Desktop, Nearline & Enterprise HDDs

What’s the difference?

Willis Whittington, Seagate Technology
The material contained in this tutorial is copyrighted by the SNIA.

Member companies and individuals may use this material in presentations and literature under the following conditions:

- Any slide or slides used must be reproduced without modification.
- The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations.

This presentation is a project of the SNIA Education Committee.
Abstract

Desktop, Nearline & Enterprise Disk Drives
What’s the difference?

For the past twenty five years the storage marketplace has been divided into two major categories namely “Desktop” and “Enterprise”. Recently, a third player variously known as “Nearline”, “Reference” or “Business Critical” has evolved to provide a low cost, high capacity storage solution for Enterprise data that no longer needs to exist in a high availability transactional processing environment but must maintain 24 x 7 availability as a reference or backup resource.

Each of these classes of drives requires a unique and specific set of attributes to fulfill its role. This presentation will explore these differences and explain why you need to use the right drive for the right application.
Agenda

- Basic Comparisons
- SAS & SATA Compatibility
- The Advantages of Nearline SAS
- Rotational Vibration
- Data Error Rate
- Error Correction Capability
- Data Integrity
- Performance
- Annualized Failure Rate

~ Q & A along the way ~
Basic Comparisons
## Comparison Table  DT / NL / MC*

<table>
<thead>
<tr>
<th>Metric</th>
<th>Desktop</th>
<th>SATA Nearline (NL)</th>
<th>SAS Nearline (NL)</th>
<th>Enterprise MC*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity (GB)</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>450</td>
</tr>
<tr>
<td>Cost</td>
<td>Low</td>
<td>Mid</td>
<td>Mid</td>
<td>High</td>
</tr>
<tr>
<td>Power Consumption</td>
<td>1x</td>
<td>1.2x</td>
<td>1.2x+</td>
<td>1.5x</td>
</tr>
<tr>
<td>MTBF (Hrs)</td>
<td>600,000</td>
<td>1,200,000</td>
<td>1,200,000</td>
<td>1,600,000</td>
</tr>
<tr>
<td>Duty Cycle</td>
<td>Low (&lt;10%)</td>
<td>Low/Medium (&lt;20%)</td>
<td>Low/Medium (&lt;20%)</td>
<td>High (100%)</td>
</tr>
<tr>
<td>Data Integrity</td>
<td>Parity (?)</td>
<td>EDC + (ECC?)</td>
<td>EDC/ECC + Proprietary Data Integrity Protection</td>
<td>EDC/ECC + Proprietary Data Integrity Protection</td>
</tr>
<tr>
<td>Unrec Error Rate</td>
<td>$10^{-14}$</td>
<td>$10^{-15}$</td>
<td>$10^{-15}$</td>
<td>$10^{-16}$</td>
</tr>
<tr>
<td>RV Radians/sec²</td>
<td>6</td>
<td>12.5</td>
<td>12.5</td>
<td>&gt;21</td>
</tr>
<tr>
<td>Error Recovery</td>
<td>SATA</td>
<td>SATA + Time Control</td>
<td>Full SCSI</td>
<td>Full SCSI</td>
</tr>
<tr>
<td>Firmware/Features</td>
<td>Standard SATA</td>
<td>SATA + Selected Nearline Features</td>
<td>SCSI + Adv. Features (Enabled by Dual CPU)</td>
<td>SCSI + Adv. Features (Enabled by Dual CPU)</td>
</tr>
<tr>
<td>Power On Hrs/Year</td>
<td>2400</td>
<td>8,760 (Low Duty cycle)</td>
<td>8,760 (Low Duty cycle)</td>
<td>8,760</td>
</tr>
<tr>
<td>Multi Initiator</td>
<td>No</td>
<td>No</td>
<td>16 Hosts &amp; Dual Port</td>
<td>16 Hosts &amp; Dual Port</td>
</tr>
<tr>
<td>Performance</td>
<td>1x</td>
<td>1x</td>
<td>1x+</td>
<td>1.4x / 2.5x (Seq / Rand)</td>
</tr>
<tr>
<td>T10 Data Protection</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Scalability</td>
<td>Low</td>
<td>Low</td>
<td>High + Dual Port</td>
<td>High + Dual Port</td>
</tr>
</tbody>
</table>

*Mission Critical

---

Key: [ ] Good  [ ] Better  [ ] Best

© 2008 Storage Networking Industry Association. All Rights Reserved.
Motor
Higher rpm than NL or DT
Tighter specifications
Less runout
More expensive
Anatomy of an Enterprise Drive

Discs
- Four platter design
- Smaller diameter than NL/DT
- Full media certification
- Fully characterized
- Variable sector format
Anatomy of an Enterprise Drive

