

# The Promises and Pitfalls of Virtualization



Storage-bound Random I/O and the Need for Performance VM Datastores

An analysis by Astute Networks



## More VMs Equal More Headaches

The promised benefits of virtualization are appealing: improved utilization, consolidation, management and performance. Then reality sets in. You've virtualized, but your benefits are throttled by current or potential performance problems. You'd like to consolidate even further, but running more than five to ten Virtual Machines (VMs) per server bogs down applications. Adding more memory, CPU cores or better Network Interface Cards (NICs) doesn't help. You've tried to virtualize your database, Enterprise Resource Planning (ERP), media, graphics or other process-intensive applications, but users grumble when performance slows. Database queries and reports take hours instead of minutes, daily batch processes chug along for a day and a half and customers complain when transactions or projects are delayed. Management is upset when Service Level Agreements (SLAs) aren't met and they want to know why your budget for new servers and storage capacity keeps escalating. They also want you to implement a Virtual Desktop Infrastructure (VDI) to save money, but you're concerned that performance will grind to a halt. The ugly truth is that virtualizing often introduces more problems than it solves, and can actually increase costs if not implemented properly.

60% of server workloads will be virtualized by 2013

An average physical server will host 10 virtual servers

15% of enterprise PCs will be virtualized by 2013

Gartner, 2011

A recent Forrester Research report revealed that server virtualization has matured extensively over the past year and is now widely adopted for use in critical production level applications. These include mission critical data-

bases, with deployments growing from 53% to 68%, and email servers, showing an increase from 29% to 51% year over year. Gartner Group estimates that by 2013, 60% of server workloads will be virtualized, and the

average physical server will host no more than ten VMs.

Virtualization is showing all the signs of being mature and well entrenched in the datacenter. Following suit, everything that surrounds the central function of processing is also maturing to support virtualization's unique requirements. Servers are sporting multi-core processors and improved capabilities that empower virtualization. NICs, network switches and "pipes," such as 10 Giga-bit Ethernet, are "upping their game" to help provide increased throughput for virtualized infrastructures. What still lags, however, is adequate storage-bound performance that is capable of keeping up in a virtualized world.

### Companies virtualize to:

Reduce workloads & enhance management ..... >



Lower hardware and energy costs ..... >



Consolidate and improve server utilization ..... >



Increase application performance ..... >



## Say Goodbye to Sequential

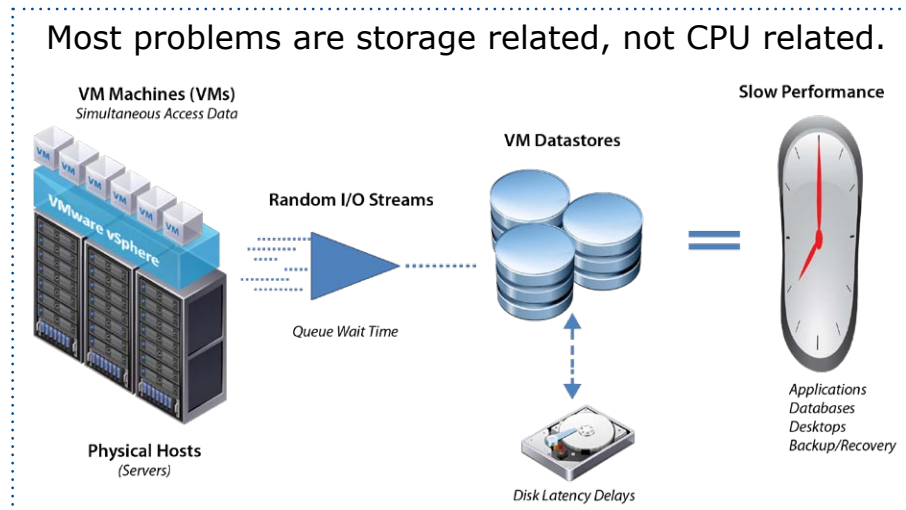
Almost every IT professional understands that while a VM sits on a host, or physical server, VM datastores—those repositories where the data resides—are housed on storage devices, usually a Storage Area Network (SAN). Server consolidation simplifies access to this storage, which may itself be virtualized, but server virtualization poses a significant challenge for storage performance.

