

# Lab Validation Report

## Permabit Albireo

### Empowering Unified Deduplication

*By Brian Garrett*

August 2011

## Contents

Introduction .....	3
Background.....	3
Introducing Permabit Deduplication Technology.....	4
ESG Lab Validation .....	6
Capacity Savings .....	8
<i>What the Numbers Mean</i> .....	8
Resource Efficiency.....	10
<i>What the Numbers Mean</i> .....	11
Scalability and Performance .....	12
Fault Tolerance .....	14
Ease of Integration .....	15
Maturity.....	17
ESG Lab Validation Highlights .....	18
Issues to Consider .....	18
The Bigger Truth .....	19
Appendix.....	21

### ESG Lab Reports

The goal of ESG Lab reports is to educate IT professionals about emerging technologies and products in the storage, data management and information security industries. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by Permabit.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. Copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482.0188.

## Introduction

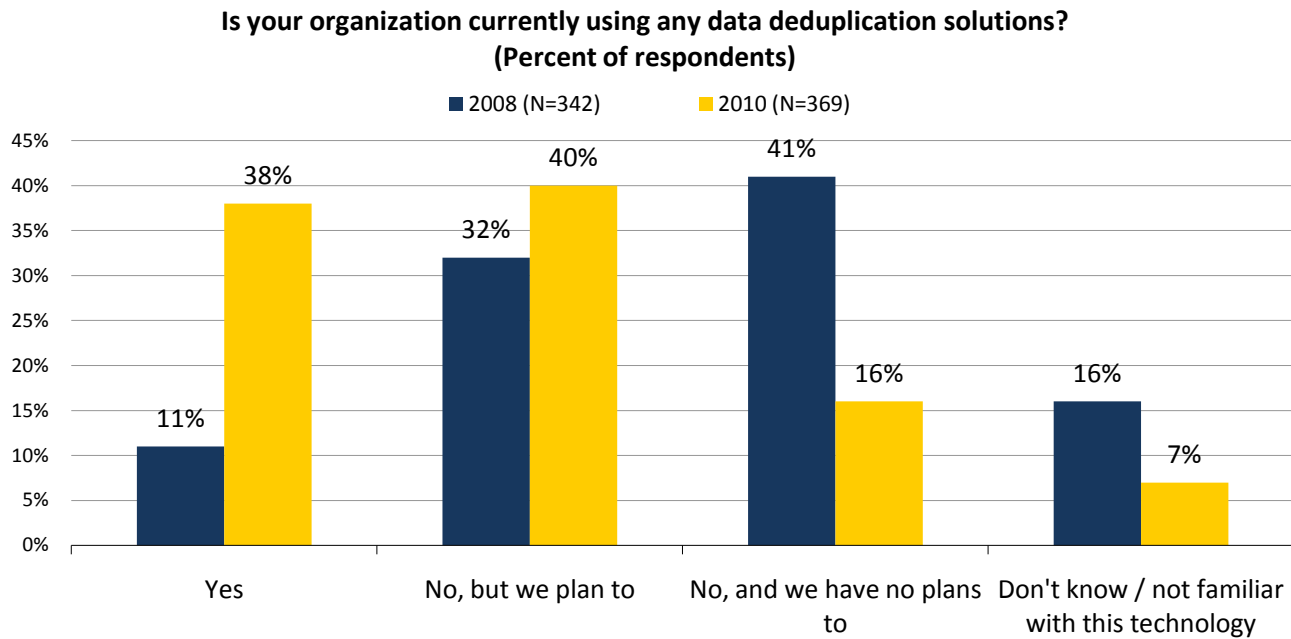
Permabit’s field-proven data deduplication engine is now available as a software library and software development kit (SDK) which enables unified data deduplication. This updated ESG Lab Validation report documents the results of hands-on testing of the Permabit Albireo SDK which focused on: ease of integration into existing solutions, capacity savings that can be achieved with real-world data, resource usage, and fault tolerance. This report also examines the progress that has been made over the past 18 months with a focus on grid scalability and performance improvements.

## Background

Back in 2001, when Data Domain was founded and EMC purchased Belgian startup FilePool for its Centera product line, few in the industry had heard of data deduplication. Fewer still were aware that Permabit, founded in 2000, was hard at work developing scalable data deduplication technology. Today, it’s hard to find anyone in the storage industry who hasn’t heard about data deduplication.

ESG research indicates that there has been a large jump in deduplication adoption in the data protection market. As shown in Figure 1, adoption in 2008 was 11%, and grew to 38% of respondents in 2010 with more concentration in enterprises (45%) than in midmarket organizations (29%).<sup>1</sup> Further, the number of organizations that have adopted, or plan on adopting, deduplication has grown to 78%. ESG expects that interest and adoption in deduplication will increase significantly over the next five years as it reaches mainstream market adoption.

Figure 1. Widespread Interest in Data Deduplication



Source: Enterprise Strategy Group, 2010.

Interest outside of the backup and recovery market is growing as well. In primary storage, data deduplication has been added to network attached storage (NAS) systems from a number of emerging and market leading storage vendors (e.g., NetApp FAS, EMC VNX). As a matter of fact, ESG research indicates that data reduction technologies are quickly rising on the list of most important features and attributes for storage solutions. This survey of decision makers within enterprise-class organizations indicates that 27% would not purchase a scale-out storage solution without data reduction and 51% would strongly prefer a solution with this attribute.<sup>2</sup> Given the dramatic data

<sup>1</sup> Source: ESG Research Report, [2010 Data Protection Trends](#), April 2010.

<sup>2</sup> Source: ESG Research Report, [Scale-out Storage Market Trends](#), December 2010.

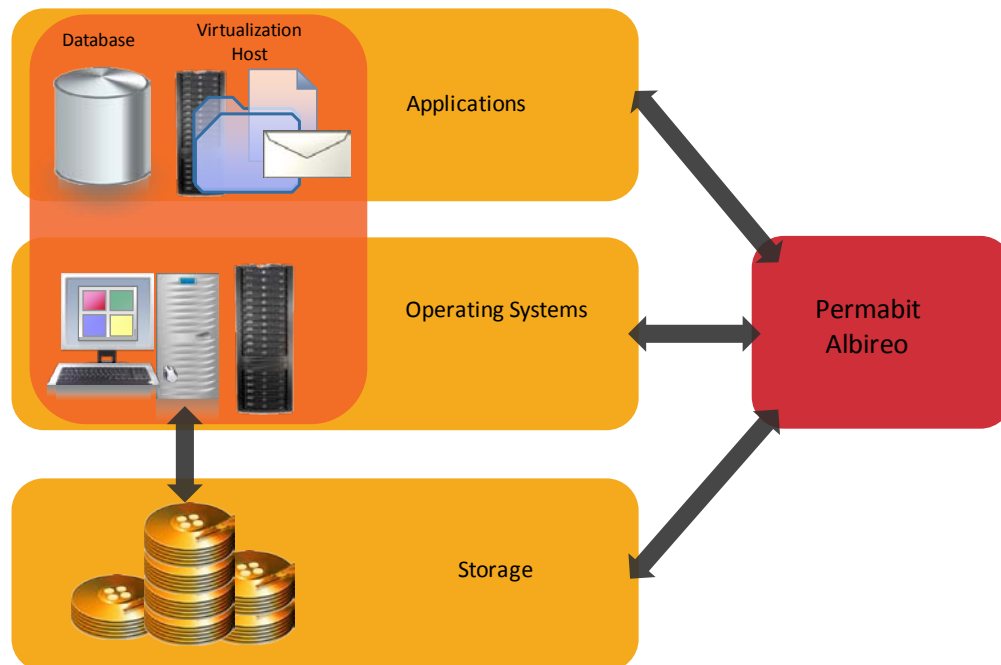
reduction benefits that have been realized with existing data deduplication technologies, ESG believes that it's only a matter of time before deduplication becomes a core component of every tier of storage including primary, backup, archive, block, file, unified, solid state disk, and cloud, *and* moves up the stack into operating systems, applications, and databases.

