



F5 White Paper

Create a Smarter Storage Strategy

Building a smarter storage infrastructure based on the business value of data enables you to address your data storage requirements efficiently and at the lowest cost.

by Renny Shen

Product Marketing Manager



Contents

Introduction	3
<hr/>	
Understanding Your Data	3
Not All Data Is Created Equal	4
The Business Value of Data	4
Change Is a Constant	5
<hr/>	
Making Your Storage Strategy Smarter	5
Determining the Business Value of Data	5
Managing by Business Value	7
Mapping Business Value to Storage	8
IT Agility and the Storage Infrastructure	9
The Importance of Automation	9
Forecasting the Future	10
<hr/>	
Enabling Technologies from F5	11
Free Trial Software Download	11
<hr/>	
Conclusion	12



Introduction

As file data continues to grow at a rapid pace, businesses are looking for more ways to cut their storage costs and lighten the management burden on IT. But where should IT departments begin? Even for those with experience in storage, the market presents an overwhelming array of choices in terms of both technologies and vendors. Instead of starting with potential solutions, it's important to first develop an understanding of your company's data storage requirements—not only as they are today but also in regard to how they are expected to change over time.

To achieve this level of understanding, you must monitor your company's file storage environments over time, tracking which kinds of files are being created, why they are being created, who is creating them, how old they are, and how much storage capacity they consume. Using this information, you can create a smarter storage strategy that uses business value as the barometer by which to match your data to storage with the appropriate performance, availability, and cost. The result is a storage infrastructure that addresses your business needs, and does so at the lowest possible cost.

This paper explores why it is important to have deep insight into your file storage environment, and offers tips on how to develop a storage strategy that best meets your organization's needs. It also considers how storage reporting and capacity forecasting tools from F5 Networks® can help you create a smarter storage strategy by automating the critically important but time-consuming process of data capture and analysis.

Understanding Your Data

Technology marketers like to talk about data growth, because growth is a convenient culprit on which to blame your IT woes. As storage managers, you cannot ignore it, but growth is something that is out of your control. The rate at which your file data grows lies in the hands of your users, your applications, and the demands of your business.

However, there are ways to manage data growth more efficiently, with fewer resources and at lower cost. The first step is to understand the data. Some kinds of data are more important to your business than others, and it is crucial to recognize how each kind of data contributes to the growth of your storage environment.



Not All Data Is Created Equal

The truth is that not all data is created equal. Some data is worth more to your business than others. For example, nobody would argue that pictures from an architecture company's holiday party are as important as the CAD drawings for an office tower it is designing. Or that three-year-old engineering designs are more important than ones for the next-generation product. But, whether it's a user's MP3 collection or the CEO's presentation to the board of directors, all data contributes to growth, and it is up to you to understand which data is most important to your business.

The Business Value of Data

How do you determine which data is important? An easy way to do this is to classify your data by business value. Business value measures the potential contribution to the success or revenue of your company. For example, design specifications for a top-selling product will have a large impact on your company's success, and hence have higher business value.

File data lends itself well to this type of classification, as files inherently contain a business context. Every file has associated metadata with various characteristics that can be used to gauge its business value. How these characteristics influence value will vary from company to company, but there are some attributes, as described below, that are consistent across all companies and storage environments.

- **Type.** Certain types of data inherently have more business value than others. One way to easily distinguish between different types of file data is by file extension. The applications you interact with on a daily basis already do this, assigning context to a file based on its extension. For example, Microsoft Outlook understands that PST files contain archived email content. If files of this type of data have higher or lower value for your business, you can easily identify them through their file extension.
- **Age.** Age is often the most useful indicator because business value typically declines over time. Data tends to be most heavily used immediately after creation, when the information contained within it is most relevant, and it depreciates quickly after that. In some industries, government regulations require companies to retain data long after its value has expired.
- **Location.** The location or, more specifically, the path of a file can also indicate its business value. For example, users often store their files in home directories at known network locations. Similarly, applications are typically configured to create and store different types of data in specific directories.



Prioritizing specific types of data created by users or applications can often be as simple as identifying the location where they are stored.

- **Name.** A file's name typically provides an indication of its contents, purpose, or relationship to other files. To identify files with higher business value, you can search for file names with high-value keywords or known patterns in the file-naming schema.

