

WHITE PAPER

The Economic Impact of File Virtualization: Reducing Costs and Improving Efficiency for File-Based Storage

Sponsored by: F5 Networks

Richard L. Villars

April 2009

IDC OPINION

The continued expansion of file-based, business-critical information within extended enterprises is changing the storage dynamic in a wide range of industries and organizations. In a series of interviews with U.S. and European enterprises, IDC found that companies are increasing their file-based storage by 40% to 120% a year and place a high priority on boosting the efficiency and reliability of their management processes for file-based information. IDC research indicates that unstructured, file-based data drove a majority of new storage capacity in all organizations' datacenters in 2008 and projects this growth to accelerate, in spite of current economic conditions. By 2012, over 75% of new storage capacity shipped will be dedicated to the storage, organization, and protection of files.

The added costs associated with this growth are forcing organizations to place a high priority on boosting the efficiency and reliability of their management processes for file-based information. This explosion in file-based data with its unique use and retention requirements is driving the adoption of new solutions for optimizing storage utilization and reducing the burdens associated with data protection and business continuity. Network-based file virtualization is a key technology for more effectively managing file-based information. By using the virtualization and data management functions included in F5's ARX system, interviewed companies stated that they could:

- ☒ Automate movement of data to lower-cost tiers based on preset policies (allowing them to defer spending on high-performance storage and reducing overall spending on disk capacity for file-based data by 50% to 80%)
- ☒ Rapidly and nondisruptively migrate data from older file servers or NAS systems to newer, more scalable file-based systems (cutting migration times by up to 90%)
- ☒ Shrink backup times for file-based assets by up to 50% to 90% while also reducing resources set aside for backup (e.g., secondary disks and/or tapes) by up to 80%

BUSINESS PRIORITIES IN A SHIFTING WORLD

Today, companies rely on an expanding set of applications to compete in this difficult business environment. They are:

- ☒ Relying on email, collaboration tools, and Web sites to communicate and conduct business with customers and business partners
- ☒ Collecting, storing, and analyzing more information about products, customers, and transactions
- ☒ Digitizing records, design documents, videos, and other types of unstructured data to boost efficiency, offer new services, and comply with evolving government regulations

This expansion in the range of information that companies are creating, accessing, and archiving is having a significant, though often overlooked, effect on how they organize their datacenters and store information. Despite difficult economic conditions, IDC sees no signs of a meaningful slowdown in data creation, so CIOs must place a laserlike investment focus on storage solutions that address four business requirements:

- ☒ Reduce and control increases in the cost of doing business by boosting storage asset utilization
- ☒ Enable more effective use of all types of information by a wider range of users
- ☒ Ensure the integrity of the business and its information assets in the face of natural disaster, systems failures, or outside regulatory oversight
- ☒ Implement strategic points of control in the infrastructure that increase agility, enabling CIOs to use IT to drive business innovation and competitive differentiation

Content Is King: Enterprises' Evolving Information Requirements

In July 2008, IDC surveyed storage administrators at 401 U.S. companies about the primary applications driving storage capacity growth in their datacenters. Transactional-type applications such as ERP, CRM, and OLTP that typically create and access structured data remain consumers of new storage capacity; however, new applications that generate and access unstructured data stored as files are emerging as much more voracious consumers of storage capacity in typical enterprises (see Figure 1).

FIGURE 1

Key Applications Driving Storage Capacity Growth

Q. Which of the following are your company's two top priorities relating to storage for the next 12 months?



n = 401

Source: IDC/Computerworld Storage Survey, July 2008

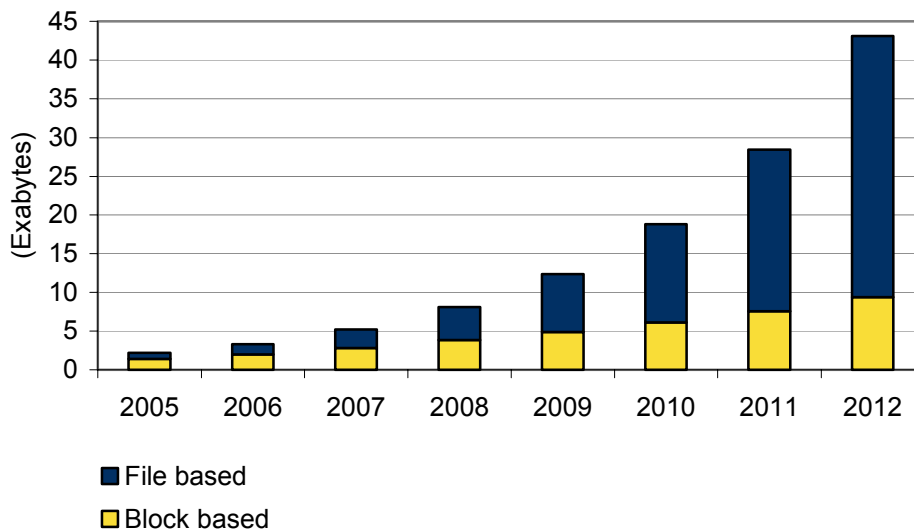
The continued expansion in the creation and archiving of unstructured information is already having a dramatic impact on organizations' storage environments. In early 2009, IDC spoke with a number of senior IT executives at major telecommunications, business services, and financial services companies where file-based storage grew from 5% of their installed storage capacity to over 40% in the past three years. The applications driving this growth included:

