Best Practices When Architecting Data Center Storage for Various Application Intensities

March 2015

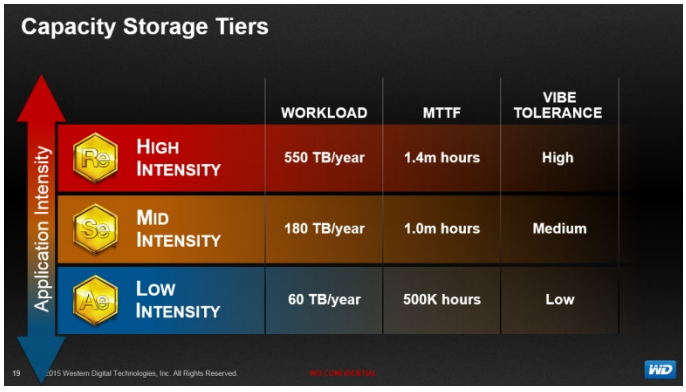
Best Practices When Architecting Data Center Storage For Various Application Intensities

by Alfredo Troche
Western Digital Corporation

Executive Summary

Until recently, storage performance has been afterthought in large enterprise storage system architectures, and was determined more by pricing than need. The more expensive the array you purchased, the more performance the array provided. Historically, the array interface gated sequential access throughput, and the size of the array cache gated I/Os per second (IOPS) for random access. The idea that different applications demand different storage performance profiles is encapsulated in the concept of "application intensity" – the amount of data that a given storage device will have to source (reads) or sink (writes) in a year, as measured in TB/year. While this is in many ways an approximation for a very complex set of processes inside storage media, it captures the factor that has the most impact on media life.

Application intensity is characterized into three "bands": high-intensity (550TB/year), medium intenstity (180TB/yr), and low-intensity (60TB/yr). While the first two of these are generally agreed on, there is some controversy on how "low" low-intensity should be. While the choice of a specific "number" for these levels may seem arbitrary, it has a huge impact on media selection, which further impacts media design, cost, performance, and TCO. For instance, utilizing medium-intensity drives for low-intensity applications doesn't just mean you pay more for the drives when you purchase them – it also means that you consume more energy, and possibly have lower-capacity drives.



Western Digital has worked with a variety of customers to understand their needs, and to develop disk drives that are optimized for the specific needs of high-intensity, medium-intensity, and low-intensity drives. Our Re/Re+ drives are optimized for high-intensity applications. Our Se drive is aimed squarely at the medium-intensity space, balancing power and performance. The Ae drive is focused on "cold storage". To fit this need, the Ae drive has exceptionally low power consumption, and is made to be repeatedly power cycled for systems that keep most of the archive drives in an unpowered state. The result is an unmatched portfolio of media choices that reduce TCO across the board, while at the same time providing the performance required for their specific application band. This is just one of the ways that Western Digital leads the way in storage for today's modern data center.

Background

Storage performance has historically been more of an afterthought in large enterprise storage systems, more related to system capacity and ultimately pricing. The bigger (and move expensive) the array, the more performance that the array provided. While the most demanding applications fifteen to twenty years ago were generally databases, the concept of optimizing storage for particular applications generally revolved around maximizing bandwidth to the storage devices, and seldom impacted media choices themselves (other than "get the fastest drives possible"). Only a handful of applications (video servers and telco database servers are two examples) actually had performance-tuned storage architectures. Even data durability was only differentiated on two levels –tape for archiving, and disk drives for online.

Today, there are massive levels of differentation in storage architectures, and even in storage media. As the amount of data being stored has exploded, "fastest" isn't always the best – faster consumes much more power, for instance, and can have negative impacts on both capacity and disk life. Figure 1 shows how this thought process has evolved over time. As you can see, the number of data "tiers" (a construct that has grown out of heirarchical storage management practices[1]), has increased significantly to accompodate the different speeds of data access required by various applications.

