



WHITE PAPER

DataDirect Networks xSTREAMScaler File Storage: Flexible, High-Performance Scale-out NAS

**By Terri McClure
With Steve Duplessie**

March, 2009

Table of Contents

Table of Contents..... **i**
The Case for Scale-Out Storage..... **1**
 Unstructured Data Growth 1
 Scale-out Architectures for Truly Scalable High Performance Storage 1
 DataDirect Networks xSTREAMScaler File Storage 2
 The Power of DDN’s S2A Architecture 3
 Putting it All Together..... 4
 Distributed LAN Client..... 4
Summary..... **5**

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188. This ESG White Paper was developed with the assistance and funding of DataDirect Networks.

The Case for Scale-Out Storage

Unstructured Data Growth

The information we store today is very different from the information we stored 30 years ago—first, it is much richer graphically, often including embedded audio and video, computer generated images, and animations. In addition, every endpoint device we use has evolved to become a content capture and creation device. Such advances have enabled faster and more efficient business processes—for example, in healthcare, 3- and 4-D imaging enable more reliable diagnosis. High definition satellite images enable more accurate weather forecasts. This all adds up to generate massive growth in file-based data.

Nowhere is the file growth impact felt more than in the audiences served by DataDirect Networks (DDN). Large-file, bandwidth intensive environments such as those deployed in the media and entertainment, oil and gas exploration, bioinformatics, and high performance computing (HPC) sectors—as well as in industries that rely on satellite imaging such as weather forecast and analysis, government intelligence, and geographic imaging services—are driving demand for scalable storage solutions.

File data grows much faster in these environments than in most others thanks largely to the huge file sizes typical of this kind of data. Chip manufacturers render multi-terabyte files. Oil and gas exploration relies on 3-D models in the hundred terabyte range. Satellite images run into the hundreds of GB range. A two hour, high definition video can be almost 14 GB compressed (numbers vary depending on compression and format) and in the multi-terabyte range uncompressed. These are extreme growth environments and with such large file sizes, the storage implications are profound. Users need a new approach to storing massive amounts of large file data; traditional storage solutions just can't scale to meet large sequential file capacity and performance requirements.

Scale-out Architectures for Truly Scalable High Performance Storage

Scale-out, the ability to independently scale and tune bandwidth, processing, and storage capacity on the fly—all while managing the file system within a single global namespace—is becoming the new backbone of file-based storage solutions. Scale-out storage architectures are significantly different than the monolithic, scale-up storage architectures (e.g., traditional NAS or SAN systems) that were developed to meet distributed computing needs.

Scale-out NAS not only meets high-performance file storage requirements, it does so cost efficiently. With independent scaling of storage capacity, processors, and bandwidth, users can grow when and as needed—without buying racks and power supplies in advance of capacity or purchasing extra spindles to stripe files across. Scale-out NAS provides “just-in-time” scalability even in extreme storage environments.

In scale-out NAS systems, adding capacity and bandwidth—as well as file system expansion—is done online with minimal system performance impact. This granular scaling capability provides a price/performance advantage as it allows users to start small and scale where needed. And, since these systems scale into the multi-petabyte range and are managed as a single entity, they can meet most users' needs without paying the management penalty associated with deploying tens or hundreds of scale-up systems.

Relative to scale-up systems, scale-out systems offer:

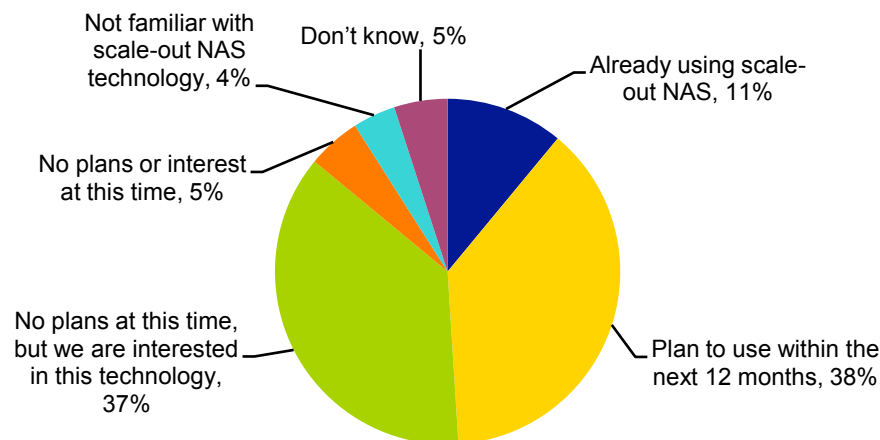
- **Low entry cost relative to monolithic systems.** The entry cost for a scale-out system varies, depending on the minimum configurations supported. Most systems start as small as two nodes and scale out from there.
- **Just-in-time scalability.** Because of the modular nature of scale-out systems, there is no need to buy (and power or cool) frames, power supplies, and mostly empty cabinets in advance of storage capacity.
- **Reduced change management planning cycles.** When one file can be multiple terabytes in size, conventional three or six month change management planning cycles are no longer effective. Requirements are unpredictable and time-to-provision is more important than ever. The modular and easily scalable characteristics of scale-out storage allow for extremely fast provisioning.

