# Analyzing Dell PS Series Storage with SAN Headquarters

Dell Storage Engineering
November 2019

A Dell Technical White Paper

# Revisions

| Date | Description |
|---|---|
| September 2010 | Initial release |
| June 2010 | Updated for SAN Headquarters v2.0 |
| November 2010 | Updated for SAN Headquarters v2.1 |
| October 2015 | Updated for SAN Headquarters v3.1 |
| November 2019 | vVols branding update |

# Table of contents

# Executive summary

The purpose of this document is to help storage administrators and other IT professionals use SAN Headquarters to monitor PS Series SANs. Real-world examples are used to provide performance-analysis techniques and methods.

SAN Headquarters provides exhaustive monitoring for performance and health of the PS Series groups. In addition, enhanced automation allows for proactive collection of diagnostic information for Dell™ support analysis and problem resolution.

This document also highlights the many benefits of Dell PS Series storage arrays, including:

- Simplified management of storage resources
- Comprehensive protection of data resources
- Load balancing and performance optimization of storage resources
- Advanced integration with server applications and operating systems

# Acknowledgements

This document was written by Chuck Farah of the Dell Storage Applications Engineering team.

# 1 SAN Headquarters overview

SAN Headquarters (SAN HQ) is a client/server application that runs on Microsoft® Windows Server®. It monitors one or more PS Series groups and uses SNMP to query the groups. SAN HQ collects data over time and stores it on the server for retrieval and analysis. The client connects to the SAN HQ server, formats and displays the data in the graphical user interface (GUI). Analyzing this data can assist in making decisions to improve performance, reliability, and allocation of resources.

In addition, SupportAssist is included with SAN HQ to provide diagnostic collection, upload, and analysis of PS Series arrays.

SAN HQ provides the following functions:

**Performance**

- Improve performance by identifying bottlenecks.
- Obtain comprehensive information about group performance based on latency, IOPS, I/O rate, I/O size, and other data.
- View real-time data for group member or volume I/O and save the results for future analysis.
- Quickly be informed of hardware-, capacity-, and performance-related problems.
- Monitor group performance data from multiple servers from a single SAN HQ Client.

**Operations**

- Allow multiple administrators to access and monitor the same performance data.
- Configure the SAN HQ Server to send SNMP notifications to an SNMP management console.
- Display Dell customer support cases created from SupportAssist diagnostic uploads and traditional support calls.
- Launch the Dell Storage Update Manager tool to update controller firmware, disk-drive firmware, and language packs.
- Obtain an automatic weekly update of the SAN HQ local repository, including newly supported disks and hardware, used for estimated IOPS calculations. This local repository is also used to verify member configuration and disk-drive support. You can also choose to update the local repository manually at any time.
- View 512 byte or 4 Kb disk sector size for each disk per member and for each volume in the group.
- Display space consumed by recoverable volumes in the recovery bin.
- Report delegated space for replication allocated by a secondary group to a primary group in multiple pools.
- See the number of Ethernet ports with active and inactive Data Center Bridging (DCB), and the number of ports incompatible with DCB.
- Specify favorite views.

**Analysis**

- Determine how the group is performing, relative to a typical I/O workload of small, random I/O operations. This information helps determine if a group has reached its full capabilities, or whether the group workload can be increased without impacting performance.
- Allocate group resources more effectively by identifying underutilized resources.
- Apply different RAID policies for a group, pool, or member to analyze the performance benefits and determine the reliability of your current RAID policy.
- View the percentage of all TCP segment packets sent by the member port.
- Visualize synchronous replication volumes and NAS volumes.
- Visualize data for groups configured for snapshot and replication compression.
- Visualize storage pools with storage containers comprising one or more VMware® vSphere® Virtual Volumes™ (vVols) mapped to a virtual machine (VM). For each vVol, list the connected vVols, identify the vVol type, and report capacity information and I/O performance.
- Show detailed space borrowing statistics including volume snapshots borrowing from replica reserve and delegated space; and replicas borrowing from other from other replica reserves, delegated space, and snapshot reserve.
- Preserve group performance data for later analysis by creating archives.
- Create customized reports of group performance data.
- Export performance data to a spreadsheet.

**Notification**

- Set and display threshold criteria for performance and capacity alerts.
- Quickly be informed of hardware, capacity, and performance-related problems.
- View events, audits, and group alerts.
- Receive email notification of alerts.

