Optimizing Storage Performance and Capacity in the Virtualized Data Center: Dell’s SAN HeadQuarters

October 2009

Virtualized storage is the foundation for datacenter virtualization. In order to fully realize the benefits of virtualization, whether for server consolidation, desktop virtualization, or both, storage virtualization is essential, and most often takes the form of a storage area network (SAN). The SAN environment must be as dynamic, flexible and mobile as the virtualized workloads it supports, but these characteristics can present new challenges and complex management issues, particularly in rapidly growing virtual infrastructures. Fortunately, there are powerful monitoring solutions available for the SAN that gather, integrate and report key performance metrics that are essential to support storage planning, troubleshooting and daily operations – and they are designed for today’s integrated server-storage virtualization operations teams.

In this Product-in-Depth, we review Dell’s SAN HeadQuarters (SAN-HQ), a no-cost storage monitoring solution designed for and included with the Dell EqualLogic PS Series iSCSI SAN family. SAN-HQ delivers consolidated and centralized storage performance, capacity, and network monitoring and can dramatically reduce the time required to identify performance bottlenecks, correlate them to events in the virtual server infrastructure, and develop remediation strategies. SAN-HQ includes storage monitoring capabilities found in costly competitive vendor offerings, and is intelligently integrated with Dell’s PS Group Manager SAN configuration software. We review SAN-HQ’s key features and benefits using a use-case approach, outlining the most common challenges in virtualized infrastructures - at various stages of maturity - and the questions most frequently asked by administrators.

Virtualization Emerges from the Sandbox

The mobility of virtual machines, and the dynamic nature of the connections they make to storage and network resources throughout the IT infrastructure, cause new types of contention and capacity problems to emerge. These are not adequately addressed by existing solutions designed for a much more static environment. In particular, storage performance issues are often the Achilles’ heel of production server virtualization deployments, and can expose a lack of visibility across dynamic server and storage domains.

Server virtualization has been proven in the development and test sandbox over the last decade, and is now considered production-ready. As enterprises deploy virtual servers into production, however, it is imperative to recognize that deployment strategies for production are not the same as they were in the sandbox. In development and test environments, workloads are often spiky
(compiles, test harness runs) and capacity is often the primary server and storage design criteria: many administrators find it easier to maintain excess capacity and to manually move workloads as required.

Capacity-focused planning and reactive contention management, however, in a virtualized production environment, may mask potentially serious underlying design problems. Early-mover enterprises that have deployed critical applications on virtual servers have exposed some unexpected storage contention problems. In addition, large virtual desktop infrastructure (VDI) deployments create additional strains on storage capacity and performance planning. As virtualization breaks out of the sandbox, it demands a comprehensive solution for storage monitoring, planning, and reporting.

**Do I Need Storage Monitoring?**

We are seeing the demand for virtual storage monitoring and optimization following the same evolutionary path observed in the virtual server and desktop domains. Initially, a core set of key capacity and performance indicators, such as CPU and memory allocations and utilization, were developed for virtual machines (VMs). These provided a coarse-grained view of individual VM performance. As the server environment grew and advanced host capabilities were added (resource scheduling, live migration, high availability, etc.), administrators needed monitoring – and correlation of monitored attributes – spanning groups of VMs, clusters of host servers, and multi-site deployments.

In the virtualized storage realm, the same requirements are needed today, driven by similar advanced technology available in virtualized storage arrays. Dell’s EqualLogic PS Series features automated, real-time load balancing across drives, arrays, connections, cache and controllers in order to optimize performance. Administrators are increasingly demanding visibility into this behind-the-scenes activity, driven by several key benefits:

**Better Visibility = Faster Diagnosis:** In a dynamic production environment, server, storage and cross-domain operations teams need storage performance data correlated across various attributes at their fingertips to quickly identify what is being used and by whom, and the ability to drill down instantly to locate a source of contention at the object level: storage groups, pools, volumes, disks, and more.

**Better Data = Faster Resolution:** Once a problem has been identified, the administrative team needs to get everyone in agreement on a plan of action. With detailed historical data on hand, it’s easier to correlate a problem in the array with an event in the server tier and to assign remediation to the right team or to develop a joint attack plan.