- **Non-disruptive technology refresh.** With most scale-out systems, the process of managing technology refreshes is faster and easier than with monolithic NAS because the global namespace maps logical mount points to physical mount points, allowing back-end technology changes to be made with little or no disruption to client access.
- **Ability to scale capacity without scaling headcount.** Essentially, it should be just as easy to manage a clustered storage system with 100 nodes as it is to manage one with two nodes. Scale-out file storage systems enable this through a global namespace, which provides a single point of management for massive amounts of file data.

This wide variety of advantages is driving more and more users to take a hard look at scale-out solutions: Recent ESG research shows that 75% of those surveyed are evaluating scale-out NAS solutions or planning to deploy within 12 months (see Figure 1).

FIGURE 1. SCALE-OUT NAS ADOPTION IN ENTERPRISE DATA CENTERS

Please describe your organization's usage of or interest in scale-out NAS systems. (Percent of respondents, N=338)



Source: Enterprise Strategy Group, 2008

DataDirect Networks xSTREAMScaler File Storage

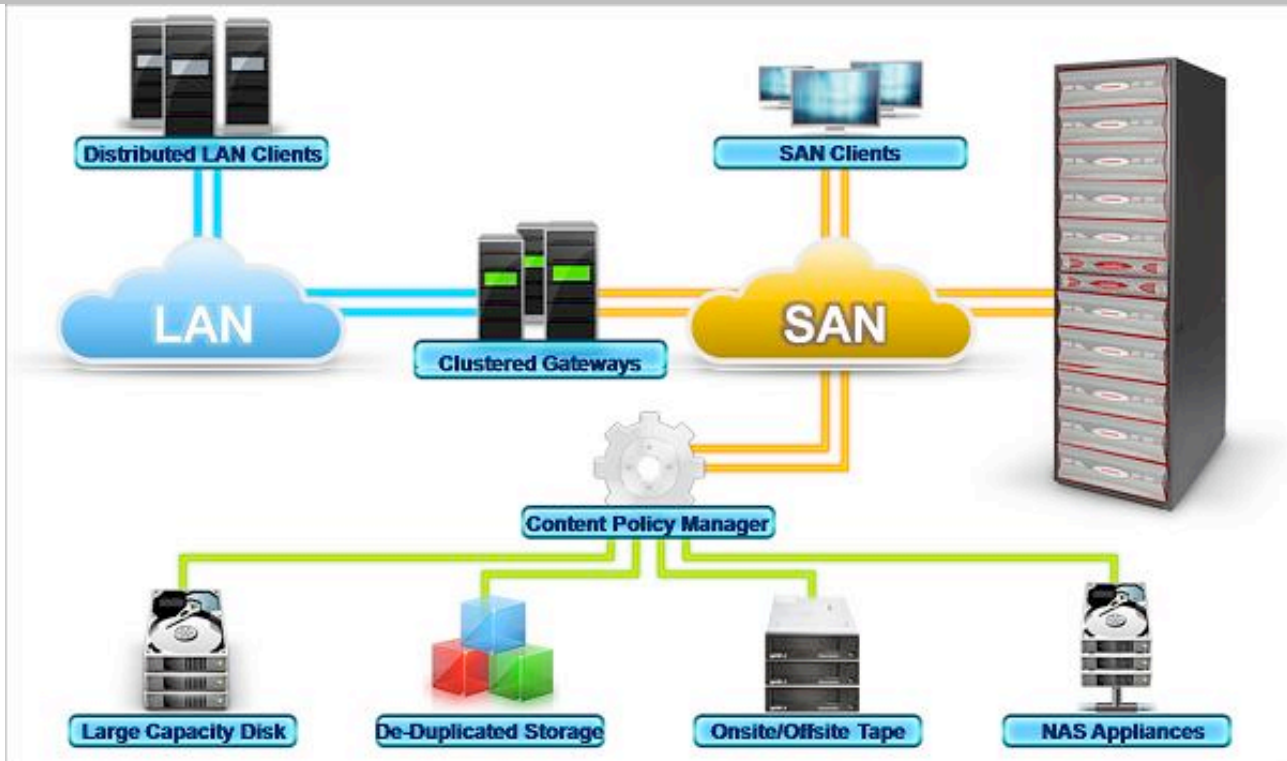
DataDirect Networks xSTREAMScaler File Storage is a scale-out file storage solution based on the Quantum StorNext File System and Storage Manager software and the DataDirect Networks Silicon Storage Architecture (S2A). The xSTREAMScaler system is an integrated, turnkey solution sold and supported by DataDirect Networks that is designed to provide scale-out file storage optimized for high-bandwidth and large-capacity environments.

