Securing High-Value Data in Mid-Tier IT

by George Crump, Lead Analyst

Despite its name, high-value data is often mishandled in many organizations. Often the backup process is counted on to secure and retain this information, but while some of the software has basic archive functionality it was neither intended nor designed to secure and retain data for an extended period. Instead, data protection is intended to return the most recent copy of data to production storage as quickly as possible. The data center spares almost no expense in meeting this requirement.

High-value data is a subset of the unstructured data set that is growing at an astronomical rate. However, high-value data has an additional set of needs. It needs to be identified, secured and retained in the most cost efficient manner possible, all while being readily accessible. The requirements of high-value data along with its association with unstructured data are what make it so challenging to secure.

What is High-Value Data?

High-value data is typically different from mission critical data. Mission critical data ensures the operation of the business day-by-day, hour-by-hour. High-value data is data that the organization has created and used but still needs to retain for an extended period. Re-creation of this information is not possible, and its retention time may be longer than any other data in the environment. The reason for the long retention has historically been to meet governmental or industry regulations, but increasingly organizations want to retain high-value data so that it can be monetized again in the future.

High-value data is often the most sensitive data that an organization will store. Most of the recent data breaches (Sony, Anthem, and Target) have been attacks that were looking for this exact type of unstructured, high-value data. Additionally once captured this data was specifically used to damage the organization. Protection from these attacks is now a critical component of properly caring for high-value data.

Mid-Tier Data Centers Cannot Ignore High-Value Data

These organizations are just as accountable to government and industry regulations. They will also have to recall this data for potential monetization opportunities or legal requests. The challenge is that these organizations lack the time, and sometimes the skill set, to identify and cost effectively maintain high-value data. Unlike an enterprise that can dedicate manpower to manage sophisticated data retention strategies, the mid-tier data center needs a fully automated and integrated approach that protects their assets but does not require a full-time focus.

High-Value Data vs. Mission Critical Data

High-value data differs from mission critical data. Mission critical data, typically data leveraged by databases and applications, has relatively small capacity requirements. Databases are both efficient and have a relatively small working data set. High-value data is typically data created by users or devices. This data can be relatively large on a per file basis.

High-value data also differs from mission critical data in its versioning and retention time requirements. There is little need to store a version of a database as it looked at the end of a particular month or year. There is also seldom a need to retain a copy of that database for years or decades to come. For the most part, all that is needed is the latest copy of the database and a few recent copies to protect against corruption. Those copies can be quickly aged out and removed.

High-value data has the opposite need. Each version of a file may need to be retained to show how it evolved as it was updated. For example, if a product develops a problem a firm may want to be able to research when a faulty component was introduced, or a design change was made. Retaining each significant version of a file is important, and demonstrating the chain of custody will be required. Retention of each version of a file also leads to additional demands for storage capacity.

Requisitions of data from the high-value data set may take years to occur. Making sure that the file is accessible and readable is critical. The requirements to keep multiple versions of a file and retain those for a long period exacerbate the capacity requirements discussed above.

Finally, data durability is also more critical for high-value data. Mission critical is essentially "tested" every day when users interact with it through an application. High-value data, after initial creation and use, may sit idle for years. Continuous error checking needs to be a critical element in any storage system that is used to store high-value data. The demand for data when and if it occurs, is often essential, and company revenue and reputation may be at stake.

Steps to a High-Value Data Plan

The first step is to identify the high-value data in the environment. Identification can often be the most challenging task and the organization may get bogged down in the analysis. Storage Switzerland recommends that the organization simply get started. There are two possible quick start methods. First, most organizations can identify a core set of data without the need for detailed analysis. They could start there and then move other data, as it is identified, into the high-value storage area later.

The other alternative is to move, or at least copy, all unstructured data to the high-value storage system as it becomes inactive on the primary storage system. Removing some of the copied data from production is also very affordable. Most high-value data storage systems are less expensive than production storage systems. Most importantly, capturing all data into the high-value storage system and setting appropriate policies later ensures that no data slips through the cracks.

The next step is to select a storage system that is appropriate to the task.

High-Value Data Storage

Typical production storage systems do not usually have the capabilities to meet high-value data's increased set of needs. Again, high-value data has to be secured and retained in the most cost effective manner possible while being readily accessible. As a result, a storage system designed for high-value data should meet a unique set of requirements.

1 - Single, Cost Effective Storage System

The storage systems must be cost efficient so that the organization can afford to store data on this system for an extended period. That also means that the system should offer sufficient scalability to meet an organization's future demands.

2 - Ease of Access

The demand for re-use or recall of the high-value data may come years after it was originally created. Access to the data needs to be simple, not requiring a special interface or software application. Ideally, access should come through the standard network access methods like CIFS or NFS. These standards, and more importantly the knowledge of how to access data through them, are likely to be in place for decades.

3 - Security

Hacking into an organization is no longer the actions of vandals looking for fun. The goal of this era of security breaches is to embarrass or extort money from organizations. An independent layer of encryption is required to make sure the data being stored on the high-value data storage system is secure. The encryption should have the option of being independent from other forms of security that the data center is employing. A multi-layered protection strategy helps insure that high-value data is protected from these security breaches.

4- Compliance

The high-value data storage system should also be able to meet governmental and industry compliance requirements. In addition to encryption, compliance features provide the ability to provide a chain of custody for a data set as well as show all successful and denied attempts to view the data.

5 - Integrity and Durability

The ability to store information for a long time has no value if once that data is recalled it is unreadable. The high-value data storage system should have the ability to perform regular data integrity checks between the 2 exact copies of the data files to make sure that the media storing the data has not degraded to the point that it is unreadable. These checks should be done automatically in the background without user interaction. In an ideal world, the storage system would have the ability to automatically replace the degraded file with a copy from the original.

6 - Retention

Retention is really the net result of a system's scalability, security, compliance controls and integrity checks but it is one of the primary purposes of the high-value data storage system. Scalability is required so that no matter how much high-value data the organization has it

can be stored. Security and compliance are required to make sure that the data is not modified, deleted nor accessed inappropriately. Integrity checks are required to make sure that data is in a recoverable state when needed.

7 - Recovery

The primary purpose of a high-value data storage system though is recovery. At some point, the organization is going to need this data again to help create new products, make decisions or respond to a legal request. As stated before, this is where ease of access becomes so vital. If the system can be navigated similarly, to how production data is accessed, then its access should transcend staff and technology changes. In this case, the lowest common denominator is the best form of access.

Conclusion

High-value data is data that is critical to an organization's survival. It is often unstructured in nature and needs to be retained for decades. Time is the single hardest problem for technology to solve. The storage system that is used to house this data has to meet a list of requirements that no other storage system will. However, identifying that system and implementing it can greatly simplify the management of high-value data, turning it from a problem into an asset.

About Imation

[Nexsan](#) storage solutions by [Imation](#) are purpose-built for the needs of small-to-mid-sized businesses: enterprise-class, easy-to-use and efficient storage solutions with uncompromising value. It's a different kind of storage experience. Imation's Nexsan Assureon can reduce storage costs and boost user efficiency, all while protecting high-value data in conformity with the toughest corporate governance and governmental regulatory requirements.