

IBM SONAS – A New Kind of Infrastructure to Meet Large-Scale File Storage Challenges

Analyst: Anne MacFarland

Management Summary

You cook and shop differently during the holidays when you are entertaining large groups of family and friends than you do when your household consists of one or two people. You buy, cook, and freeze ahead, and do everything possible to optimize the eating experience for the cooks as well as the guests – unless, of course, you use a caterer. Caterers and restaurants standardize to optimize, because they have to provide a superb gastronomic experience and still make enough money to support their restaurant infrastructure – and themselves. Most businesses follow this pattern of industrializing the processes that are their infrastructure in order to grow. Most know precisely what scale of operations is needed to support their business model (a minimum), and what opportunities they do not have the resources to address (a maximum). Similarly, in the data center IT standardizes processes into repeatable services to drive down costs and ensure quality.

These days, transactions may be done by calculations (on structured data), but businesses are using unstructured information in many new ways. Business usually is done with collections of data, email, letters, and other forms of communication – all which end up being stored as files. This is like a restaurant supporting multiple cuisines – inherently challenging. There are four challenges of using unstructured information.

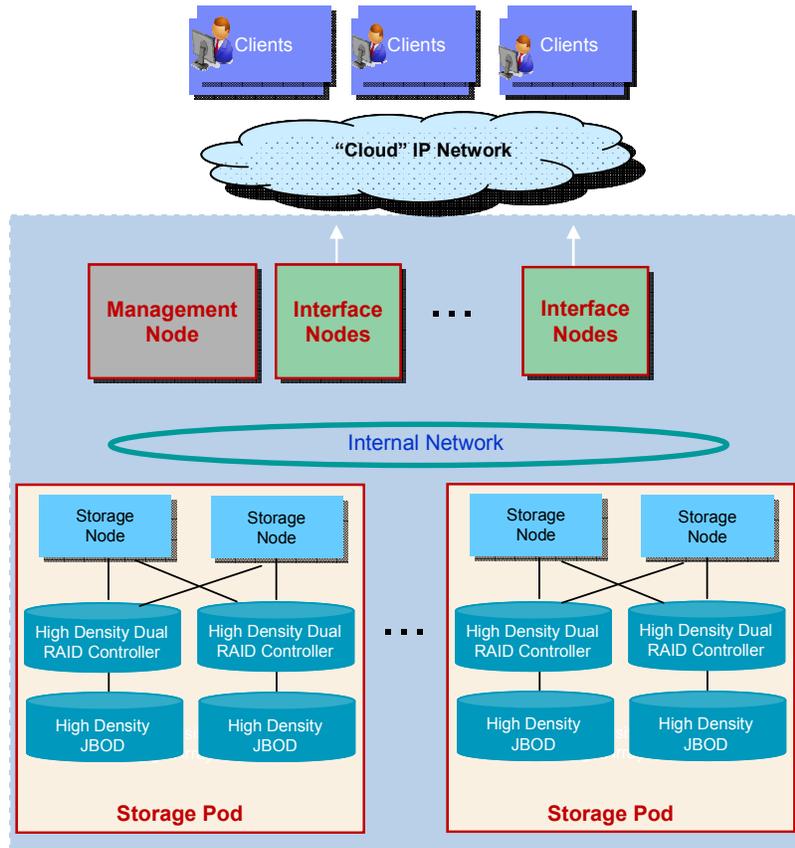
- **Transparency** – Transparency, in a business information environment, is not exposing all information to all parties. *Exposing it all* is, inherently, not a good business strategy. However, removing barriers to finding what you need, or what a process needs, promptly (often ASAP) is a great strategy.
- **Change Management Agility** – The increasingly high rate of change faced by most businesses brings with it many changes in IT and also exacerbates the already-difficult problem of stale or obsolete information. This is not merely a matter of versioning but also includes dealing with the data links and indirectly-associated data sources that give business data its relevance. Managing informational change cannot be an ad hoc process, as the consequences of using the wrong information can be severe.
- **Security and Access Control** – This can no longer be a matter of assigning silos, which tend to be artificial (because organizations and peoples' roles change), perceivable, and, thus, hackable. Access control must be more granular (i.e., role based), easy to evolve, and, in some cases, enforced by the data itself, not just by the application.
- **Cost Control** – If you cannot change the inexorable nature of unstructured data growth, you must change the economics of supporting the use of that data. This calls for standard approaches that support more automation of what now are human tasks of overwhelming proportions.

