Grid Storage Technology and Benefits

Silverton Consulting, Inc. StorInt™ Briefing

Introduction

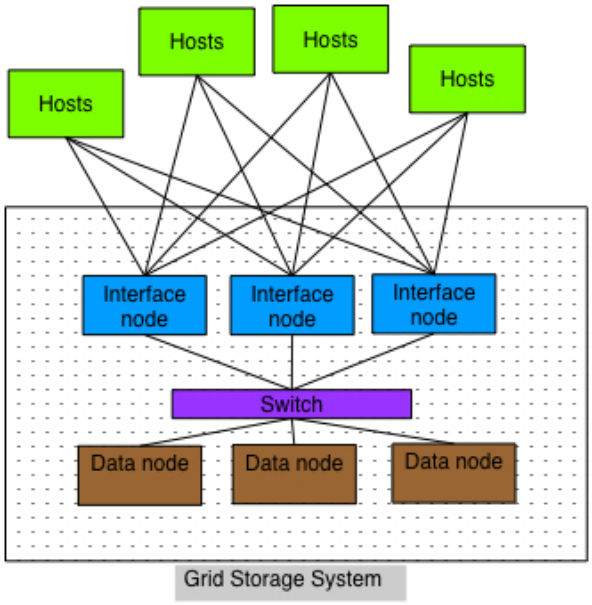
Distributed computing took off earlier this century primarily as microprocessors and inter-cluster communication links became more affordable and fast enough for such systems to process highly concurrent workloads, like storage input/output (IO) activity. Grid-based storage systems, an outgrowth of distributed computing, have in turn focused on data storage services.

The popularity of this data storage technology has led to a number of vendor solutions that qualify as grid storage. IBM® in particular has been an innovator in grid storage architectures over the last decade, with their XIV® storage and IBM Spectrum Accelerate™ software-defined storage, which came out this past year.

Grid-based processing for data storage offers significant technical advantages such as ease of deployment, scale-out IO performance, and tuning-free storage that are ideal for many data center environments such as virtualisation, cloud services and data analytics. We will discuss these virtues at length, but first we will explain how grid storage systems work and how they differ from traditional storage.

Grid-scale storage architecture

Grid storage systems are composed of a number of components called modules or nodes. All the nodes in the grid interoperate in a clustered fashion to provide storage services. Grid storage systems use one of two deployment alternatives:



1. **Software-defined grid storage**, a data storage system deployed as a software-only solution on customer commodity server hardware using direct attached storage as backend capacity. In software-defined grid storage, each storage instance or virtual machine (VM) acts as its own controller or node.

2. **Hardware-defined grid storage**, a data storage system deployed on vendor-supplied controller hardware using vendor-supplied data storage as backend capacity.

Each grid storage node connects to the other modules in the system with an inter-cluster network such as InfiniBand or high-speed Ethernet so that data can move quickly from grid module to grid module as the need arises.

In a typical grid storage system, all nodes have the same processing, networking, number of data drives and storage capacity. Some grid modules also have host interface connections. All modules in a grid participate in host IO activity, which enables grid storage to supply massively parallel or concurrent IO operations.

To facilitate the grid's concurrent IO processing, host volume data is split into segments, which are distributed in a pseudo-random manner to all grid nodes for storage. In this manner, host data is striped widely across all nodes and data drives in a grid storage cluster. Thus, all modules and data drives are potentially active when servicing host IO.

For instance, when a host writes data to a volume, its data enters the grid via a host interface node and is gathered in that node's cache. At destage time, the interface node splits the data into segments and performs a calculation to determine which node will hold the data. The segment is then sent to the identified module, where it is written to one of the node's data drives. A corresponding process occurs at read time. As such, grid storage 'scatters' write data and 'gathers' read data to service host IO activity. Moreover, because host data is striped across all data drives, to define new storage volumes, all an admin need do is to insure there's enough capacity available to satisfy the request.

In grid storage, data protection customarily involves data mirroring, where host data is automatically written to two or three distinct drives residing in separate modules. As a result, no single drive or node failure can impede access to host data.