### Introducing Permabit Deduplication Technology

Permabit's field-proven data deduplication engine is available as a software development kit (SDK), called Albireo, which enables unified data deduplication. The Albireo SDK is unique in its ability to provide a wide variety of deduplication services. The Albireo SDK can be used to:

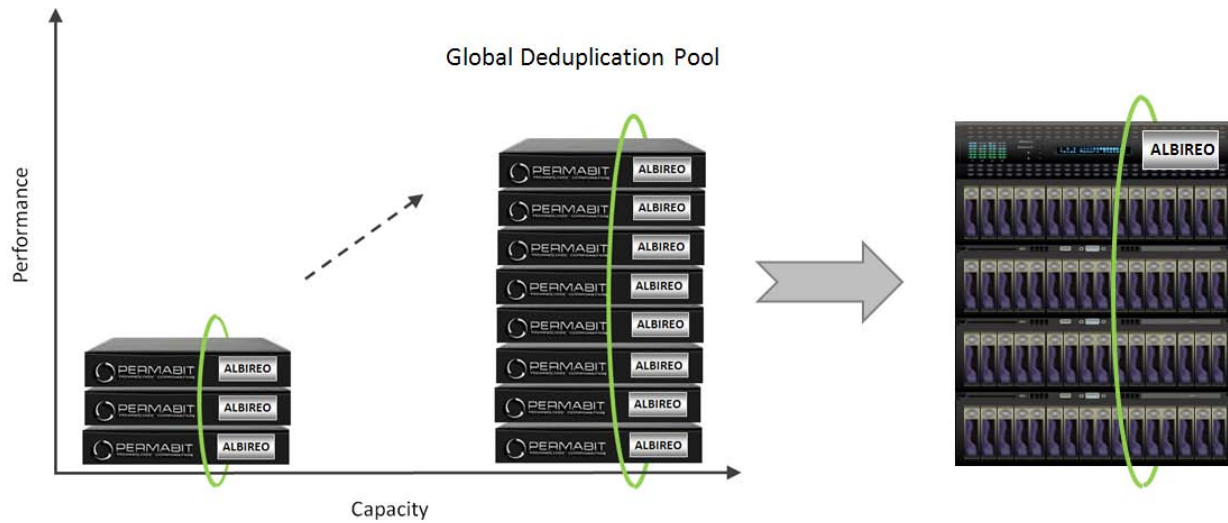
- Reduce capacity within file, block, or unified storage systems (e.g., NAS appliances or FC disk arrays).
- Process objects at the sub-file or single instance storage (SIS) level.
- Remove duplicates in real-time (inline) or later after data has arrived (post-process) or a hybrid approach which Permabit refers to as "parallel."
- Scale from a single storage controller to a cluster of storage controllers or a cluster of deduplication appliances communicating with a storage solution over an industry standard Ethernet interface.
- Provide deduplication services that are "end-to-end and everywhere" in the IT stack—from applications and databases running on physical and virtual hosts through operating systems and hypervisors to storage solutions.

*Figure 2. End-to-End and Everywhere Deduplication with Permabit Albireo*



Permabit was founded in 2000 by MIT engineers with a goal of developing a scalable enterprise-class archive product with built-in data deduplication. That product, now known as the Permabit Enterprise Archive, is shown toward the left in Figure 3. Industry standard servers are arranged in a grid, with each server acting as either an access or storage node. The storage nodes are packed full of high capacity SATA drives and presented as a network attached file system (NAS). Servers can be added to the grid for increased performance and capacity.

Figure 3. Introducing Permabit Albireo Deduplication Advisory Services



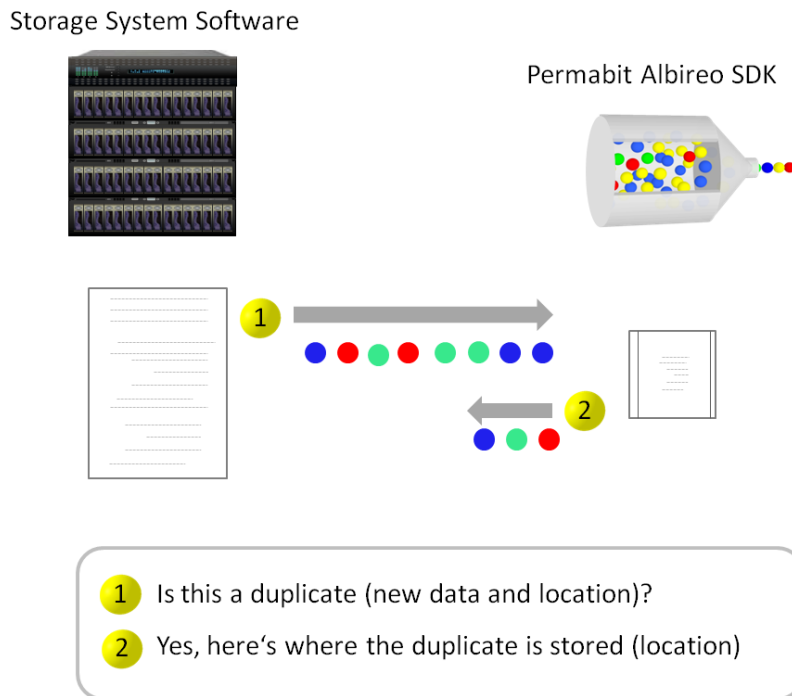
In the Permabit Enterprise Archive product, a global pool with deduplication technology is implemented within a grid of servers. Permabit has extracted this core deduplication technology to create the Albireo SDK. Albireo deduplication advisory services are accessed through an application programming interface (API) provided by Permabit. A storage system with Albireo running within a single controller attached to a number of drive enclosures is shown toward the right in Figure 3.

Albireo indexes chunks of data and provides advisory notifications when duplicates are detected. The software running in the storage system decides whether or not to take the advice. If the advice is taken, it is the responsibility of the storage system software to update references to duplicate data. The Permabit Albireo API supports both block and stream-based access methods. The stream-based API provides content-aware segmentation to optimize deduplication based on file type. Duplicate advisory services can be provided synchronously, or asynchronously using a registered callback. Taken together, the Albireo SDK provides a flexible and powerful array of deduplication services.

## ESG Lab Validation

ESG Lab initially evaluated the Albireo SDK during two days of hands-on testing at Permabit headquarters in Cambridge, Massachusetts in late 2009 and performed an audit of resource and performance improvements in 2011. The evaluation began with an overview of how Albireo deduplication advisory services are used within an existing storage system. As shown in Figure 4 the software running within a storage system uses an API call to send new data and addressing information to Albireo. Albireo ingests the data and performs a SHA-256 hash. A two-stage lookup (memory and, if needed, disk) is performed to see if the data has already been stored. If the data is a duplicate, the API returns the location of the pre-existing data. If the data is new, Albireo remembers the SHA-256 hash and the chunk's location so that it can advise of duplicate data in the future.

