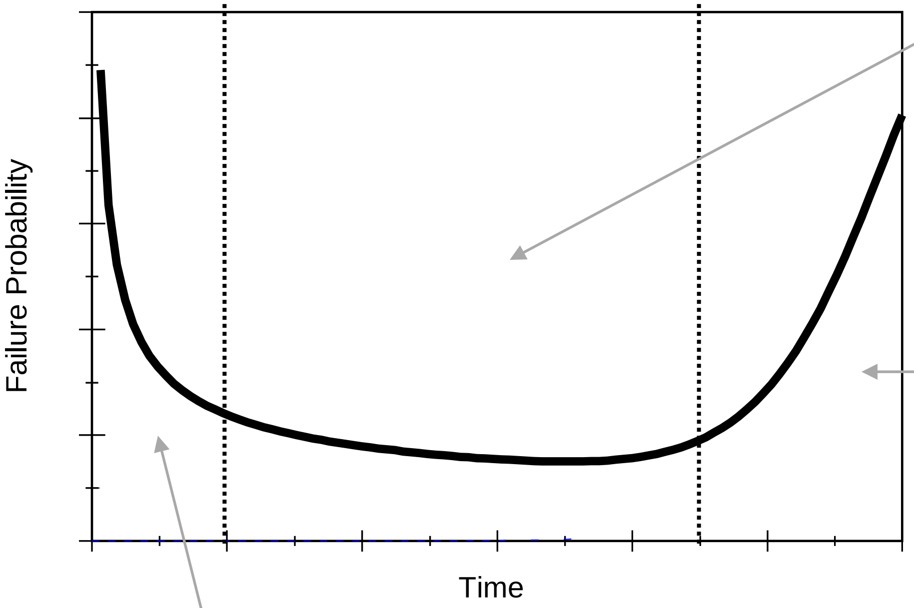

- **Head example: contact stress can result in wear of the pole tip**
  - Corrective action required stricter process control measures
  - Process optimization of $sp^3$ / $sp^2$