xSTREAMScaler File Storage is made up of gateways that support either LAN-based clients or a unique SAN client that allows users to leverage higher bandwidth connections. For SAN and LAN-based clients, metadata servers manage permissions and locking—bringing the best of both SAN and NAS worlds together. It has block-level locking, distributed client caching, and full namespace sharing between SAN and NAS.

The system is performance-tuned with near linear scalability. Performance options can be matched to user requirements and all managed within the namespace. There are three performance options to choose from, and multiple options can be supported simultaneously within a namespace:

1. **StorNext Distributed LAN Client (DLC, see sidebar):** For clients outside the reach of the SAN, the distributed LAN client is NAS on steroids. Just like it sounds, it is a client that resides on the LAN, but accesses files with a proprietary protocol that significantly reduces the overhead typically found with NFS and CIFS clients. ESG Lab has tested the file system and found near wire speed performance.¹
2. **SAN client access:** DataDirect Networks xSTREAMScaler File Storage takes scale-out a step further by leveraging a SAN client to further boost performance. Introducing a SAN client into the mix allows users to take advantage of Fibre Channel or even infiniband bandwidth. The client has extremely low overhead compared to NFS and CIFS, with no processing of network packets or TCP/IP protocol overhead. Such a combination of higher-bandwidth data paths reduces processing overhead and virtually eliminates latency.
3. **Standard NFS/CIFS clients:** For those who prefer to stick with a standards-based approach, the xSTREAMScaler File System supports standard SAN and NAS clients accessing files over a LAN.

FIGURE 2. EXTREMESCALER FILE STORAGE ARCHITECTURE



Source: DataDirect Networks, 2009

The Power of DDN's S2A Architecture

The combination of the xSTREAMScaler file system and DataDirect Networks S2A is powerful. The DataDirect Networks S2A is optimized to reliably store and process large data files and has a track record of performance in bandwidth-intensive IT environments. There are a number of contributors to S2A's success, but several areas that particularly stand out:

- **The S2A has predictable performance, even during rebuilds.** DDN's S2A architecture features a hardware-accelerated parity engine that enables it to serve very large file data sets with very predictable performance. It buffers and manages system congestion in real-time, ensuring data is kept protected without a significant performance impact. Both reads and writes benefit from the hardware-assisted parity engine. ESG Lab tested the S2A system in early 2008 and confirmed that the DataDirect Networks hardware-accelerated S2A architecture provides consistent levels of throughput during reads,

¹ Source: ESG Lab Validation Report, *Quantum StorNext 3.0: High Performance Workflow and Digital Archiving*, May 2007.

writes, and mixed workloads—even during hardware faults. Sustained throughput in excess of 2.8 GB/sec was observed for large block sequential reads and writes on the S2A 9950, the previous generation of the S2A-based storage appliance. ESG Lab also verified a number of S2A recoverability and resiliency features that provide enterprise-class performance and data integrity using standard, off-the-shelf SATA drives. Based on ESG Lab's experience testing dozens of storage systems, the S2A9550 is the first block-based parity protected system that can write as fast as it reads, regardless of RAID level. Thanks to hardware-assisted RAID capabilities, the S2A system maintains performance levels even in degraded mode during a rebuild.