**Better Data = Accurate Planning:** With detailed storage capacity and performance available both historically and in real time, administrators can establish meaningful baselines, estimate available headroom, and build accurate trending models. This enables the environment to grow only as much as needed, when needed, and to take advantage
of the modular design of virtualized arrays to avoid both under- and over-provisioning.

Who Should Explore Storage Monitoring?

Wherever you are on the road to server and storage virtualization, there are compelling reasons to implement comprehensive storage monitoring. As their virtualization maturity evolves, we find that companies typically find themselves faced with one or more of the following virtual infrastructure situations, which we'll call profiles. For each profile, we'll define the situation, review how storage monitoring in a dynamic virtualized environment provides proven benefits, and highlight some of the key questions that such storage monitoring will answer.

Profile 1: Migration
Deploying New Virtual Storage

If you're just building out an initial virtual server environment and getting started with virtual storage, you are most likely focused on maintaining adequate storage capacity as you provision new virtual machines, and finding the best strategy for load balancing VMs across storage ports to maintain adequate throughput for I/O operations. You may also be a server administrator who has taken on some responsibility for virtual storage and you'd like to become more familiar with the key performance indicators in a SAN environment.

**Storage monitoring benefits:**
- Track storage capacity utilization in real time to plan for future storage needs;
- Assess the impact of additional VMs on storage to manage the pace of roll-out;
- Understand the I/O impact of moving VMs to different physical hosts/clusters;
- Learn the dependencies between events in the virtual server and storage tiers.

**Key questions addressed:**
- What is the impact of adding or load-balancing VMs on storage utilization and I/O performance?
- How has capacity demand grown over the last week, month, or quarter? When and where will I need to order more?
- What is my aggregate network load? What events cause load spikes and how will load balancing shift these spikes?

Profile 2: Growth
Rapidly Growing Virtual Environment

If you’re past the initial learning phase and have embraced virtualization, it’s likely you’ll start to experience both virtual server and storage sprawl. Virtual machines are much easier to provision and allocate than physical servers, and it can quickly become operationally overwhelming to rein in growth. You’ll also likely start to leverage the mobility and load balancing features of virtual machines (such as VMware’s VMotion and DRS) and it’s critical to understand the impact of these processes on storage loads.

Monitoring is particularly important in VDI deployments, which can scale rapidly into the hundreds or thousands of virtual desktops. New advances in VDI technology such as VMware’s Linked Clones, in conjunction with PS Series thin provisioning, can dramatically reduce the storage footprint for each virtual
desktop, but I/O must still be tracked carefully to ensure adequate user experience.

Storage monitoring benefits:
• Correlate storage issues with mobility or maintenance events in virtual servers;
• Ensure adequate capacity for VDI Linked Clone desktop pools;
• Track thin provisioning to maintain adequate unreserved (unallocated) space;
• Eliminate the need to build your own scripts or tools to gather SNMP data.

Key questions addressed:
• How many additional desktops can I deploy, and on which storage pools, while maintaining adequate storage reserves?
• What are the optimal minimum and maximum Linked Clone pool sizes I should configure?
• Do I have enough capacity, I/O, or network headroom to add new VMs? On which servers and storage pools?

Profile 3: Production
Virtualizing Critical Tier-1 Applications

To virtualize critical business applications, real-time performance monitoring and time-based correlation across various monitored attributes are necessary components of your operations toolkit. You will need immediate visibility into load spikes, and the ability to quickly replay historical log files to look for spike patterns across multiple attributes. Critical applications also require higher degrees of data protection and resiliency, so you’ll need insight into your snapshot capacity and replication/failback operations. And you’ll need the ability to analyze performance data across any time period to understand how recurring events such as end of month processing impacts your virtual storage environment.

Storage monitoring benefits:
• Track free and reserved space for volumes, thin volumes, snapshots, and replication across storage groups/pools over time, to optimize load balancing;
• Monitor I/O latency and distribution of read/write operations and track the impact of load balancing, to size storage properly for different workload types;
• Measure network load and link percentage, to understand how much of your total bandwidth is in use over time.

Key questions addressed:
• What is the best snapshot/replication strategy for the dynamic environment to maximize utilization while ensuring adequate data protection?
• Do I have the most efficient balance of workload types across storage groups and pools? Have I defined the best mix of RAID types across my storage volumes?
• Have I sized network throughput to handle expected peak loads for the next 6 months, based on historic trends?