# Time to Failure: The “Bathtub” Reliability Model

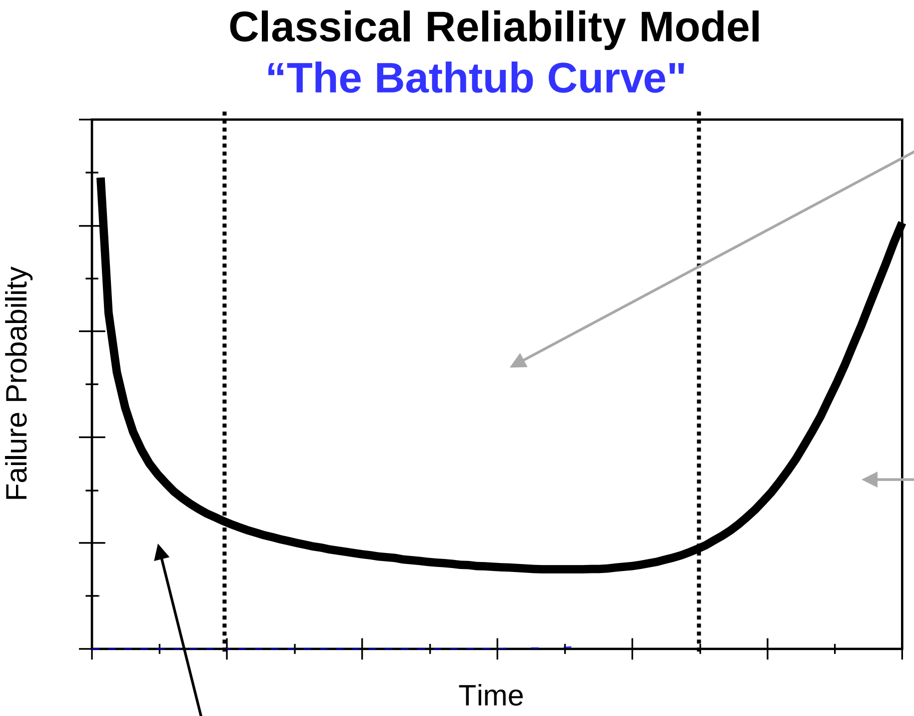
**Classical Reliability Model**

**“The Bathtub Curve"**




- **Steady State region**
  - After the weak drives are removed from the population, the failure rate reaches a fixed value for the service life of the drive
- **Wear-out Region**
  - At long times, one enters the wear-out region where normal wear and tear of the system components results in an increasing failure rate with time
- **Infant mortality region**
  - Failure rate decreases with increasing time
  - Result of defects etiher designed into, or inadvertently built into a product
    - Indicative of quality “escapes”
    - Marginal materials
    - Drives with the least margin for some critical design tolerance.
    - Manufacturing anomaly

# Time to Failure: The "Bathtub" Reliability Model



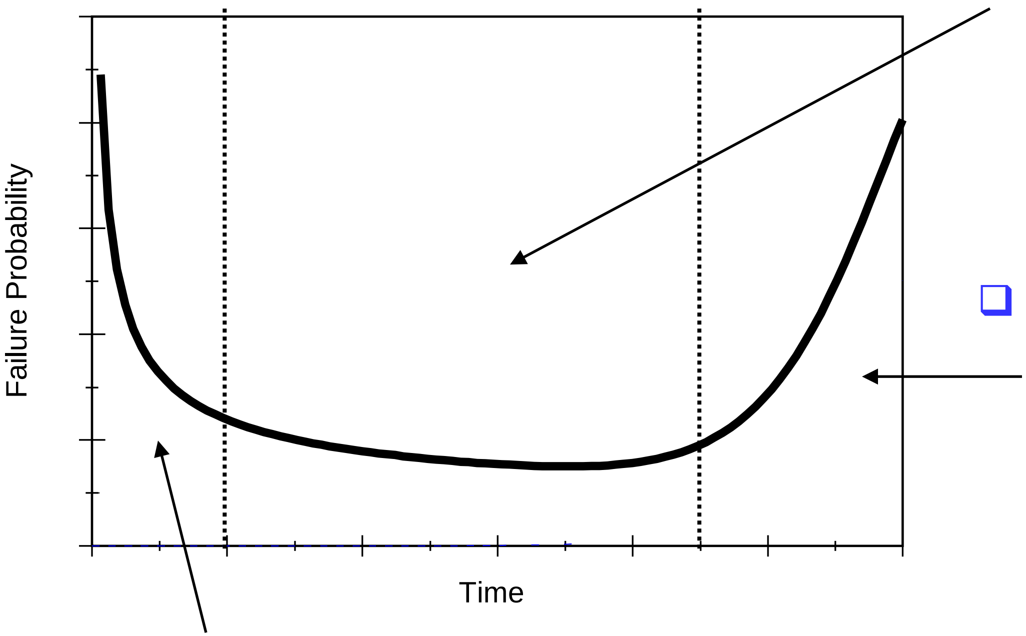

- **Steady State region**
  - After the weak drives are removed from the population, the failure rate reaches a fixed value for the service life of the drive
- **Wear-out Region**
  - At long times, one enters the wear-out region where normal wear and tear of the system components results in an increasing failure rate with time
- **Infant mortality region**
  - Failure rate decreases with increasing time
  - Result of defects either designed into, or inadvertently built into a product
    - Indicative of quality "escapes"
    - Marginal materials
    - Drives with the least margin for some critical design tolerance.
    - Manufacturing anomaly



absolute!