Figure 4. Permabit Albireo Deduplication Advisory Services

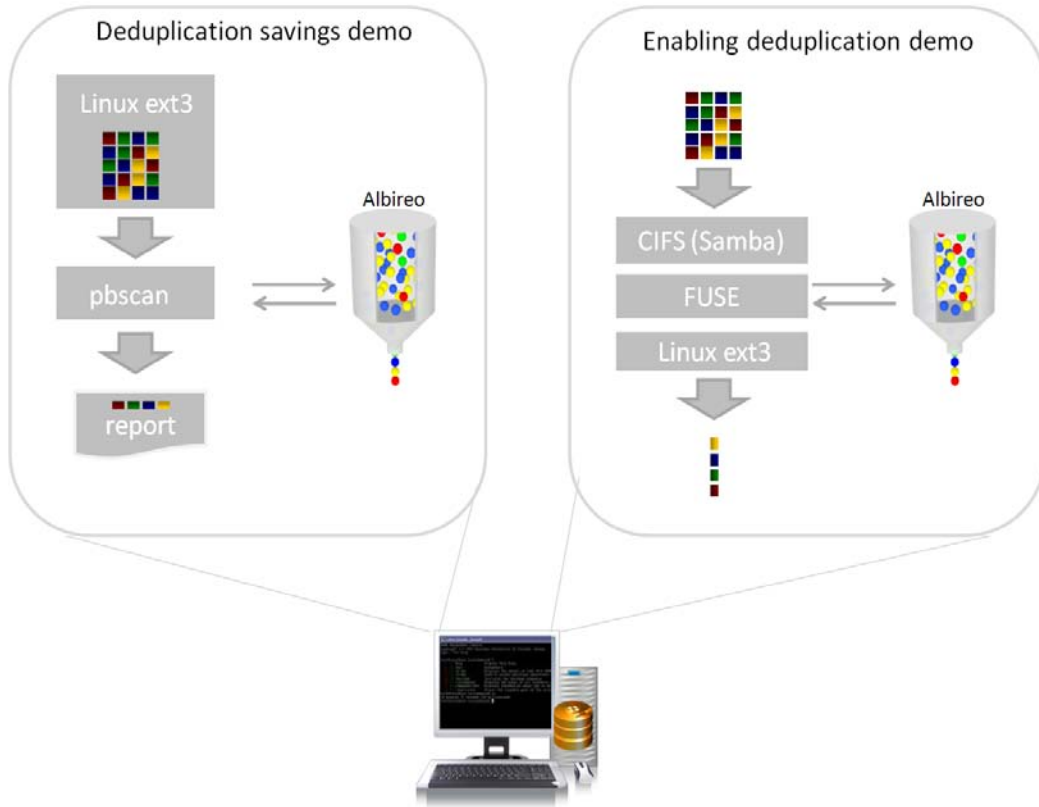


ESG Lab evaluated the Albireo SDK using a pair of programs developed by Permabit:

- **PBSCAN:** This utility uses Albireo advisory services to determine the savings that can be achieved with deduplication. The program is routinely used at prospective customer sites to determine the benefits of deduplication with real-world data sets.
- **DD2FS:** An open-source user-space file system was modified to illustrate the ease of integrating Albireo into an existing file system.

The test program was implemented as a single process with the Albireo engine running as a separate service. The pbscan utility was implemented with multi-threads working in parallel. As shown in Figure 5, the test bed used for this phase of the evaluation consisted of a single quad-core Intel Xeon 3 GHz processor with 4 GB of RAM. For test purposes only, and to more easily discern the resource utilization by Albireo, all but one of the processor cores were disabled in BIOS.

Figure 5. The ESG Lab Test Bed



While ESG Lab testing was performed with a single processor core, it should be noted that the extreme scalability of the Albireo engine has been proven in production customer environments. Later in this report, we'll take a look at how Permabit's underlying deduplication technology is routinely deployed within a grid of multi-core servers.

## Capacity Savings

ESG Lab used the Albireo-enabled pbscan utility and real-world application data collected from Permabit’s production IT servers to evaluate the capacity savings that can be achieved with Permabit deduplication advisory services. The results for two data sets are summarized in Figure 6 and Table 1.

Figure 6. Capacity Savings

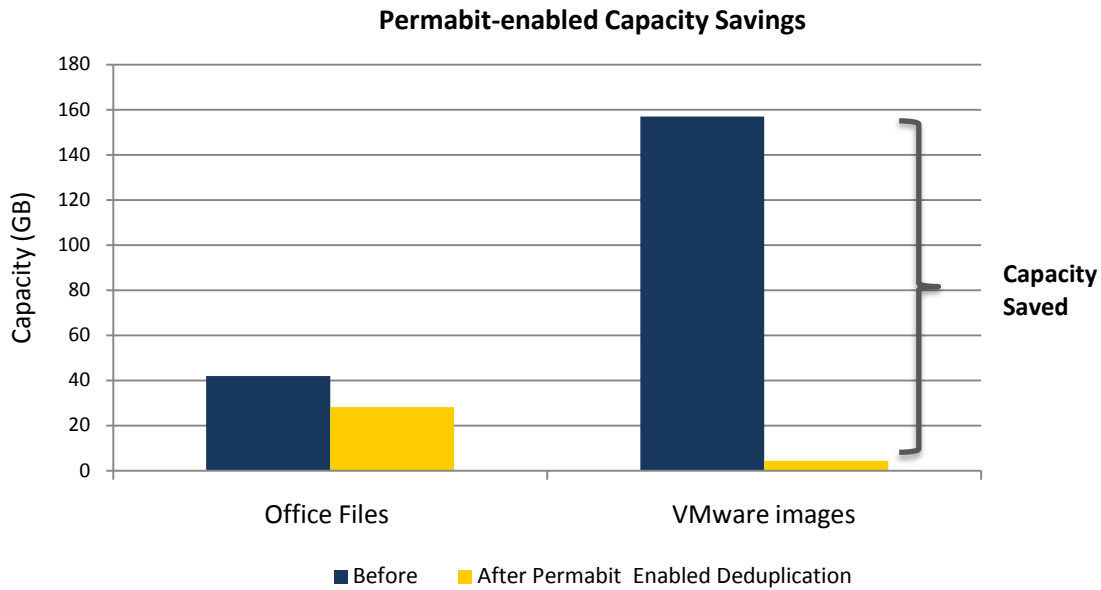


Table 1. Capacity Savings

Data Type	Before (GB)	After (GB)	Deduplication Rate	Deduplication Ratio
Office Files	42.0	28.3	33%	1.5
VMware images	157.0	4.3	97%	36.2

### What the Numbers Mean

- A 43 GB set of common office productivity files (e.g., documents, spreadsheets, presentations) was reduced by 32.67%.
- Four VMware virtual server images with a total capacity of 157 GB were reduced to only 4.3 GB. VMware virtual server images are often highly redundant, especially when each virtual machine is running the same guest operating system as was the case for this test. In this example, 97% of the capacity required for the VMware images can be saved with Albireo-enabled deduplication.

The experiment was repeated for a pair of Microsoft Hyper-V images and a week’s worth of Microsoft Exchange backup images. One of the backup images was a full backup; the other images were incremental. The results for all of the data types evaluated by ESG Lab are summarized in Figure 7 and Table 2. Note that the savings are shown as a deduplication ratio instead of a percentage of capacity saved.