Profile 4: Cloud
Building a Private or Public Cloud Service

For those enterprises who are comfortable with virtualization and are exploring a shared internal IT service (private cloud) or planning to host a virtualized environment with an external service provider (public cloud), it’s critical to establish a clear picture of current service levels in order to have a baseline for validation. Without a solid
understanding of your daily, weekly, monthly, etc., capacity and performance requirements, it’s difficult if not impossible to establish a meaningful service-level agreement with your internal customers or your external cloud provider.

These requirements apply to all flavors of cloud services: Software-, Infrastructure-, or Platform-as-a-Service. The earlier you begin to collect multi-attribute storage data, and integrate real-time, correlated storage monitoring into your daily operations, the more confident you will be that you can adequately test and validate service-level claims in the future.

Storage monitoring benefits:
• Identify, diagnose, and resolve issues faster; streamline your support processes;
• Achieve higher service levels both before and after a cloud service deployment;
• Validate how you’ve performed over time against service-level agreements.

Key questions addressed:
• How many 9’s can I deliver today? Can my cloud service provider match my current levels? Can they prove it?
• How can I take advantage of the advanced features of my virtual storage array to automate capacity and data protection operations?
• How much initial capacity and network bandwidth do I need to source from a service provider? Based on recent trends, when will I need more?

It’s a well-worn adage in IT operations that you can’t manage what you can’t measure. Storage performance monitoring delivers the visibility needed to design efficient IT operations processes for the dynamic virtualized data center.

SAN HeadQuarters in Depth
SAN HeadQuarters (SAN-HQ) is a storage monitoring solution designed for and shipped at no cost with Dell EqualLogic PS Series iSCSI storage arrays. The PS Series includes automated load balancing across drives, RAID sets, connections, cache and controllers to optimize performance and utilization. SAN-HQ complements this dynamic, virtualized architecture and provides IT managers with clear visibility and insight into performance and capacity trends throughout their storage infrastructure.

Architecture and Configuration
SAN-HQ is deployed as a standard client-server application and consists of a Monitor Service, Monitor Client(s) and Log Files containing the data collected from any number of PS Series Groups. SAN-HQ provides a single-system view of multiple PS Series Groups via a unified GUI. Installation is quick and requires minimal configuration, consisting of just a few steps:

1. Perform standard Windows application installation (less than 10 minutes)
2. Configure of Group network (IP) address for the Monitor Service in order to issue SNMP requests to Group(s)
3. Identify the SNMP community name, configured in the PS Series Group Manager application.
Data Management

The Monitor Service maintains a set of log files for each monitored Group. Maximum log files sizes are configurable, and once a maximum is reached, older data is overwritten. Over long periods, data is sampled at increasingly longer intervals – starting at a default of 2 minutes - allowing SAN-HQ to efficiently store very large data sets. In addition, log files can easily be exported at any time interactively, via a command line interface (CLI) or by using a Windows task scheduler.

User Interface Overview

The SAN-HQ GUI presents a familiar two- or three-pane navigation style, with an object tree at the left, object detail in a reading pane, and a top-level timeline slider indicating the full, period, and selected data range on display. Alerts and alarms are available via a lower panel which can be minimized, expanded, and pinned.

SAN-HQ's GUI is efficient and makes effective use of screen real estate, without the clutter of similar systems. Configuration options are kept to a minimum: the user can add new Groups, configure e-mail notifications, and select color schemes (or define a new one). Simplified data drill-down and roll-up is an integral part of SAN-HQ GUI design and is evident throughout the product. As shown in Figure 2, once a Group is selected from the left-side tree, a summary screen is displayed in the reading panel that includes Group information, IOPS, I/O, and
network summary information for the period currently selected in the timeline range selector. The menu underneath the title bar offers drill-down into the objects that make up the selected Group: Pools, Members, Volumes, Inbound Replicas, Network Ports, and Disks.

All information is time-synchronized and controlled by simple adjustment of the timeline range selector. From the available data, the user can select predefined ranges (1 hour, 2 days, etc.) and slide the bar to the desired start date, or manually stretch the slider width to define a custom range. Time-synchronized data for the three data types in SAN-HQ — capacity, I/O performance, and network performance — provides rapid correlation and easy replay of comprehensive Group storage performance, and is one of the immediate differentiators for the novice user.

