drobo

Automated Data-Aware Tiering

White Paper

Drobo's revolutionary new breakthrough technology automates the provisioning, deployment, and performance acceleration for a fast tier of SSD storage in the Drobo B1200i array. Customers can now take advantage of traditional HDD capacity plus SSD acceleration across all of their applications *automatically* and *immediately*.

Introduction

Storage has traditionally been difficult to understand and manage, especially for non-storage experts. Tasks such as pooling drives, scaling capacity, and protecting from data loss are complex tasks for individuals and businesses that require high-capacity and reliable storage. Also, the explosion of data at every level of business requires more sophisticated digital storage. *The simple fact is that legacy solutions are typically too complex or too limited for current needs.*

Unique Drobo technology delivers breakthrough ease-of-use, affordability, and value so that individual professionals and businesses can have the best storage experience ever. The first Drobo innovation, known as BeyondRAID™, provides a unique combination of simplicity and storage sophistication. BeyondRAID breaks down the barriers of traditional RAID in order to greatly simplify storage deployment and management and to protect data from a drive failure. Automated Data-Aware Tiering works seamlessly with BeyondRAID to optimize data placement automatically based on the type (or "tier") of data being stored.

SSDs Enable Automated Data-Aware Tiering

SSD (solid state drive) support in the Drobo B1200i creates a faster tier of storage for data that is transactional in nature, the type of data that benefits from the faster read and write speeds that SSDs provide. But how do small and medium business (SMB) IT administrators effectively use SSDs to accelerate their applications? They can't. Ether the array is too expensive (costing more than $50,000) or, even if the SMB-focused arrays allow SSDs to be used, they must be in their own pool and separate from the rest of the data. Data must then be manually configured to be sent to the fast volume (SSD) or to the normal volume (hard disk drives, or HDDs).

Physically separated storage tiers force administrators to segregate their traffic and direct it to different types of storage, which is both challenging and time consuming. But how can IT administrators figure out what data is transactional in nature, that is, what data should be sent to the fast volume? Often data has mixed characteristics, making it even more confusing and complicated. To get the most out of expensive SSDs, the solution must be intelligent and fully automatic—*data just needs to go to the right place automatically.*

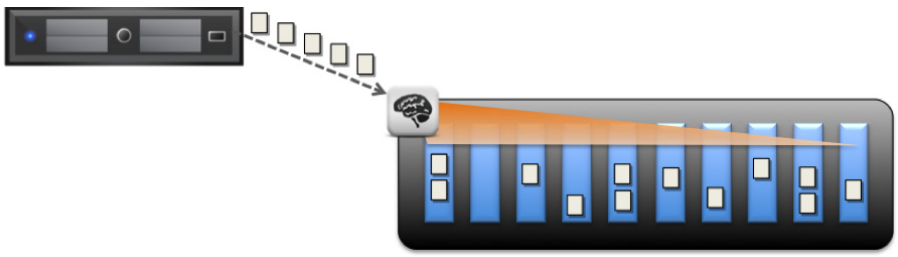

Figure 1:
Data is analyzed in-flight as it enters a Drobo and is placed intelligently on the drives.

What Is Data Aware?

Drobos are data aware, that is, they intelligently handle data for added resiliency, performance, and management. Drobo is data aware in two areas. First the Drobo knows the difference between data and non-data; traditional RAID does not. This allows Drobo to provide benefits that are unique in the industry including an array-level capacity gauge, virtual hot-spare, proportional rebuild times, automatic space reclamation for deleted data, and much more.

Second, Drobo is aware of the difference between streaming and transactional data, which enables instantaneous and intelligent tiering of data. As data is written to the Drobo, it is analyzed in flight and intelligently placed on the correct media (SSD or HDD) to maximize performance. This increases the benefit of SSDs by eliminating the lag that can occur with other tiering solutions that must analyze I/O patterns before placing data on the appropriate tier.

A Drobo is a single pool of drives by design. Users do not need to think about pooling drives or creating RAID groups. Also, drives of different sizes and types can be used in a Drobo, making Drobo ideal for helping SSDs and HDDs work together optimally and automatically in the same chassis.