- Digital imaging and archiving systems
- Collaborative applications (e.g., email, SharePoint, or digital asset management)
- Video-intensive applications such as video surveillance and digital entertainment

As part of IDC's ongoing monitoring of storage deployments, we found that 2008 was the first year in which unstructured, file-based data drove a majority of new storage capacity in all organizations' datacenters (see Figure 2). Even with current economic difficulties, the growth in file-based information will only accelerate in the coming years. By 2012, over 75% of new storage capacity shipped will be dedicated to the storage, organization, and protection of files.

FIGURE 2

Worldwide Block-Based Versus File-Based Capacity Shipments, 2005–2012



Source: IDC, 2009

A growing number of IT executives now realize that managing such file-based, unstructured data will become the primary task for storage administrators in their datacenters. These administrators need storage solutions that can:

- Enhance the monetary value of unstructured data by making the sharing of business content more cost-efficient and secure
- Support faster access to significantly more and larger files for much longer periods of time
- Reduce the capital and operational costs of storing and protecting growing volumes of file
- Increase infrastructure flexibility so new or emerging technologies can be easily integrated as the business requires, now and in the future

The Economic Perils of Managing File-Based Information

The rapid growth in file-based information 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 explosion in file-based information has a direct economic impact on enterprises in a number of ways:

- ☒ The use of dispersed file servers leads to severe underutilization of storage capacity, often at levels below 15%. In addition, enterprises require an increasing allocation of backup resources to meet higher business continuity standards. A quick review of the actual backups reveals that they consist primarily of older, unchanging files that have already been backed up many times or unneeded copies of extraneous files (e.g., employee MP3 files). These aged or low-priority files — which often account for 80% to 90% of all files being backed up — are responsible for driving backup times beyond acceptable backup windows and higher backup costs.
- ☒ 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 to a year. The factors driving these migrations include rapid data growth, reducing server footprints through consolidation, centralized control of data to meet privacy regulations, or lease rollovers on older equipment.
- ☒ In organizations experiencing rapid content growth (e.g., digitization of medical records, digital delivery of media content, or a shift to disk-based video surveillance), it is necessary to constantly integrate and rebalance data sets across many independent file structures. This often leads to severe and ongoing disruptions, driving down employee productivity, jeopardizing revenue streams, and increasing the burden on IT staff.
- ☒ The long-term retention of files creates a growing performance and cost mismatch in terms of NAS systems deployed and data protection policies followed.

Dealing with Long-Term Storage: The Tiering Opportunity

Managing long-term retention, what IDC calls the "archive" problem, will become increasingly critical in the coming years as more companies and industries move to a digital-based information environment. At its simplest, this problem is a consequence of end users refusing to delete older files.

IT executives now acknowledge that the problem goes well beyond simple user resistance. Companies in sectors such as media and Web services that deliver content to customers must economically support access to the hottest new information and to older, less frequently accessed files. Other requirements including eDiscovery and other regulatory mandates make long-term access and retention of seldom used files critical. A growing number of IT staffs now recognize that 90% of file data being stored at any given time doesn't require the access performance or data protection levels associated with high-end NAS systems.

Companies need to manage files throughout their life cycle, from creation to archive to possible destruction. The inability to automatically, but nondisruptively, migrate information from high-performance storage to lower-cost storage leads to excessive spending on systems that cost three to four times as much to acquire and administer.

The Next Generation of File-Based Storage

Given the high growth in file-based information and the critical need to more reliably and cost-effectively manage these information assets, enterprises are adopting new strategies for building and managing their file-based storage assets.