# Time to Failure: The “Bathtub” Reliability Model

**Classical Reliability Model**
**“The Bathtub Curve"**




- **Steady State region**
  - After the weak drives are removed from the population, the failure rate reaches a fixed value for the service life of the drive

- **Wear-out Region**
  - At long times, one enters the wear-out region where normal wear and tear of the system components results in an increasing failure rate with time

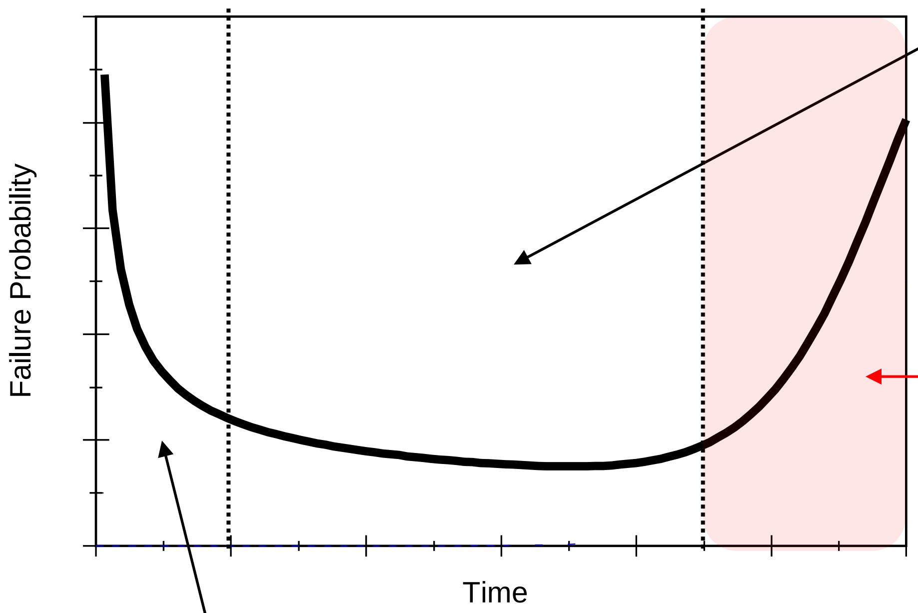
- **Infant mortality region**
  - Failure rate decreases with increasing time
  - Result of defects etiher designed into, or inadvertently built into a product
    - Indicative of quality “escapes”
    - Marginal materials
    - Drives with the least margin for some critical design tolerance.
    - Manufacturing anomaly



absolute!

# Time to Failure: The "Bathtub" Reliability Model

**Classical Reliability Model**
**"The Bathtub Curve"**




- **Steady State region**
  - After the weak drives are removed from the population, the failure rate reaches a fixed value for the service life of the drive
- **Wear-out Region**
  - At long times, normal wear and tear of the system components results in an **increasing failure rate with time**
  - This type of behavior results in costly excursions to both WD and our customers
  - **This regime must be avoided at all costs**
- **Infant mortality region**
  - Failure rate decreases with increasing time
  - Result of defects either designed into, or inadvertently built into a product
    - Indicative of quality "escapes"
    - Marginal materials
    - Drives with the least margin for some critical design tolerance.

# Classical reliability theory – time dependence

- Standard Weibull Treatment
  - Gives time dependence of unreliability, F(t)
  - Where: t = time
    η = time constant
    β = shape parameter

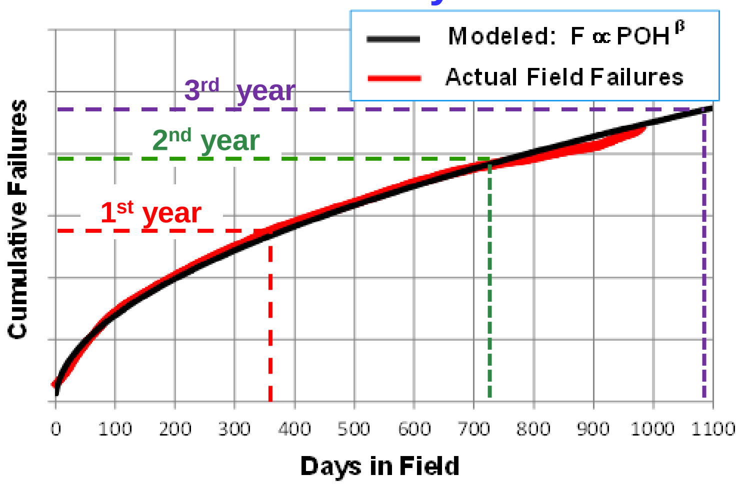
$$F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]$$