Physical servers usually have dedicated storage Input/Output (I/O) access to separate storage systems or specific allocations within a storage system. I/O requests for data residing on that storage is more sequential, or serial in nature. The performance required by each physical server, measured as Input/Output Operations per Second (IOPS), is usually moderate. Storage resources can be configured to provide high throughput to particular servers that need it, and I/O bottlenecks are usually not an issue. Virtualizing changes all that. Where sequential order and structure once ruled the day in a physical server datacenter, random chaos now reigns supreme.

Server virtualization increases resource utilization by running each application or workload in its own virtual server. CPU utilization goes up, and power, cooling and floor space are conserved. Each virtual server shares the system resources available to the physical server hosting the virtual servers. But, with all these servers requesting access to storage, available storage I/O becomes a bottleneck. Adding a significant number of virtual servers to a physical server translates to a much larger storage I/O requirement and results in more random data accesses—sometimes by an order of magnitude. Traditional hard drives, designed to handle sequential data requests, simply can't keep up. Even the best SAS drives running at 15,000 RPM still suffer from latency delays of more than three milliseconds, which is an eternity for a process-

intensive virtualized application.

Although individual virtual servers may have modest storage I/O needs, when a dozen or so virtual servers are hosted on a single physical server, the demand for storage I/O skyrockets. Typically, a server hosting fifteen VMs may require, on average, 2,250 IOPS or more. As the number of virtual servers per physical server increases, so do the IOPS required. A midsize data center with twenty-five physical host servers can require 56,250 IOPS for the hosted VMs. Other physical servers, not dedicated to



virtual hosting but to a particular high I/O application, can drive these requirements even higher. Note that these assumptions do not account for high email traffic, backup or other data protection activities, which are also IOP hungry.

Virtualizing desktops exacerbates this problem as VDI causes "I/O storms" when every morning users log-in to their virtual desktops and overwhelm storage system performance.

## Higher VM Count Means Bigger Bottlenecks

The upside of increasing the number of virtual servers hosted per physical server is that the cost-effectiveness of the virtualization environment goes up. The downside is that server hosts become more limited by available storage I/O. Anticipating the IOPS that a virtualized infra-

| Item                       | IOPS          |
|----------------------------|---------------|
| SQL Database               | 1,800         |
| Oracle Database            | 1,200         |
| ERP Application            | 1,000         |
| MS Exchange                | 2,000         |
| 10K Emails/Day             | 10,000        |
| 6 Perf Apps                | 6,000         |
| 1,000Desktops              | 25,000        |
| 75 VMs                     | 11,250        |
| <b>TOTAL IOPS</b>          | <b>58,250</b> |
|                            |               |
| <b>SAN Disks</b>           | <b>IOPS</b>   |
| <b>40 x 500GB (20 TBS)</b> | <b>8,000</b>  |
|                            |               |
| <b>IOPS Deficiency</b>     | <b>50,250</b> |

|                                       |               |
|---------------------------------------|---------------|
| Number of Desktops                    | 3,500         |
| Launch Rate (in seconds)              | 0.5           |
| Login IOPS (Win7=~100)                | 100           |
| Workload IOPS (Login VSI Medium = ~5) | 5             |
| Desktop Login Time (in minutes)       | 0.5           |
| <b>Peak IOPS</b>                      | <b>28,800</b> |
| <b>Steady-State IOPS</b>              | <b>17,500</b> |
| <b>Estimated Boot IOPS (+/- 10%)</b>  | <b>77,000</b> |

Sources:

- 1200 IOPS per Oracle DB from VMware Capacity Planner data
- 1800 peak IOPS per SQL DB from Sep 2009 Pivot Point study
- 2000 IOPS per MS Exchange from EMC Virtualized Architecture for Microsoft Exchange Server Whitepaper, July 2009
- 150 IOPS average per VM and 200 IOPS maximum per SAS 15K drive from Storage Switzerland article, Sep 14, 2010
- 25 IOPS per medium VDI user from Citrix VDI IOPS Estimate, Nov 2010
- 1000 IOPS per ERP or performance app from Aventi Group user study Q1 2011

structure will need, to help prevent performance issues, is a required prerequisite to any virtualization project. One of the best ways to accomplish this is to create an IOPS calculation worksheet like the one shown below.

## How Do We Solve the Problem?

Attempting to accommodate ever-growing demand for storage capacity has come at the cost of performance I/O and throughput. Random data access, in particular, suffers under the strain of growing server virtualization, unstructured data, deduplication and other technologies that have randomized I/O.

There are several techniques available that attempt to remedy the I/O bottleneck. These include:

- Adding more hard disk drives or storage systems
- Hard drive short-stroking
- Adding solid-state flash drives as a cache or “tier-zero” storage

## Adding Hard Drives

The initial response for most IT professionals battling an

I/O bottleneck problem is to load up on more and faster spinning disks. This fails to solve the need for improved random access, and is costly in terms of power, cooling, and data center space. The performance gains from each additional drive are marginal and decreasing, especially as the storage controller starts to become the primary bottleneck, and the wasted non-utilized storage can be significant. Keep in mind that even a fast SAS drive can't achieve more than 200 to 250 IOPS, even by short-stroking the drives. If you have an IOPS deficiency of 50,000 IOPS, as shown in the example above, you'll need something like 250 hard drives.

Of course, your storage vendor would be thrilled to sell you an additional 100 or 200 terabytes (TBs) of storage to tackle your IOPS deficiency, but will your budget support all that wasted capacity? And let's not forget that storage management complications and hassles will increase. Bottom line: the ROI of adding more spinning disks just doesn't make sense.

## Short-stroking

An alternative to buying more storage is hard drive short-stroking. The chief factor determining random IOPS

performance is the random seek time of the disk drive. Short-stroking improves random IOPS by reducing the distance the read/write head must move when seeking new random data. By limiting the number of tracks used on each hard disk drive to just the outer tracks on each platter, random IOPS can be improved by up to 50%.

However, short-stroking leads to a sizable reduction in drive capacity. Usually, only the outer third or less of the platter tracks are made available, reducing capacity by 66 to 90%. Power and cooling requirements remain the same for each drive, and as more drives must be acquired to meet overall capacity requirements, power and cooling costs will go way up, along with the need for more space, cables, etc. Ultimately, capacity is limited by the size of the storage system, which again puts the data center on the path to overbuying storage systems to try to defeat the I/O bottleneck. Although this tactic can work for some data centers, short-stroking's tradeoff is high cost and complexity.

## SSDs

High-performing Flash SSDs (solid-state drives) are ideal for random access I/O. They're fast, with random IOPS in the thousands or more, and can be as reliable as hard disks. In most cases, however, they're also lower in capac-

ity and more expensive than hard drives. Still, a Jetstress test completed by Principal Technologies showed that only three SSDs outperformed twenty-four HDDs by almost 200% for delivering sustained IOPS.