- **Parallel data access puts the entire system to work to meet high bandwidth requirements.** The S2A is a true cluster; true clustered storage provides a single logical storage system that allows any application to access any piece of data stored on it through any storage controller in the cluster. A true cluster provides for the aggregation of all hardware resources into a single system. The result is that host ports, CPU, memory, internal bandwidth, and capacity are all logical pools of physical resources. Consider the power of having multiple CPUs working concurrently to run various system and I/O operations: host ports that create a virtual pipe that reaches 100s of gigabits of bandwidth and a pool of disks with hundreds and even thousands of actuators all working in concert reading and writing data.
- **The system scales to extremes in both capacity and performance.** The S2A scales to meet the massive storage requirements of high-bandwidth environments where single files can be in the multi-terabyte range. The architecture supports up to 1,200 disk drives within a single system, or 1.2PB (raw), in only two floor tiles. Facing massive data growth and pressure to significantly reduce operating costs, IT managers cannot afford to deploy tens or hundreds of scale-up systems to store a petabyte of data—the operational requirements, including system management, space, power and cooling, are cost prohibitive. With many system architectures, performance tails off as the system grows. This is not the case with DDN's S2A. Thanks to its clustered architecture, the system can attain near linear scalability. Further capacity scalability in the same physical space is expected with next-generation disk technology later this year.

Distributed LAN Client

DataDirect Networks xSTREAMScaler File Storage Distributed LAN Client is client software installed on each of the LAN clients that communicates with one or more Clustered Gateway servers, providing transparent failover and a load balanced performance boost. These gateway servers running the distributed file system are attached to a heterogeneous pool of SAN attached storage. This provides the best of both worlds: affordable fan-out and fast, fault tolerant connectivity for lower cost LAN attached clients and high speed SAN access for throughput-intensive applications.

Instead of using Fibre Channel host bus adapters (HBAs) and switches to connect a server to the file system, servers are able to access the file system over an IP network. The distributed LAN client reduces the cost of connecting servers to the shared file system, increases availability using a clustered gateway approach and provides a performance boost of up to four times when compared to traditional network attached storage.

Note that the LAN and SAN access methods can be used at the same time on the same file system. This flexibility is ideally suited for workflows requiring high throughput by a few content creators (SAN Clients) and lower cost, scalable access by many users for content review and distribution (Distributed LAN Clients).

Applications that have benefited from the scalability and performance of the S2A architecture include video and film production, broadcasting, computer modeling and simulation, web content storage cluster and supercomputing, and a variety of backup and VTL applications. Storage systems from DataDirect Networks have been deployed by leading companies in the media, oil & gas, web, financial, and government sectors as well as in a growing number of Fortune 1000 enterprises with large scale backup to disk deployments. S2A systems are attached to several of the fastest compute clusters in the world.

Putting it All Together

DataDirect Networks xSTREAMScaler File Storage solution brings all the benefits of scale-out NAS to the table and then some. For starters, it has a low entry cost—users can start out small and grow as required. It has a scalable single namespace that spans across NAS and SAN-attached clients, providing maximum flexibility, cost-effectiveness, and administrative efficiency.

In addition, the system can support up to 128 SAN-attached nodes. Suffice it to say that deployments with more SAN clients exist, but are

not the norm. There is, however, no scalability limit within the xSTREAMScaler File Storage namespace itself, but it is limited to what supported operating systems can handle. xSTREAMScaler File Storage will scale beyond tens of petabytes as operating system limits evolve.

The system is a fault-tolerant clustered architecture, combining high availability with high scalability. The architecture can withstand losing a drive or even an entire disk enclosure without any data loss and minimal performance impact during rebuild.² In addition, a variety of features have been added to increase availability outside of “faults,” including the ability to dynamically grow namespaces/file systems and migrate to new storage while data is being accessed.

DataDirect Networks xSTREAMScaler File Storage provides mechanisms to dynamically migrate content/data from one generation of storage to another, enabling full technology upgrades without impacting accessibility. Customers may still choose to plan for maintenance “downtime” to perform the task, but if data is needed in the middle of the migration process, it’s available and accessible—without any visible change—to the user.

The system offers a number of features that optimize delivery of large files. Granular locking provides an extremely high level of concurrent file access, eliminating long waits while ingest and rendering complete before operations can be performed. User-defined block transfer sizes reduce the number of IOPS the system needs to perform to transfer a file, aligning transfer sizes with stripe sizes for optimal file streaming.