- In the limit of small failure rates (<10%) the failure rate vs time simplifies to a power function of time:
  - POH = power-on hrs

$$F(t) = \left(\frac{t}{\eta}\right)^{\beta} \propto POH^{\beta}$$

- Since failure rate assumed to depend on time alone, **MTTF has been used to quantify the intrinsic reliability of HDD's**
  - MTTF defined by the time required for 63.2% of all drives in a given population to fail
- **But…..use of MTTF alone is meaningless**
- **Reason behind the workload specs introduced by all HDD manufacturers**

**Field Reliability vs Time**




**β < 1: Failure rate decreases with time**

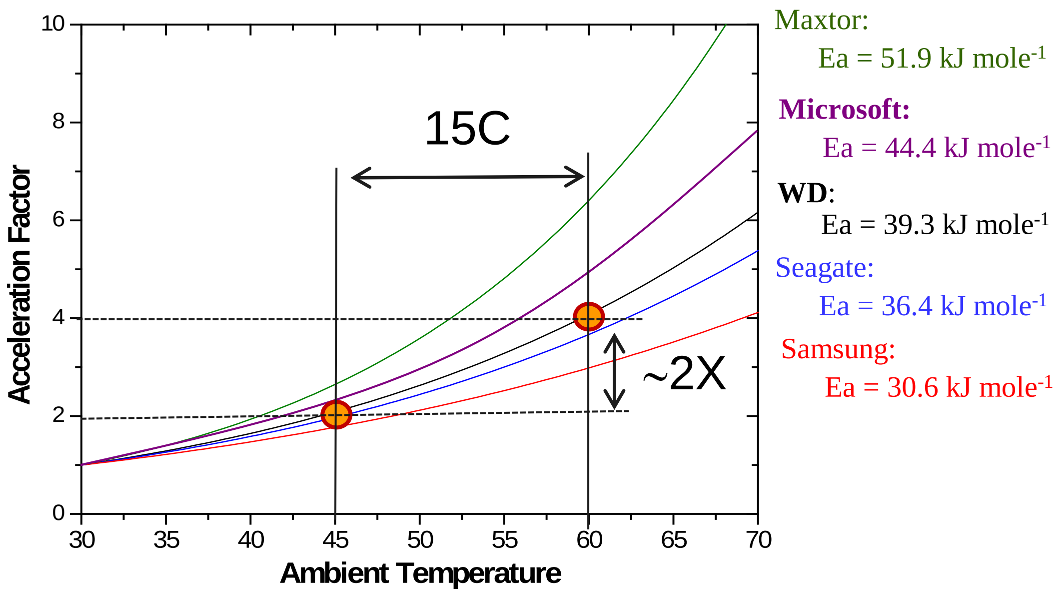
# Historical reliability model – impact of temp

- Numerous studies conducted on HDD failure rates at elevated temperature
- Consistent reports of HDD failure rates increasing exponentially with increasing temperature

  - Arrhenius equation:

$$F(T) = Ae^{\left[-\frac{Ea}{RT}\right]}$$

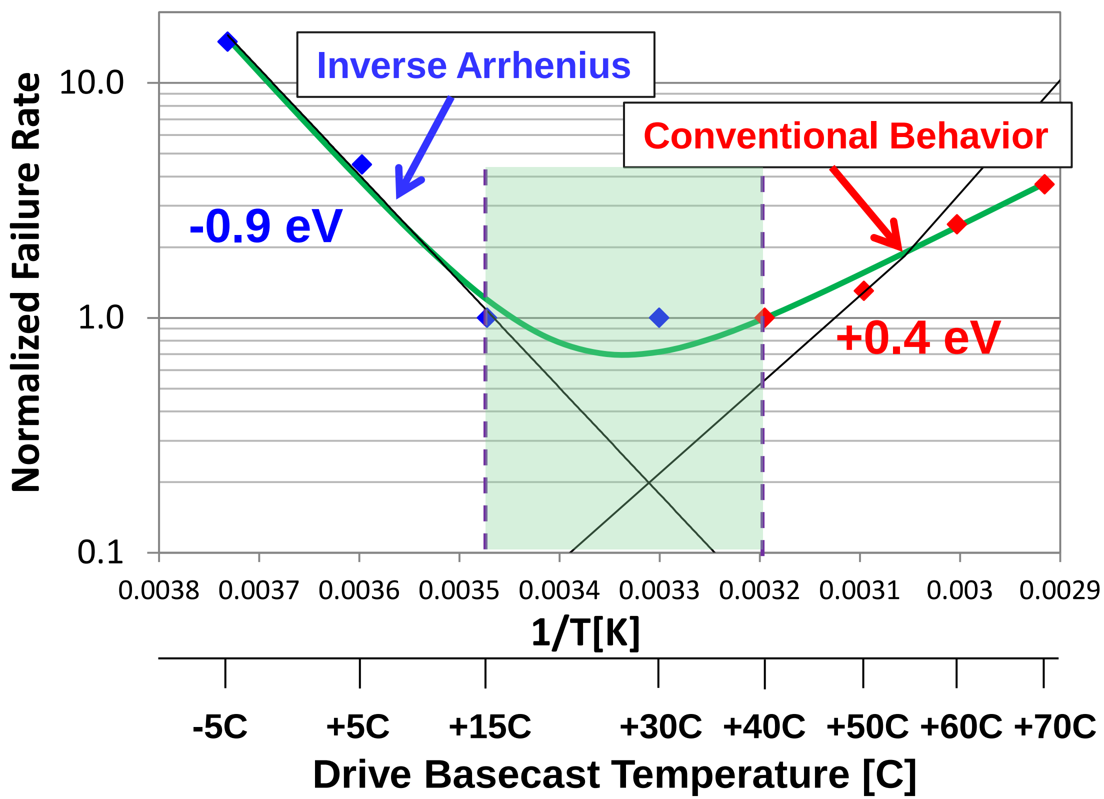
  - $E_a$ = activation energy
  - A, R = constants
  - T = Temperature (kelvin)




- Rule of thumb:
  - **Failure rate doubles for approximately 15 C increase in temp**
  - Note: We make use of this acceleration by testing all products at the upper temp spec limit

# Failure Rate vs Expanded Temp Range

- **Failure rate dependence at cold shows inverse Arrhenius behavior**
  - Magnetic challenges due to higher media coercivity
  - Stronger adhesive forces between heads and media/lube
- **Failure rates are relatively independent of temp between 15 – 40C**




$$F(T) \propto \exp\left(-\frac{E_a}{kT}\right)$$

$E_a$ = activation energy
$k$ = 8.62 x $10^{-5}$ eV/K
$T$ = absolute temperature (Kelvin)

# Failure Rate increase with temperature at both ends

## Risks of Running at Cold

- Overwrite challenges due to higher media coercivity
- Stronger adhesive forces between heads and media/lube
  - Normal forces leading to HFW
  - Lateral forces leading to off-track issues
- Operating regime of head/disk spacing is squeezed at both ends
- Permanent write damage to tracks
- Reading does not cause damage but error rate high
- Backoff has to be calibrated across the entire operational temperature range

## Risks of Running at Hot

- Key Issues
  - Backoff (Clearance) calibration and control
  - Outgassing and evaporation
  - Degradation of heads, oils, and other components
- **High temperatures can be designed for with appropriate trade-offs**