Change Is a Constant

The last thing to understand is that data is continuously changing. Users and applications are constantly churning out new data, and existing data is used less over time. What is important today might not be important tomorrow, and even the rate at which your users create new data can change over time. This change might reflect shifts in your company strategy or just the normal progression of data in the business lifecycle. As storage managers, you need to be aware of, and adapt to, this constant change in order to keep the storage infrastructure fully optimized for your business.

Making Your Storage Strategy Smarter

By using business value as the foundation of a storage strategy, you can better align your storage resources with business priorities. For example, you will be able to place your business-critical files—those with high business value—on storage devices that have high performance or availability, while moving files with lower business value to storage devices that trade lower performance or availability for lower cost. The end result will be a storage infrastructure that addresses the storage requirements of your business in the most efficient and economical manner.

Determining the Business Value of Data

The first step in creating a smarter storage strategy is to classify your data by business value. For most organizations, it is not practical to manually classify all of their file data. Not only is the volume of files too great, but the rate at which their values constantly change makes this an endless task. This is where an automated storage reporting and forecasting tool such as F5® Data Manager™ can really help.



Reporting and forecasting tools help you look at your data from the perspective of business. They monitor your file systems and enable you to generate customized reports on your file data, breaking it down by various statistics to give you insights into what kinds of data you have, how it is being accessed, and who is generating the most data.

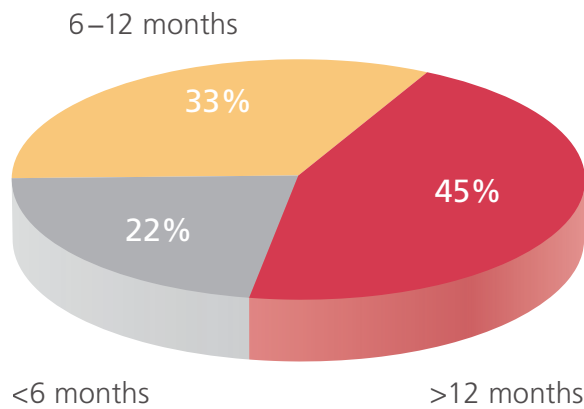


Figure 1: Capacity by modify time

Modify Time	Total Files	Total Capacity
<6 months	375,149	3,608.6 GB
6-12 months	532,787	5,412.9 GB
>12 months	749,735	7,381.3 GB

Figure 2: Capacity and file count by modify time

Figures 1 and 2 demonstrate how data age—the amount of time since its last modification—can be a useful barometer for assessing business value. Excerpted from a Data Manager report for a rapidly growing technology company, these charts help the company understand the amount of storage capacity consumed by inactive data. For this company, this type of aged data is less important than active data. Figures 1 and 2 classify data into three categories based on when it was last modified. They highlight a key problem the company needs to address: Three-quarters of its data has not been modified in more than six months, and almost half of the data has not been modified in more than a year. Although the chart and table use the default Data Manager ranges of less than 6 months, 6 to 12 months, and more than 12 months, these values can be customized to reflect your business’s value classification.

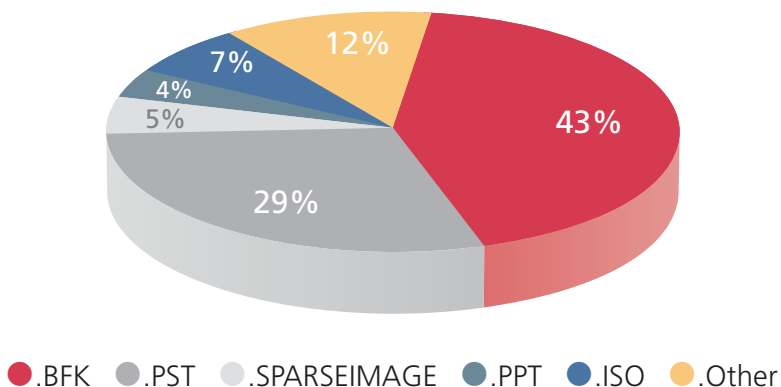


Figure 3: Top file extensions by capacity

Figure 3 demonstrates how the business value of the same data can be determined in a different way. It shows the top file extensions based on the amount of capacity consumed. In this example, the vast majority of capacity on this file system is consumed by three file types: .BKF, .PST, and .SPARSEIMAGE. Because these files contain backup and archived content, they would likely have less business value than other forms of live data.