Figure 7. Deduplication Ratios

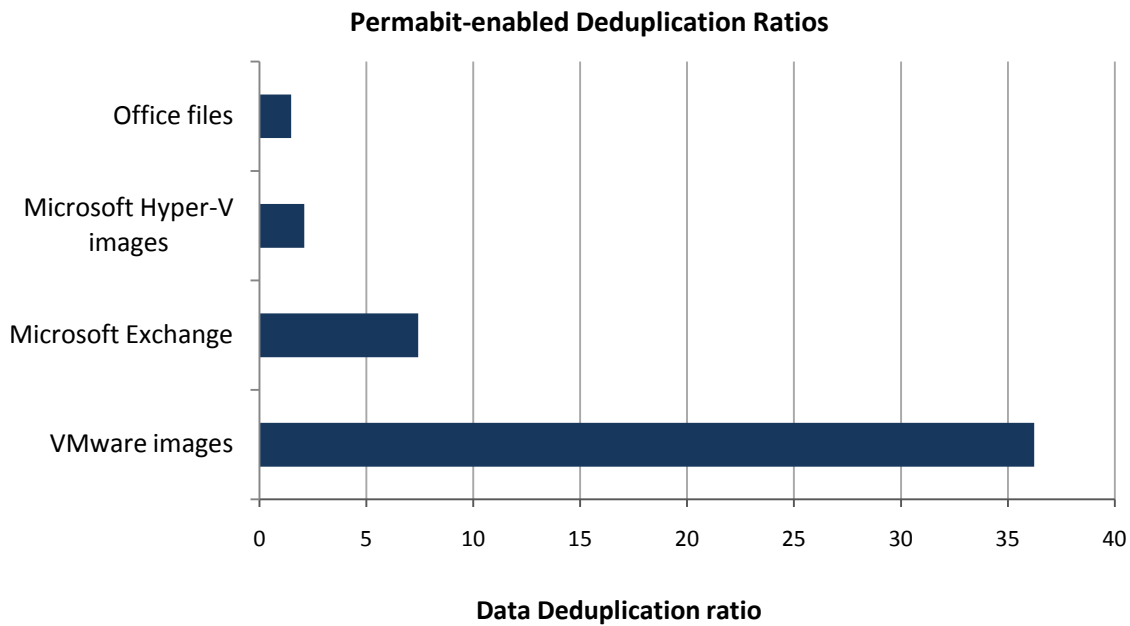


Table 2. Deduplication Ratios

Data Type	Before (GB)	After (GB)	Deduplication Rate	Deduplication Ratio
Office Files	42	28.3	33%	1.5
Microsoft Hyper-V images	5.3	2.5	52%	2.1
Microsoft Exchange	2,203	501	86%	7.4
VMware images	157	4.3	97%	36.2

**What the Numbers Mean**

- A pair of Microsoft Hyper-V images was reduced by a factor of 2.1:1.
- A single week’s worth of Microsoft Exchange backups (weekly full, daily incremental) was reduced from 2,203 GB to 501 GB, providing an excellent deduplication ratio of 7.4:1.
- A deduplication ratio of 36.2:1 was recorded for four VMware images.

**Why This Matters**

Data proliferation is a challenge for IT professionals within organizations of all sizes. By eliminating redundant data, data deduplication can significantly reduce capacity requirements, which reduces the cost of storing and protecting data.

ESG Lab testing with real-world application data has confirmed that Permabit Albireo deduplication advisory services can be used to reduce capacity requirements for primary storage arrays, disk-based archives, and backups. ESG Lab recorded an outstanding deduplication rate of 97% (36.2:1) for four VMware virtual server images.

## Resource Efficiency

Identifying duplicate data is a resource-intensive operation. First, data needs to be fed into a deduplication engine. Moving a lot of data can consume a lot of bandwidth. Next, the engine needs an algorithm which can be used to quickly and accurately identify duplicates. Most deduplication solutions use cryptographic hashing functions to identify duplicates, but hashing a lot of data can consume a lot of CPU horsepower. Last, but not least, the deduplication solution needs to maintain an index of previously processed data to find and keep track of duplicates. Most deduplication solutions use a two stage lookup, with the first lookup occurring in memory and the second occurring on disk. Indexing lots of possibly duplicate data can consume a significant amount of memory. Each of these resource issues can have an effect on the overall performance of a deduplication solution.

Permabit supports two modes of indexing in the latest version of Albireo. Traditional dense indexing is a full chunk index approach that keeps all index entries available in memory to determine if that data chunk has previously been seen. At scale, this approach limits the efficiency of indexing and degrades performance. If the amount of memory is limited, then the ability to scale out is also limited by the amount of available RAM. Sparse indexing improves memory efficiency using a sampling technique that exploits the inherent locality within data streams. It addresses large scale (e.g., hundreds of terabytes) data chunk lookup bottleneck problems as it avoids the limitations of traditional dense indexing. Permabit has developed a hybrid approach using both sparse and dense indexing with a goal of increasing overall deduplication efficiency by 10X or more with only 0.1 bytes of RAM required for each chunk of indexed data.

ESG Lab’s analysis of CPU and memory efficiency began with a review of Permabit’s patented two-stage deduplication detection algorithm.<sup>3</sup> The patent describes a highly efficient two-stage index residing in memory and on disk. This novel approach uses a combination of bit sampling and byte differencing to provide a first stage memory lookup that executes very quickly, consumes very little memory, and has a very low probability of false positives.

A 42 GB set of office productivity files (documents, spreadsheets, presentations, PDFs, etc.) was processed by an Albireo-enabled deduplication utility. The utility was used to quantify the speed and resource overhead of Permabit deduplication advisory services. The utility opened and passed all of the data within each file to the Albireo API. The Albireo API was used to identify and keep track of chunks of data in a Permabit index of duplicate candidates.

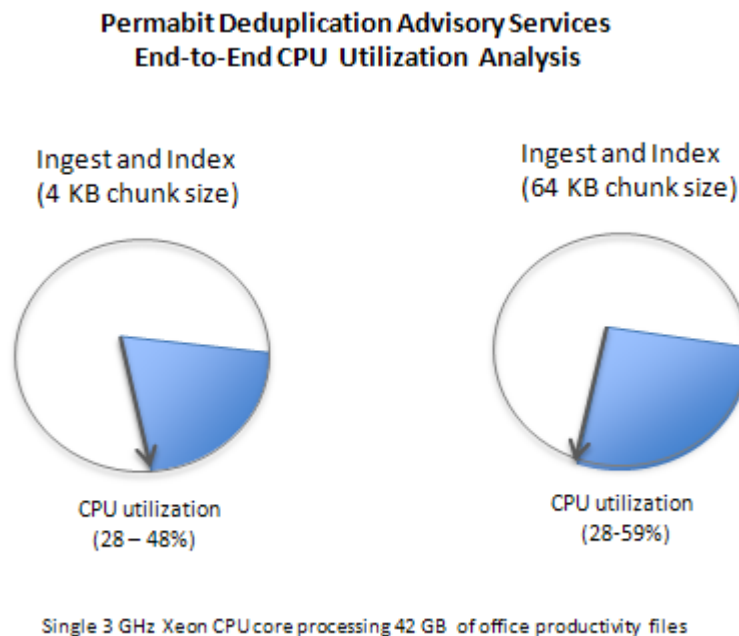
Linux utilities were used to record CPU and memory utilization with Albireo services running on a single Intel Xeon processor. Trace output was used to record the average latency for each Albireo deduplication fingerprint API lookup. In addition, trace data was used to record the end-to-end latency associated with Albireo-enabled deduplication including the SHA-256 ingest, the Albireo fingerprint API, the first pass memory index, and a second order index operation for likely duplicate candidates. The results are summarized in Table 3 and Figure 8.

*Table 3. Albireo Resource Usage*

Chunk Size	Fingerprint Latency	End-to-End Latency	RAM	End-to-End CPU
4 Kilobytes	9-17 microseconds	43 microseconds	4.7 MB	28% - 48%
64 Kilobytes	9-17 microseconds	480 microseconds	0.38 MB	28% - 59%

<sup>3</sup> US Patent 7457813, Storage System for Random Blocks of Data, Nov 25, 2008.