Figure 3 shows the always-available alerts panel, maximized and pinned, and the context-sensitive help pop-ups available on every sub-panel throughout the GUI.
For a new storage administrator or a virtual server administrator new to storage performance management, integrated help features dramatically reduce the learning curve by eliminating the need to consult reference manuals for explanations of terminology. In addition, SAN-HQ’s pre-configured alerts and alarms (more than thirty are included, ranging from “Caution” to “Critical”), ensure that the administrator is informed of transient or persistent array conditions that should be resolved in order to restore or maintain performance.

Using SAN HeadQuarters

SAN-HQ’s GUI design allows navigation along two complementary paths. First, for each Group, five information categories are presented when the group is expanded in the left-side tree:

- Hardware/Firmware: configuration detail
- Capacity: utilization & allocations
- I/O: storage performance metrics
- Network: network performance metrics
- Combined Graphs: combined metrics

In addition, for each information category (Capacity, I/O, Network), the user can drill down into each object in the Group. This two-dimensional navigation is intuitive, requires virtually no prior experience to master in a few minutes, and succeeds at presenting a large amount of data in a way that’s easy to comprehend and at no point seems overwhelming. Switching between views is fast, the GUI is responsive, and the user quickly develops a broad, clear picture of storage activity and overall health.

Distributed administrative teams with differing levels of storage expertise should benefit from this level of baseline performance data, to quickly build a consensus concerning the source of a bottleneck.

Capacity Monitoring

Group-level capacity information is presented as shown in Figure 4, including a used capacity trend chart, seven detailed capacity breakdowns, and pie charts summarizing allocations, unused space, RAID level distributions, and delegated space (capacity reserved for site-to-site replication).

Capacity trending is crucial in rapidly growing environments, as we’ve discussed. SAN-HQ provides a centralized view of capacity trends across multiple groups, taking the guesswork out of trending and capturing a historic record for justification of additional storage purchases, when required. Volume Reserve, Count, and Unreserved Space are important data elements to monitor when utilizing thin-provisioned volumes, where reserve is dependent on usage patterns. Snapshot and Replica Reserves, along with Delegated Space, are crucial elements for planning data protection strategies, as they indicate the number of snapshots that can be retained and the amount of space required for replication and failback operations.
I/O Monitoring

Group-level I/O data is presented somewhat differently, to reflect the nature of I/O performance information. Several averages are displayed: I/O rate, latency, IOPS, and size. For each average, the raw data for the selected time period is shown in the main panel (Figure 5). While Average I/O rate (both reads and writes) is not a critical performance indicator, its value can indicate the primary types of workloads currently utilizing the Group. IOPS measurement is very important for load balancing, and along with latency provides an excellent gauge for measuring storage load.

If an application (virtualized or not) is experiencing high latency, a common diagnostic starting point is to determine the amount of overall application latency that’s due to the storage component of the application stack.
Network Monitoring
SAN-HQ goes beyond capacity and I/O monitoring by measuring and reporting on network performance as well, which is displayed using the same time-synchronized slider. This capability should be of particular interest to any administrative team new to the iSCSI protocol and/or iSCSI storage arrays. SAN-HQ gathers SNMP data on all storage network traffic, including iSCSI operations, replication operations, SNMP requests from SAN-HQ or other applications, and SAN-HQ’s own GUI operations. Group-level data is presented, as a set of Group averages and historic charts (Figure 6).

The Network display includes averages and historic charts for: Number of iSCSI connections; Load (including retransmission percent); Rate (aggregate throughput); and Link Percentage, displayed as a simple dial.