# Duty Cycle

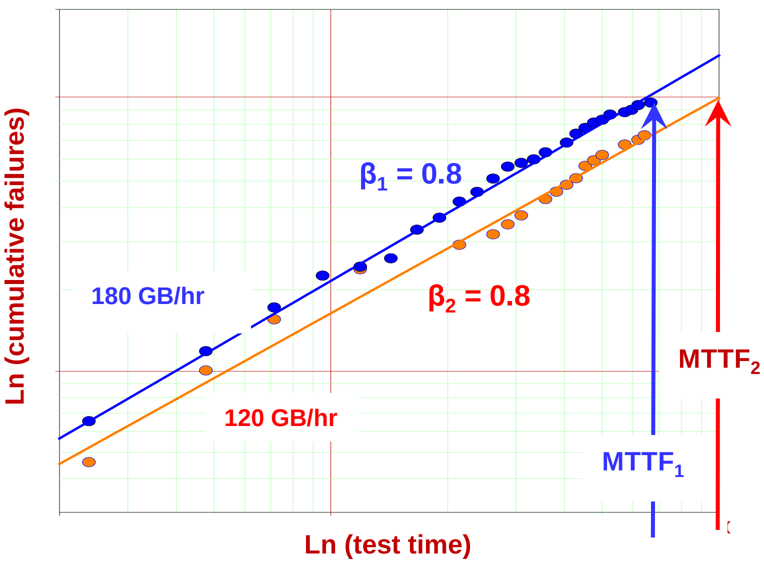
- **Is the concept of "Duty Cycle" valid?**
  - DOE with same drives built at the same time
  - Two tests with equivalent duty cycles (>95%)
  - ……..but differing workloads (1.5:1)

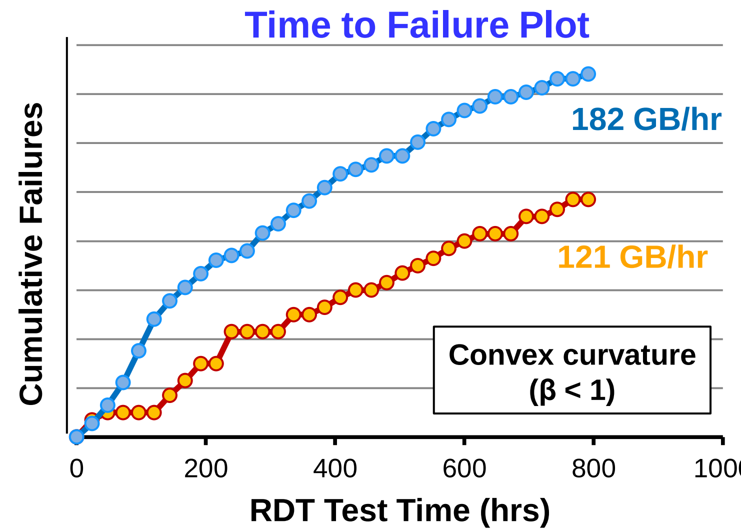
Usage — Thermal Term

$$F(t,T) \propto DC^{x} \text{x } POH^{\beta} \times \exp\left[-\frac{E_A}{kT}\right]$$

- Results clearly show that **failure rates do not scale with duty cycle**

- **Standard (time-based) Weibull Analysis**




Same drive / same DC / two workloads = different MTTFs

- **Conclusions**:
  - **MTTF typically used to specify reliability of HDDs**
  - **Since MTTF is not uniquely defined……**
  - **MTTF alone is an insufficient measure of drive reliability!**



absolutely™ WD

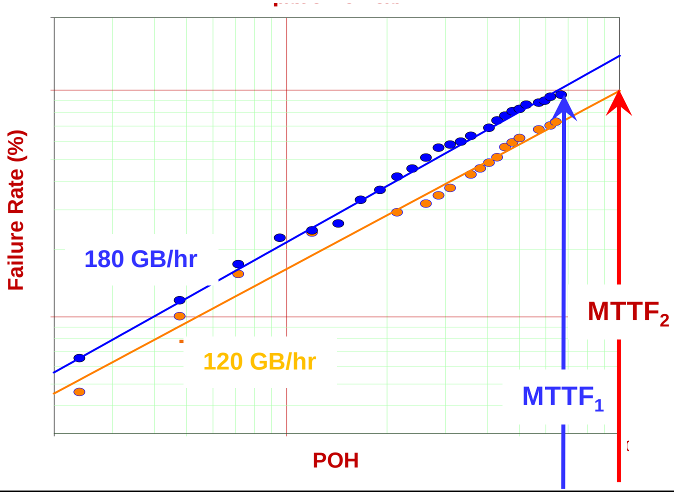
# Validation of Workload Impact on HDD reliability