Figure 8. End-to-End CPU Utilization



### What the Numbers Mean

- The Albireo fingerprint API delivered fast in-memory deduplication lookups that completed in 9 to 17 microseconds.<sup>4</sup>
- A SHA-256 hash and index with a 4 KB chunk size incurred only 43 microseconds of latency. That's less than 1% of the overhead associated with a typical disk access latency of 5 milliseconds.
- Albireo consumed less than 3.5 bytes of RAM per index entry with traditional dense indexing and 0.1 bytes per entry with sparse indexing.
- A sparse index overhead of 0.1 bytes and a 4 KB chunk size allows deduplication up to 40 TB of data with 1 GB of RAM and up to 2.56 PB of data with 64 GB of RAM.
- Albireo-enabled deduplication, including SHA-256 hashing, consumed approximately half (28% to 59%) of a 3 GHz Xeon processor core.
- A larger chunk size consumed slightly more CPU. This is due to the fact that bigger chunks of data were passing through the CPU-intensive SHA-256 algorithm.

### Why This Matters

Data deduplication is a resource-intensive operation that can have a dramatic impact on the overall cost and performance of a storage solution. ESG Lab is confident that resource-efficient Permabit deduplication advisory services can be used to architect a cost-effective solution that provides a virtually limitless pool of globally deduplicated capacity using industry standard server hardware.

<sup>4</sup> Confirmed via a review of traces of code execution through the in memory index code path.

## Scalability and Performance

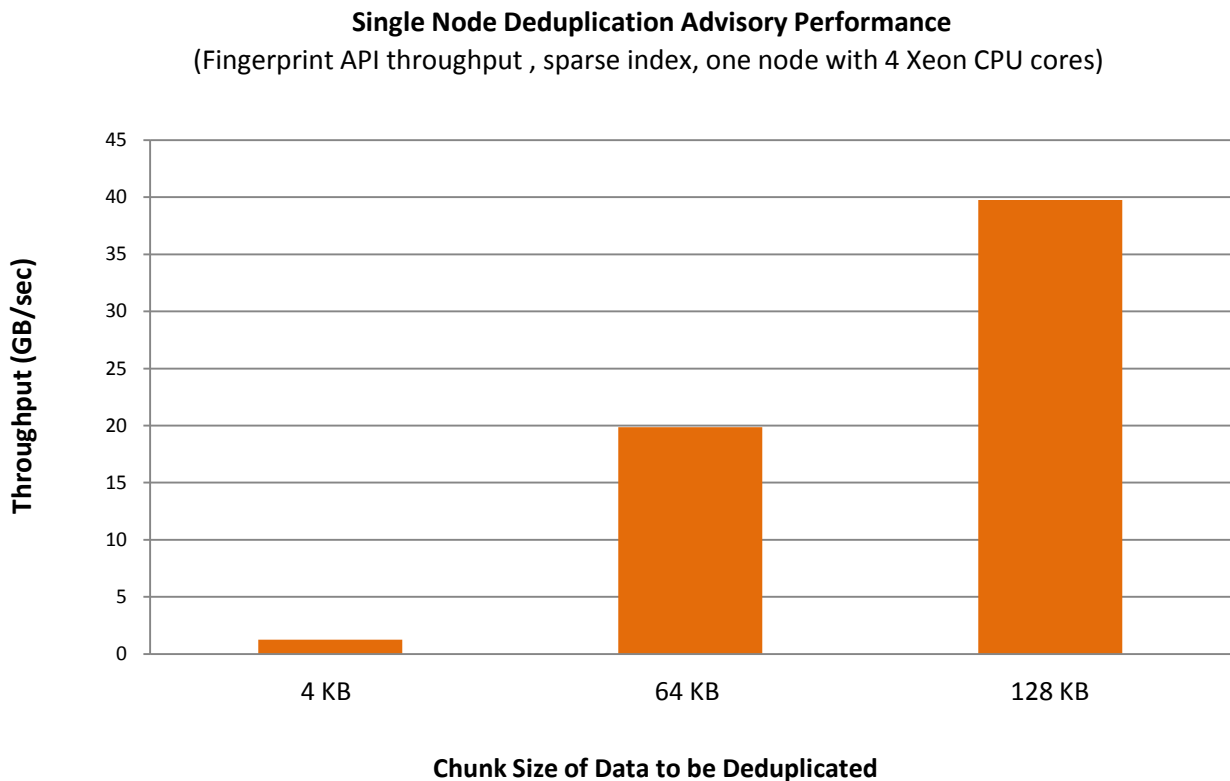
Permabit’s global deduplication algorithm is designed to run within a grid for maximum performance and capacity scalability. Permabit currently supports up to 4.6 PB (4,600 TB) in a single pool of capacity. The Permabit Archive product is continuously tested with Albireo deduplication services running on a grid of three or more servers. A typical entry-level grid at a customer site, comprised of access and storage nodes, has deduplication services running on 11 nodes. The largest grid deployed in production at a customer site is 38 nodes.

The performance scalability of a scale-out deduplication solution is dependent on its ability to handle two resource intensive tasks: hashing and index lookups. The Albireo architecture was designed to maximize the efficiency and scalability of both. The Albireo architecture can support cryptographic hashing (e.g. SHA-256) distributed over multiple CPUs or offloaded to special-purpose hardware, and delivers scalable index lookup performance using an engine that can be deployed across multiple CPUs within a single server or a grid of Ethernet connected servers.

ESG Lab audited the results of nightly regression tests to assess the performance scalability of Albireo index lookups. A Permabit developed utility was used to process data that was deterministically unique (i.e., no duplicates) during this test. This type of data was used with a goal of maximizing the stress on the Albireo deduplication advisory engine.

The performance of the Albireo fingerprint API was tested with 4 KB chunks on a single server. The number of chunks processed per second was multiplied by the chunk size to calculate the effective throughput of the Albireo deduplication advisory engine. Similar calculations were performed to assess throughput when processing 64 KB and 128 KB chunks of data. As shown in Figure 9, increasing the chunk size increased deduplication index lookup throughput to 39.74 GB/sec on a single server with four Intel Xeon processor cores.

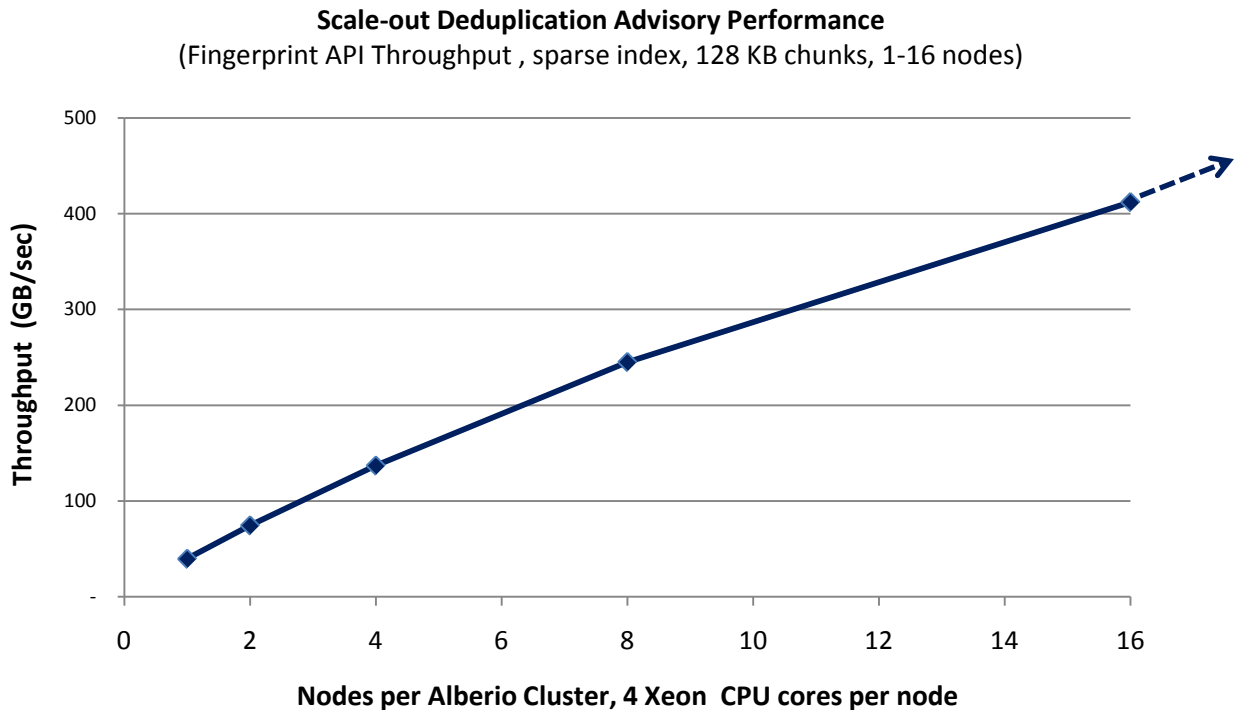
*Figure 9. Albireo Deduplication Advisory API Performance Analysis*



