

Albireo Software Development Kit (SDK) for SOHO Storage Solutions



Data storage for the small office and home office (SOHO) is now reaching tens of terabytes in scale. All of the data being created by imaging and audio applications, coupled with the existing data growth, makes SOHO a perfect opportunity for data deduplication. Much of the data created is copies of copies and the duplicate rates can be very high. Eliminating these duplicates with Albireo deduplication technology dramatically reduces overall storage costs by lowering data storage requirements by 5-35X. In a SOHO implementation, that cost reduction is especially significant relative to the overall budget.

Cloud backup offerings have emerged as viable, cost-effective options for the SOHO market. Applying Albireo Bandwidth Optimization (source deduplication) to the SOHO reduces bandwidth usage, resulting in faster backups. It also lowers the cost of cloud services because the data is optimized (reduced) prior to transmission so it consumes less overall storage at the service provider. A ubiquitous deduplication engine that delivers data reduction seamlessly, such as Albireo, will provide OEMs of SOHO-based storage solutions with technology they must have to stay competitive in this market.

Permabit Albireo is the industry’s first purpose-built OEM data deduplication software designed to meet the needs of hardware, software, and service providers who wish to expand their existing solutions without negatively impacting existing differentiating capabilities or reducing overall performance. Albireo delivers deduplication at the sub-file level and can be flexibly integrated into current or next-generation storage and platform architectures. Albireo deduplication is seamlessly deployed in primary, archive, and backup storage across the data center and the cloud. With Albireo, OEMs leverage their existing R&D investments while accelerating time to market for must-have, industry leading data optimization capabilities.

An initial deduplication implementation for SOHO storage can be completed in a matter of days using the Albireo SDK. As shown in Figure 1, Albireo operates outside the data path of the storage application software as a duplicate advisory service. This ensures that data integrity is never at risk and that there is zero performance impact.

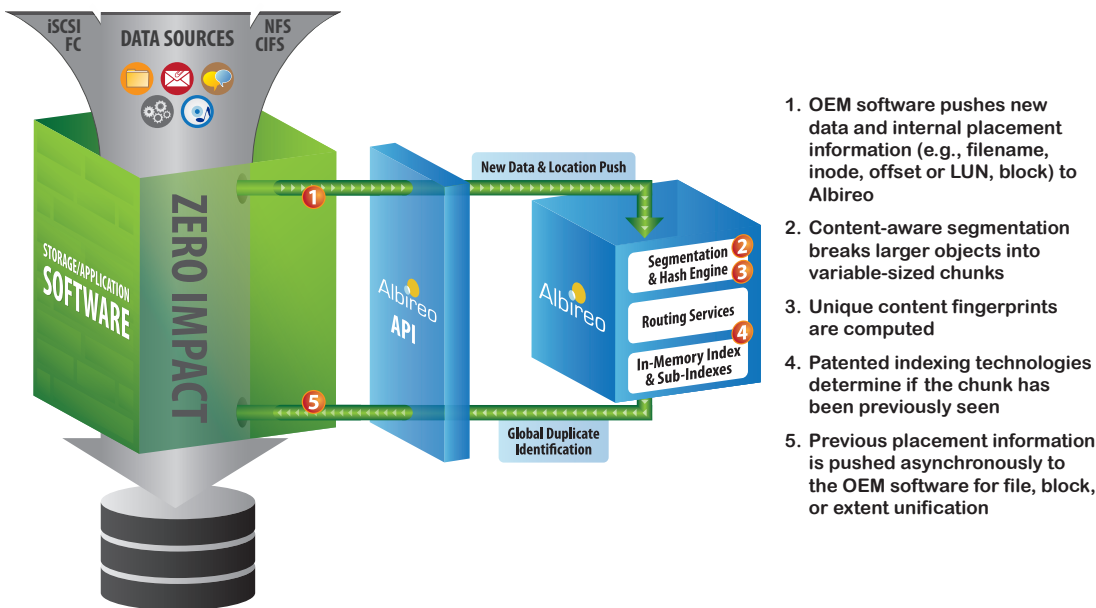


Figure 1: Albireo Process Flow

Deduplication Efficiency

The Albireo architecture supports both fixed and variable block deduplication. Data can be intelligently segmented into chunks of variable length based on analysis of its content. Albireo uses content “scanners” to identify and optimize deduplication of objects within specific compound data formats (e.g. Microsoft®

“The Albireo technology from Permabit will save an OEM 18-24 months getting to market... This stuff is so far ahead in its capabilities and performance I can’t see why you would want to do it yourself, unless you already have it baked.”

— Steve Duplessie
 Founder & Sr. Analyst
 Enterprise Strategy Group

Business Benefits

- Reduces storage growth and associated costs
- Works seamlessly at the server level without need for user intervention
- Improves quality of service and increases customer satisfaction
- Reduces the space required for storage devices
- Makes add-on services such as backup-to-the-cloud more attractive

Technical Benefits

- Background server operation delivers storage reduction without user intervention
- Increases efficiency by lowering the effective cost per GB stored
- Enables data growth without expanding the physical space required
- Reduces bandwidth requirements for cloud backup

Office® documents, ZIP, PDF, tar). An API is available for vendors to implement their own application-specific scanners for further savings. Independent validation has demonstrated that Albireo offers 7X data reduction for general purpose environments, while data reduction rates of over 35X are seen for virtualization images generated by products such as VMware®.

High Performance Index

The Albireo index utilizes advanced techniques to identify duplicates across massive storage pools using memory-resident information over 99% of the time, avoiding costly disk access and eliminating the largest single bottleneck in storage deduplication. Index operations average less than 10 *microseconds* — orders of magnitude faster than other deduplication solutions. This enables sustainable ingestion rates of 11 GB per second with a single processor core (64 KB chunks and hardware-accelerated hashing) and scales out linearly by leveraging Albireo grid technology.

Flexible Deployment Options

Albireo is integrated by the OEM as an inline, post-process, or parallel solution, depending on which is best suited to address the customer requirements for the vendor's architecture. Source-based deduplication deployment (where data is deduplicated at the source for bandwidth optimization) is supported along with the more traditional target-based approach.

Complete SDK for Easy Implementation

Albireo is delivered as a Software Development Kit (SDK). The SDK contains the Albireo software library, full API documentation, code samples, and application notes for integration with primary, archive, and backup storage solutions. Permabit provides technical assistance to ensure rapid and seamless integration. OEM customers typically implement Albireo deduplication (and begin performance optimization and testing) in under 5 days.

Features and Benefits

Albireo Feature	Albireo Benefit
High Performance Index Engine	Albireo handles heavy workloads and can identify duplicate data in <i>microseconds</i> across petabytes of data
Solid State Disk Optimization	Provides optimal performance for flash-based storage environments
Hard Disk Optimization	Provides optimal performance for hard disk-based environments
Bandwidth Optimization (Source Deduplication)	Improves bandwidth efficiency for writes both within the local network and into the Cloud
Portable Source Code	Allows for ports of the SDK to custom embedded devices

Find out More

To learn more about Permabit Albireo technology, or to license our products, visit our website at www.permabit.com or call us directly at 617.252.9600.

About Permabit

Permabit is a recognized leader in data efficiency technology. We enable OEMs to leverage their R&D investment, increase margin, accelerate time to market and achieve competitive advantage. Permabit Albireo software massively improves performance and efficiency of data creation, transmission and storage. Solutions built with Albireo are being delivered by leading hardware, software, and service providers.



Ten Canal Park
Cambridge, MA 02141
Phone: 617.252.9600
FAX: 617.252.9977

info@permabit.com
www.permabit.com