the (451) group

IBM turns GPFS into an appliance with Scale Out NAS

Analyst: Simon Robinson

IBM has upped the ante in the battle to win customers with the most demanding capacity and performance requirements. Big Blue this week announced Scale Out NAS (SONAS), a new high-end NAS system that encapsulates many of its existing file storage software and service capabilities for HPC and high-end commercial environments into an integrated appliance. IBM believes that by delivering high performance and high capacity in such a form factor, it can appeal to organizations looking to support applications requiring the most demanding storage capabilities.

The 451 Take

The announcement of SONAS by IBM is noteworthy for a few reasons. First, it underscores IBM's ability and desire to take organically developed technology and repackage it for new and emerging markets; expanding its customer base in storage in particular has been a big focus area for IBM recently. Second, it's the latest example of how scale-out storage software is being paired with commodity hardware – especially running on x86 architecture – to develop new offerings, a phenomenon we have referred to as the 'serverization' of storage. IBM's entry into any new market segment is noteworthy in and of itself, since it helps validate the space. Any organization looking to implement a scale-out storage system now has another option to consider.

Context

IBM is often overlooked as a player in the NAS market, despite having had some skin in the game for some time. For sure, its N series systems that are sold into the mainstream enterprise NAS market are essentially rebranded **NetApp** boxes, but Big Blue also claims it has been a pioneering force in the HPC NAS space since 1995 with its General Parallel File System (GPFS), which today includes some very large clusters (some exceeding 2,500 nodes, others with multiple petabytes of capacity). Since 2007, IBM has made a more concerted effort to take GPFS into more commercial environments by selling GPFS in conjunction with its Global Technology Services group; an offering it calls Scale Out File Services (SOFS), for which it claims a couple dozen engagements. One of the best case studies for GPFS is IBM itself; it has powered IBM's Global Storage Architecture since 2001 – supporting over 100,000 employees worldwide – and maintained by one small IT team.

IBM says that through its SOFS engagements it has encountered repeated customer requests for a more integrated offering that is capable of delivering the required performance, but with much of the complexity removed to make management easier. These requests, twinned with the continuing rampant growth of unstructured data in particular – along with a perceived rise in the strategic nature of that information to the enterprise – persuaded IBM to develop a dedicated NAS appliance tailored specifically for scale-out environments. Indeed, IBM notes a growing set of use cases that increasingly require a easy-to-manage yet high-performance and scalable file storage platform: digital media, Web 2.0- and 'cloud'-type content stores, high-performance analytics and computer-aided engineering, in industries such as media and entertainment, financial markets, aerospace and defense, energy/geo-science markets, high-tech/engineering, government and healthcare.

What were the design requirements of such a product, and how do they differ from other, existing enterprise NAS systems, many of which also purport to be 'scalable'? IBM notes a few fundamental differences between the two. SONAS employs a single namespace for all files, can scale to large numbers of nodes in a single cluster, performance of individual nodes can be aggregated to deliver parallel access to date, while capacity is also aggregated to promote better utilization. Meanwhile, performance and capacity can be scaled independently, providing multiple GB/sec of performance and multiple petabytes of capacity, while high availability is ensured via an N+1 architecture.

Aside from the architectural requirements, IBM also emphasizes the management features that are desirable in a scale-out NAS platform, particularly around policy-based data management (for things like data migration and backup) across the entire filer pool. This contrasts with traditional enterprise NAS, which scales up to a certain point, after which a new filer then has to be added, and managed separately. Establishing policies that manage data across multiple filers is often prohibitively complex.

Products

IBM's SONAS architecture can be broken down into two broad categories: hardware and software. From a hardware perspective, there are two basic components. Interface nodes provide access to clients over the network via CIFS, NFS and other protocols via 1GigE or 10GigE ports. The interface nodes connect to storage nodes via a high-speed (20Gbps) InfiniBand cluster data network. Workload and data is evenly distributed across all nodes and storage (to eliminate hotspots, with healthy nodes taking over should a node fail), while a 'shared everything' architecture ensuring that all nodes continuously serve all files, which IBM claims helps deliver near-linear aggregate performance and capacity scaling.

The storage nodes themselves are packaged into 'pods.' Each pod consists of two storage nodes, each of which is connected to a dual RAID controller (for high availability) with 60 HDDs, which can be either 15K SAS drives (configured as RAID 5) or 1TB SATA (RAID 6) drives; support for 2TB SATA and solid state storage is on the roadmap. There's also the option to connect up to two high-density JBOD expansion units (also 60 HDDs each) within each pod, so each pod can be populated with up to 240 HDDs.

There's a separate node for management. The standard SONAS appliance itself consists of two management nodes, two to six interface nodes, one storage 'pod' and connecting switches. However, the appliance can be expanded for either performance or capacity. An I/O dense configuration supports up to 30 interface nodes, with a storage dense option offering up to 30 storage pods (for a total of 7,200 HDDs); this currently offers up to 7.2PB of capacity in a single system with 1TB drives, though this will double to 14.4PB when support for 2TB drives is added (planned before mid-year.)

The real smarts for SONAS, however, are in the software layer. Indeed, IBM notes that it uses commodity components wherever possible (all nodes are built from x86 servers, for example). The core engine is, of course, GPFS, which IBM claims is a 'true' parallel file system, where all nodes can process both metadata and data in parallel (though rivals such as Lustre claim to do this also, IBM notes that this is often achieved through nodes dedicated to processing either metadata or data.) The upshot of its own architecture, IBM claims, is that it's more scalable than any other scale-out file system. IBM also notes that GPFS is good for small files as well as large files – meaning it can support more transaction-oriented processes and applications such as **Oracle** RAC – while it is also increasing the volume of metadata that can be associated with each file's extended metadata attributes to 64KB.