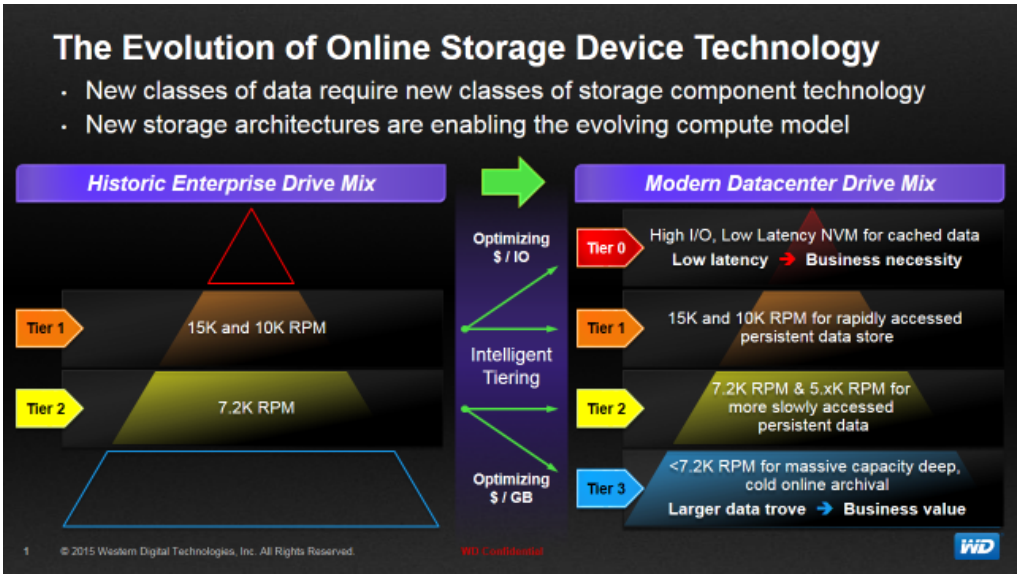


Figure 1: Evolution of Enterprise Storage Tiers

The Concept of Application Intensity and Drive Life

One of the outgrowths of the evolution of enterprise storage tiers is the concept of "application intensity". Application intensity attempts to quantify the demands of various storage tiers on the storage media itself, and how these demands affect the life of the storage media. There are three primary factors that affect drive life are drive workload, mean time to failure (MTTF), and vibration. Of these, drive workload has the greatest impact for a given class of hard disk drives.

Historically, power-on-hours (POH) had been the pacing metric for the life of a hard disk drive. This was largely driven by the motor life, which could be characterized by the classic "bathtub curve". The concept that the bathtub curve embodies is that the probability of failure is high early in the life of the component (this is known as "infant mortality"), and also near the design life of the product. In between these two ends of the spectrum, the probability of failure is relatively low. If this curve is plotted as cumulative failures over time, it looks like the curve shown in Figure 2. As can be seen from the data, the actual failure rate reasonably fits the model. However , as motor design/production processes improved and head fly heights (the nominal spacing between the read/write head and the disk surface) continued to decrease, drive manufacturers noticed that the life curves were significantly impactd by drive reads and writes, as shown in Figure 3. This was the genesis of the concept of drive workload, which is also known as "application intensity".

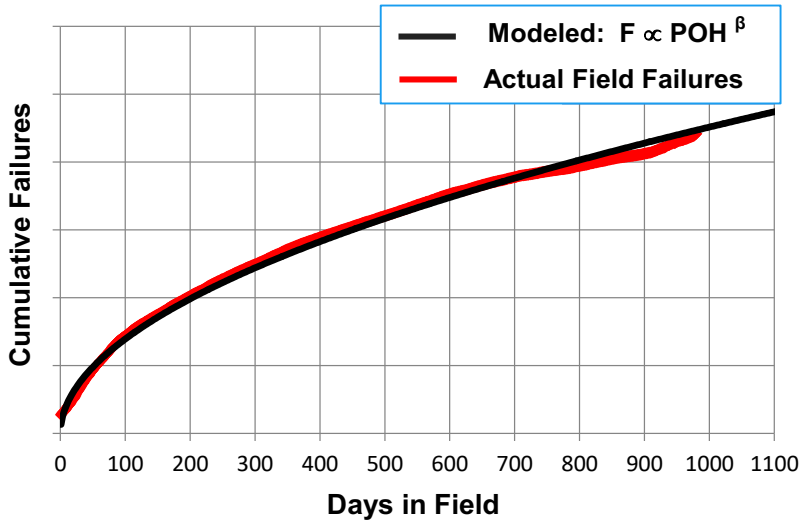


Figure 2: Modeled vs Actual Field Failures Based on POH[2]

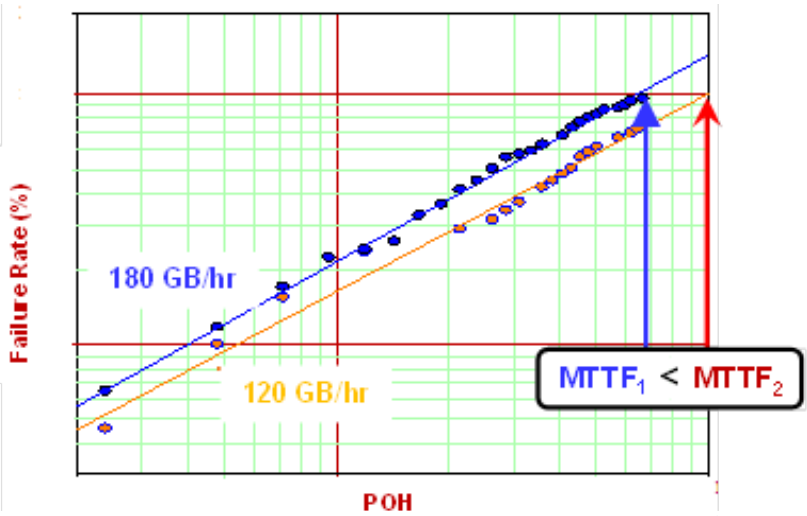


Figure 3: The Impact of Different Access Rates on Drive Life[2]