In addition to data drives, grid storage can take advantage of flash storage as an extension of its data cache. Using flash like this increases the amount of host data that can be rapidly read out of cache.

As all nodes are the same, customers can easily grow the grid system's capacity and performance by adding more modules. When the grid detects a new node, it automatically redistributes host data segments across the newly enlarged grid, rebalancing host IO activity across all the modules.

How grid systems differ from traditional storage

Conventional, highly available storage systems make use of **dual-controller architectures** with attached data storage. In such architectures, each controller has host interfaces and a data cache.

In a non-wide striping, dual-controller architecture, host volume data resides on a specific set of drives assigned to a single controller. To configure host volumes, customers must decide which controller to use for that data. Furthermore, within dual-controller architectures data protection is usually based on various RAID levels, meaning that host volume data is assigned not only to a specific controller but also to a specific RAID group set of data drives behind that controller.

These assignments usually last until an operator needs to move the data because this configuration of host volume data can lead to 'hot drives,' data drives whose performance suffer from too much IO activity. When IO performance problems are detected in a dual-controller architecture, storage administrators must tune the system by moving host volume data to other RAID groups or **to the other controller**.

By extension, in some configurations one of the controllers can suffer 'hot controller' performance problems. In this case, an operator may need to move the data from an entire RAID group from one controller to the other or **to another storage system** altogether to fix the problem.

Unlike all-inclusive grid storage operations, dual-controller architectures support either 'active-active' or 'active-passive' operations. In active-active systems, both controllers service host IO activity all the time, whereas in active-passive systems only one controller services IO while the other one sits idle. When a controller failure occurs in active-active configurations, the alternate controller takes over access to the failing controller's data drives. The alternate controller thus takes on all host IO but only has use of a single controller, roughly dividing system IO performance in half.

Dual-controller storage performance usually cannot be scaled out. While data drives, cache and host interfaces can be added to these systems, new processing capabilities cannot be added. Moreover, because data is forced to reside within a single RAID group and on one or the other controller, IO performance for a single host volume does not improve when capacity is added. Additional cache and host interfaces may increase performance slightly, but adding capacity does not increase IO performance unless host volume data is moved.

Advantages of grid storage

The technical advantages of grid storage over traditional dual-controller storage include the following:

- **No hot spots** - Grid storage automatically stripes and distributes host data across grid modules and drives. Because host volume data does not reside on a single node or drive, there are no problems with hot drives or hot controllers.

- **Autonomic performance tuning** – Grid storage system administrators no longer have to move data around to tune storage performance because the data is continuously striped across all modules and data drives.

- **Consistent IO performance** – Grid storage systems provide uniform, predictable IO performance across widely varying application workloads because of the no hot spot/self tuning nature of grid storage systems.

- **Effortless volume configuration** – Grid storage host volume configuration is much simpler as customers no longer have to decide which controller, RAID group and data protection to use for host volume data.

- **Faster rebuild times** – Grid storage can return the system to full fault tolerance faster because all drives and nodes participate in reading mirrored data for the failed drives or nodes and writing that data to new locations.

- **Linear performance scalability** – Grid storage IO performance can scale-out linearly when more modules are added because compute and capacity are added together in contrast to dual controller systems that only add capacity while increasing compute load. With expanded processing power, cache and drives, the grid has the additional resources needed to improve the performance of any and all host IO activity.

The technical advantages described above translate into real customer benefits like worry-free data storage that is simple to deploy, configure and manage. With grid storage's scatter/gather IO, the system automatically handles data location, eliminating data placement and performance concerns (e.g., determining which physical storage to associate with which host volume data and which controller or RAID group is being overtaxed).

In addition, unlike classic dual-controller systems, customers with grid storage can easily scale-out performance with capacity. As previously discussed, a node added to a grid includes processors, cache and data drives, and storage IO performance improves linearly as the system automatically redistributes host data across new nodes.

Grid-scale storage at IBM

IBM offers three grid storage solutions: IBM Spectrum Accelerate™, IBM XIV™ and IBM FlashSystem A9000/IBM FlashSystem A9000R™ all-flash grid storage. All IBM grid storage products share similar functionality, though their performance differs. All of IBM's grid functionality originally came from XIV but is now implemented in Spectrum Accelerate.