Head Stack
Eight head design
Low mass, high rigidity
Voice coil designed for:
  - optimal performance
  - 100% duty cycle
Higher cost design
Anatomy of an Enterprise Drive

Environmental Control Module
Humidity Control
Chemical Absorption
Multi-point filtration
Windage Design
Anatomy of an Enterprise Drive

Misc Mechanical
Powerful Voice Coil Motor
Stiffer Covers
Air Control Devices
Faster Seeks
High Servo Sample Rate
Low RV susceptibility
Anatomy of an Enterprise Drive

Electronics
- Dual processors
- Multi host queuing
- Dual port
- Twice the memory of NL/DT
- High rpm control
- Command scheduling
- Superior error protection
- Superior error correction
- Smart servo algorithms
- Perform. optimization
- Data integrity checks
- Sequential h/w assist
SAS & SATA Compatibility
I/O Connectors for SAS & SATA

For SAS, the key-way is filled in and its flip side is used for the 2nd Port. This prevents a SAS drive from being plugged into a SATA cabinet slot.
The Advantages of Nearline SAS
NL SATA Compared to NL SAS

Stepping up to SAS provides Mission Critical Compatibility

NEARLINE SATA

- Full EDC & ECC
- 100% Phy Compatible
- End-to-End Data Integrity
- Variable Sector Size
- Multiple Host Support
- Full SCSI Command Set
- Enterprise Command Queuing
- Concurrent Data Channels
- Full Duplex (Bidirectional) I/O

NEARLINE SAS

- Dual Port

Mission Critical Features

Desktop, Nearline & Enterprise HDDs – What’s the difference?
© 2008 Storage Networking Industry Association. All Rights Reserved.
SAS/SATA NL Physical Differences

SAS Electronics

SAS Port “B”

Nearline Head/Disc Assy.

SATA Electronics

Desktop, Nearline & Enterprise HDDs – What’s the difference?
© 2008 Storage Networking Industry Association. All Rights Reserved.
SAS/SATA NL Differences

**SAS**
- **Full Duplex**
  - Xmits in both directions at one time
- **Dual Port**
  - Two Data Channels
  - 2 Concurrent Writes
    - OR
  - 2 Concurrent Reads
    - OR
  - 1 Write + 1 Read

**SATA**
- **Half Duplex**
  - Xmits in one direction at one time
- **One Data Channel**
  - 1 Write
    - OR
  - 1 Read

---

Desktop, Nearline & Enterprise HDDs – What’s the difference?
© 2008 Storage Networking Industry Association. All Rights Reserved.
Multi Host Command Queuing

Drive Queue Supports 16 Hosts

Drive NCQ* supports a single Host

Interposer to handle Q’ing for 2nd host

SAS

SAS SATA

EXPANDER

EXPANDER

Desktop, Nearline & Enterprise HDDs – What’s the difference?
© 2008 Storage Networking Industry Association. All Rights Reserved.
Both SATA and SAS Nearline drives are designed for use in Enterprise Mission Critical environments.

SAS Nearline drives have additional advantages which are made possible by the Serial SCSI interface and enterprise electronics:

- Full system interface compatibility at the protocol, physical (“phy”), and command level
- Enterprise error recovery and performance optimization controls
- Full data integrity protection both within the drive and at the system level with DIF\(^1\) support.

\(^1\)Data Integrity Feature also known as T10 PI (Data Protection Information).
Rotational Vibration
PS drives are not designed for backplane (JBOD/SBOD) use and are not equipped to cope with the effects of RV.

Neighboring Drive’s Servo needs to compensate for externally induced RV.

HDA subjected to rotational forces.

**RV is Proportional to Seek Current**

*Scope Picture, Seagate Prod. Dev.*
Rotational Vibration

Impact on Performance*

*Source: STX Competitive Analysis.
**RV in 33 Different Cabinets**

**Rotational Vibration**

- **Acceptable for Enterprise**: 22 radians/sec$^2$
- **Nearline**: 12 radians/sec$^2$
- **Desktop**: 5 radians/sec$^2$

More stringent RV spec. needed for SATA cabinets

RV aggravated by system fans, random access and “bursty” workloads
Data Error Rate
The UER for SATA desktop is 1 in $10^{14}$ bits transferred
   • $10^{14}$ bits = 12½ terabytes

A 500 Gbyte drive has $1/25 \times 10^{14}$ bits

Rebuilding a SATA drive in a RAID 5 set of 5 drives means transferring $5/25 \times 10^{14}$ bits = 1/5 of UER spec.
   • 20% probability of an Unrecoverable Error during the rebuild.

Better odds would be available with RAID 1 or 6
   • RAID 1 rebuilds from a single mirror drive
   • RAID 6 can tolerate a second error during the rebuild.