**Support**

- Deploy Dell SupportAssist for diagnostic data collection on a weekly basis, on-demand as needed, or when critical events occur. Diagnostic data collections are automatically uploaded to Dell Support for analysis; if you disable automatic uploads, the data is encrypted and stored locally.

SAN HQ does not disrupt access to group storage or degrade performance on the hosts or groups.

## 1.1 SAN HQ architecture

SAN HQ uses a client/server model which includes:

**SAN HQ Server**: This is the server which runs the Monitor (*EQLxPerf*) service and communicates to the PS Series groups. SNMP requests are issued to collect configuration, status, and performance data. In addition, a syslog server may be configured to log hardware alarms or performance alerts (*syslog* is typically the same IP as the SAN HQ server and is configured in the PS Series Group Manager CLI or GUI).

**SAN HQ Client**: This provides a GUI for managing the SAN HQ environment and viewing data collected by one or more SAN HQ servers. A SAN HQ Client accesses the log files maintained by the SAN HQ Server and formats the group data into tables and graphs. You can run the SAN HQ Client on multiple computers.

SAN HQ stores the data in group log files. Computers running the SAN HQ client access the files and display the data in the SAN HQ GUI.
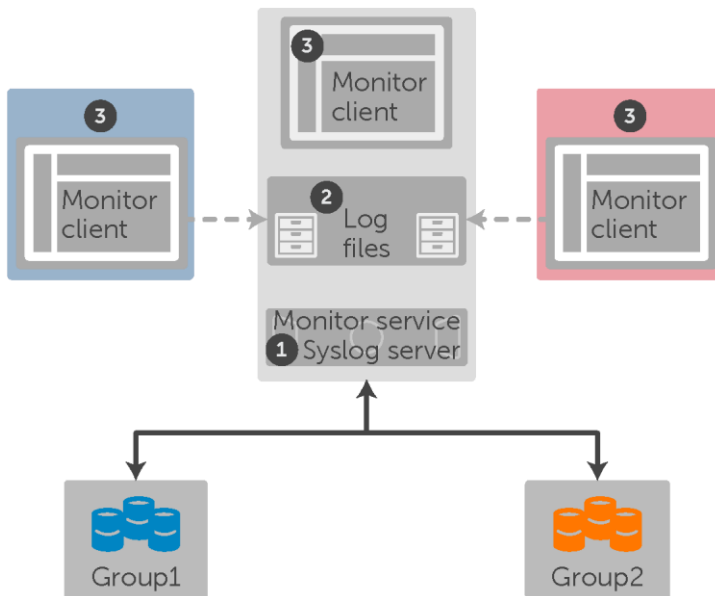


Figure 1        SAN HQ architecture

| Element | Description |
|---------|-------------|
| 1 | The computer running the Monitor Service issues a series of SNMP requests (polls) to each group for configuration, status, and performance information. The Monitor Service also includes a syslog server to which a PS Series group can log events. |
| 2 | When the first set of SNMP requests returns from a group, the Monitor Service stores this baseline information in the log files for that group. The Monitor Service issues subsequent SNMP requests at regular intervals of time (default: two minutes). |

| Element | Description |
|---------|-------------|
|  | To obtain a data point, the Monitor Service averages the data from consecutive polling operations. |
|  | After a year (by default), the SAN HQ Server overwrites the oldest data. |
| 3 | Each computer running a Monitor Client accesses the log files maintained by the Monitor Service and displays the group data in the SAN HQ GUI. |
|  | **Note:** The computer running the Monitor Service also has a Monitor Client installed. |

## 1.2 SupportAssist architecture

SupportAssist is an optional component of SAN HQ that enables you to collect and optionally upload PS Series group configuration diagnostic data to Dell Support for analysis and diagnosis. SupportAssist is bundled with every SAN HQ Server and configured from the SAN HQ Client. In order for SupportAssist to be enabled all groups need at least PS Series Firmware 6 and SAN HQ 2.5.

### 1.2.1 How SupportAssist Works

SupportAssist communicates directly with a PS Series group to collect diagnostic data. When the data collection completes, SupportAssist automatically uploads the data to the Dell Support data center.