- Failure rates scale with the total TB transferred

$$AFR \propto (TB)^{\beta}$$

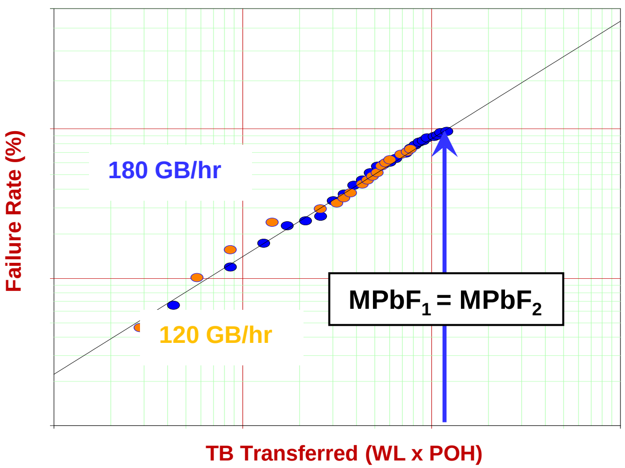
- Weibull Analysis

**Standard (time-based) Treatment**




**Same drive + 100% DC / two workloads = different MTTFs**

**Workload-based Treatment**




**Same drive +100% DC / two workloads = unique MPbF**

- Results demonstrate that TB transferred is the critical reliability parameter….not time POH
- **Natural reliability metric: Mean Petabytes to Failure (MPbF)**

  ➔ **This naturally leads to a DWM (Drive Workload Monitor) (like an odometer)**

- **Minimum requirement**: Simultaneously define **max workload spec** **and** **MTTF**
  - **This is now done by all HDD manufacturers**

# Workload is a critical driver of HDD reliability

**Workload** is a much better measure for drive usage

- Defined as the total data transfer (sectors written + read)
- Tightly coupled to the dominant failure modes in test and field
- Quantifiable from internal drive logs, thereby facilitating accurate field AFR predictions
- **All HDD manufacturers now explicitly specify workload ratings**

# Workload is Measurable

```
**********************************************************************
*                    Drive Odometer Metrics                          *
**********************************************************************
==========  Statistics  ==========
 Active power-on time, lifetime              594 hours (24 days, 18 hours, 17
minutes)
 Sectors transferred to host               82,612,289,007 (0x133C137DEF)
 Sectors transferred to drive              48,993,354,239 (0xB683B41FF)
```

- HDD logs record sectors read and written, as well as active hours
- Quantitative, statistical data, Best Units: TB/year

# Reliability Landscape and Trends

AFR = Annualized Failure Rate

$$AFR \propto test\ \ failure\,rate \times \left( \frac{field\,TB\,/\,yr}{test\,TB} \right)^{0.6} \times \exp\left( \frac{0.4eV}{kT} \right)$$

Typical Workloads for different market segments –

- **Re: Near line Enterprise (550 TB/yr)**
- **Se: Cloud (180 TB/yr)**
- **Ae: Cold Storage (60 TB/yr)**
- **Desktop usage (55 TB/yr)**
- **Mobile (25 TB/yr)**

**10-20x increase in usage from typical Mobile/DT drive to Enterprise**

# WD's health monitor – capabilities under development

- WD is pursuing parametric data collection for data center fleet management
- Potential to –
  - Provide a real-time health dashboard
  - Identify poorly performing drives / heads before actual failure, so appropriate action can be taken

***Capabilities under development:***

- Assign a Health Index to each drive – as a function of workload and environment
- Do 'Drive-less FA' – diagnose (not fix) root-cause of degradation as it happens in real time
- Enable a closed loop: Monitor → Analyze → Predict → Act

- ***Internal Benchmarking***:
  - Measurement of parametric data now routinely performed in Reliability testing (2+ years experience)
  - Natural progression will be to deploy something similar in a live DC environment.