These challenges rule out many strategies that are perfect at a small scale or for simple, read-only information presentation. IBM's new scale-out NAS, called *SONAS*, addresses all of them. The way that *SONAS* addresses them presents an

IN THIS ISSUE

➤ A Tipping Point for NAS.....	2
➤ The New Thinking in IBM SONAS.....	3
➤ Conclusion.....	5

Exhibit 1 — IBM's Scale Out Network Attached Storage (SONAS) Architecture



Source: IBM

improved NAS value proposition in many ways. Some customers who sampled the approach via IBM's *Scale out File Services (SoFS)*, a services-based offering) asked for a less custom system that would change the economics of both procurement and management. The resulting SONAS also will become the basis of IBM's public cloud offerings. For more details of how the technological innovations of SONAS can change business economics, please read on.

A Tipping Point for NAS

At a certain scale, the variety of data center components (be they physical or virtual) and the interrelatedness of key business processes, which can work against each other, complicate management and create process brittleness. Within the domain of file serving called *NAS*¹,

¹ *Networked Attached Storage*, onto which folders (directories) and files are stored. Not to be confused *SANs (Storage Area Networks)*, which are used for block storage. See these compared and contrasted in the "classic" issue of The Clipper Group Explorer entitled *SAN versus NAS - The Holy War Not*

this can show up as limitations on file sharing, silos of data, and a lot of expensive importing and exporting of files to support interrelated processes. When there are multiple file systems and multiple file trees, search is not straightforward.² Metadata³ and/or indexing can solve this – but these approaches become additional elements to be managed, arbitrated between different sources, and evolved. This can lead to a digital form of *committee purgatory*.

The NAS challenge is not just the bulk of files that infests all organizations, the velocity of their growth, or the variety of their context and content. It is what we want to do with

Worth Fighting (dated September 27, 2000) and available at <http://www.clipper.com/bulletins/sanvnasfinal.pdf>

² It might seem that "webifying" documents via http (an arduous, but a now familiar process) might address this, but most business use requires more sophisticated tools.

³ Metadata is data about data. Classical examples are the information in tables of contents and indices. Metadata tends to be somewhat structured (well-defined fields) but also can be open-ended (e.g., there may be a long list of keywords associated with a field or file).

the information. Corporate sustainability requires knowing more about markets – and about markets not addressed that might be profitable. It requires knowing more about operations and those of partners. Re-thinking is needed.

The New Thinking in IBM SONAS

In SONAS, many points of innovation work together as a whole. They are the basis for changing the economics and extending the utility of unstructured information. Not all are new, but in combination, they present any organization with some interesting new operational possibilities.