Data transmission occurs in any of the following ways:

- Weekly: This is according to a user-specified day and time.
- Event driven: This occurs whenever SAN HQ receives a critical event notification through the syslog mechanism from a PS Series group. Event-driven data collections occur once every 24 hours. If multiple critical events occur during the 24-hour period, events are reported but no additional data collection occurs.
- On-demand: This is user-driven on an as-needed basis. You can run an on-demand data collection from the SAN HQ Client.

Diagnostic data uploaded by the SAN HQ Server enables Dell Support personnel to easily view the data when working on a support case. Dell Support contacts you only regarding issues discovered when analyzing the data collected by SupportAssist or when you specifically request support.

SupportAssist does not fully replace the E-Mail Home feature or the **pssdiag** script for diagnostic reporting to Dell Support. Instead, it reduces the need to use these reporting methods on a frequent basis.
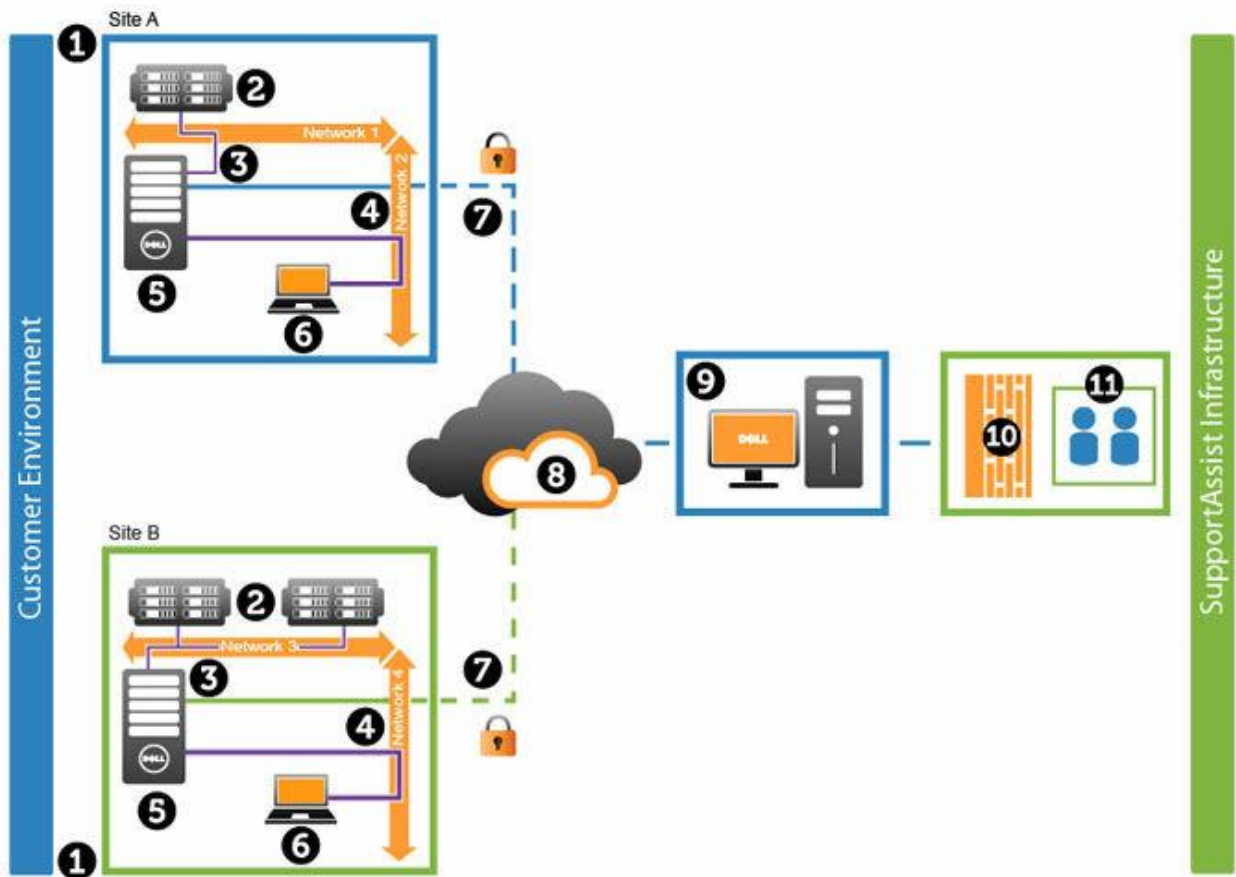
Figure 2    SupportAssist components and process

| Element | Description |
|---------|-------------|
| 1 | PS Series group at Site A (blue) and Site B (green) |
| 2 | PS Series SAN arrays |
| 3 | SAN networks (orange) |
| 4 | LAN networks (purple) |
| 5 | SAN HQ servers |
| 6 | SAN HQ clients |
| 7 | SSL Internet links |
| 8 | Internet |
| 9 | Secure SupportAssist web server |

| Element | Description |
|---------|-------------|
| 10 | Dell data center firewall |
| 11 | Dell Support and other departments |