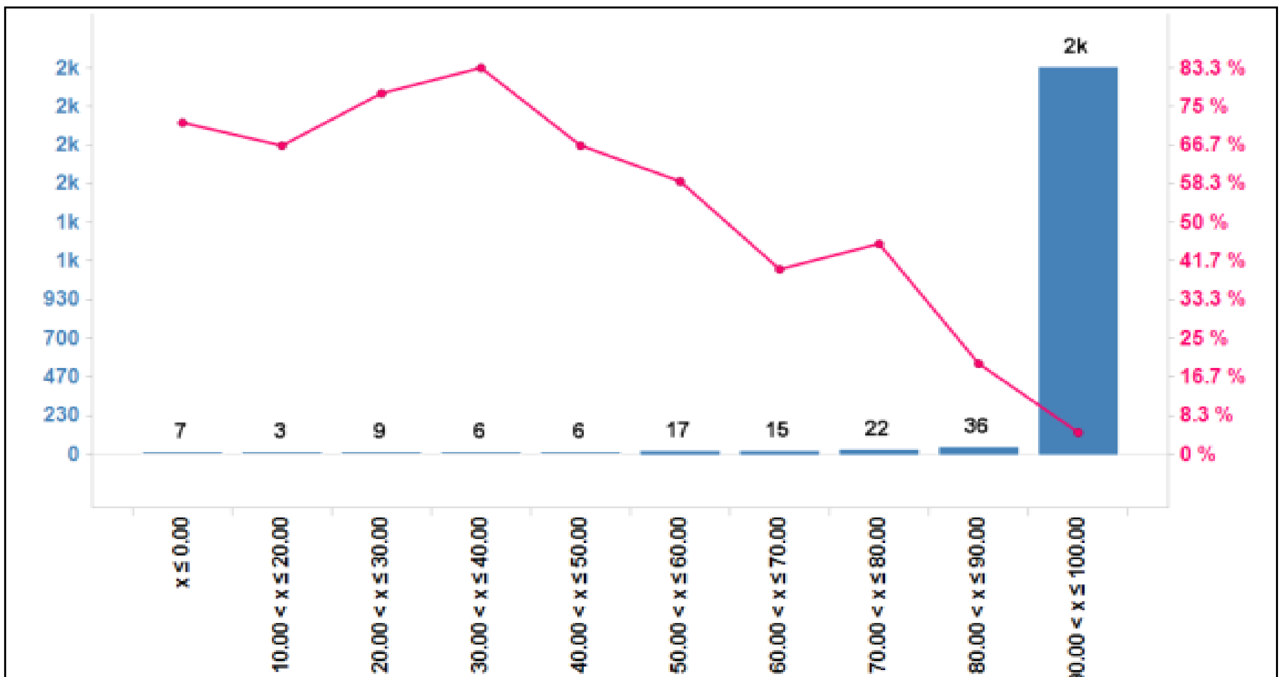
# WD’s health monitor – capabilities under development

**The In-Field Health Monitoring (IFHM) system will measure, and store, critical parameters periodically throughout the HDD lifetime, and output a relative health index.**

***A ‘Gas Gauge’ - Failure Probability Increases as ‘Health Index’ gets worse***

**Low Score (Bad) ⟵ High Score (good)**

**Drive Count**




**Real Failure Rate**

**IFHM Health Index**
**(derived from critical parameters)**

# Meaningful Predictive Modeling

- **Powerful models that predict failures of individual drives need to make use of signals beyond SMART**

- **One of the main motivators for SMART was to provide enough insight into disk drive behavior to enable such models to be built**

- **IFHM + workload monitor + SMART may effectively overcome the shortcomings of current predictive tools.**

# Thank You