File-based storage solutions must address the following requirements:

- ☒ Allow organizations to deploy (and consistently manage files across) a wide range of disk storage tiers with different performance, capacity, availability, and cost characteristics, without user disruption and without requiring wholesale displacement of existing assets
- ☒ Provide a common set of scalable and highly available data migration, data life-cycle management, and data protection services across these multiple storage tiers

The remainder of this white paper looks at one of the key technologies, network-based file virtualization, that many organizations (large and small) are already using to meet these requirements and more effectively manage their file-based information assets. It highlights comments from a number of enterprises that are using F5 Networks' ARX intelligent file virtualization solutions to address key business requirements:

- ☒ Boosting the effective utilization of installed storage assets without jeopardizing performance or information availability
- ☒ Accelerating the consolidation of dispersed corporate information while reducing the IT management time needed for planning and reconfiguration
- ☒ Achieving storage cost-cutting goals by making it easy to automatically and nondisruptively move file-based information between tiers of storage based on file age, type, or other preset policies
- ☒ Reducing the time and cost of backing up and restoring file-based information
- ☒ Simplifying and automating data management to free up IT resources for other business initiatives

ENSURING BUSINESS GROWTH WITH FILE VIRTUALIZATION

Virtualization is a major trend throughout the IT industry. For most IT managers, it means consolidating many x86-based applications onto a single server.

Virtualization is an equally important trend in the storage market; however, the focus is almost exactly the opposite. Rather than consolidating multiple servers onto one piece of hardware, network-based storage virtualization is most often about allowing storage managers to manage multiple storage systems (e.g., SAN-attached arrays, NAS systems, or file servers) as a common pool of capacity. Such systems allow IT managers to nondisruptively migrate information between storage, boost the utilization of installed capacity, implement tiered storage policies to manage costs, and better scale data protection processes.

F5 Networks is a global provider of network solutions that make applications secure, fast, and accessible. With F5's ARX solution, which adds file-level intelligence and manageability into the network, organizations can optimize their data storage assets. F5's current customers include enterprises, service providers, and Web 2.0 content providers worldwide.

F5 ARX Series: Intelligent File Virtualization

The ARX series of intelligent file virtualization devices provides a powerful suite of storage management capabilities that eliminate the business and application disruptions associated with data migration, storage tiering, data replication, and storage load balancing. It virtualizes all of the file-based storage infrastructure that is already in place through delivery of a common global namespace, irrespective of underlying file system, platform, vendor or protocol (NFS or CIFS). Specific capabilities include:

- ☒ Enabling adoption of tiered storage architectures that make it possible to reduce or even eliminate spending on high performance/high cost NAS systems as well as backup systems through use of intelligent and automated data movement and replication services
- ☒ Accelerating information migration and consolidation for dispersed file servers or legacy NAS systems through use of network-based data movement services that work across heterogeneous storage systems
- ☒ Boosting the aggregate performance and utilization of virtualized devices through intelligent load balancing and enhancing the availability of all information on a growing range of virtualized storage systems

Interviews with F5 Customers

F5 has a wide range of customers in industries including finance, manufacturing, technology, media and entertainment, telecommunications, life sciences, energy, and government. Its systems are deployed in the United States, Europe, and Asia.

As part of IDC's efforts to assess the impact of the explosion in file-based information on enterprises' IT environments, we conducted a series of interviews with a number of F5's current customers in January and February 2009 (see Table 1). The amount of file-based capacity installed ranged from 6 terabytes (TB) to over 200TB, and all customers were experiencing growth rates in their file-based storage of 40% to 120% a year.

The companies IDC interviewed all placed a very high priority on reducing future spend on file-based storage assets without reducing the rate of information creation. Most also placed a high priority on boosting the efficiency and reliability of their backup and recovery processes for file-based information.

The goal of these interviews with senior IT executives was to gain qualitative and quantitative information about how the deployment of file virtualization helped companies better meet business objectives and manage growth in IT expenditures.

TABLE 1**Adoption of File Virtualization: Company Profiles and Interview Results**