Combined Graphs & I/O Analysis
SAN-HQ also provides a preconfigured “Combined Graphs” data view (Figure 7), which displays the most commonly tracked capacity, I/O and network data elements, time-synchronized. Figure 8 illustrates a new feature in Version 2.0: enhanced I/O analytics, including estimated busy percent, estimated maximum IOPS, and degraded maximum IOPS.
New Features in Version 2.0

Dell is committed to the SAN-HQ solution. As if to dispel any fears that a no-cost, bundled tool might not receive the attention or investment required to satisfy power users, the company will release version 2.0 less than a year after the product was first shipped. We are impressed by the addition of a broad set of pre-built reports which summarize critical capacity and performance metrics. These include:

- Performance Report
- Host Connections Report
- Group Configuration Reports with Alerts
- Capacity Utilization and Trending Report
• Thin Volume Status Report
• Replication Status Report
• Alerts Report for All Groups
• Hardware Summary Report
• “Top 10” Reports:
  ✓ Top 10 Volumes by Capacity
  ✓ Top 10 Busy Volumes
  ✓ Top 10 Hosts by Connection
  ✓ Top 10 by IOPS

All reports are filterable and available for online viewing, printing, or export as .pdf, .htm, .rtf, .xls, .csv or .txt). Reports can be created in the GUI, via CLI, or automatically via the Windows task scheduler.

In addition to the new reports, version 2.0 also updates several core capabilities:

• Enhanced Analytics: calculation and display of maximum theoretical IOPS; degraded performance limits; estimated percentage-busy IOPS for groups, pools, and members; IOPS vs. latency; and volume queue depths
• Broader Data Collection: additional metrics on iSCSI sessions, outbound replication, volume collections, network ports and controllers, including temperature charts
• New Alerts: including those for unreachable member ports, port status, member disk mirroring issues, and disk alternate signatures
• Better Communication: finer-grained control over e-mail notification and group e-mail settings
• Improved Usability: single sign-on, new wizards, simpler navigation, and chart enhancements

Customer Testimonials

Taneja Group conducted several independent interviews with current SAN-HQ and found overall satisfaction levels to be very high. Our discussion with the network manager of a midsized US manufacturing company (6,000 employees) was typical:

“This customer bought an EqualLogic array to support a server virtualization project, and he had little prior experience with either virtualization or storage management.

“Once my users saw how easy it was to create new virtual machines, the number I had to manage doubled in just the first year, and we started to see storage performance problems.”

He installed SAN-HQ and was impressed by how easy it was to configure and learn. Now,

“when things slow down on the array, I can immediately narrow it down to the volume, then look at that volume in [VMware’s] vCenter Server and pinpoint which virtual machines are responsible. And with the new experimental analysis data in version 2.0, SAN-HQ has confirmed my suspicion that we are consistently over-utilizing our array. This gives me the hard data I need to justify a new capacity purchase before I take on any new projects.”

Confident in his ability to manage resource contention with SAN-HQ, he is comfortable supporting both production and development workloads on his shared servers.
Powerful, Multi-Site Storage Resource Management

Our overall impression of SAN-HQ is that the Dell EqualLogic team has invested significant resources in the delivery of an efficient, well-designed, intuitive and high-performing tool for both the novice and experienced storage administrator.

SAN-HQ works virtually out of the box, requiring very little installation or configuration effort. Storage capacity, I/O, and network data – the three primary dimensions of iSCSI array performance - are presented clearly and can be filtered to multiple levels of depth, a key benefit for a centralized monitoring solution designed for administrative teams of varying skill levels. Data across all monitored groups is available at a glance, including alerts ranging from basic health status to more sophisticated latency and threshold metrics.

The product’s rich historical data, combined with its intuitive, time-synchronized display of large data sets, should eliminate the need for storage teams to develop their own complex scripts or custom monitoring software. Indeed, PS Series customers often cite the well-designed, pre-configured monitoring capabilities of SAN-HQ and the time saved by not having to develop tools in house as primary reasons for their satisfaction with the product.

Customer Testimonials

One customer was impressed by SAN-HQ’s GUI design, noting that “when IOPS are through the roof; you can open the I/O view, filter by Volumes, and sort by Avg. I/O – in just a few seconds, you have the information you need to balance the disk I/O load.”

Another customer, who directs the storage practice at a European IT integration consultancy, has deployed SAN-HQ at over fifty customer sites and also relies on it to manage an in-house array that supports testing and training.

“SAN-HQ is an excellent product for the novice storage administrator,” he said. “It’s easy to install and upgrade, and we especially like the integrated e-mail feature to automatically notify the support team when there’s a problem. In our client work, the most valuable feature is the way network performance is presented in a clear summary view, with pop-up help and descriptions.”