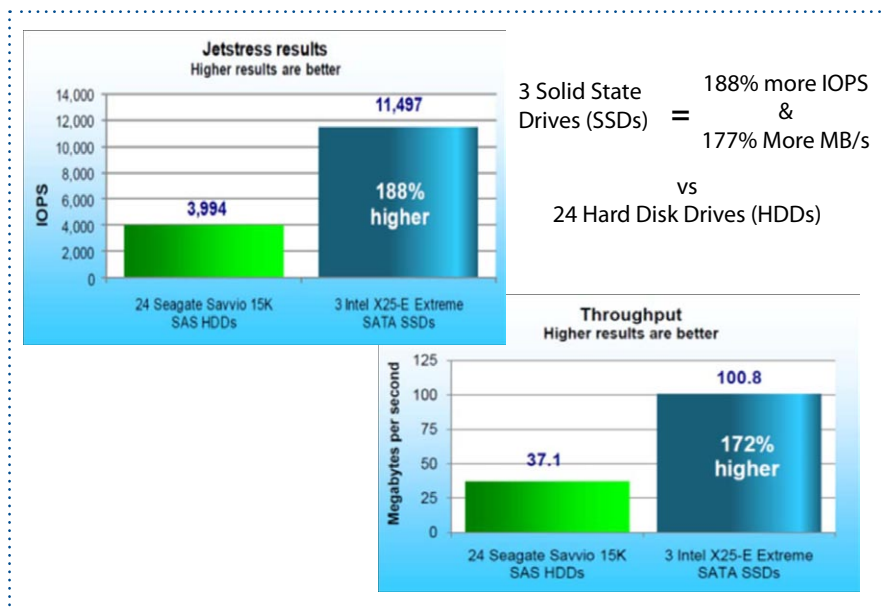
Even a year ago concerns about SSD reliability and endurance prevented mainstream storage vendors from adopting SSDs, but newer SSD technologies now offer enterprise-class endurance with improved error checking. Most storage vendors today prefer to use SSDs as a cache in attempting to improve IOPS performance. Data is written first to the cache, and this is where the storage system first looks for data on reads. Using SSDs as a cache can provide a strong boost to IOPS performance, but it's important to note that this is typically burst IOPS and not sustained IOPS. When it comes to a virtualized environment, the difference is critical. Remember, virtualization results in an order of magnitude more random I/O, which requires sustained IOPS. Simply adding a solution that provides an occasional IOP boost will be like asking a race car driver to constantly pull in for a pit stop.

Another factor to consider is cost. Even when limited to cache applications, adding SSDs to a SAN can increase system costs substantially. In response, system architects try to minimize the SSD cache used. To get a cache hit rate of at least 50% implies that a cache should comprise

10% or more of total storage. If a 100TB storage system has less than 10TB of cache, the hit rate falls off exponentially. Moreover, managing a cache can be complex and time-intensive, further raising costs.

## Tier-Zero

An alternative approach is to use SSDs for the primary storage tier, or tier-zero. This tier contains the most recently stored and accessed data, and data used by applications requiring high random access performance. Over time, if not needed, the data is moved to lower storage tiers using



less costly, slower drives. The storage vendor is able to reduce the number of high performance high-cost hard drives, instead using a smaller number of high capacity lower performing hard drives together with fast SSDs to provide the performance needed. Power, cooling, and space requirements are all reduced for further savings.

The key to making this scheme work is the data migration process, which must either be done manually or through storage tiering software. Automated storage tiering can be expensive, is usually unable to handle dynamic workloads, and lacks granularity. Additionally, this process tends to migrate data from faster tiers to slower tiers, and cannot be counted on for the precision and accuracy required to ensure that data in the primary storage tier is available when needed. In reality, IT professionals will often take the costly step of over-provisioning each tier to get the performance they need.

There are other issues with automated tiering. Because data is frequently moving from one tier to another, data requests can be stalled while waiting for the move to complete. Many storage managers like to have better control over what data is stored on various tiers. Automated storage tiering also tends to increase power usage substantially, as disks spend more time spinning.

Lastly, for automated tiering or caching schemes to be effective, they must be predictive. That is, they need to be able to predict what data will be needed in advance and place it on the SSD tier. This can work fine in a physical server data center where I/O is sequential, but is not very effective in a virtualized environment where I/O requests are highly randomized.