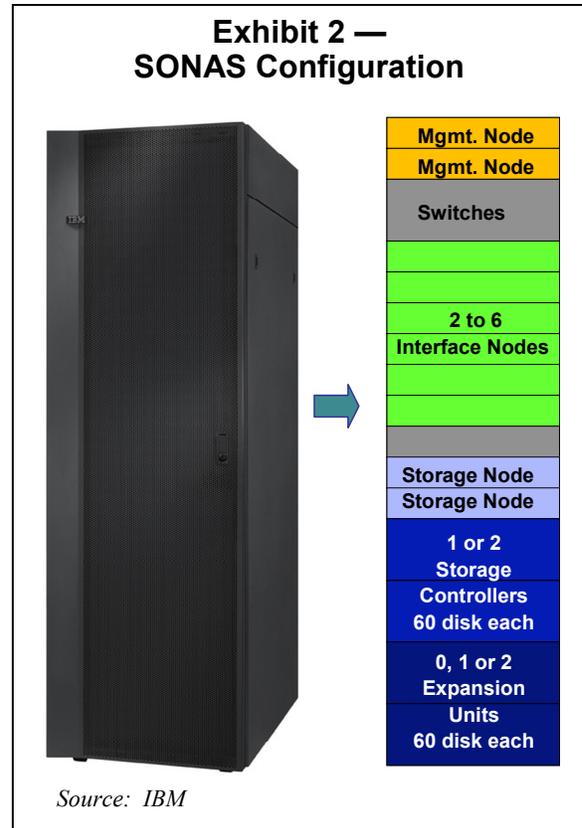
A Grid of Nodes

Management nodes⁴, interface nodes⁵, and storage nodes (as shown in Exhibit 1, on previous page) are instantiated on the same base hardware.⁶ This facilitates clustering for high availability and supports easily adding more kinds of nodes over time. These nodes are connected with users, applications, and IT management by 1 and 10 Gigabit Ethernet interfaces, as is appropriate. They are connected to each other with Infiniband. When paired with the common SONAS software, this creates an almost mainframe-like coordination of process.

Each node runs the SONAS operating system. Upgrades to the OS will be accomplished via DVD or downloads in a rolling, non-disruptive fashion⁷. Upgrades will not incur a software charge. This is refreshing done right.

Two storage nodes are the points of access for each storage pod. However, data is striped widely across all available pods and is not tied into a specific pod. (As shown in Exhibit 1. Also see Exhibit 2, to the right.) They front high-density storage controllers, similarly cross-connected and 60 drives in a 4U enclosure. The drives can be SAS (with RAID 5) or SATA (with RAID 6).⁸ The configuration is set at the factory.

The node approach allows easy and safe



expansion of new applications and user groups wishing to use the system. Grid implementation supports both high availability and more profuse sharing. Workloads and files are shared across all storage pods (as shown in Exhibit 1) and all nodes see and can access all storage. Nodes are mutually aware and have the internal communications to keep the cluster coherent and optimized. SONAS' *Clustered Trivial Data Base* oversees distribution and fulfillment of requests. It uses *byte range locking* to support multiple processes (people) working on the same data and uses tokenization to arbitrate use between multiple processes working on the same byte range. This is more granular than check-in/check out and supports heavier use by many.

Multidimensional Scaling

Capacity and Performance

The SONAS architecture supports independent scalability of both capacity and performance. This is key to satisfying a variety of demand streams while controlling costs. Customers do not want adding capacity to reduce performance; nor do they wish to pay for performance they cannot leverage for business use.