These results not only highlight the efficiency and speed of the deduplication engine running on a single server, they also demonstrate the power of supporting a programmable chunk size. It also provides the flexibility to choose between maximizing deduplication rates (smaller chunks) or minimizing index overhead and maximizing the throughput of the deduplication engine (larger chunks). A programmable chunk size can also be used to simplify integration with an existing architecture.

Similar tests were performed as nodes were added to an Albireo grid. As shown in Figure 10, performance scaled in near linear fashion to 412 GB/sec for a sixteen-node grid processing 128 KB chunks of data.

*Figure 10. Albireo Deduplication Advisory API Performance Analysis*



**Why This Matters**

If a deduplication solution can't scale to provide the capacity or performance required to meet the needs of the business, then costs are incurred and capacity savings are lost as multiple islands of deduplication are deployed within an organization. These islands also increase the ongoing costs associated with deploying and maintaining a deduplication solution.

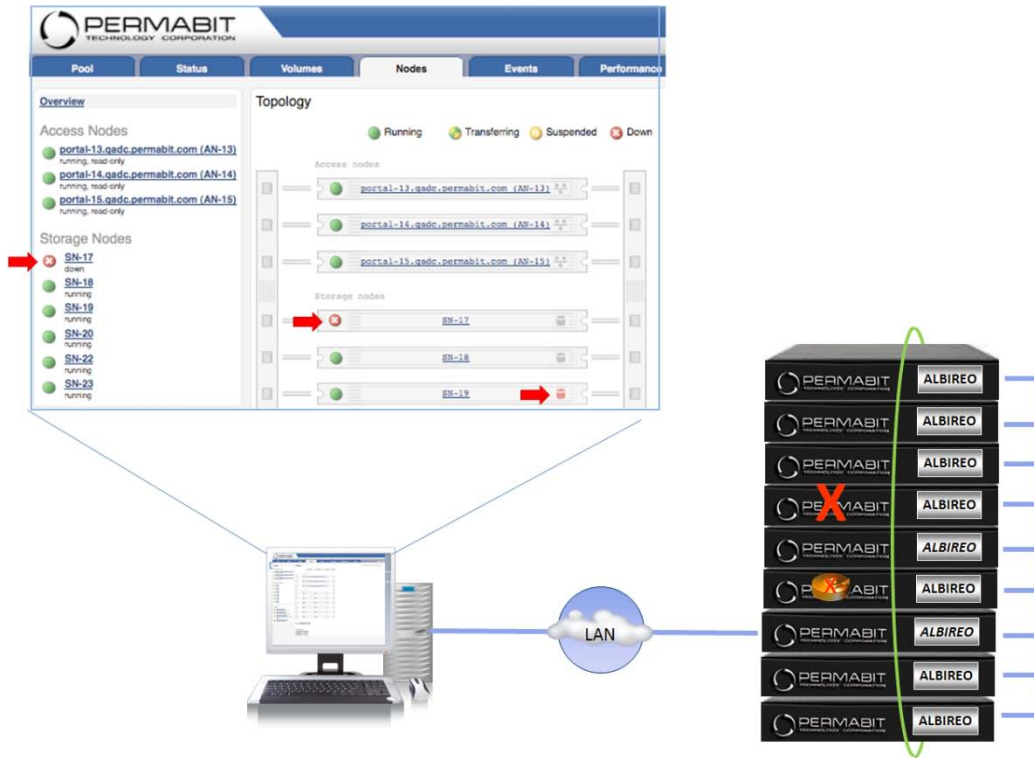
ESG has confirmed that the algorithms accessed via the core Albireo fingerprint API are extremely fast, efficient, and scalable. A sixteen node grid of servers, each with a quad-core Intel Xeon processor processing Albireo fingerprint API calls, delivered an outstanding deduplication lookup rate of 412 GB/sec.

ESG Lab's experience with nearly all of the vendors offering data deduplication solutions indicates that scaling a large pool of global deduplication that resides within multiple storage controllers or servers is a difficult task that can take years to complete. ESG Lab has confirmed that the field proven deduplication technology at the heart of the Permabit Albireo SDK has been deployed on a grid of up to 38 servers in production environments and validated on a grid of 96 servers in Permabit's test lab.

## Fault Tolerance

ESG Lab performed a series of tests on a nine node Permabit Enterprise Archive system to determine whether Albireo deduplication advisory services continue running after multiple hardware failures. A long running directory level file copy operation was started. As shown in Figure 11, a node was powered off and a drive was removed on a different node. The system remained available and the copy operation completed without error.

Figure 11. Validating Permabit Fault Tolerance



### Why This Matters

Errors within a deduplication solution can cause application downtime or data corruption. These impact end-user satisfaction and productivity—and, in the worst case, lead to a loss of revenue. As the size of a scale-out deduplication system grows, the risk and potential impact of downtime increases dramatically.

ESG Lab has confirmed that the scale-out Albireo architecture running on multiple server nodes can be used to deliver fault tolerant scale-out deduplication services. An error injection test by ESG Lab proved that deduplication services remain available after a drive and a server failure.

## Ease of Integration

ESG Lab evaluated the ease of adding deduplication to an existing storage solution using the FUSE open-source user-space file system framework.<sup>5</sup> The FUSE file system is built over the native ext3 Linux file system. There's no need to patch or recompile the kernel. Permabit created a file system using FUSE named dd2fs. The dd2fs file system was modified to use Albireo to identify and eliminate duplicates. Six Albireo API calls and 52 lines of supporting code were added to the 1,563 line FUSE demonstration program. Synchronous (inline) and asynchronous (post-processing) deduplication modes were demonstrated.

As shown in Figure 12, the `update_block` function passes data to be written to the Albireo `uds_index_block` API. A return value of 1 indicates that the block is a duplicate and that it can share storage with the canonical chunk. After a bit of bookkeeping and error checking, the storage associated with the duplicate block is freed. Other than this key function call, the majority of the Albireo-related code changes were isolated to initialization, shutdown, and handling callbacks when operating in asynch/post-process mode.