Nonetheless, automated storage tiering paired with SSDs can provide a real I/O benefit for non-virtualized data centers. Unfortunately, most process-intensive applica-

tions require that data reside on tier-zero at all times, sidelining the potential benefits of automated tiering. Given the typically high costs for storage tiering solutions, the hoped-for savings usually fails to materialize.

One last point to consider: while it's tempting to turn to your favorite storage vendor for a "blended" SSD + HDD solution, or have them slap in some SSDs into your current SAN, taking this approach is like trying to drive that racecar on a dirt road. Traditional SANs weren't designed to support high-IOP and high-bandwidth SSDs. They were built to support slow-performing HDDs. As such, backplanes, controllers and architectures for these systems induce bottlenecks that can actually throttle SSD performance. To try and overcome this, even to a small degree, SAN vendors usually inject a high degree of cost,

which is transferred to you. This is especially true for data centers using iSCSI.

Most IT professionals are well aware of the differences and relative benefits of Fibre Channel (FC) versus iSCSI versus Fibre Channel over Ethernet (FCoE). The synopsis is

that FC often supports higher throughput but comes with cost premiums. iSCSI over 1GbE was introduced as the slower option but at a much lower cost. With the advent of 10GbE, the performance gap has narrowed, and in some cases has even eclipsed FC. In fact, the proliferation of FCoE is a strong indicator that Ethernet is starting to encroach on traditional FC territory.

For the purposes of our discussion, it doesn't matter if your SAN is FC or iSCSI, or if you have 1 GbE or 10 GbE. As we've discovered earlier, to overcome IOP deficiencies in a virtualized environment, you need tens of thousands of sustained (not burst) IOPS. The best way to gain IOPS is via SSDs, but caching, auto-tiering or other "slap-on" schemes are expensive and ineffective in resolving VM datastore bottlenecks. So where does that leave us?

***Adding SSDs to traditional SANs is like driving a racecar on a dirt road.***



*Backplanes, controllers & architectures for traditional SANs were designed for low-IOP HDDs, not high-IOP SSDs*

## A Balanced Approach

Wouldn't it be great if there was a simple, efficient and elegant solution that solved our performance problems without breaking the bank or requiring a forklift? A purpose-built SSD VM Acceleration Appliance that cost one dollar sign versus ten? What if we could simply plug this iSCSI appliance into an Ethernet port, turn it on and configure everything within minutes? What if this excit-

tion and on-going administration.

ViSX is compatible with any virtualization suite, including VMware, Hyper-V and Citrix. With VMware, ViSX is "carvable" as a vSphere datastore. You can migrate VMware disk files using Storage vMotion, and can boot an ESXi server from a ViSX iSCSI volume for stateless operation, fast recovery of failed VMs, and improved VM management. For faster disk-based backup and recovery, you can

### Fully Integrated Solid State VM Acceleration Appliance

- Optimized for VMware
- Small footprint 3U chassis
- 1.2TB, 2.4TB, or 4.8TB of lightning-fast MLC Flash
- Optional 1 or 2TBs SAS for disk backup & configuration
- Patented Athens multi-core storage processor
- Validated iSCSI and TCP offload performance
- Properly balanced latency, IOPS, and bandwidth
- Multi-level RAID 0,1,5,6,10 data protection
- High-speed network ports (2x10GbE or 2x1GbE)
- Converts DAS islands into shared networked storage
- Integrated system management console
- Hot-swap Flash drives; Hot-swap power supplies
- Zero-footprint capacity upgrade
- Installs and configures in minutes



ing new appliance could deliver 80,000 IOPS and 670 Megabytes of bandwidth in real-world performance tests without breaking a sweat...or the bank? Would it be too much to ask for 10GbE and VM datastore optimization?