Managing by Business Value

A smarter storage strategy recognizes that the intrinsic value of data to your business changes over time. Classifying data based on its business value is the first step to being able to manage it by value. But what is management by value? Rather than treat all of your data equally, by understanding your data's intrinsic value, you can apply more resources to managing high-value data, and fewer resources to data with lesser value. In terms of your storage, this means:

- Placing high-value data on high-performance or highly available, but also high-cost, storage resources, and placing low-value data on resources with lower performance and availability, but also with lower costs. A fully optimized infrastructure will apply different levels of performance, availability, and cost to different kinds of data based on their business value.
- Dedicating more IT time and personnel to managing high-value data and fewer IT resources to managing low-value data. For example, you might not need to spend as much time or resources backing up low-value data as you do backing up high-value data.



Mapping Business Value to Storage

When researching which storage platforms are right for business, companies are typically concerned about the attributes of performance, availability, and cost. A traditional storage environment is often based on a single storage platform. While this simplifies storage management, it also means that all of your data is mapped to a specific mix of attributes, regardless of business value. And because it has to accommodate the highest common denominator—for example, data that is being actively used by the business—the attribute mix tends to be skewed toward higher performance and/or availability, with correspondingly higher costs.

Instead of having a single homogeneous approach to all data, a smarter storage strategy maps business value to different tiers of storage. Once you have classified the different kinds of data that you have, you can assign each data class to a different storage tier based on its specific performance and availability requirements. Each storage tier can then be optimized for the mix of attributes that is appropriate for that level of business value.

For example, the findings from Figures 1 and 2 suggest that a storage environment incorporating three different storage tiers would be ideal.

- **Tier 1.** Files that are being actively used (last modified within the past six months) should be moved to storage that emphasizes higher performance and availability.
- **Tier 2.** Files that are inactively used (last modified between six months and a year ago) can be moved to storage that balances performance and availability with lower costs.
- **Tier 3.** Files that are rarely or never used (last modified more than a year ago) should be moved to low-cost storage.

This design can be further refined with other kinds of data classification. For example, you can incorporate the findings from Figure 3 for a tiered implementation that takes into account both the amount of time since a file has been modified and file type.

- **Tier 1.** In addition to files being actively used, those that have a .VMDK extension should always be moved to this tier because of the performance requirements of virtual server images.
- **Tier 3.** In addition to files that are never used, those that have a .BKF, .SPARSEIMAGE, or .BAK extension should be moved to this tier because they are older, or provide backup data, and, therefore, are rarely accessed.



As Figures 1, 2, and 3 demonstrate, only a small percentage of data typically has high value to the business, regardless of the method of classification. Most data can be placed on a lower-cost storage tier without any impact to the business, and the difference in cost for capacity between storage suitable for Tier 1 and Tier 3 can be significant. By mapping business value to storage, organizations can dramatically reduce their storage costs and utilize their budget more effectively.

IT Agility and the Storage Infrastructure

The concept of storage tiering highlights the need for improved IT agility in the storage infrastructure. In traditional file storage environments, users and applications are statically mapped to physical file shares on back-end storage devices. Moving files to a different location breaks the static mappings and disrupts access to those files. This inflexibility presents a major obstacle to implementing a storage strategy predicated on moving data to reduce costs.

Intelligent file virtualization helps overcome this obstacle by enabling the movement of files without disrupting access to those files. It accomplishes this by introducing a layer of virtualization in the network between user and application clients and the physical storage devices. Commonly referred to as a global namespace, this virtualization layer decouples the logical access to a file from its physical location on back-end storage. Instead of mounting physical shares presented by back-end storage devices, clients mount logical network shares presented by the Global Namespace. The Global Namespace then routes any logical file access to the right location.

By shielding users and applications from physical changes, file virtualization devices such as F5 ARX®, the enterprise-class standard for intelligent file virtualization, give you the flexibility to move files without disrupting users or applications. This means that you now have the freedom to match individual files to the appropriate storage and start optimizing your storage infrastructure based on business value.

The Importance of Automation

Not only does data age, but the value of different kinds of data to your business will ebb and flow over time. To keep your storage environment optimized to the business, you must continue monitoring your file storage environment over time and refine your data movement policies as the business value of your data changes. But the burden of continuously identifying and manually moving individual files can be another major obstacle in implementing a tiered storage infrastructure.