## 1.3    Information provided by SAN HQ

### 1.3.1    All Groups Summaries

SAN HQ allows for monitoring of multiple PS Series arrays and provides summary information across all the groups monitored (Figure 3).

To view this information, click **All Groups Summary** in the **Servers and Groups** tree.



Figure 3        All Groups Summaries

**Note:** All summary information is refreshed and averaged over the refresh period, which is typically two minutes.

## 1.3.2 Group information

Once an individual group is selected, the informational dashboards shown in Figure 4 are available:
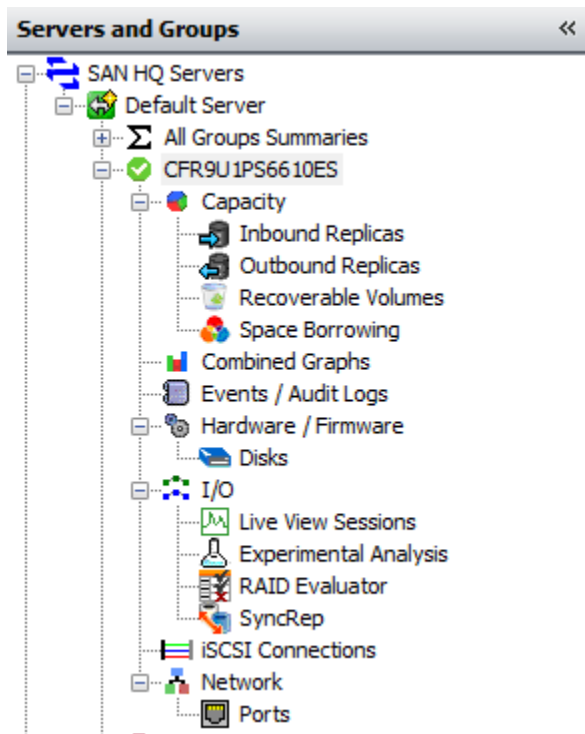


Figure 4　　Available functions by PS Series group

Details for each selection are provided in the document, *Dell EqualLogic SAN Headquarters Version 3.1 Installation and User's Guide,* available on eqlsupport.dell.com (login required). This document focuses on several panels to illustrate methods for analyzing PS Series data.

## 1.4 Exporting data

SAN HQ enables data to be exported in csv format for one or more monitored groups. Open the csv file in a tool such as Microsoft Excel® for external analysis. Specify the time range, groups, pools, members, volumes, network ports, replicas, and disks when exporting the data.

## 1.5 Archiving data

SAN HQ allows data to be archived for one or more monitored groups. This enables you to use SAN HQ to view and analyze the archived data offline, without needing access to the monitoring service.

You can also archive data to preserve detailed data for a particular time period.

Typically, a data archive is more valuable than exported data when working with Dell support to resolve issues since the archive retains the original view of the data for offline analysis.

Normally, each SAN HQ Client connects to the monitoring service to obtain and format the latest group performance data. Archiving the data allows you to analyze data when the SAN HQ monitoring service cannot be accessed.