The Astute Networks StorPoint ViSX VM Acceleration Appliance is a 3U rack-mount system with SSD Flash capacities of 1.2TB, 2.4TB or 4.8 TB. You can optionally blend in up to 2 TBs of SAS hot-swap removable drives for fast VM backup/recovery. ViSX connects via 1GbE or 10GbE iSCSI, and features Astute's patented multi-core iSCSI TCP offload storage processor. The result is a solution that is end-to-end balanced with unparalleled performance at an unmatched price to performance ratio. ViSX comes bundled with a secure GUI-based user-friendly point-and-click management console that simplifies configura-

tion and on-going administration. ViSX is compatible with any virtualization suite, including VMware, Hyper-V and Citrix. With VMware, ViSX is "carvable" as a vSphere datastore. You can migrate VMware disk files using Storage vMotion, and can boot an ESXi server from a ViSX iSCSI volume for stateless operation, fast recovery of failed VMs, and improved VM management. For faster disk-based backup and recovery, you can

Within hours after installing StorPoint ViSX, which usually takes less than thirty minutes, most customers report astonishing performance results. Servers can handle 10X the number of VMs and VDI is 10X faster. Batch processes, database queries and reports are completed in one-tenth the time. Applications, including OLTP, OLAP, Business Intelligence, media/graphics, CAD, analytics and others hum along at then times their former speeds. Power, cooling and rack space requirements are reduced by 90% or more.

With the Astute Networks StorPoint ViSX, the promise of virtualization can finally become a reality, and compared to other alternatives, this powerful and simple solution is one-tenth to one-twentieth the cost. To find out how StorPoint ViSX can improve application performance in your data center, visit Astute's online calendar today to schedule an appointment, or send an email for more information.

**Online Appointment Calendar:**

[www.supersaas.com/schedule/Aventi\\_Group/astute\\_networks](http://www.supersaas.com/schedule/Aventi_Group/astute_networks)

Email: [info@astutenetworks.com](mailto:info@astutenetworks.com)

Toll-free Phone: 866-673-7701

## Summary

- Virtualizing creates a more randomized input/output (I/O) environment
- VMs are on servers, VM datastores (data that apps need) are on storage devices
- Multiple VMs on servers are randomly competing to get data to/from SAN
- Spinning disk drives in SANs don't like randomized I/O and have latency delays (slow)
- The only way to solve this is to increase your Input Output Operations per Second (IOPS)
  - » A 10 TB SAN with spinning disks is limited to 2500 IOPS on average
  - » Most virtualized apps, like SQL or Exchange, require almost 2000 IOPS alone
  - » Adding up IOPS requirements usually shows a 10K to 70K IOP deficiency
- The best way to increase IOPS is with solid state drives (SSDs)
  - » SSDs are now very reliable, but most vendor solutions are still expensive
- Putting SSDs into a standard SAN is like driving a racecar on a dirt road
  - » Typical SANs are built for hard drives, not SSDs
  - » Their backplanes, architectures and controllers can't handle high IOPS
  - » Auto-tiering or cache SSD solutions are not suitable for random I/O VMs

- Astute Networks has the only purpose-built SSD appliance optimized for virtualization
- 3U appliance, 300 watts power, plugs into any Ethernet port, configures in 20 min
- Has patented storage processor for iSCSI offload (allows higher bandwidth and IOPS)
- Can connect to a fibre channel network via Ethernet port on switch
- Comes with 1.2, 2.4 or 4.8 TB SSDs and up to 2 TBs SAS drives for disk to disk backup
- Delivers 80,000 IOPS and 670 MB/s bandwidth
- Developed over 10 years, 3rd generation, used in military/telecom environment
- VMware certified and optimized for virtualization (VMware, HyperV, Citrix, etc.)
- Cost per IOP is one-tenth that of other solutions
- ROI is immediate
  - » If you are buying new servers within 6 months, you can double or triple your VMs instead. Saving one server = the cost of one ViSX
  - » If you are buying more storage within 6 months, you can improve performance by 10X and your cost per TB with ViSX is comparable to that from most SAN vendors.

