Get Data Deduplication that Includes Complete End-to-End Security.

Data deduplication is a key to managing bandwidth impact and storage growth. But not all deduplication is the same: many deduplication processes expose data to security risks and, rather than improving bandwidth, actually further restrict it. The result is that many organizations are forced to choose between the benefits of deduplication or the continuous security and privacy offered by data encryption.

The problem is in the way many data deduplication processes work with data encryption.

Data deduplication is implemented through an algorithm that analyzes data blocks across an entire data store. The process identifies duplicate blocks in the target archive and eliminates them, on the client side, before the data is transmitted. The remaining data blocks are then reduced in size (compressed) to minimize storage and reduce bandwidth consumption. To further reduce the amount of data being transmitted, the deduplication process may then be applied across multiple data sources.

But data deduplication can run into a problem when it encounters data encryption—a process that ensures information can be read only by authorized users. When data is encrypted on the source, using different encryption keys, the encryption renders them differently than the typical deduplication process, creating device-specific data blocks unrecognized by the source as duplicates. Traditional data deduplication processes do not eliminate these encrypted duplicates, thus rendering the process ineffective.
Secure, Efficient Data Deduplication for Endpoint Protection

As noted above, this leaves organizations with a choice: get the full benefits of data deduplication to minimize bandwidth impact and manage storage growth, or get the full benefits of data encryption. But not both.

EVault Secure Deduplication Addresses Gaps in Security, Reduces Bandwidth Impact

EVault Endpoint Protection from Seagate includes a client-side data deduplication process that goes several steps further than other deduplication processes to ensure data safety and minimize bandwidth and storage impact. Its goal is to ensure end-to-end data security and privacy while using the most efficient backup processes possible.

The solution’s unique encryption and deduplication processes remove the traditional vulnerabilities by deduplicating data after encryption—data does not have to be decrypted in order to be deduplicated across the target environment—and by implementing additional safeguards for encryption keys. Its block- and device-level encryption system never allows for data exposure; data remains encrypted from device to final storage.

EVault Endpoint Protection assigns encryption keys based on "scope"—data blocks within an assigned company, team, or user group that enable the deduplication algorithm to identify data blocks across multiple sources right at the device. This secure global deduplication process never exposes data and it minimizes impact on the network and storage.

Additionally, these scoped blocks can be identified only by a unique encrypted encryption key for each data block. No single or shared key exposes data, minimizing vulnerabilities across the organization.

Each data block is completely processed on the client side. To maintain security, no data analysis is even possible on the server; it just files the data in the data store. Other data deduplication technologies decrypt data on the server (compromising security, privacy, multitenancy, and so on), deduplicate data only from individual data sources (which lacks the benefits of global deduplication), or introduce key management vulnerabilities (single or shared keys that expose data to unauthorized entities).

How It Works

EVault Endpoint Protection secure deduplication processes disassemble each file into a set of variable-length blocks that are then processed as follows:

1. **Selective decompression of data.** Each data block is selectively compressed using standard compression techniques. The process is “selective” in that compression is applied only if the original version of the block is larger than the compressed version. The smallest version of the block is then passed onto the next stage.

2. **Associating environments (scoping).** The compressed data block is then “scoped” based on the configuration of the environment set by the administrator. Scoping allows for explicit data separation between organizational boundaries in a shared environment. It can be configured from enterprise to individual levels, drawing required boundaries in the vault where data is stored. Configurations may include teams or departments.

3. **Encryption key assignment of data blocks.** After scoping rules have been applied, a unique block encryption key is generated based on the scoping rules. This key is then used to apply AES 256-bit encryption to the block. The result is an encrypted data DNA block.

4. **Encrypting the encryption key.** The block encryption key is then itself encrypted and any clear text representation of the key is removed.

5. **True global deduplication.** All encryption processing thus far has been based on scope, and all data DNA blocks are identical within the scope. This allows the data to be deduplicated across the target archive, so only a single instance of any particular data block is ever sent to the archive.

6. **Indexing.** After data duplication, each file can be represented by a simple index that associates a list of unique data blocks with their order of arrangement, and identifies the block encryption key required to completely reassemble an instance of the original data. Every data source will maintain its own unique index of its data, but will share all the encrypted data DNA blocks (subject to scoping rules). No further analysis of the data can be performed once it has been dispatched by the client.
Secure, Efficient Data Deduplication for Endpoint Protection

The Only Endpoint Protection Solution That Doesn’t Trade Deduplication for Security

With EVault Endpoint Protection, you don’t need to choose between the economic benefits of data deduplication and your security requirements. You get it all.

As bandwidth limitations continue to rise and compliance requirements tighten, minimizing bandwidth impact, maximizing storage savings, and maintaining privacy all become increasingly important. Only the secure global data deduplication used by EVault Endpoint Protection can deduplicate encrypted data, providing enterprises and MSPs with the full economic benefits of data deduplication (freeing up bandwidth and storage capacity) without sacrificing data security or privacy.

Take the Next Step
To learn more about Seagate hybrid cloud solutions, visit www.seagate.com/cloud-backup.