Company Description	Current Environment	Challenges	File Virtualization Use Case
Global provider of Web-based services	<ul style="list-style-type: none"> • 4 NAS systems (6TB) at two sites • Snapshots to disk then backup/archive to tape • Capacity demands increasing >40% a year 	<ul style="list-style-type: none"> • Reduce spending on expensive, high performance storage capacity by implementing a tiered infrastructure • Reduce cost and improve reliability of data protection 	Using ARX for >2 years <ul style="list-style-type: none"> • Tiered older data sets on lower cost platform (80% lower \$/GB) • Delayed purchase of high-cost storage for 18 months • Reduced backup windows for remote snapshots by 80% and reduced tape use by 25%
Regional division of global automotive manufacturer	<ul style="list-style-type: none"> • 3 NAS systems (6TB) • Snapshots to disk then backup/archive to tape • New data creation increasing >50% a year 	<ul style="list-style-type: none"> • Transition from 10 Windows file servers to more scalable NAS systems • Boost utilization levels on NAS systems • Reduce cost and time required for backup/recovery 	Using ARX for >2 years <ul style="list-style-type: none"> • Completed migration in 10% of the time originally estimated • Implemented automated file aging solution that delayed new storage acquisitions for two years and counting • Reduced backup times by 90% and tape consumption by 50%
New media company	<ul style="list-style-type: none"> • 3 NAS systems with both high performance and high capacity (SATA disks) • 100TB of customer-generated, revenue-generating, file-based data (growing >50% a year) • Continually rebalancing content files across systems to meet changing performance needs 	<ul style="list-style-type: none"> • Excessive cost and performance problems associated with rebalancing files across systems • Boost effective utilization rates of systems (previously under 30%) • One-size-fits-all storage environment couldn't meet performance needs for high-value content or cost needs for low-value content 	Using ARX for >1 year <ul style="list-style-type: none"> • Implemented automated data movement that reduced planned site downtime by >90% (eliminating workload of one FTE) • Added a high-performance tier (25% performance boost for high-value content) • Added a low-cost tier that reduced cost of new storage equipment spending by 66%
Legal services firm	<ul style="list-style-type: none"> • Multiple NAS systems • 80TB of file-based data • Growing 6TB per month • Wants to offer vaulting services for eDiscovery 	<ul style="list-style-type: none"> • Migration of records from multiple Windows file servers and older NAS systems to more scalable, reliable NAS would require two weeks of application downtime • Enable launch of cost-effective eDiscovery vaulting services for clients 	Using ARX for 1 year <ul style="list-style-type: none"> • Completed consolidation with zero downtime for end users • Added low-cost tier that was 70% less expensive • Eliminated backup from primary storage (no downtime for backups)
Global publishing company	<ul style="list-style-type: none"> • 3 NAS systems (20TB) • Capacity demands increasing >50% a year 	<ul style="list-style-type: none"> • Control escalating storage costs for archiving of image-intensive files • Backups exceeding backup windows and consuming more tape resources 	Using ARX for >4 years <ul style="list-style-type: none"> • Implemented automated file aging that reduced backup times 90% and tape consumption >55% • Delayed purchase of high-cost storage for 3 years despite 50% data growth

Source: IDC, 2009

Leveraging ARX to Respond to Current Economic Challenges

For all of the interviewed companies, finding ways to more efficiently and better manage the explosion in file-based information was a high priority. The three overarching themes (see Table 2) related to use of F5's file virtualization solution include:

- ☒ Enabling automated movement of data based on preset policies as part of a dynamic tiered storage implementation to delay purchase of high-cost storage and take advantage of lower-cost storage options (e.g., SATA-based storage, data deduplication storage, energy-efficient storage, or cloud-based storage)
- ☒ Reducing the time and resources that must be set aside for backup (cutting backup times by 50% to 80% and often reducing spending on tape capacity by 50% to 80%)
- ☒ Enabling rapid, nondisruptive migrations of data from widely dispersed, hard-to-manage, and expensive-to-expand departmental file services to more scalable and cost-effective NAS systems (cutting migration times by up to 90% and eliminating staff time needed for migration planning and implementation)

The importance of implementing or extending the use of an automated data migration and tiering strategy was a high priority for all companies interviewed. In these difficult economic times, reducing capital expenditures is the highest priority for IT departments. Tiering for file-based storage through use of the ARX platform is seen as a low-cost, nondisruptive way to both delay spending on expensive, high-performance NAS storage and reduce the future costs of NAS storage acquisitions by 50% to 80%.