To support drives of different sizes, the Drobo virtualization layer sits on top of the physical drives (find out more about Drobo BeyondRAID at www.drobo.com/how-it-works/beyondraid). Because Drobos are data aware, data is placed intelligently in the best place on the drives.

Being Data Aware

As data enters the Drobo—even while it is still in flight—it is already being examined, but note that *performance is not impacted*. An additional check determines whether or not the data is transactional in nature. If it is transactional, then the data is written to the faster SSD tier.

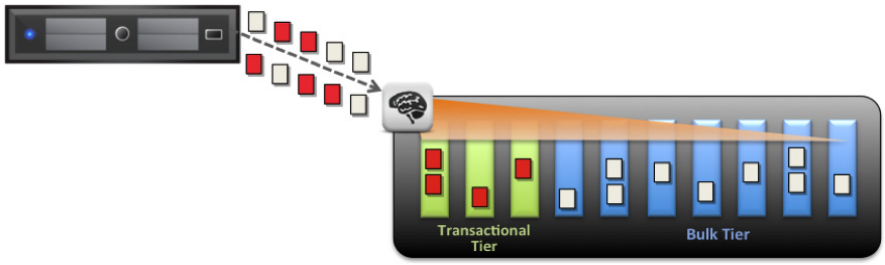

Figure 2:
As data is written to the Drobo, transactional data (shown in red) is written to the transactional tier (SSDs) and the remaining data is written to the bulk tier (HDDs).

Even leading enterprise storage arrays that offer tiering capabilities and typically cost more than $50,000, are not always aware of what type of data being stored. Often data is first written to bulk storage and then analyzed before it is sent to a high-performance tier—and this takes time. *With Drobo, transactional data goes directly to SSDs immediately, which ensures that reads AND writes are optimized.*

Drobo's tiering functionality allows you to store any type of data on the Drobo and the type of data that can by optimized by SSDs is automatically optimized in flight. This means that all data—mixed data types from VMs in a cluster, Exchange databases/datastores, backup data, or files on a file server—automatically goes to the right place.

Streaming vs. Transactional Performance

When streaming data to and from storage (that is, writing and reading files or storing backup data) the data set is large and operations are often sequential. When small bits of data—such as metadata or fields in a database—are sent to and from storage operations are often random. *Therefore, the data and access patterns of streaming and transactional performance require different metrics* from the storage array.

- When *streaming* large amounts of data, the metric MB/s (number of megabytes per second that can be written to or read from storage) is the most important.
- When storing data that is *transactional* in nature, the metric IOPS (number of inputs and outputs per second that can be written to or read from storage) is the most important.

When you are evaluating storage to ensure that it can meet the needs of your applications, you need to take into account both metrics—MB/s and IOPS. The most intelligent way to have the best of both worlds in business-critical environments is to use storage that has automated tiering capabilities.

One Pool, Multiple Tiers

BeyondRAID technology allows the Drobo B1200i to behave as a single pool of storage. Whether you put 3 drives,12 drives, or any number in between, they are all managed as a single pool in the Drobo.

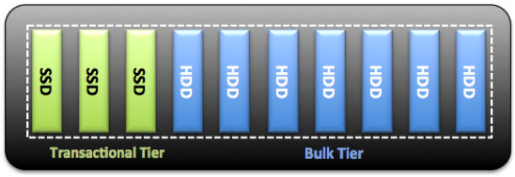

Figure 3:
Two tiers of storage automatically configured in the same pool of storage on the Drobo.

If you have 8 x HDDs and 2 x SSDs or 9 x HDDs and 3 x SSDs, they will all be utilized appropriately without any preliminary configuration. All the administrator has to do is put the drives in the Drobo.

Fully Automated Tiering

After data is written to storage, data patterns often change. Sometimes a large amount of data is written to storage and not read for weeks or even months. And it is not uncommon for a small amount of data to be written and accessed very frequently. Traditional storage arrays treat all data the same way, so if the storage is optimized for streaming performance, then it will most likely not be optimized for transactional performance (see "Streaming vs. Transactional Performance").