For example, if you start SAN HQ, but do not have access to the monitoring service, simply choose **Ignore** when launching SAN HQ, as shown in Figure 5. This allows SAN HQ to start in offline mode and then import archive files.
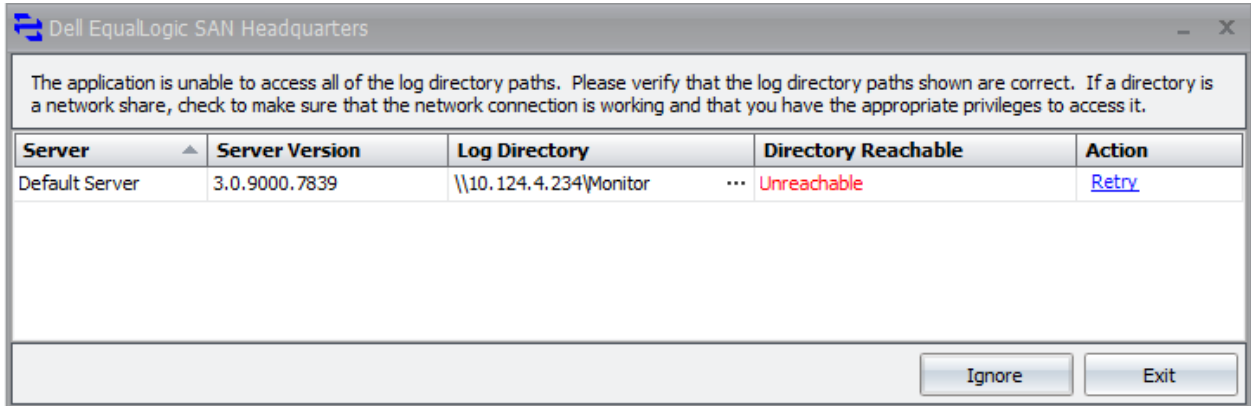


Figure 5      Click **Ignore** to continue in offline mode

## 1.5.1 Importing data

To open an archive, select **SAN HQ** > **Open Archive** from the SAN HQ menu bar. A new SAN HQ session appears with the data from the selected archive file. The data can be viewed and analyzed as if it were connected to the monitoring service.
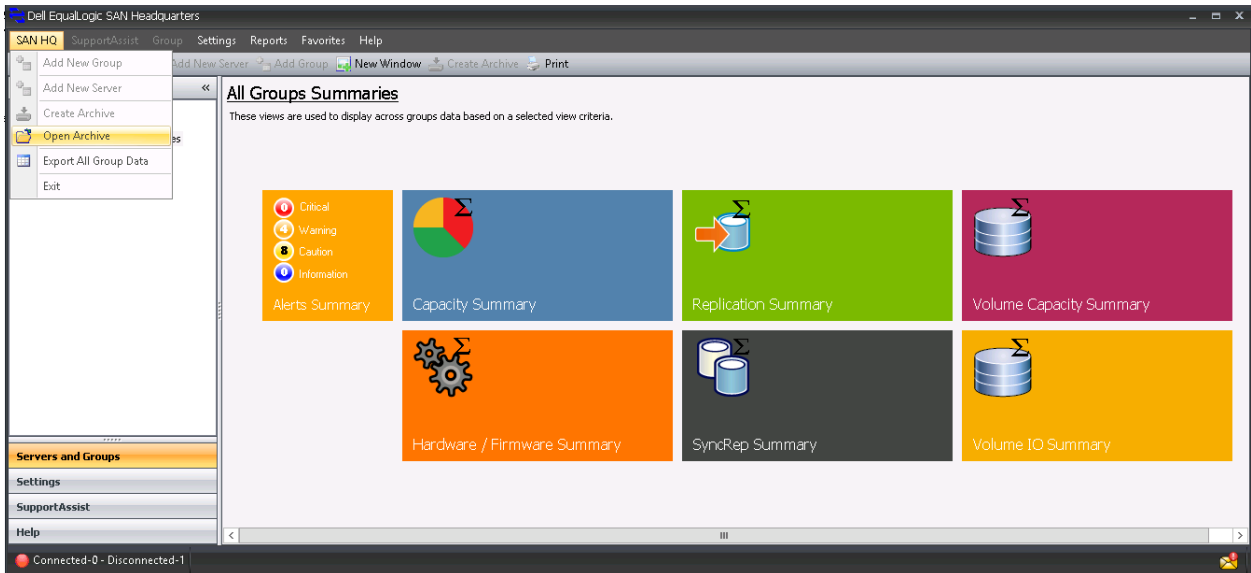


Figure 6      Click **Open Archive** to view and analyze SAN HQ data offline

**Note:** Most of the examples in this document use the Open Archive method.

# 2 Using SAN HQ to find performance bottlenecks

SAN HQ is designed to help identify hardware bottlenecks within the PS Series infrastructure. Many factors should be considered when troubleshooting or isolating the root cause. Hardware issues or misconfigurations should be identified and eliminated first.