The system is extremely flexible, supporting multiple tiers of storage to economically meet user needs. As mentioned above, the system can migrate files between tiers with no client interruption to access. The system also supports optional policy-based storage tiering. Files are automatically moved between storage tiers based on administrator-defined policies, including when to move files and where to store them, as well as the number and location of data protection copies. To the end-users, however, the tiering and data movement are completely transparent—the files always remain in the same namespace, regardless of physical location on tiers such as disk or tape. The identification of candidates for moving to more economical storage tiers is based on always-available file system metadata, eliminating the need to “walk” a file system to determine “aged” or “stale” content. There are multiple options for storage tiers; all can be used at the same time:

- High performance SAS
- MAID (SAS or SATA)
- Deduplication
- High-density bulk storage archive (e.x., 1 TB SATA disks)
- Tape (onsite or offsite)

Containing file server sprawl has turned into a huge challenge for IT. Analysis of responses to a recent ESG survey indicates that storage consolidation is one of the top two IT priorities in 2009. Due largely to multiple client deployment options, tiering, and manageability, DDN is an excellent platform for consolidation; the S2A platform has market-leading density and energy efficiency, packing 1.2 PB of storage into two floor tiles.

Summary

File data is growing exponentially. ESG research indicates that the vast majority of corporate digital assets are stored in unstructured files. Unstructured files—which include digital images, audio, and video files—accounted for 77% of global digital archive capacity in 2007 and is expected to constitute the bulk of digital assets for the foreseeable future.³ Users need to explore new options to store that file data—continuing to buy monolithic, stove-piped systems is no longer an option. Deploying many scale-up systems to meet performance demands for large file storage leads to a proliferation of underutilized systems, increasing management, power, and

² For more information on DataDirect’s S2A architecture, see ESG Storage Systems Brief: *DataDirect Networks*, April 2008.

³ Source: ESG Research Report, *Digital Archiving Survey*, November 2007

cooling requirements and limiting effectiveness of deduplication efforts as the technology cannot span the stovepipes.

Keeping up with data growth driven by new types of applications, richer media types, and the ubiquity of content capture devices requires a new approach to keeping storage costs in check. New rich media content is being created for everything from research and development, to training, to marketing, and is becoming a mandatory component of everyday business. Whether it's blogs, video, or HD imaging, content is easier than ever to create—and management will become harder than ever without significant changes.

Scale-out NAS has a compelling value proposition relative to scale-up systems. Enterprises that deploy scale-out NAS solutions can get more value, dollar-for-dollar, from their infrastructure investments. Its lower infrastructure costs, power efficiency, and management efficiencies should put scale-out solutions on the short list for anyone deploying new NAS capacity.

Users evaluating scale-out need to ask vendors the right questions in order to separate the pretenders from the true scale-out leaders.

- Does the system scale bandwidth, processors, and capacity independently?
- What kind of processor/capacity ratios does the vendor support?
- Does the file system have a global namespace?
- How large can the system scale while still being managed as a single entity?
- Does the system support load balancing across NAS heads?
- Does it support all operating systems in the environment, including Linux, Windows, UNIX or even MacOS X?
- Does it support a variety of connectivity options to meet performance needs, including LAN, SAN and Infiniband?

Beware of clustered systems without global namespace capabilities—the global namespace simplifies management for a better administrator/capacity ratio. Beware of systems with global namespaces that only support linear scaling, where a processor must be purchased every time a storage node is added, which can add significant storage costs. Look for solutions that allow you to adjust the ratio of storage to processors to meet *both* bulk storage/infrequent access files (about 80% of what you store) and bandwidth intensive active files while keeping both domains in a single namespace. And beware of systems that can't scale to extreme capacities—16 TB is a starting point in the rich data era, not an end point.

DataDirect Networks xSTREAMScaler File Storage is a flexible solution that allows users to meet multiple storage performance profiles in a single namespace—from high-performance SAS-based systems for active files all the way down to tape for long term storage. And DDN brings its scale-out solution one step further with its SAN and Distributed LAN clients to meet bandwidth-intensive performance demands without the overhead associated with NFS and CIFS.

Users can't afford to wait. With storage consolidation on the top of the IT priority list for 2009, the one thing they know for sure is that things need to be done differently. Get started today and beat the economic pressures with smart, flexible, just-in-time scale-out storage.



20 Asylum Street
Milford, MA 01757
Tel: 508-482-0188
Fax: 508-482-0218

www.enterprisestrategygroup.com