

I D C V E N D O R S P O T L I G H T

Got Files? Get Cloud!

November 2010

Adapted from *State of File-Based Storage Use in Organizations* by Richard Villars, IDC #221138

Sponsored by F5 Networks

The explosion of file-based unstructured data is creating a number of new challenges for organizations in terms of both storage asset growth and efficient data management. At the same time, many companies are considering the use of cloud-based storage services to provide a more incremental and more elastic platform for long-term information archiving. This Vendor Spotlight examines the operational challenges that arise as more storage is consumed for file-type data and as new policies for the long-term retention of file-type data emerge. The paper also discusses the most visible challenges revolving around controlling the costs associated with data migrations, ongoing data management, tiering to the cloud, and escalating storage facility costs. It also examines an emerging set of solutions that allow companies to more easily and safely take advantage of cloud-based storage services as part of their file-based storage (FBS) environment. The paper looks at the role of F5 Networks as a provider of intelligent file management and cloud integration capabilities in this important market.

Growing Diversity in Data Use Cases Is Reshaping FBS Requirements

The creation, distribution, and archiving of unstructured, file-based data are now major drivers of storage growth in many organizations. For organizations in industries such as biosciences, healthcare, entertainment, design, and business services, the efficient storage, organization, long-term retention, and timely retrieval of files are at the core of the business. The growth in the number, size, and importance of files, however, is not relegated to specific industries. Small and medium-sized businesses (SMBs) as well as large enterprises are experiencing a flood of file data growth.

The rapid growth in file-based information within organizations is exposing the weaknesses of existing IT management practices for file-based storage. Many companies continue to deploy large numbers of dispersed, underutilized file servers in support of different workgroups and geographically remote sites. Others have consolidated file-based storage on larger NAS systems but are still struggling to deal with rapid data growth and rapid technology changes.

The creation, use, and long-term retention of file-based data within organizations also pose significant challenges around optimizing storage utilization for file-based data, reducing the burdens associated with data migration and backup of FBS assets, and creating and maintaining an accurate inventory of organizations' file-based information. Some of the more significant challenges are the following:

- **Inefficient storage utilization and backup.** Rapid but variable file growth leads to wide imbalances in file server utilization levels and disrupts backup processes. The actual backups often consist of older, unchanging files (that have already been backed up) or unneeded copies of extraneous files (e.g., employee MP3 files). These low-priority files often account for 80–90% of all files being backed up and can drive backup times beyond acceptable backup windows.
- **Time-consuming and expensive data migration.** The use of dispersed file servers leads to severe underutilization of storage capacity, but efforts to migrate and consolidate data from even a limited number of file servers onto larger, more centralized NAS systems can often take six months or more.

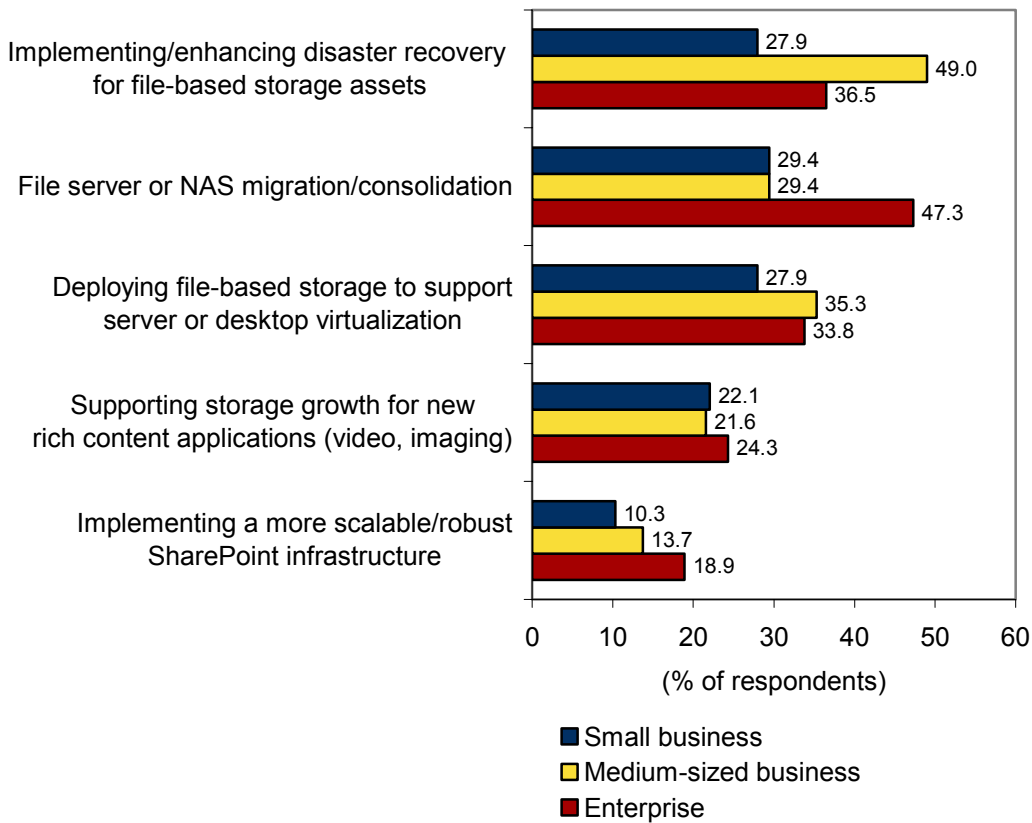
- Increased uncertainty and risk in data protection/retention.** The growing diversity of file-based data pools makes it difficult for IT staffs to effectively catalog and organize file-based data. Without knowledge about what is being stored and how it is being used, IT staffs can't respond to performance and cost mismatches or ensure that data protection policies are being followed.