TABLE 2

The Business Value of Network-Based File Virtualization

Use Case	Key Findings	Interview Highlights
<p>Dynamic tiered storage: Deploy a solution that allows the movement of old, infrequently accessed, and backup copies of data to NAS systems optimized for capacity and low cost</p>	<p>Challenges: End users refuse to eliminate old files and refuse to manually move data to separate archival drives. Compliance policies require long-term retention of all files. In many cases, >80% of stored information is infrequently accessed.</p> <p>Impact of file virtualization: Companies were able to reduce spending on disk storage hardware by 50% to 80% and delay spending on expensive, high-performance storage.</p>	<p>"We had 60 or 70 file systems and there was this treadmill of constantly trying to use manual scripts that migrate data around. Before we implemented the ARX, we had one of our senior software engineers spending a couple hours a day doing nothing but playing catch-up to migrate data around." (New media company)</p> <p>"Using [ARX] also allowed us to continue buying SATA storage for the bulk of our growth. 90% of the data rarely gets accessed, but we still have to keep it online. The 10% that gets hit all the time [the scratch pool or working set] need 15K Fibre Channel drives that meet all of the performance needs. The Fibre drives are about three times the price per GB compared to the SATA drives. On 80% of our storage, we've saved 66%." (New media company)</p>
<p>File migration and consolidation: Consolidating dispersed file servers or legacy NAS systems onto a smaller number of NAS systems to boost utilization, enable greater data sharing, and improve backup/data retention processes</p>	<p>Challenges: Migration times for even a limited number of file servers can take six months to a year to plan and complete, consuming large amounts of IT management resources and planned downtime hours while severely disrupting end-user productivity.</p> <p>Impact of file virtualization: Companies were able to reduce total migration times (including both planning and downtime) to less than one month while dramatically reducing disruptions for end users.</p>	<p>"We are now able to manage our storage without affecting our clients [our clients being my attorneys, my lawyers] who are very sensitive, and they work all hours of the day. I honestly don't think we could have accomplished the migration without it. Just the amount of outage and the amount of downtime that we would have had to take to accomplish this migration would have made it unacceptable for the business." (Legal services firm)</p> <p>"Last year we were serving all the file shares off a single Windows box with tape at the back end. We introduced the ARX as part of the environment along a NetApp System 3040 for better scalability. Without the ARX, migration would have taken six months, but we completed it in two weeks." (Regional division of global automotive manufacturer)</p>
<p>Intelligent and rapid backup: Enable reliable, timely, and cost-effective backup of file-based data, including remote replication</p>	<p>Challenges: Backing up infrequently used data increases backup times, leads to massive duplication of data on tapes, and increases the risk of data loss in a disaster.</p> <p>Impact of file virtualization: Companies improved recovery times for files and email from days to hours and dramatically reduced the time and costs of backup to tape. File virtualization provides a consistent and cost-effective solution for enabling remote data replication.</p>	<p>"Instead of having everything snap vaulted to a remote site, we were able to adjust our backup schedules to a minimum. We modified our policies from having it backed up as frequently as we used to because now the data exists on multiple storage environments at any given point in time." (Global provider of Web-based services)</p> <p>"In the last two years we moved to backup to disk first [then to tape] by leveraging the ARX for tiering of old data. Today our full backup of active data on Friday takes about 25 minutes." (Global publishing company)</p>

Source: IDC, 2009

Challenges/Opportunities for F5

F5 currently offers a growing portfolio of data management and replication capabilities that complement its file virtualization functions. As that horizon broadens, the company must continue to expand the range of heterogeneous storage and application sets that it supports. F5 must add support for new low-cost tiered storage solutions such as file-oriented data deduplication systems and cloud-based NAS appliances. F5 must also continue to enhance the ability of its solutions in the area of automated data life-cycle management and integrated access control for security. This effort includes expanding partnerships with content management, archiving, and data classification solution providers.

Finally, F5 must further educate its customers on the benefits of setting up an environment that takes advantage of the automated data migration and data management capabilities for tiering and compliance. This coordinated storage architecture will ensure that F5's ARX platform can meet rapidly expanding and increasingly demanding file-based storage requirements.

CONCLUSION

As comments from the interviews with F5's customers make clear, enterprises are struggling to better manage their file-based information assets. Those that deploy a network-based file virtualization solution such as F5's ARX quickly achieve significant economic benefits. They can:

- Reduce overall storage costs and simplify data management
- Automate movement of data to lower-cost storage tiers and defer spending on high-cost storage
- Accelerate migrations of data from file servers or older NAS systems to more scalable and cost-effective NAS systems
- Reduce the time and resources set aside for file backup (improving information availability for file-based information)

The key to attaining these benefits, of course, is effective implementation of the solution, including the adjustment of related processes (e.g., backup processes) and the leveraging of new efficient storage options such as data deduplication. A well-thought-out implementation (e.g., tiered storage) also makes it easier to expand use of the solution into new user bases and for additional services such as disaster recovery. Finally, a sound implementation makes it easier for an enterprise to react to changing business conditions and new information needs.

When evaluating solutions such as F5's ARX, IT managers must place equal focus on the ability of the solution provider and its business partners to design and deploy an effective and adaptable solution.

The companies that IDC interviewed stated that the ARX solution was easy to deploy and extend as they activated additional capabilities for data consolidation,

data protection, and dynamic data life-cycle management. Those that acquired the solution through an F5 business partner also stated that their partner's ability to provide a set of complementary implementation and planning services was critical in ensuring that the solution met current and future needs of the business.

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2009 IDC. Reproduction without written permission is completely forbidden.