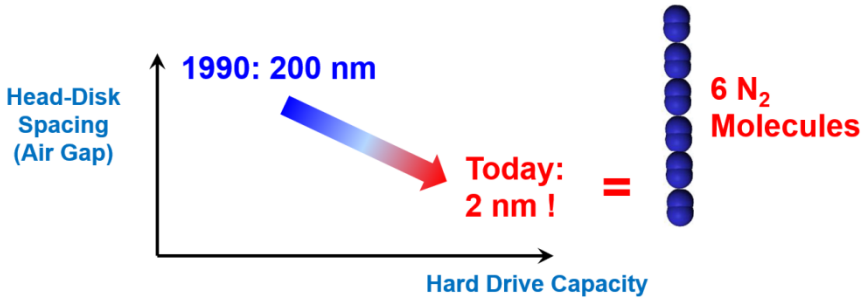


Figure 2: Head-Disk Spacing Reduction for HDDs

Drive workload or aplication intensity measures the amount of access to a drive, regardless of if the access is reads or writes, and is measured in terabytes per year (TB/yr). The reason that this factor now has such a disproportionate impact on drive life is due to the extremely low "fly height" (spacing beween the read/write head the the disk surface) in today's state-of-the-art disk drives, as shown in Figure 4. Over the past twenty-five years, the spacing has been reduced by a factor of one hundred (100!), with a spacing today of only 2nm. To put this in perspective, 2nm is the size of six nitrogen molecules stacked on top of each other. As can be expected, this increases the likelihood of head/platter interaction, which has a negative effect on error rates (and hence useful drive life).

"Bands" of Application Intensity

Once the concept of application intensity is accepted as the primary predictor of drive life, the next obvious questions are "how intense are different classes of applications?", and "how do disk drive manufacturers optimize drives for these different applications?". The industry has characterized application intensity into three "bands" from the perspective of hard disk drives:

High-Intensity: "Always-on", rapidly-accessed persistent storage, with a workload metric of 550TB/yr. Because of the constant access patterns on these drives, the drive enclosures tend to see (relatively) high vibration, which the drives have to be able to tolerate.

Mid-Intensity: Also "always-on", but generally utilized for nearline applications. Mid-intensity drives are accessed roughly a third as much as high-intensity drives (180TB/yr). The need for vibration tolerance in mid-intensity drives is also significantly reduced.

Low-Intensity: These drives are typically utilized for archival purposes, and are often powered off until they need to be accessed. Manufacturers are split on the workload metric for these drives, with target values between 60TB/yr and 120TB/yr.

Choosing The Right Drive For Applications Is Critical to Data Center Optimization

As drive manufacturers optimize drive models for different application intensities, they make tradeoffs across a handful of parameters. Chief among these are:

- **Capacity**, which is influenced by the number of platters, and the areal density (AD) of the data stored on the platter (higher provides more capacity in both cases).
- **Throughput**, which is influenced by the spin speed of the drive, the number of platters, the AD, and the drive's I/O components (higher/better provides more throughput in all cases).
- **Power utilization**, which is influenced by the spin speed of the drive and the number of platters (higher takes more power in both cases).
- **Cost**, which is influenced by the spin speed of the drive, the number of platters, the AD, and the drive's I/O components (higher/better is more costly in all cases).

Obviously, the influencing factors play into multiple parameters simultaneously, meaning you can't build a drive that simultaneously has the highest capacity, highest performance, lowest power consumption, and the lowest price. Thus, drive manufacturers must make tradeoffs to provide the optimum drive for a given application intensity. Also entering the equation are the deployment patterns for drives in each of these application intensity bands: (in general), the further down in application intensity you go, the more drives are used. As an example, low-intensity drives are used in very large numbers in "cold storage" systems which archive vast amounts of data that is seldom accessed. Conversely, high-intensity drives are utilized only for high-performance applications that can justify the cost associated with them. Because of this,

high-intensity drives generally are optimized for performance, while low-intensity drives are optimized for capacity, low power consumption, and low cost, and mid-intensity drives represent a balance between performance, capacity, power consumption, and cost.