The diversity and scope of file-based sets also mean that IT organizations are deploying a wide array of FBS solutions, including file servers, NAS/data deduplication appliances, and NAS gateways, but the diversity of FBS assets deployed, as well as the growing diversity of use cases, poses a number of challenges for IT organizations. The biggest pain points vary depending on the volume of files that need to be dealt with, the nature of the files, the sources that produce them, and the geographical dispersal of producers, consumers, and retainers of the data.

Two overarching themes are shaping organizations' investment in FBS solutions: asset consolidation and improved file availability/disaster recovery (see Figure 1).

Figure 1

Top 5 File-Based Storage Investments in 2010 by Company Size



n = 195

Source: IDC's 2009 Trends in File-Based Storage Survey

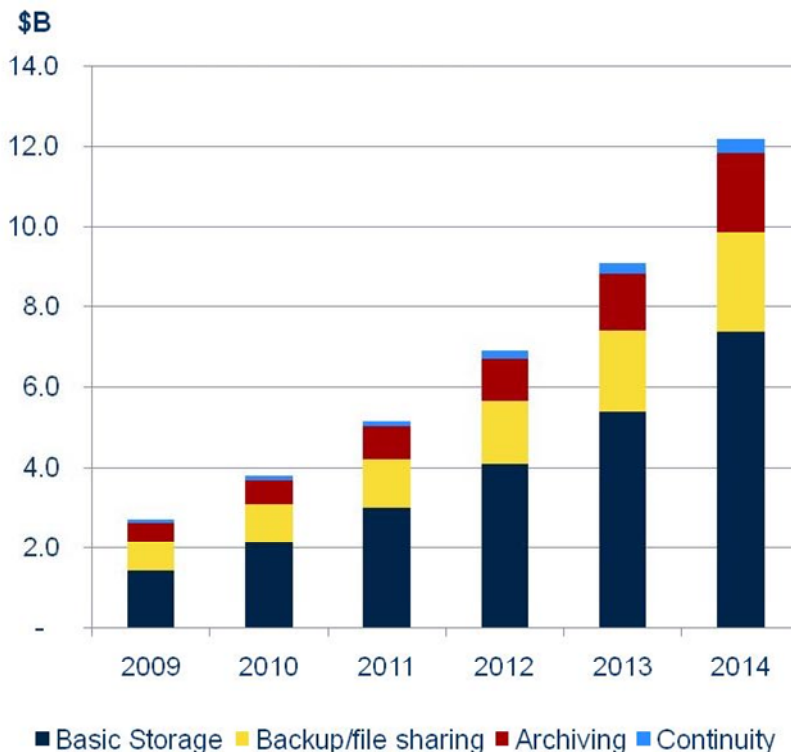
Rapid Data Growth and the Need for More Flexible IT Investment Are Spurring Interest in Storage in the Cloud Services

Continued budgetary and IT staffing pressures are also driving many organizations and/or business units within those organizations to evaluate and use cloud-based solutions for basic infrastructure, new application development, and business applications. Many of the use cases for cloud services are data intensive or content intensive in nature, and this circumstance promises to have a major influence on organizations' interest in cloud-based storage services.

In 2010, organizations around the world will spend \$3.8 billion on all types of public cloud-based storage services, ranging from basic storage capacity for storing large volumes of rich content to advanced storage services for cloud-based backup/recovery and archiving. By 2014, spending on public cloud-based storage services worldwide will approach \$12.2 billion (see Figure 2).

Figure 2

Worldwide Public Cloud Storage Services Revenue, 2009–2014



Source: IDC, 2010

Concurrently, many large organizations that have data security and privacy concerns will construct private storage clouds to help them more cost-effectively address the extreme storage demands of large file archives.