Figure 12. Integrating Albireo

FUSE demo, `update_block()`:

```
void
update_block(struct dd2fs_context *ctx,
             struct dd2fs_inode *inode,
             uint32_t blkno)
{
    struct dd2fs_block_metadata *block_metadata = get_block_metadata(ctx);
    uint32_t blkno = inode->blocks[blkno];
    struct index_metadata new_metadata = {
        .blkno = blkno,
        .generation = block_metadata[blkno].generation
    };
    uds_block_address_t *address = malloc(sizeof(uds_block_address_t)
                                         + sizeof(new_metadata));
    uds_block_segment_t segment = {.len = BLOCK_SIZE,
                                   .block_address = address};
    uds_block_segment_t canonical_segment;
    uds_chunk_name_t *chunk_name;
    int ret;

    address->size = sizeof(new_metadata);
    memcpy(address->data, &new_metadata, sizeof(new_metadata));

    ret = uds_index_block(ctx->uds,
                         &segment,
                         find_block(ctx, blkno),
                         &canonical_segment,
                         &chunk_name);

    free(address);
    if (ret) {
        size_t old_metadata_len = canonical_segment.block_address->size;
        struct index_metadata *old_metadata;

        if (old_metadata_len != sizeof(*old_metadata)) {
            errx(1, "unexpected metadata return length: expected %u, got %u",
                sizeof(*old_metadata), old_metadata_len);
        }
        old_metadata =
            (struct index_metadata *)canonical_segment.block_address->data;
        if (block_metadata[old_metadata->blkno].generation
            == old_metadata->generation) {
            merge_block(ctx, inode, block, old_metadata->block);
        } else {
            uds_update_block_mapping(ctx->uds,
                                    chunk_name,
                                    &segment);
        }
        free(canonical_segment.block_address);
    }
    free(chunk_name);
}
```

Pass write data to Alberio to look for duplicates:

```
ret = uds_index_block(ctx->uds,
                    &segment,
                    find_block(ctx, blkno),
                    &canonical_segment,
                    &chunk_name);
```

A duplicate has been identified. Merge so two data blocks share the same storage:

```
merge_block(ctx, inode, block, old_metadata->block);
```

The FUSE demo was run to see Albireo-enabled deduplication in action. As shown in Figure 13, a 64 KB file full of random data was written to an empty Albireo-enabled file system. A copy of the file with a new name was added. The Linux `df` and `du` utilities were used to verify that the user's view of the file system included the space consumed by both files, yet only a single file's worth of disk capacity was consumed. The Linux `md5sum` utility was used to verify that the files were the same. As each file was deleted, the file system capacity and underlying disk capacity were checked.

<sup>5</sup> fuse.sourceforge.net

Figure 13. Validating Albireo-enabled Deduplication

Validating the FUSE demo:

```

>./runjava.pl format
>./runjava.pl albireo
Albireo running in daemon mode. Ctrl-C to exit
>sudo dd2fs/dd2fs -s -o allow_other /tmp/dummy
>df -h /tmp/dummy
Filesystem Size Used Avail Use% Mounted on
dd2fs 4.0M 4.0K 4.0M 1% /tmp/dummy
>du -hs /tmp/dummy
4.0K /tmp/dummy
> cp one /tmp/dummy
> ls -l /tmp/dummy
total 64
-rw-rw-r-- 1 root root 65536 Aug 19 23:52 one
> df -h /tmp/dummy
Filesystem Size Used Avail Use% Mounted on
dd2fs 4.0M 68K 3.9M 2% /tmp/dummy
> du -hs /tmp/dummy
68K /tmp/dummy
> cp two /tmp/dummy
> ls -l /tmp/dummy
total 128
-rw-rw-r-- 1 root root 65536 Aug 19 23:52 one
-rw-rw-r-- 1 root root 65536 Aug 19 23:52 two
> df -h /tmp/dummy
Filesystem Size Used Avail Use% Mounted on
dd2fs 4.0M 68K 3.9M 2% /tmp/dummy
> du -hs /tmp/dummy
132K /tmp/dummy
> md5sum /u1/64k /tmp/dummy/one /tmp/dummy/two
82a3aeb3fb8b2bc770aef8d0e5d129cc /u1/64k
82a3aeb3fb8b2bc770aef8d0e5d129cc /tmp/dummy/one
82a3aeb3fb8b2bc770aef8d0e5d129cc /tmp/dummy/two
> rm /tmp/dummy/one
> ls -l /tmp/dummy
total 64
-rw-rw-r-- 1 root root 65536 Aug 19 23:52 two
> df -h /tmp/dummy
Filesystem Size Used Avail Use% Mounted on
dd2fs 4.0M 68K 3.9M 2% /tmp/dummy
> du -hs /tmp/dummy
68K /tmp/dummy
> rm /tmp/dummy/two
> ls -l /tmp/dummy
total 0
> df -h /tmp/dummy
Filesystem Size Used Avail Use% Mounted on
dd2fs 4.0M 4.0K 4.0M 1% /tmp/dummy
> du -hs /tmp/dummy
4.0K /tmp/dummy

```

**Start the Albireo indexer and the file system**

**Confirm that the file system is empty**

**Copy a 64 KB file to the file system**

```

> df -h /tmp/dummy
Filesystem Size Used Avail Use% Mounted on
dd2fs 4.0M 68K 3.9M 2% /tmp/dummy

```

**Copy the same file with a new name**

**Duplicate capacity has been eliminated**

```

> df -h /tmp/dummy
Filesystem Size Used Avail Use% Mounted on
dd2fs 4.0M 68K 3.9M 2% /tmp/dummy
> du -hs /tmp/dummy
132K /tmp/dummy

```

**Delete the first file**

**No change in capacity**

**Delete the second file**

**Empty file system as expected**

### Why This Matters

Deduplication is valuable technology that can be frustratingly hard to develop and debug. The field-proven and patented deduplication technology at the core of Permabit Albireo API incorporates over 25 man years of development. Integrating Albireo using six well-documented API calls consumed only 52 lines of code. ESG Lab is confident that an experienced storage systems architect working alongside a Permabit engineer can complete a proof of concept integration in two weeks—or less.

## Maturity

ESG Lab performed a high level assessment of Permabit's software development processes to understand the maturity of the Albireo SDK. The bulk of the code within the Albireo SDK has been used within Permabit's shipping products for more than seven years. Permabit has been using agile software development processes for more than eight years.

Agile software development refers to a group of software development methodologies based on iterative development where requirements and solutions evolve through collaboration between self-organizing cross-functional teams. The term was coined in the year 2001 with the formulation of the Agile Manifesto.<sup>6</sup> Agile methods generally promote a disciplined project management process that encourages frequent inspection and adaptation; a leadership philosophy that encourages teamwork, self-organization, and accountability; a set of engineering best practices that allow for rapid delivery of high-quality software; and a business approach that aligns development with customer needs and company goals.

The code at the core of Albireo SDK is continuously integrated and tested on a two week iteration cycle. Stories captured in an online Wiki are used to manage requirements. Unit, functional, and stress tests are highly automated and run continuously. Developers write the majority of unit and system level tests. All changes are peer reviewed. Face-to-face cross functional interaction is embraced. Reducing code complexity is valued and recognized.

### **Why This Matters**

Data deduplication is complicated technology. When it works as designed, it saves capacity and money. When it fails, it can corrupt data. Storage system vendors looking to add deduplication technology to an existing product must be absolutely sure that their deduplication algorithm will not fail. Rigorous design processes and continuous testing are needed to ensure that the deduplication implementation is bug free.

While a more rigorous review is recommended for organizations considering a partnership with Permabit, ESG Lab is very impressed with the maturity and stability of Permabit's software development and QA processes.

---

<sup>6</sup> [Agilemanifesto.org](http://Agilemanifesto.org)

## ESG Lab Validation Highlights

- ☑ The Permabit SDK identified potential capacity savings ranging from 33% to 97% for real-world applications including office productivity files, virtual server images, and e-mail backups.
- ☑ Fast and resource efficient deduplication indexing was confirmed. Less than 0.1 bytes of memory per index and deduplication advisory throughput of 412 GB/sec was measured on a 16 node quad-core Xeon grid.
- ☑ An open-source user-space file system was modified to use Albireo deduplication advisory services using six API calls and only 52 lines of code.
- ☑ Synchronous and asynchronous APIs were used to implement inline and post-process data deduplication, respectively.
- ☑ Permabit's agile software development processes were audited.
- ☑ Systems in the lab and in the field confirm that Permabit global deduplication advisory services have been deployed on grids of up to 38 servers and validated on grids of 96 servers in Permabit's test lab.
- ☑ An error injection test confirmed that the Albireo deduplication services running within a field-proven Permabit Enterprise Archive solution remain available after both a drive and a server failure.