Of the three intensity bands, the one with the most "controversy" in defining it is the low-intensity band, with workload metrics from different manufacturers of between 60TB/yr and 120TB/yr. As can be seen from the discussion above, choosing the "right" number has a huge impact on the drive's design targets. If one designs a drive with access numbers in mind that approach 120TB/yr (2/3 of the workload of mid-intensity drives), one may have to increase head fly heights to increase media life, which impacts AD (the closer the head is to the disk, the higher the AD generally is). This forces the drive manufacturers to utilize more platters to achieve the same capacity as a higher-AD drive, which impacts power consumption and cost. From a customer history standpoint, the largest cloud storage providers have seen workload rates that generally are below 50TB/yr, and in some cases are actually below 20TB/yr, in their cold storage pools, which would tend to argue towards a 60TB/yr metric for low-intensity apps.

Western Digital – Leading the Way in Optimized Application Intensity Choices

As can be seen above, utilizing drives that are optimized to the application workload that they are servicing is critical to optimizing the overall cost and performance of the modern data center. This is an area that Western Digital has made significant investment in, and which has resulted in a family of drives that are uniquely matched for the three workload bands, offering the right mix of capacity, performance, power utilization, and cost needed for each workload, as shown in Figure 5, and described below:

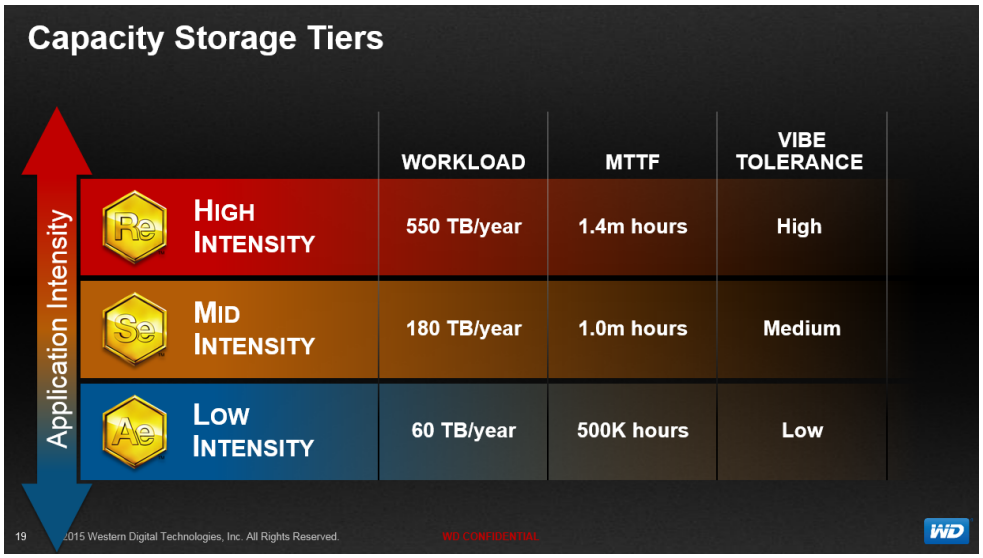


Figure 5: Western Digital Drive Offerings for Various Application Intensities

High Intensity - Western Digital Re/Re+ Drives: The Western Digital Re and Re+ drive families are both optimized for high-intensity applications in today's modern data center. With both drives providing 6TB of capacity today, the Re provides 225MB/s in performance, while the Re+ drive reduces power consumption by 40% (vs the Re drive), while still providing 175MB/s in performance.

Mid Intensity – Western Digital Se Drives: The Western Digital Se drive provides optimum performance for scale-out architectures that do not require the performance levels of the Re/Re+ drives.

Low-Intensity: Western Digital Ae Drives: When it comes to low-intensity, cold/archival storage applications, the Western Digital Ae drive offers very low power consumption, and industry-leading capacity/watt and performance/watt. It is also the first of Western Digital's drives to provide Progressive Capacity, which allows customers to realize progressive increases in capacity without having to wait 18-24 months for a new drive family to come out, and go through expensive drive qualifications prior to use. With progressive capacity, meaningful capacity growth happens every quarter on drives models that are already qualified.

Our multiple drive families for various application intensities is only one of the innovations that Western Digital is developing for the modern data center. Western Digital has the optimum drives for your modern data center, from archival drives with the highest capacity and performance per watt, to high-efficiency and high-performance drives for optimized nearline storage, Western Digital is leading the way. If your need is to store data efficiently, look to Western Digital to help you get there.

---

1 - http://en.wikipedia.org/wiki/Hierarchical_storage_management

2 – George Tyndall, "Why Specify Workload?" Western Digital (2013)

WD, a Western Digital Company
3355 Michelson Drive, Suite 100
Irvine, California 92612
USA