The High Performance Alliance: SAN, SSD and Virtualization

Introduction

Companies in nearly every business and discipline are deploying Storage Area Networks (SANs) as a powerful means of controlling the escalating cost and complexity of administering, managing, and moving data. Initial SAN implementations have been primarily focused on issues such as server-less backup, storage consolidation, ease of management, and/or high availability operation. While each of these benefits is positive for a company, there are incremental advantages to be gained by configuring a SAN to provide maximum performance to an ever-wider range of applications. As this paper explores, impressive application performance gains can be realized by incorporating solid-state disks (SSDs) into SANs, with maximum results achievable in a virtualized SAN environment.

SANs: An Efficient Use of Resources

A typical enterprise storage plan requires projecting storage capacity and building in extra capacity to avoid emergency situations where applications could fail due to insufficient space. This surplus capacity typically ranges from 10% to a whopping 50% of total storage capacity, depending on how predictable the application is and what risks the storage administrator chooses to take.

In IT environments where each server has its own direct-attached storage (DAS), that 10% to 50% of unused space quickly adds up when multiplied by the number of servers in use. Aggregated, the cost of the surplus gets very expensive very quickly as a result of the extra hardware purchased and the necessary personnel to administer it.

One of the driving reasons to implement a SAN is to mitigate the headaches of direct-attached storage and make more efficient use of storage resources and administrative personnel. SANs allow enterprises to consolidate individual storage requirements into a storage farm, managed separately from the server farm. Instead of dealing with storage per server, storage administrators manage the entire storage farm for the enterprise. This one-to-many configuration allows for more efficient capacity provisioning, economical use of storage administrators' time, and improved reliability, availability, and scalability. While a SAN environment can sometimes add complexities of its own, sophisticated storage management tools have been developed that alleviate these complexities. One of these tools is storage virtualization, a technology that is addressed later in this paper.

SSD in the SAN

Storage devices most commonly deployed in SAN implementations include cached disk arrays, JBOD, and tape drives. But a key SAN benefit is the ability to make more efficient use of any and every kind of storage resource. Companies can leverage strategic storage resources on an application-by-application basis to any server on the storage network. Specialized storage devices like solid-state disks (SSDs) or file cache accelerators can be positioned in the SAN fabric and manipulated in the same manner as traditional SAN storage resources.

SSDs have historically been used to successfully respond to I/O bottlenecks affecting application performance on a single server, single application basis. Typically, 3% to 5% of application data accounts for approximately 50% of all I/O activity. Isolating these "hot files" on an SSD typically improves individual file-level performance 200-800% and commonly increases overall application performance by 30-40%.

There are numerous applications in the broader market that could benefit from such a high-speed storage device, and SSDs are proven capable of delivering major improvements in performance in a variety of environments. However, justifying an SSD as a dedicated device for the broader spectrum of applications has been historically difficult. This is exactly where the synergy of SANs and SSDs occurs. Leading SSD devices support up to 8 or more Fibre Channel ports for switched SAN environments. They

also are partitionable into numerous logical unit numbers (LUNs) to effectively present volumes to multiple applications on multiple servers simultaneously. And finally, they can be managed using standard off-the-shelf SAN and SRM tools. These capabilities allow the performance benefits of SSD to be effectively leveraged across the SAN and specifically directed to areas of a SAN where performance acceleration is most needed, when it's needed. This powerful capability enables SSDs to be increasingly used to supercharge applications when only periodic high speed disk is desired for peak times, as well as more traditional applications that need full time performance enhancement, such as database logs or message queuing systems.

Storage Virtualization: Getting the Most Out of SANs

Ironically, while the intended purpose of SANs is to simplify matters, there is often a high degree of complexity inherent in the implementation and subsequent maintenance of a traditional SAN solution. And even though basic storage devices may be getting cheaper, complexity is not just frustrating; it's also expensive. The solution is Storage Virtualization. A superior storage virtualization solution simplifies storage provisioning and reduces administrative overhead. It also enables and simplifies the targeted provisioning of resources, so that the fastest storage (e.g., SSD) can be provisioned to those applications that need it, when they need it, for maximum performance.

What is storage virtualization? Simply defined, it is a layer of abstraction that separates the representation of storage from the physical devices themselves. It provides the ability to consolidate multiple physical storage devices of various interface protocols (such as SCSI, iSCSI, Fibre Channel) into a logical "storage pool." From this generally available storage pool, virtual storage devices can be created and provisioned to application servers and end users, providing the following benefits:

1: Easier storage resource management for the system operator
2: Maximized utilization of all physical storage space
3: Improved storage network performance
4: Elimination of downtime to add and reconfiguration storage resources
5: Automated capacity-on-demand for assigning storage based on policy-driven strategies

The amount of surplus storage capacity that must be maintained to avoid out-of-space conditions is radically reduced compared to non-virtualized SANs (which is already much less than DAS environments). Idle or unused storage is eliminated as a result of the pooling of *all* the available storage capacity—and subsequent monitoring and utilizing of it as a whole unit. Administrators can be proactive, rather than reactive, regarding necessary changes to storage. In advanced installations, the response can even be instantaneous where policies add storage capacity automatically when pre-defined thresholds are reached.