## Issues to Consider

- ☑ The Permabit Albireo SDK detects duplicate data, but it does not actually remove it. Removing duplicates and maintaining pointers to duplicate data is implemented within the storage system that uses the Permabit SDK. Data structures, which map and keep track of duplicate data references, are needed to take advantage of the Permabit SDK. This is a trivial consideration for NAS systems which use an inode map to keep track of data on disk. For modern block-based disk arrays this service is often available for thin provisioning or virtualization, but it may be an issue that impacts the complexity and resources associated with Albireo integration.
- ☑ An Albireo-enabled storage solution can continue operating even if Albireo becomes unavailable. If and when Albireo becomes unavailable the system is unable to detect new duplicates, but data access is unaffected. This is due to the fact that Albireo only provides deduplication advice, and the storage system uses that advice to eliminate duplicates while maintaining data integrity.
- ☑ While the Permabit architecture has been designed to use any hashing algorithm, Permabit has been using the SHA-256 algorithm for years. While the 256-bit SHA2 hashing algorithm virtually eliminates the risk of deduplication induced data corruption due to a hash collision, it does add CPU overhead compared to less rigorous 128-bit algorithms (e.g., MD5).
- ☑ The API integration and capacity savings results presented in this report were collected using relatively simple test programs running on a single server. Estimating the effort required to evaluate, architect, and implement a solution using a production storage system is beyond the scope of this report. Similarly, estimating the savings that can be achieved with your customer's data is beyond the scope of this report. Testing in your lab, with your storage solution, and with your data is strongly recommended.

## The Bigger Truth

One of ESG Lab's first projects was a 2004 validation of a disk-based backup appliance with built-in data deduplication from Data Domain.<sup>7</sup> Since then, data deduplication has evolved into the hottest, most paradigm shifting technology to hit the storage industry since the UC Berkeley RAID papers were published in 1989. Like RAID, data deduplication quickly permeated the storage market due to its outstanding value proposition. Storage administrators struggling to finish backups within shrinking windows were able to reduce the capacity required to retain backups on disk by 90% or more. The value of this new technology was clearly compelling: data deduplication reduced the cost of disk-based backups, putting it on equal, or better, footing than tape. Backups that finish within a shrinking window and quick ad-hoc restores from disk had suddenly become economically feasible.

In recent years, data deduplication has begun to permeate the storage industry. A number of startups, including Diligent, Sepaton, and Exagrid, followed Data Domain into the disk-based backup appliance market. Since then, all of the major systems vendors have added deduplicating backup appliances to their portfolios. More recently, all of the major backup software vendors have added deduplication to their offerings. Content addressable disk-based solutions with embedded data deduplication technology were introduced in the archive market. Permabit was among the first vendors to enter this growing market. Deduplication has been used to reduce WAN traffic within primary and secondary storage replication solutions. And last, but not least, data deduplication is beginning to take hold in the primary storage market, with NetApp and Microsoft leading the way (Deduplication for FAS and Single Instance Store within Windows Storage Server, respectively). Deduplication within primary storage solutions has proven to be extremely effective in virtual server and virtual desktop deployments that tend to have lots of data in common. In ESG's opinion, it's simply a matter of time before data deduplication gains wide market acceptance throughout the primary storage market.

Data deduplication technology has driven a number of strategic acquisitions. ADIC purchased deduplication technology from Rocksoft for \$63M (ADIC has since been acquired by Quantum). IBM acquired Diligent in a deal rumored to be worth between \$160M and \$200M. Those acquisitions were dwarfed by EMC's acquisition of Data Domain, where a bidding war with NetApp drove the value of the deal up to \$2.1B in cash.

So what's the big deal with deduplication? It's actually rather simple: deduplication reduces storage capacity requirements up to 99%. In other words, IT managers can squeeze up to a hundred times more out of each dollar they spend on disk capacity. Even as data deduplication becomes more of a feature than a product, the value is clearly compelling. While one could argue that the hype and valuations of deduplication solutions in the backup arena have gotten a bit ahead of the market, it's clear to ESG that we haven't seen the peak in the archive and primary storage markets yet.

The deduplication market has begun to mature in recent years. As the feature becomes more of a check-off item, vendors are leveraging differences in architectures and implementations to grow market share. Aside from the usual delineations based on price and performance, vendors are competing based on the finer differences between deduplication solutions: object vs. block-based, inline vs. post-process, fixed vs. variable length, and global vs. islands of deduplication. Permabit has extracted the core deduplication technology from a field-proven archiving solution with a goal of delivering a deduplication algorithm that can be used to architect solutions with any of these attributes in mind. In other words, instead of arguing the merits of one implementation versus another, Permabit enables a vendor to implement multiple alternatives and just say "yes".

ESG Lab has confirmed that Permabit Albireo deduplication advisory services work as advertised. Inline and post-process deduplication support was added to a user space file system with only six Albireo function calls and 52 lines of code. The capacity of real-world data sets were reduced between 33% and 97%. The patented deduplication lookup and indexing algorithm was fast and efficient. Permabit deduplication was observed running on more than one server for a scalable global pool of deduplication and fault tolerance. ESG Lab saw no interruption in access when errors were injected on a nine node Permabit Enterprise Archiving grid.

---

<sup>7</sup> See: ESG Lab Report, *The Data Domain DD200 Restorer*, February 2004.

ESG Lab's experience with nearly all of the vendors that have brought data deduplication solutions to the market indicates that correctly implementing data deduplication is a difficult task that requires man-years of effort. Performance, resource efficiency, and scalability have proven to be particularly challenging for a number of vendors. On top of the technical challenges, this relatively new and valuable technology has a growing number of patent portfolios that need to be navigated.

Speaking of patents, Permabit has been awarded a total of 26 patents covering diverse areas in data protection and archive and many more filings are pending in similar areas. The growing portfolio includes patents in the areas of hash-based deduplication for scalable file and object data storage, encrypted deduplication, memory-based snapshots, and many other features of Permabit's product line. ESG Lab was particularly impressed with the resource efficient two-stage indexing method described in Patent 7,457,813, Storage System for Random Blocks of Data. Highlights of that well-claimed patent are summarized in the resource efficiency section of this report.

ESG Lab is confident that the flexibility provided by the Albireo SDK is unique in the industry. ESG Lab has confirmed that Albireo can be used with file or block storage. It provides deduplication services at the object or sub-file level. It supports inline and post-processing programming models with minimal performance and resource impact. Running over a grid, it can be used to create a global pool of deduplication with predictably scalable performance and rock solid reliability. It provides deduplication capacity savings that are far greater than can be achieved with compression. Stream support can be used to provide content aware data deduplication for objects that are misaligned with the block boundaries of an underlying file system. This allows data stored in container formats (e.g., TAR and ZIP files) to be intelligently deduplicated.

Last, but not least, the Permabit Albireo SDK was designed with quick and easy integration in mind. Based on hands-on experience with the Permabit Albireo SDK, ESG Lab believes that Permabit deduplication can be tested within an existing storage system in a matter of weeks. Given the growing size of the market for capacity reduction and the high cost of developing a deduplication solution, organizations considering the merits of adding data deduplication to an existing storage solution should seriously consider a test drive of Permabit's field proven, patent protected deduplication algorithm.

## Appendix

Table 4. The ESG Lab Test Bed

Server	Quad core 3 GHz Xeon with 4 GB of RAM Linux 2.6.18
Storage	RAID-0 over four internal 7.2K RPM SATA drives



Enterprise Strategy Group | **Getting to the bigger truth.**