With Drobo automated tiering, optimization of data does not stop when the data reaches a transactional or bulk storage tier. As the data stored on the Drobo is being read, it is analyzed to look for patterns over time. If data on the bulk tier is frequently read and begins to look more like transactional data, it will be migrated to the transactional tier. If data on the transactional tier becomes "cold" with very few or no read requests, it will be migrated to the bulk tier. These migrations occur in the background when the storage is not under high load and do not require any administrative interaction.

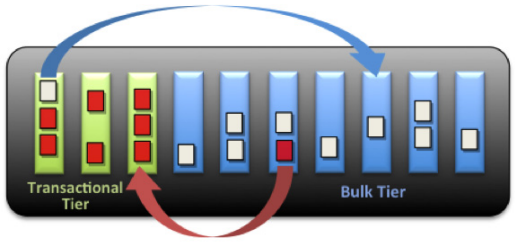

Figure 4:
When data gets "cold" it is migrated to the bulk tier.
When data becomes "hot," it is migrated to the transactional tier.

Tiering Without SSDs

Even without SSDs, Drobo still puts its automatic tiering feature to work. Without SSDs, the advantage of tiering is to eliminate the write penalty when parity data is created. One of the characteristics of BeyondRAID is that a Drobo can automatically have zones of different types. When all HDDs in the Drobo are the same, Drobo lays out zones differently on the transactional tier than it does on the bulk tier.

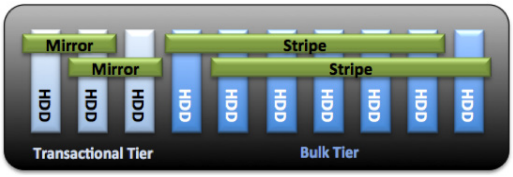

Figure 5:
Drobo creates different zone types on transactional and bulk tiers to tune the storage appropriately for hot and cold data.

For example, mirror zones are optimized for writes and stripes are optimized for capacity. Leveraging BeyondRAID technology, zones are created automatically and altered on the fly so that the Drobo is always optimized for the type of data stored on it.

Just Add SSDs

If your disk pack contains only HDDs, they will be optimized automatically. If your applications need even more performance, all you have to do is introduce at least two SSDs (for single-disk redundancy) or three SSDs (for dual-disk redundancy) to the Drobo B1200i. The SSDs then become the new transactional tier and all of the HDDs are delegated to the bulk tier. This migration occurs automatically, so the only administrative task is to insert new drives into t he Drobo.

NOTE: Introducing SSDs into a Drobo increases both write and read performance because SSDs eliminate drive seek time. This performance boost is critical in database and email environments, which are characterized by frequent read-write requests that are small in nature.

Summary

Drobo was created to solve challenges inherent in storage technologies in SMB and enterprise workgroup environments. Those who were previously familiar with storage assumed that RAID, which allows drives to be aggregated and protected, was the norm. But it is no longer the norm. BeyondRAID is a huge advance for users who do not want to have to learn about storage in order to protect their data and for storage administrators who desire to manage storage more efficiently.

Automated Data-Aware Tiering takes storage automation to the next level, allowing Drobo to automate the once-complex tasks of data protection, capacity planning, and performance tuning (tiering with SSDs).

Choosing the Right SSD

Not all drives are the same and if you do not choose the right one, performance may not meet your expectations.

- Although not true in the past, enterprise-grade drives are now available that use MLC flash, significantly driving down the per GB cost of the drive.
- Enterprise-grade SSDs often have a higher level of overprovisioned space, increasing performance and extending the life of the SSD.
- Enterprise drives are tuned for extended use to handle more write cycles and use reclamation features such as TRIM and UNMAP to optimize capacity and performance.
- Similar to HDDs, SSD drives are available with SATA or SAS connectivity. Only enterprise-grade SSDs are available with SAS connectivity . (The Drobo B1200i supports both SATA and SAS drives.)

For these reasons, use enterprise-grade SSDs and HDDs in your Drobo B1200i for best performance.

drobo
2460 North First Street, Suite 100, San Jose, CA 95131 • www.drobo.com • 1.866.97.DROBO

WP--0109-00 • March 2012