He has not yet tested version 2.0 yet, but is eager to see performance data plotted against benchmarks so that he can help his clients estimate their performance headroom.

Every EqualLogic customer we spoke to felt that SAN-HQ was a key selling point and differentiator for Dell in the storage market. One administrator had a message for PS customers:

“If you’re not using this yet, you should be.”
Finally, and it's worth repeating, SAN-HQ is free for all EqualLogic customers under warranty or a service agreement, lowering the adoption-cost barrier and enabling a wider range of end users to benefit from comprehensive storage monitoring.

By bundling SAN-HQ monitoring and performance software with the array, Dell helps its customers take an important first step toward establishing best practices for storage resource management (SRM). SRM is a broad discipline, covering storage monitoring, planning, provisioning, configuration, problem diagnosis, service-level management, and reporting. But each of these operational activities depends first and foremost on a solid foundation of comprehensive, historical data. SAN-HQ collects this SRM baseline data out of the box, and stores and presents it so that it's immediately available to novice and expert storage administrators alike.

Several independent vendors of multi-vendor SRM software, and most of the storage vendors themselves, offer comprehensive SRM software solutions. These product families include EMC's Ionix ControlCenter, NetApp's SANscreen, HP's Storage Essentials, and IBM's TotalStorage Productivity Center. Each of these suites offers several components, typically licensed separately, and each requires multiple components in order to monitor and manage the full range of storage capacity, I/O, and network performance. SAN-HQ is a centralized and integrated solution delivering these key metrics without additional effort or license costs.

To be clear, SAN-HQ is not intended to be a complete SRM solution; it collects element-level data via SNMP passively, and is not designed to enforce policy, execute remedial actions, or automate provisioning or configuration tasks. But it goes well beyond what many of its competitors provide in expensive, bundled monitoring, performance and reporting modules. Packaging these key features of an enterprise SRM solution at no cost with the array can significantly lower total cost of ownership, shorten implementation times, and reduce ongoing operating costs.

**Taneja Group Opinion**

Successful virtualization demands higher levels of capacity and performance management, based on correlated cross-domain data, than was required when most servers were physical and most storage was static and directly attached. The mobility and flexibility of virtualized assets make them inherently more difficult to define, monitor, correlate, control and maintain.

IT operations teams, wherever they are on the road to the fully virtualized data center, increasingly demand that their server and storage vendors deliver much better visibility into the run-time metrics needed for accurate capacity, performance, and availability management and planning. And, it's impossible to optimize just one side of the equation when server and storage virtualization are deployed together and changing in real time.
For customers who are new to virtualized storage, SAN-HQ makes the management learning curve less steep, with an intuitive GUI, painless installation, well-designed charts, and context-sensitive help.

For virtualization-savvy customers who are facing a rapidly growing environment, SAN-HQ provides detailed historic capacity data to dramatically simplify trending analysis, detailed information about thin-provisioned volumes (a critical technology for efficient desktop virtualization), and correlated performance data to help quickly identify growth-related I/O or network problems.

If the challenge is how to virtualize business-critical production applications, which demand high reliability and availability, SAN-HQ provides detailed snapshot and replication metrics along with I/O and network performance data, at a glance.

Finally, for those customers exploring internal or external “cloud” virtualization, SAN-HQ delivers the foundational performance data needed to define, monitor, and maintain adequate service levels across a shared storage infrastructure.

Dell’s SAN-HQ delivers a comprehensive response to the demand for better visibility in the virtualized storage tier. The no-cost, bundled solution provides multi-site, multi-SAN monitoring and reporting for the three critical dimensions of storage resource management: capacity, I/O and network performance.

In our view, customers who select a Dell PS Series iSCSI SAN for storage virtualization enjoy an added bonus: enterprise-class storage monitoring without the need to develop in-house scripts or tools, license any third-party add-ons, or invest up front in an over-configured SRM product suite.

---

NOTICE: The information and product recommendations made by the TANEJA GROUP are based upon public information and sources and may also include personal opinions both of the TANEJA GROUP and others, all of which we believe to be accurate and reliable. However, as market conditions change and are not within our control, the information and recommendations are provided without warranty of any kind. All product names mentioned herein are the trademarks of their respective owners. The TANEJA GROUP, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors which may appear in this document.