THE CLIPPER GROUP

Navigator™

Navigating Information Technology Horizons^{SM}

BlueArc Does Scalable Storage Right

Analyst: Anne MacFarland

Management Summary

For many enterprises, IT support of business has become an athletic-style endeavor involving massive amounts of stored data. The increasing reach of IT systems means more data is captured and produced. The extended active use and re-use of this data means that this *more* is kept longer. Government regulations further exacerbate the growth. The nature of most business practices demand that access be prompt, regardless of how large the data or its repository. For even a few terabytes of data, storage scalability – the ability to grow capacity and performance to high levels quickly and easily – is mandatory. Without this scalability, management becomes complex, and its costs insupportable. Management of the information is, at best, inefficient, and, at worst, inadequate. Automation alone can hurt as much as it helps if the storage environment is not architected for scalability.

Scalable architectures have two-and-a-half main characteristics. First, the elements – all of the elements – have a degree of independent functionality (sometimes known as intelligence or autonomics). With this intelligence, they can manage themselves and report exceptions, something that makes managing a large environment much easier. Programmable intelligence also allows the elements to be optimized for the task at hand at a granular level. Second, these elements also must have a capacity to be federated into teams, something enabled by virtualization and by open connectivity standards. Then, the environment must be able to evolve, which requires an extension to this ability to federate, (which is why it is only a half characteristic). Evolvability requires that all elements have the ability to be a contributing part of a still larger environment – not just as peers, but also as subordinates where that is necessary. Flexibility of roles is what makes good athletic teams great. In storage, such adaptability builds an environment that is somewhat forward compatible.

This is a change from the traditional deep hierarchies and hands-on IT controls of last century. Hierarchies become less effective as the rate of change increases (which, in today's large IT environments, is inevitable), because the overhead of discovery becomes unsupportable. Anyone who has been part of a large organization or colonial-style bureaucracy will recognize this situation.

Basically, a scalable storage environment looks less like a tree and more like a well-practiced football team that can parallelize the efforts of all components to good effect. BlueArc, a company based in San Jose, CA, was founded with a vision of massive data throughput. With the *Titan 2000 Series*, it accessorizes the throughput with navigation, management, and utilities of a similar scalability, and the ability to attach third party storage to support converged data storage. It always had the strong-armed quarterback. Now it has the rest of the team needed to make it a good tool for enterprises that possess of an overwhelming amount of stored data. Please read on for more details.

The Unified Networked Storage Approach

With BlueArc, up to 512 TB[1] of networked storage is parallelized behind a *Titan* gateway. This massively parallel architecture is similar to that of the Internet, but on an enterprise scale. BlueArc's *Parallel RAID* striping can distribute I/O across all the storage arrays Titan fronts. Customers can choose what type of technology (Fibre Channel (FC) or SATA), how many spindles, controllers, etc., to stripe across to get the performance characteristics each application needs. Stripes can be replicated for greater throughput when that is needed. BlueArc's *Virtual Volumes* aggregate storage capacities from different arrays for data management and protection. BlueArc's architecture supports both block service and file service. This means that the file service can use the finer granularity of block-level locking, access control, and space allocation. It also means storage administrators have a single set of tools for data protection, management, and virtualization that will work across the entire storage environment.

With the hardware accelerants[2] for all the bulky protocols like CIFS, NFS, iSCSI, and NDMP, multi-protocol support for TCP (versions 2 and 3) and UDP(versions 2 and 3), Titan can handle all kinds of storage, from solid state storage, to Fibre Channel, SATA, SAS, and tape. And yet, Titan also lets you characterize *Virtual Volumes* to meet the needs of the business use of applications.

Titan gateways are usually implemented as a cluster, with an active-active HA configuration with full redundancy of paths and components. At present, the cluster limit is two. Larger clusters will be available later this year.

This cluster uses ten Gigabit Ethernet (10GbE). Since the architecture distributes memory coherently across all the FPGAs, there is no memory contention. The 10 GbE

---

[1] The architecture will support over 2 PB of data.

[2] Acceleration is implemented in BlueArc's Field-programmable Gate Arrays (FPGAs) because they do the job more effectively than a general-purpose processor with a multi-hop data path. BlueArc has implemented the entire TCP/IP stack in hardware. One might say that Titan has enormous TOEs.

Titan Cluster Interconnect means that a misdirected client request can be automatically redirected, through the cluster interconnect, to the correct node. The locking, routing, and access controls are managed similarly. This meta-management layer is what lets Titan storage scale and perform so well.

BlueArc Clustered Name Space

BlueArc's *Clustered Name Space* (CNS), a global name space, is object-based. Since it is implemented by the coherent cluster of FPGAs, the name space does note become a bottleneck, the way a separate appliance can. Indeed, since CNS allows administrators to create multiple file systems and have them appear under a common directory structure and single mount point, from a cluster of Titans, users can get double the performance from a request to what appears to be a single directory.

The object orientation, in the future, will let the name space be augmented with attributes as desired, not just universally relevant attributes. This will permit added functionality (such as indexing) to be added as a feature of the unified networked storage environment. It is a very potent metadata server, and can parallelize traditionally tedious processes like discovery and search.

At the same time, like a good athlete, the BlueArc Titan CNS can also be a node in a still larger file system. Existing enterprise storage can become part of the BlueArc environment, or it can federate with it as a separate entity.

Utilities and Management

BlueArc offers a comprehensive set of core storage utilities.

- Transparent Data Migration
- Block and File Level Replication
- Snapshots and File Restores
- Hard and Soft Utilization Quotas
- Virus Scanning
- Role-Based Management.

BlueArc takes a systems management

approach because, as a front-end to storage arrays, BlueArc is managing a system. Titan 2000 is a point of aggregated management. Titan is the basis for invocation of data utilities like replication and migration. Titan, by its architecture, supports storage virtualization and thin provisioning. A global name space, like CNS, can also be a good vehicle to manage data by content. The clustered memory will handle the mapping to physical locations. With 10GbE as a cluster interconnect, the whole responds with faster throughput. For enterprise with considerable real-time demand, this kind of responsiveness is needed to let the business - and its revenue - grow unimpeded.

BlueArc, like any NAS box, converges file serving, but, as a gateway, BlueArc converges much more. Its architecture means the convergence does not become a bottleneck.

Conclusion

Storage often involves a tradeoff between performance and scalability, and is always gated by cost considerations. BlueArc is devoted to high performance, in both throughput and IOPS[3]. If you have massive data storage, and need to use it without worrying about whether you are asking too much of your storage arrays, BlueArc Titan is an array you will want to consider.

---

[3] Input/outputs per second.

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 Director of Data Strategies and Information Solutions 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 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**, 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.

Disclosure

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.

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.