In order to reduce the burden on IT, a tiered storage implementation must provide for intelligent management policies that can automate this process. These policies must be able to act in two ways:

- **In real time.** The process of matching data to storage starts at creation. Policies must be able to act in real time to ensure that newly created files are placed on the proper storage tier based on their initial value.
- **Scheduled.** In order to respond to changes in business value over time, policies must be able to act on a scheduled basis to ensure that existing files are moved to the appropriate tier as their value changes.

Forecasting the Future

It is important to recognize that implementing storage tiering is not the end of the road. Matching business value to storage can optimize your infrastructure based on the most current conditions. However, just as data changes over time, so do the needs of your business. Understanding what is happening today and how it is changing can help you plan for the future, so that you can keep your storage infrastructure aligned with your business over the long term and more effectively manage storage expenditures.

When monitoring your data, you always should be on the lookout for new trends that challenge the assumptions on which your tiering policies are based. Therefore, it's important to keep the following issues in mind:

- **Data growth.** While implementing storage tiering will reduce the amount of money your organization spends on storage capacity, it cannot entirely eliminate these expenditures. However, you can mitigate the impact of data growth on your storage budget by carefully adjusting your tiering policies, increasing the amount of data to be moved to lower-cost storage, or modifying the classification of data that is moved.
- **Trends in data value.** The business value of any kind of data will fluctuate over time. By monitoring your data, you can keep abreast of trends that can alert you to these fluctuations, including higher growth among certain classes of data, and increased access from users and applications.
- **Changing storage technologies.** Storage technology is constantly evolving as vendors both improve existing storage technologies and create new types of storage that offer different mixes of performance, availability, and cost. Storage tiering enables you to incorporate new technologies—such as



solid-state drive (SSD), deduplication, and cloud—into your environment where they can meet your business needs.

Enabling Technologies from F5

As we've discussed, building a smarter storage strategy that can respond to your changing business requires several components. First, you need the ability to monitor your file data to keep abreast of its changing storage requirements. Then you need the infrastructure flexibility to match data with the appropriate storage tier without disrupting users or applications. And finally, you need the intelligence to automatically orchestrate this continual process over time. For organizations looking to implement this strategy, F5 offers two products—Data Manager and ARX—to help you move it into production.

Data Manager is a powerful software tool that gives you the ability to look deep inside your storage environment. It monitors your file data and discovers information you need to proactively manage your storage. Rich data profiling capabilities and powerful reporting tools help you identify trends in your file data so you can improve capacity planning and forecasting, create effective file management policies, and uncover optimization opportunities.

As an intelligent file virtualization device, ARX gives you the ability to move file data without downtime or disrupting user access to that data, so you can make changes in your storage environment in respond to business needs. ARX devices include “set-and-forget” data management policies that automate the movement of data. ARX is available in four hardware platforms, each with the same virtualization and automation capabilities, for performance and scalability for any storage environment.

To learn more about Data Manager and ARX, please visit www.f5.com/products.

Free Trial Software Download

F5 offers a free trial of Data Manager that you can download and install in your environment. This trial version will provide you with detailed information about your own file storage environment, enabling you to understand which file types are being created, who is creating them, how quickly they age, and which resources they consume. With this information, you can create a storage strategy that best supports your business, at the lowest possible cost. To download the free trial version of Data Manager, go to <http://www.f5.com/products/data-manager/trial.html>.

Conclusion

A smarter storage strategy recognizes the unique value of data to your business. The fundamental tenet is that the business value of different kinds of data varies and constantly changes. Therefore, the first step in implementing such a strategy requires understanding the business value of your data, not only as it exists today, but also as it has changed over time and how it will be in the future. With F5 Data Manager, you can analyze data through several different file characteristics to assess its value according to business needs.

Once you understand your data, the next step is to execute on that understanding and map business value to different types of storage. This requires flexibility to not only place files on specific storage tiers as they are created, but also to move them as their value changes. F5 ARX gives organizations the ability to move file data wherever and whenever needed without disrupting users and applications. Coupled with automated management policies, ARX enables an agile intelligent storage infrastructure that can respond to change.

By understanding the business value of data and automatically matching it to storage with the right mix of performance, availability, and cost, a smarter storage strategy helps you to optimize your storage infrastructure to your business. The result is an infrastructure that provides users and applications with access to the data they need, when they need it, and with the lowest complexity, latency, and cost.