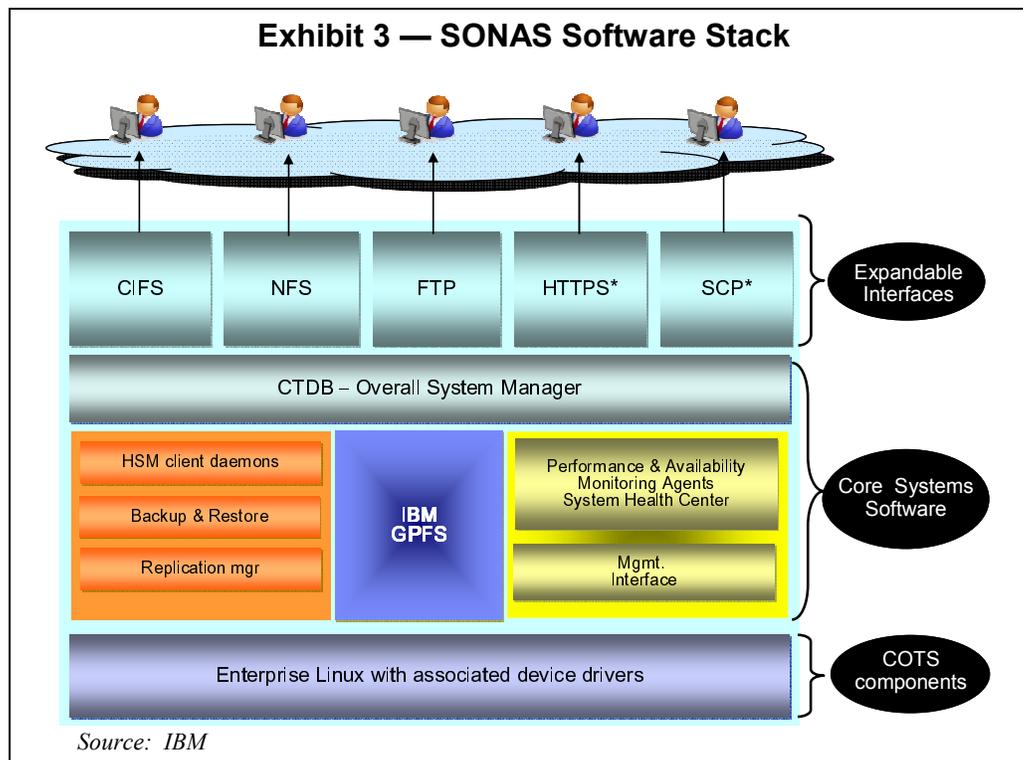
⁴ Management nodes are not counted in the per-node flat pricing of SONAS.

⁵ There must be a minimum of two interface nodes for high availability, with a maximum of 30 interface nodes available at launch. They yield over 25 GB/s of aggregate throughput. Since the architecture is non-blocking, the aggregate can be trunked and exploited.

⁶ This decreases sparing costs.

⁷ The multiplicity of nodes also facilitates testing.

⁸ Solid-state Disk is on IBM's roadmap for SONAS. The proper use of SSDs in a file environment is a topic of fervent discussion.



Scaling in multiple dimensions requires well-thought out architectures. Too many tiers create complexity, particularly if they do not communicate and share a common software basis⁹. (SONAS is based on *Linux*.) Too few tiers can create brittleness.

Much of this capability is provided by the IBM General Parallel File System (GPFS) that IBM developed a decade ago for the demands of high-performance computing grids. (See Exhibit 3, on the next page.) It leverages ILM and HSM heritages, together with full support of POSIX semantics, together with full support of POSIX semantics to handle and parse billions of files with a methodology that has no single point of failure.¹⁰

Change and High Availability

Today's real-time information demands require support for both a high rate of change and high availability – two optimizations that traditionally have been incompatible. GPFS supports a quickness of scan that then allows backups of changed data to be done in a parallelized fashion (quick to finish) and non-disruptively (no performance penalty for high priority active processes). Once again, asynchronicity (or a separation of concerns)

allows a reconciliation of what was previously deemed incompatible. This multidimensionality is supported by the policy sophistication that SONAS supports.

Multidimensional Policies

For very large environments of any size, using intelligence is not always a simple thing. **SONAS has policy sets for placement, migration, deletion, back-up/archive, and restore/retrieve.** Note that the first three categories are more business facing, while the latter two are the concern of IT. In the future, the node paradigm could allow nodes to be added for more business or application control of a particular domain of files.

SONAS supports multiple domain associations. For example, a file may be in a stringent domain for replication, and a more mundane domain for migration. SONAS software supports SQL-like commands for sophisticated file handling. Files in the same file directory can be intentionally placed in different physical storage pools comprised of different pods. Note also that backup can be highly parallelized – no large backup windows are needed for this large-scale environment. An administrator can use all the data placement tactics that are needed – then SONAS automates and carries out the implementation.

⁹ For the common software basis, see Exhibit 3.

¹⁰ Unlike *Hadoop* and *MapReduce*.

One must remember that this is a first release. Exhibit 3 (on the previous page) shows the specific protocols and products supported. IBM says that these will expand significantly, in the short term, and greatly, over time.