The other key software innovation within SONAS is the inclusion of the open source Clustered Trivial Database (CTDB) as the overall system manager. This provides a means to achieve the gnarly and complex task of CIFS clustering with no client-side changes; hence making it scalable. The net result on this latter point is that SONAS can scale to deliver 700Mb/sec with both CIFS and NFS – indeed, both can run simultaneously, sharing ACLs – with a single, 64-bit **Intel** server.

CTDB has helped by eliminating contention within the system through a combination of a large, distributed in-memory database and intelligent file locking. Coupled with the GPFS metadata engine, the SONAS management system also acts as the foundation for a series of sophisticated data management features – for initial file placement, subsequent migration, storage pooling/ILM and scalable backup – via a automated policy engine that runs inside the system. The traditional problem here is that the sheer volume of files makes it very difficult to identify the appropriate file candidates on which to perform such an action. IBM claims that the scanning engine it has built into SONAS can scan billions of files in a matter of minutes (it claims a billion files in under 15 minutes, for example), because it separates metadata from data and can execute a fully parallel scan through the metadata to find changed data.

As a result, storage managers can establish policies to achieve a variety of 'integrated' ILM capabilities; for example, established 'gold', 'silver,' and 'bronze' storage pools for different sets of files, either at creation (i.e. CEO's files always reside in gold pool) or for migrating files between pools over time (based on age or access frequency, for example). This can also be used for data deletion, as well as for a variety of file-based backup/restore policies.

IBM has also integrated the Tivoli Storage Manager (TSM) clients into the SONAS interface nodes, which it claims will drive faster and easier backups (especially when using TSM's 'incremental forever' feature); support for other backup applications, such as **Symantec**

NetBackup, is also planned. There's also support for up to 256 snapshots per file system (and up to 256 file systems in a cluster). Asynchronous replication will be delivered in the third quarter of 2010, when the ILM/policy management features will also be available.

IBM is also aiming to make the pricing and packaging of SONAS as straight-forward as possible. It's delivered as an integrated rack-mounted appliance with a minimum number of interface nodes, controllers and drives (starting at 22TB, though IBM expects average orders to range between 100TB and multiple petabytes). Moreover, the SONAS software (which is factory pre-installed) will be charged only on a 'per node' basis (referring to the number of storage/interface nodes inside the system, not client attachments), with no additional charges for functions such as CIFS/NFS, snapshots, replication or per-terabyte. Specific pricing isn't being disclosed at this point, though prospects can expect it to be in the band 9-10 range. SONAS will be generally available in late March, both direct from IBM as well as from certain IBM business partners. IBM GTS will be offering implementation, tuning and customization services.

Competition

The scale-out NAS market has been active for a number of years as other vendors have tapped into the opportunity to provide systems better tailored for a growing set of performance- and capacity-hungry applications. Many of these vendors started life in the HPC space or other niche markets, but most are now starting to target the broader market. For example, **Panasas** and Lustre (the latter purchase by **Sun** and now owned by Oracle) are both very focused on HPC Linux clusters, while **Isilon Systems** first established itself in the media/entertainment space but now is focused on a variety of verticals.

Other scale-out players (or at least, 'scalable' NAS) include **IBRIX** (acquired by Heweltt-Packard in 2009), **BlueArc** and **Exanet** (though it is struggling and rumored to be an acquisition target for **Dell**). Symantec is also starting to target the NAS market more aggressively with its FileStore software, and is planning to target the higher end of the market later this year with a project currently known by the codename 'S4.' **EMC** is another major player in the mainstream enterprise NAS market with its Celerra product line, while its new **Atmos** platform is targeted at high-capacity environments, though it's not aimed at the performance scaling market. As noted above, IBM believes not only that GPFS is the most scalable system available, but also that its ILM-type features and ability to perform data protection tasks (such as backup) efficiently will make it attractive to enterprises that have been put off by the complexity of such systems in the past.

Of course, the other player potentially affected by IBM's SONAS is NetApp. IBM believes SONAS is complementary to, rather than competitive with, the N series systems it sells from its partner, noting that its entry-level configurations start off where its N series systems – which are aimed solidly at the mainstream enterprise market – start to max out. That may be so, but NetApp still regards the NAS market as home turf and to some extent sacred ground. Indeed, it can hardly be delighted that IBM has chosen to pursue the upper end of the market, especially since the latest version of NetApp's core Data OnTAP platform now finally incorporates the clustering capabilities it acquired from **Spinnaker**, though it's true that this is more focused on 'operational scale' more than 'performance' scale.

SWOT analysis

Strengths	Weaknesses
IBM has made significant investments in the scale-out file system market for over a decade, which makes it a leading authority on how to deal effectively with the largest and most complex customer challenges in the scale-out arena. IBM's substantial market presence and influence goes without saying.	Some of the most interesting features of SONAS – especially the ILM-type capabilities – won't be available in the first release, but will follow approximately one quarter later.

Opportunities	Threats
The number of organizations facing complex challenges around managing, protecting and accessing their file storage estates is only going to grow.	IBM has some work to do to establish SONAS and its underlying GPFS technology as a platform that says 'enterprise friendly,' rather than 'HPC complexity,' and there are plenty of rivals to keep an eye on.

For additional information on The 451 Group or to apply for trial access, go to: www.the451group.com