The Dell Storage PS Series is well suited for highly virtualized iSCSI connected environments. Configure the network according to Dell best practices for PS Series iSCSI networks. More information is available in the PS Series Configuration Guide.

PS Series arrays provide a virtualized architecture which enables maximum utilization over all disks in the pool. Volumes will spread across the pool boundary of disks and provide for the aggregate I/O and capacity.

SAN HQ exposes the different layers of virtualization and helps maximize resource utilization and troubleshoot bottlenecks.

## 2.1 General performance considerations

There are several opportunities for performance issues within data processing infrastructures. Hosts may have processor or memory saturation limits. Applications may have long running queries or low tolerance for latency. When focusing on the hardware and virtualized stack, the most crucial component is the application's ability to access data. That data may reside on mechanical spinning disks, which in comparison has a significantly higher access time than the servers' main memory (DDR RAM). The relationship of access to different hardware stores is generally illustrated in Figure 7.
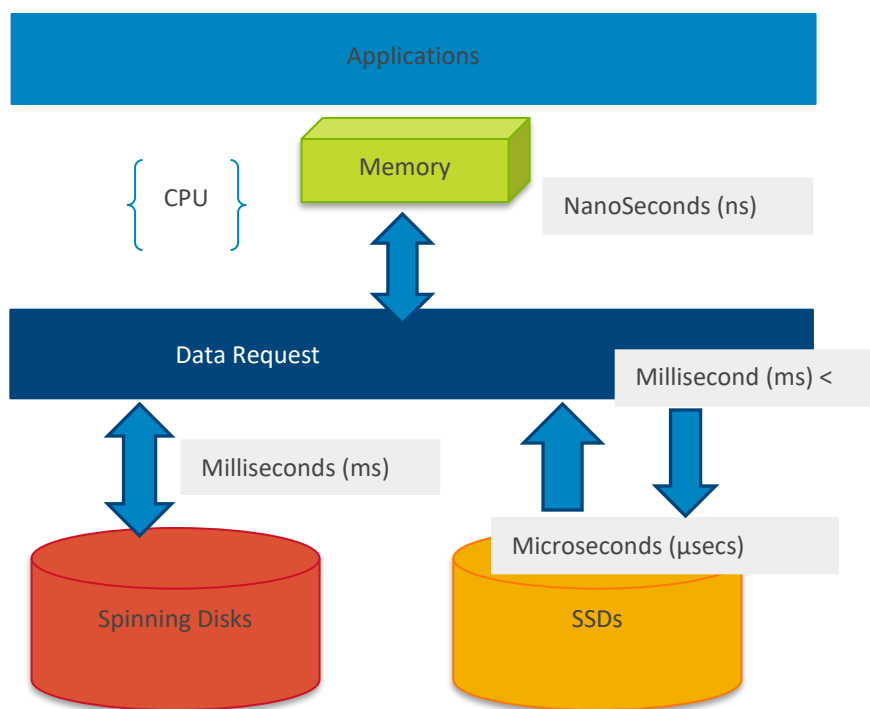
Figure 7    When observing latency relationships between components, DDR memory is much faster than spinning disks

Looking at the previous figure, all resources should be understood from a utilization perspective. If the server has saturated processors or memory, latency suffers overall. However, if the overall acceptable response time to and from the storage subsystem is more than the application's tolerance, the disk should be the focus of the investigation.

## 2.2    Understand the PS Series scope of monitoring

Storage subsystems have improved volume latency by aggregating disks together which effectively scales the total amount of disks in the set. For PS Series storage, disks are logically grouped together in pools which aggregates all the disks in that pool to support the application's need for data. For instance, a single 15K disk capable of 180 IOPS will be able to provide near 3960 IOPS in a 22-disk set (22 x 180). This aggregate capability of multiple drives working together is the basis for determining the overall potential of a storage subsystem.

Since the PS Series arrays protect disk sets with RAID, the actual supported host IOPS is less than the disk IOPS (sometimes referred to as backend IOPS). For the purpose of designing a solution or troubleshooting an existing solution, the overall backend IOPS and host-sourced IOPS needs to be fully understood. For example, an application or volume showing 1,000 IOPS at 50 percent reads may actually be pushing 2,500 IOPS to a RAID 50 protected set of disks (500 writes x 4 + 500 reads).

