

How HDD Workload Impacts TCO

Technology Paper

Written by
Andrei Khurshudov, PhD
Cloud Modeling and Data Analytics
Seagate Technology

For Demanding Data Centers

Introduction

The amount of digital data produced every year is increasing exponentially and thus increases the demand for storage. Despite recent popularity of solid state storage (SSD) devices, the overwhelming majority of digital data is still stored on magnetic recording media—namely, hard disk drives (HDD)—which are the very foundation of almost every data center. In addition, large data centers can now be found in a wide variety of industries, such as healthcare, retail and manufacturing. They power online search, shopping, social networks and other solutions offered by the IT industry.

Independent of whether a data center is built as a collection of high-end storage systems with hardware RAID data redundancy or constructed using lower-end hardware with software data redundancy (provided by a global distributed file system), drive failures and replacements are costly and can measurably increase the data center's total cost of ownership (TCO). Industry analysis suggests that it costs between US\$100 and US\$300 for each incident involving hardware failure, maintenance, repair, replacement, etc.^{1,2}

¹ *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines*, Luiz André Barroso and Urs Hölzle, 2009

² *Characterizing Cloud Computing Hardware Reliability*, Kashi Venkatesh Vishwanath and Nachiappan Nagappan, SoCC'10, June 10–11, 2010, Indianapolis, Indiana, USA.



How HDD Workload Impacts TCO



In an environment of increasing demand for computing performance and storage capacity, the TCO becomes the primary metric for almost any data center operator. TCO typically accounts for all costs involved in data center construction and operation, such as capital and operating expenses, cost of hardware and software, as well as data center administration, maintenance and repairs. HDD reliability, along with the reliability of other data center hardware, has a strong influence on the operational expenses relating to data center maintenance.

Contrary to some observations, HDDs are among the most reliable hardware components of a data center. Storage and compute servers, for example, have many other components that would limit system reliability before HDDs will. Cooling fans typically have MTBF values on the order of 100K hours. Server power supplies are usually rated with an MTBF of 400K hours. Those components are far less reliable than a typical nearline HDD rated at 1000K+ MTBF hours.

Of course, there could be many more drives in a data center (or, inside a typical server) than there are fans or power supplies. Larger numbers of drives will naturally increase the probability of one of them failing at a given time, which will prompt the need for a replacement.

Fortunately, there are several factors that could help the data center operator.

First, Seagate experience suggests that HDD reliability is strongly dependent on the operating conditions that are defined and controlled by the data center operator. One of two seemingly identical drives could potentially suffer a five-fold decrease in reliability when placed in a harsh operating environment. This gives the data center operator the ability to adjust the operating environment for better reliability with the lowest total operational cost.

Second, Seagate manufactures different types of drives designed to excel under different operating conditions, those typical for desktop, nearline and mission-critical environments. Seagate understands what is essential for higher reliability and has a set of recommendations that will assure the best possible reliability.

There are many differences in the way the drives are used (and stressed) inside an actual data center. The essential parameters of HDD stress are usage time, operating temperature and user workload. Each of these parameters is typically a strong function of both the data center architecture (including its topology, server design, overall data center storage capacity and its utilization, virtualization, workload balancing, etc.) and by the end-user's applications (the total amount of data transferred bidirectionally, data rates over time, etc.). Let us analyze the importance of usage time, operating temperature and user workload on reliability independently.

Usage Time

The *usage time* impact on HDD reliability is fairly easy to understand.

Mathematically, the simple equation shown here tells us how the usage time and the product's reliability, expressed as mean time between failures (MTBF) combine for a cumulative component failure probability. As the usage time increases, the cumulative failure probability increases as well.

$$\text{Cumulative Failure Probability (Rate)} = 1 - e^{-\text{time} / \text{MTBF}}$$

Intuitively the less one keeps the device *on*, operational and utilized, the lower the chances of its failure.

Realistically, we expect that the usage time for HDDs in the desktop environment is, on average, about 2400 power-on hours/year, corresponding to about 6.5 hours/day. Our expectations for the nearline or mission-critical environments are that the drive will be used 100% of the time (24 hours/day) resulting in 8760 power-on hours/year. Clearly, we expect that the nearline and mission-critical drives will operate under higher usage time stress. Therefore, when HDDs are developed and tested, their design and test protocols are selected in accordance with their anticipated future operational conditions, including time, temperature and workload.

Temperature Stress

High temperature also has a negative effect on the reliability of nearly all electronic and electro-mechanical devices, including HDDs. Failure rate typically increases rapidly with temperature following what is generally referred to as an Arrhenius dependence. Temperature's impact on reliability and MTBF is understood relatively well and is always considered in the drive design and testing process. The rule of thumb is to keep HDDs as cool as possible while remaining within the range specified for the product. A typical operational temperature range for HDDs is from 5°C to 60°C independently of the type of drive selected. Any data center plan for increasing HDD reliability should include efforts to provide efficient cooling.

Workload Stress

Understanding the impact of workload on reliability is somewhat more difficult.

By definition, the primary function of HDDs is to store and retrieve data, storing hundreds of Gbits of data in every square inch of storage surface. They are capable of recording and retrieving data at sustained data rates on the order of 200MB/s or more.