- **IBM Spectrum Accelerate** – software defined grid storage with a number of advanced storage features such as two or three way mirroring, differential snapshots and asynchronous or synchronous replication. Spectrum Accelerate has a world class, easy-to-use graphical user interface with a RESTful API for data center automation and provides extensive support for VMware and Microsoft storage services. Spectrum Accelerate is also available in a hybrid cloud solution with IBM SoftLayer services.

- **IBM XIV storage system** – hardware defined grid storage providing disk-only or hybrid disk-flash cache storage with InfiniBand intercluster networking and uninterruptible power supplies. XIV supports Real-time Compression™ and self-encrypting drives for greater storage efficiency and security.

- **IBM FlashSystem A9000 and IBM FlashSystem A9000R storage systems** – hardware defined grid storage providing high IOPS, low latency flash-only storage with Spectrum Accelerate grid nodes and world class, high performing IBM FlashCore™ technology, flash enclosure nodes. IBM FlashSystem A9000 has a single flash enclosure node with three grid controllers, while IBM FlashSystem A9000R supports multiple flash enclosures and grid controllers.

Application environments that benefit from IBM grid storage

Virtualisation environments

IBM Spectrum Accelerate provides support for VMware's advanced storage services such as VMware Virtual Volumes (VVOLs), storage control with vSphere APIs for Storage Awareness (VASA), and vStorage APIs for Array Integration

(VAAI). Such extensive support for VMware services makes any IBM grid storage solution a natural fit for VMware VM storage. Equally important, VMware comingles IO streams from a number of VMs into one server's IO requests. Eliminating hot spots through widely disbursed grid storage with self tuning, consistent IO performance and linear performance scalability can help to service this demanding, mixed IO environment. Comprehensive support for Microsoft Hyper-V replication automation with System Center Virtual Machine Manager (SCVMM), Microsoft Azure Site Recovery, and system monitoring through Microsoft System Center Operations Manager (SCOM) together with Hyper-V's similar IO performance characteristics, qualifies IBM grid systems as ideal storage for Microsoft Hyper-V customers as well.

Cloud and service provider environments

In contrast to dual-controller architectures, the linear scaling of performance with capacity, autonomic tuning and multitenant support of IBM grid storage enables cloud environments, managed service providers and other 'anything as a service' (XaaS) organisations to rapidly and easily deploy host volume data to IBM grid storage. Moreover, the ease of storage management and RESTful API support make IBM grid storage much simpler to set up, automate and use on an ongoing basis. These benefits and the grid's consistent IO performance are especially important for cloud and service provider environments, which have constantly changing application and data workloads.

Analytics environments

The widely striped IO functionality of grid storage performs better than dual-controller architectures for customers that want to run analytics and applications on the same host data. No hot spots mean analytics IO will interfere less with applications IO. Grid storage IO performance is at its peak when applications and analytics data are deployed on IBM FlashSystem A9000/IBM FlashSystem A9000R all-flash grid storage. Further, with Spectrum Accelerate storage, customers can deploy their analytics applications on the same servers as the grid storage modules.

Summary

Grid storage has many characteristics that make it an ideal enterprise, cloud and service provider storage architecture. From a technical perspective, the inherent ease of configuration, scale-out performance and autonomic tuning of grid storage can't be beat by classic dual-controller architectures. Such advantages translate into measurable customer benefits like linear performance scalability and worry-free data storage.

Equally important, with the software-defined Spectrum Accelerate, XIV disk-only and hybrid (disk-flash), and the IBM FlashSystem A9000/IBM FlashSystem A9000R all-flash storage systems, IBM has implemented the most comprehensive range of grid storage technology alternatives available on the market today.

Indeed, the fact that IBM has added the world-class FlashSystem technology to their grid storage offerings is testament to the enduring virtues of this storage architecture. In the end, enterprise data center and cloud/service provider customers can't go wrong with one of IBM's grid storage solutions.