Replication is a case in point. SONAS supports snapshots and synchronous replication¹¹ at launch. Asynchronous replication, already available as part of SoFS, is undergoing further replication testing and will be released in 3Q 2010 – soon enough for all but the most rabid of SONAS customers.

SONAS also must work well as part of a bigger environment. It includes a *Tivoli Storage Manager* client on every node. It also supports Symantec and other third-party storage management software products.

Simpler Procurement

SONAS is pretested, preconfigured, and its software preloaded at the factory. It comes as a single SKU. Software pricing is on a per node bases, with no extra charges for extra functionality, such as snapshots.

SONAS possesses massive functionality. IBM calculates the sweet spot of use for customers to be in the hundreds of terabytes. SONAS' smallest configuration is 27 terabytes, at a price of mid-six figures. It can expand to 30 storage pods, 7,200 hard drives, and a total capacity that will double in the second quarter of 2010 (with the introduction of two TB drives) to 14.4 petabytes. Financing is available through IBM Global Finance.

As previously mentioned, SONAS also will be available via IBM's public cloud, which will be sold by IBM directly and also indirectly via IBM partners. One can imagine that IBM partners with specific industry expertise will have a lot of value to add for their customers via SONAS.

How SONAS Can Change How Business Operates

SONAS is very different from a general-purpose storage product. For any business working with files, one coherent file environment with fine-grained access controls and automated migration and placement policies is preferable to juggling multiple content stores with multiple sets of policy structures.

One can envision much richer support for

certain stakeholders, such as sales and customers, with far fewer redundant instances of files and far better disk utilization. GPFS's support of ILM and HSM and the granular nature of access controls will let one virtual repository support multiple audiences as well as multiple applications.

Real-time information, often delivered for immediate use in a specifically targeted process, can be sampled and fed into analytic processes and reporting streams. As tools like search are added (via nodes), files become more than just the destiny of information – they become a proactive resource for reuse.

More data sources make for better analysis. SONAS can support the analysis that businesses need to support prudent decisions in the multi-factor situations they all face.

Conclusion

SONAS is targeted at extreme file use. While it is a first generation product, customer experience with SoFS makes it in some ways more like a second-generation product. If, looking to the future, you see file-based information as both the basis of your business and one of its biggest challenges, consider SONAS. It is both an integrated solution that leverages the SoFS experience and a flexible infrastructure with an aggressive roadmap. SONAS can meet a broad range of enterprise demands with the lower-cost economics of a scale-out approach. If all of this gets you excited, then you should take a closer look at SONAS.



¹¹ Note that, due to the limits of Infiniband technology, synchronous communications is limited to 50 meters or less in this first release.

About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- **The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.**

About the Author

Anne MacFarland is a Senior Contributing Analyst for The Clipper Group. Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group in 2001 after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- **Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 128. (Please dial “128” when you hear the automated attendant.)**

Regarding Trademarks and Service Marks

The Clipper Group Navigator, The Clipper Group Explorer, The Clipper Group Observer, The Clipper Group Captain's Log, The Clipper Group Voyager, Clipper Notes, and “*clipper.com*” are trademarks of The Clipper Group, Inc., and the clipper ship drawings, “*Navigating Information Technology Horizons*”, and “*teraproductivity*” are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosures

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

After publication of a bulletin on *clipper.com*, The Clipper Group offers all vendors and users the opportunity to license its publications for a fee, since linking to Clipper's web pages, posting of Clipper documents on other's websites, and printing of hard-copy reprints is not allowed without payment of related fee(s). Less than half of our publications are licensed in this way. In addition, analysts regularly receive briefings from many vendors. Occasionally, Clipper analysts' travel and/or lodging expenses and/or conference fees have been subsidized by a vendor, in order to participate in briefings. The Clipper Group does not charge any professional fees to participate in these information-gathering events. In addition, some vendors sometime provide binders, USB drives containing presentations, and other conference-related paraphernalia to Clipper's analysts.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.