In non-virtualized storage farms, there may be physical co-location of storage devices - but each device is managed, configured, and licensed separately, which commonly results in vendor incompatibilities in management capabilities. In comparison, in-band storage virtualization offloads virtualization and other storage services from the servers and storage devices and puts the intelligence into the network. Storage Virtualization allows enterprises to consolidate all their storage (JBODs, RAID arrays, tape drives/libraries, SSDs etc.)… into a *unified pool of heterogeneous storage*, independent of hardware type, vendor, and protocols. The result is a centralized storage solution (including services, configuration and management) that delivers accelerated performance and boundless scalability.

Optimizing SANs by Using Solid-state Disks in a Virtualized Environment

Storage virtualization eases the burden on the SAN administrator not just by providing a more manageable interface but also by granting the ability to assign storage resources based on performance characteristics and application performance requirements. This empowering capability opens up valuable possibilities for "SAN tuning" to address Quality of Service (QoS) issues within the SAN. Since the SAN infrastructure shares common components (switches, directors, and storage resources), a valid concern

is ensuring that high priority data requests are responded to promptly and not queued behind pending low-priority data requests (i.e. Quality of Service). Virtualization facilitates the development of a flexible, custom-designed hierarchical storage model within the SAN that responds to these QoS issues. SSDs can be deployed into the virtualized SAN to dynamically deliver top performance and a statistically consistent level of data accessibility to those applications at the top of the hierarchical ladder that stand to gain maximum benefit from such devices.

Conceptually, adding solid-state disk to the virtual storage pool can open up some very beneficial scenarios. The virtualized SSD becomes an architectural resource that can be directed within the SAN to address specific performance issues and seamlessly migrated between platforms, operating systems, and applications for true point-and-click performance resolution. For example, the SSD can be utilized on January 31st to speed up month-end closing on an NT based financial application. On February 1st, the SSD (or a portion thereof) can be allocated to help the engineering department with a Unix-based OS conversion dynamically by the virtualization management capabilities. This kind of mobility and scalability allows the organization to leverage the performance benefits of solid-state disk across a broader application spectrum and recognize the benefits in virtually every facet of the enterprise, thus reducing the total cost of ownership.

The ultimate goal for the administrator is to not have to worry about SAN tuning at all. The virtualization engine monitors all data requests to the storage devices and gathers statistics on the location of the data hot spots. Eventually, the ability to dynamically act on this information and move the most requested data to the fastest devices will provide automatic SAN tuning. This automatic SAN tuning ensures that the entire spectrum of applications will have access to the right storage at the right time.

Conclusion

SANs are an enabling technology that is in the late adoption phase of the technology lifecycle. The networking aspects of SANs lend themselves to deploying other enabling technologies in an increasingly more powerful and broad-based role across the organization. With innovative storage virtualization tools that continue to evolve to better help manage storage environments, companies can focus on tuning the SAN for maximum benefit to all applications. Designing solid-state disk into the virtualized SAN provides a powerful and cost-effective resource that can be used to significantly improve performance, productivity, and profitability across the enterprise.

****

In 1978, Imperial Technology revolutionized the storage industry with the first MegaRam storage accelerator. Today, Imperial is the leader in solid-state accelerators and intelligent caching appliances for supercharging storage systems and database applications. These devices allow users to eliminate I/O bottlenecks and dramatically improve the performance of their systems without the substantial expense of re-writing applications or upgrading to faster servers. CPUs, now unhampered by slow disk response, can run applications much faster. Results can be seen immediately in better terminal response times, faster queries, and quicker completion of batch jobs.

In 2001, FalconStor Software revolutionized the storage industry again with its IPStor storage networking infrastructure software, whose advanced virtualization engine accomplishes all that storage virtualization has promised to be—and much more. FalconStor engineered IPStor based on the knowledge that it is not enough to simply enable a virtualization process. The true value of virtualization can only be realized within a full-featured future-proof storage solution that leverages current IT infrastructure and provides not just virtualization but a comprehensive suite of storage services, top performance (fast data storage and access), ease of use, unified SAN and NAS provisioning and management, virtualization *across* cabinets (not just within one storage cabinet), seamless support of heterogeneous storage environments, and boundless scalability. This can only be achieved with an open architecture software solution based on open standards (SCSI, Fibre Channel, IP/iSCI, and the emerging InfiniBand) that offers total freedom of

choice in hardware (SSDs, JBODs, RAID arrays, tape drives/libraries, etc.), device vendors, interfaces, connectivity protocols, and OS platforms. IPStor works with all management tools, reporting utilities, and maintenance procedures.

IPStor allows storage administrators to flexibly choose which "hot" files to move onto Imperial's MegaRam accelerator, whenever they choose, so that read/write operations can be performed at memory speed for application optimization. MegaRam solid-state disk (SSD) accelerators and IPStor work together seamlessly to remove the performance imbalance between the high speed of the CPU and the relatively slow speed of traditional storage mediums.