Risks can be reduced with good error management
   • Intelligent rebuild (ignore unused capacity)
   • Background media scan (dynamic certification)

*Unrecoverable Error Rate
Desktop, Nearline & Enterprise HDDs – What’s the difference?

**Unrecoverable Error Rate**

### DT / NL / MC UER*

- **Desktop Drives UER = 10^{-14}**
- **Nearline UER = 10^{-15}**
- **Enterprise UER = 10^{-16}**

*Unrecoverable Error Rate*
Error Correction Capability
Standard vs Reverse ECC

(Write Command)

**Standard ECC**

- User Data
- ECC Generator
- Randomizer
- RLL Encoder
- Encoded

**Reverse ECC**

- User Data
- Randomizer
- RLL Encoder
- Encoded
- ECC Generator
- Encoded

Desktop, Nearline & Enterprise HDDs – What’s the difference?
© 2008 Storage Networking Industry Association. All Rights Reserved.
Standard vs Reverse ECC

(Read Command)

**Standard ECC**
- User Data
- ??
- ECC Correction
- De-randomize
- 120 Bit Error (Propagates in Decoder)
- RLL Decoder
- 20 Bit Error

**Reverse ECC**
- User Data
- De-randomize
- No Error Propagation
- RLL Decoder
- ECC Correction
- Error Corrected on the fly
- Error Corrected on the fly
- 20 Bit Error
The Sync Field is used to get the read channel in frequency sync with the data recorded on the media.

The Sync Mark is used to define the beginning of the User Data Field.

Failure to recognize the Sync Mark (due to a thermal asperity or a grown media defect) means the User Data Field is not delineated and the data is lost.
Sync Mark Errors on SAS Drives

- **Read Operation**
  - **Sync Field**
  - **User Data Field**
  - **ECC**
  - **Media Flaw**

- **Sector Format**
  - **Sync Mark**
  - **Sync Mark 2 embedded in the data field**
  - **n bytes**

- **Read Channel**
  - Realizes the first Sync Mark is missing:
    - Loads Buffer with n zeroes
    - Starts searching for SM2

- **Before ECC**
  - Data Buffer: 00000000000000

- **After ECC**
  - The missing n bytes are recovered using the ECC: 11010111010011

- **Sync Field**
  - 01101101010011

- **User Data Field**
  - 1010001101001100110001110100110001101010010100100111

- **ECC**
  - 0011011000011110101011000011010100110110001010101000111

- **Sync Mark Errors on SAS Drives**

- **Sync Field**
  - Data Buffer: 00000000000000

- **User Data Field**
  - 0011011000011110101011000011010100110110001010101000111

- **ECC**
  - 0011011000011110101011000011010100110110001010101000111

- **Read Channel finds SM2 and reads the data following it into the Buffer, starting at location n+1**

- **The missing n bytes are recovered using the ECC**

- **Desktop, Nearline & Enterprise HDDs – What’s the difference?**

© 2008 Storage Networking Industry Association. All Rights Reserved.
Data Integrity
Performance
Performance Comparison

Enterprise Sequential Access

Desktop / NL Sequential Access

Vendor Range

Sequential Transfer Rate $\propto$ RPM x Disc Dia x Bit Density

Enterprise Transfer Rate

Desktop Transfer Rate

\[
\frac{\text{Enterprise Transfer Rate}}{\text{Desktop Transfer Rate}} = \frac{\text{RPM} \times \text{Disc Dia} \times \text{Bit Density}}{\text{RPM} \times \text{Disc Dia} \times \text{Bit Density}} = \frac{15000 \times 65}{7200 \times 95} = 142% \\
\text{(independent of seek time & Latency)}
\]
Performance Comparison

- **Enterprise Sequential**
- **Desktop / NL Sequential**
- **Enterprise Random**
- **Desktop / NL Random**

Desktop, Nearline & Enterprise HDDs – What’s the difference?
© 2008 Storage Networking Industry Association. All Rights Reserved.
SPC-1C comprises I/O operations designed to demonstrate small storage subsystem performance (1-16 drives) while performing the typical functions of a business critical application.
AFR
(Annualized Failure Rate)
SATA drives in Enterprise applications run hotter, at higher duty cycle, and for more Power-On-Hours than in desktop applications.
In Conclusion…..

Although technological advances, driven by Enterprise research, will be leveraged into SATA products, there will continue to be functional limitations imposed on these devices by the overriding metric of Low $/GB Storage.
Q&A / Feedback

Please send any questions or comments on this presentation to SNIA:
trackstorage@snia.org

Many thanks to the following individuals for their contributions to this tutorial.

Craig Parris
Daniel Dummer
Willis Whittington
Wolfgang Rosner

SNIA Education Committee