How HDD Workload Impacts TCO

In order to achieve this high recording density and high data throughput, magnetic read and write components are kept at physical separation of several nanometers (1nm = 0.001µm) from a fast-moving rotating media. This is a complex technical design task, necessitating that drives are designed, tested and classified for a specific work environment characterized by the range of usage time and customer workload, among other factors.

Workload is an engineering term used to define the amount of work stress the drive is exposed to during normal operation. For example, Drive A could be reading and writing several GBs of data every day while another drive of the same design, Drive B, could be reading and writing several hundreds of GB of data per day. In this case, we would say that Drive B is operating under much higher workload stress.

In order to get an idea of how much workload is *too much*, let's review three typical scenarios (drives A, B and C):

Let's consider a 4TB Seagate® Enterprise Capacity 3.5 HDD.³ This drive is capable of a sustained data transfer rate of about 175MB/s. Let's imagine three of these drives all operating in similar conditions (and assuming the same server). The first drive (Drive A) is consistently transferring 5MB/s (or transferring an annual average of 158TB/year) while the second one (Drive B) is transferring 10MB/s (an average of 315TB/year). Finally, the third drive (Drive C) is, in this example, transferring 100MB/s (an average of 3150TB/year).

It is easy to see from the above scenarios that Drive B is exposed to 2× higher workload stress than Drive A, and that Drive C has 20× higher workload stress than Drive A.

Assuming linear dependence, the next reasonable conclusions would be to presume that Drive B will have 2× higher failure rates than Drive A, and Drive C will have 20× higher failure rates of Drive A. However, Seagate data suggests that assumption of linear scaling of failure rate with workload is incorrect.

Years of research and experimentation has allowed Seagate engineers to understand the complex effects of workload on drive reliability and to come to the following conclusions:

- Every HDD type has some *safe* threshold of workload that is now defined as the workload rate limit (WRL).
- As long as the workload doesn't exceed the WRL, workload stress has very little to no impact on this product's reliability and failure rate.
- When the workload exceeds the WRL, the reliability of this product will begin to decline.

Therefore, it is very important to understand the workload stress of an actual data center and select HDDs accordingly. Table 1 gives a summary of Seagate recommendations for selecting the most appropriate drives for different data center environments.

Table 1. HDD Recommendations by Workload

Workload	Recommended Product Class	Workload Rate Limit, TB/year
Light	Nearline light	<180
Heavy	Nearline	<550

Assuming that drives A, B and C were all nearline drives, we'd expect on average drives A and B to have very similar reliability (both workloads are below the WRL of 550TB/year). Drive C, on the other hand, with its average workload rate of 3150TB/year, will significantly exceed the recommended WRL for a nearline drive and will be exposed to higher risk of failure.

The table allows data center operators to select the right type of HDD for the right workload. Following the recommendations should ensure the highest possible reliability of HDDs used and lower the long-term TCO.

In Figure 1, one can see that drives A and B belong to the same safe zone and have no failure acceleration due to contributions from workload. Alternatively, Drive C operates well outside the recommended WRLs and could show decline in reliability.



Figure 1. Reliability Risk Due to Workload Stress

3 www.seagate.com/internal-hard-drives/enterprise-hard-drives/hdd/enterprise-capacity-3-5-hdd/

How HDD Workload Impacts TCO



Conclusion

TCO is one of the primary metrics for nearly all data center operations.

HDD reliability can negatively affect overall TCO if the drives used are not matched properly with the data center operating conditions. In addition to usage time and temperature, data center operators are strongly advised to consider the anticipated workload and its effect on reliability when selecting drives.

Seagate provides clear guidelines for proper selection of HDDs for any data center workload environment. Keeping HDDs at a low temperature within the specified range and within their usage time and workload specifications are necessary conditions for long-term drive reliability and improved TCO. Following these guidelines should ensure the best possible drive reliability and the lowest possible cost associated with HDD replacement, maintenance and testing.

www.seagate.com

AMERICAS	Seagate Technology LLC 10200 South De Anza Boulevard, Cupertino, California 95014, United States, 408-658-1000
ASIA/PACIFIC	Seagate Singapore International Headquarters Pte. Ltd. 7000 Ang Mo Kio Avenue 5, Singapore 569877, 65-6485-3888
EUROPE, MIDDLE EAST AND AFRICA	Seagate Technology SAS 16-18, rue du Dôme, 92100 Boulogne-Billancourt, France, 33 1-4186 10 00

© 2013 Seagate Technology LLC. All rights reserved. Printed in USA. Seagate, Seagate Technology and the Wave logo are registered trademarks of Seagate Technology LLC in the United States and/or other countries. All other trademarks or registered trademarks are the property of their respective owners. When referring to drive capacity, one gigabyte, or GB, equals one billion bytes and one terabyte, or TB, equals one trillion bytes. Your computer's operating system may use a different standard of measurement and report a lower capacity. In addition, some of the listed capacity is used for formatting and other functions, and thus will not be available for data storage. Seagate reserves the right to change, without notice, product offerings or specifications. TP648.2-1309US, September 2013