Critical Requirement for Intelligent File Management

As file-based unstructured data becomes a critical challenge for IT organizations, storage administrators need to leverage intelligent file management solutions that drive greater efficiency from the hardware investments and improve the effectiveness and efficiency of storage and data management. Such a solution makes it possible for organizations to:

- Deploy (and consistently manage files across) a wide range of disk storage tiers with different performance, capacity, availability, and cost characteristics, without requiring wholesale displacement of existing assets
- Use a common set of scalable and highly available data migration, data life-cycle management, and data protection services across these multiple storage tiers
- Take advantage of data/file discovery/indexing capability so that IT administrators can inventory existing FBS assets, identify mismatches in usage and/or value, and enable proactive policies for intelligent storage tiering and data retention

Do You Want Some Cloud with That Storage? Accessorizing FBS Systems for Growth and Disaster Recovery

In the past six months, one of the most interesting developments is the coming together of cloud-based storage services and FBS solutions. These solutions have three common elements:

- Gateway services for translating local file protocols such as NFS and CIFS to a range of cloud storage service APIs from major cloud services providers
- Intelligent file management for improved asset utilization and automated data tiering for movement of files between local FBS systems and cloud-based file storage repositories
- Integration of wide area file/application acceleration technologies (compression, deduplication, IP traffic shaping, etc.) to enable faster, more cost-effective movement of data between local sites and the cloud

These solutions enable use of local storage pools based on SSDs and HDDs, but they also include the ability to automatically move older, less frequently accessed data and backup copies to a "cloud storage tier" from major service providers. Given typical use patterns for file data, this approach can provide excellent performance for critical data (e.g., new files, VM images) while seamlessly migrating inactive and archive data to a highly elastic and remote cloud tier.

This automated tiering of data and use of cloud for automated capacity expansion is very interesting for both SMBs and enterprises with dispersed locations. Many are struggling with rapid growth in rich content but don't want to suddenly undertake massive up-front investments in storage for disaster recovery and long-term archiving. Automated tiering also has an important advantage over simply using a cloud-only approach. The management of the name space, encryption keys, and replication processes (all things that fall into the "security" concerns many organizations have with cloud services) is maintained "on premise," not solely "in the cloud."

Considering F5 Networks

F5 Networks is a global provider of network solutions that make applications secure, fast, and available. The company's ARX solution, which adds file-level intelligence and manageability into the network, enables organizations to better optimize their data storage assets. F5's current customers include medium-sized and large enterprises, service providers, and Web 2.0 content providers around the world.

The ARX Series of intelligent file management products provides a powerful suite of FBS management capabilities that eliminate the business and applications disruptions associated with data migration, storage tiering, and storage load balancing. It virtualizes the entire FBS infrastructure that is already in place, through delivery of a common global name space irrespective of underlying file system, platform, vendor, or protocol (NFS or CIFS). Specific capabilities include:

- Enabling adoption of tiered storage architectures that reduce spending on high-performance/high-cost NAS systems as well as backup systems
- Accelerating information migration and consolidation of dispersed, heterogeneous file servers and legacy NAS systems (often cutting migration times by over 70%)
- Boosting the aggregate performance and utilization of storage assets through intelligent load balancing
- Enhancing the availability of all information on a growing range of FBS systems (by cutting backup times and capacity requirements by over 90%)

In November 2010, F5 further extended the value of the ARX platform by introducing a cloud gateway facility, F5 ARX Cloud Extender (ARX CE), and an advanced API. ARX CE works in concert with the automated storage tiering capabilities of F5's ARX file virtualization, making it possible for customers to seamlessly add a "cloud" tier to their intelligent FBS pool. Customers can take advantage of the elastic scale and variable cost benefits of cloud-based storage services without sacrificing predictable performance and data security.

Challenges

These features represent important new capabilities for better managing FBS assets, but F5 must continue to add support for new cloud-based storage service offerings from a growing range of service providers as well as for private cloud deployments. A single access standard would make this easier, but at present, no such standard is in place. Part of F5's role, therefore, is to provide greater "peace of mind" for customers who worry about "lock-in" with current cloud storage services.

F5 must also create an ecosystem of partners who leverage its API for automated data life-cycle management and integrated access control for security. This ecosystem must include more content management, archiving, and data classification solution providers.

Conclusion

The continued expansion of file-based, business-critical information and the growing desire to incorporate cloud-based storage into overall storage strategies are changing the storage dynamic in a wide range of industries and organizations. Those that deploy an intelligent file management solution that also incorporates cloud-based storage can quickly achieve significant economic benefits. They can:

- Automate movement of data to lower-cost storage tiers and defer spending on high-cost storage

- Accelerate migrations of archival data from file servers or older NAS systems to more elastic and scalable cloud-based storage environments without losing control of metadata and security policies
- Reduce the time and resources set aside for file backup (improving information availability and enabling cost-effective disaster recovery for file-based information)

A B O U T T H I S P U B L I C A T I O N

This publication was produced by IDC Go-to-Market Services. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Go-to-Market Services makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

C O P Y R I G H T A N D R E S T R I C T I O N S

Any IDC information or reference to IDC that is to be used in advertising, press releases, or promotional materials requires prior written approval from IDC. For permission requests, contact the GMS information line at 508-988-7610 or gms@idc.com. Translation and/or localization of this document requires an additional license from IDC.

For more information on IDC, visit www.idc.com. For more information on IDC GMS, visit www.idc.com/gms.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com