SAN Headquarters is ideal for monitoring both the application IOPS requests as well as the actual IOPS to the disks as depicted in Figure 8.
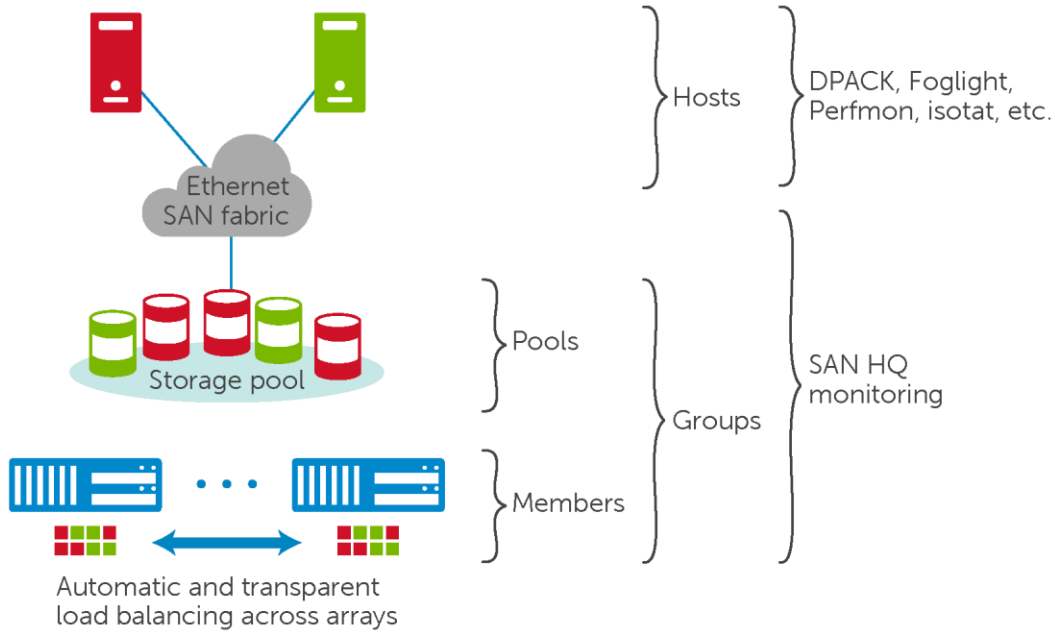
Figure 8　Performance and monitoring relationships: SAN HQ monitors iSCSI SAN for PS Series groups

SAN HQ provides the resources needed to monitor and determine where bottlenecks may occur in the PS Series infrastructure. Server resources attached to the SAN such as processor and memory consumption may be understood using other tools such as Dell Performance Analysis Collection Kit (DPACK), Windows Perfmon, Linux iostat, or Dell's Foglight.

## 2.3　Performance analysis goals

For faster problem resolution, it is important to understand the overall goal before initiating a performance review.

**Performance goal types**

- Benchmark: Focuses on the common tests and the hardware differences between solutions; helps understand the RAID configurations, virtualization differences, and disks types; and helps understand the criteria for success.
- Diagnostics: Focuses on the context, times of complaints, or high latency.
- Planning or designing: Determines the additional load a current infrastructure may support or the appropriate configuration for new workloads.

**Performance metrics**

- Input/output per second (IOPS): The measurement of actual units of work (based on I/O size) that are processed by the subsystem. Randomness usually requires more mechanical movement in the case of spinning disks.

- Throughput: Typically measured in megabytes per second (MB/s), this is the amount of bandwidth needed for the application. This may be derived by multiplying the I/O size by the number of IOPS. Understanding saturation points at the port level is important for high bandwidth needs.
- Queue depth: Indicates the number of requests that are lined up waiting on processing.
- Latency: The amount of time a request takes to arrive at the application. Typically, latency is the most telling symptom of an issue.

## 2.4 Performance analysis overview

For the purpose of this document, the focus will be on the scope that SAN Headquarters provides. This will include the storage virtualization layers, the IOPS to the volumes, pools, members and groups. In addition, the backend IOPS to the actual disks within each member will be considered.

### 2.4.1 Hardware considerations

Hardware issues should be resolved appropriately and all connectivity should follow best practices. Documentation that addresses best practices for switches, RAID selection, and advanced PS Series features is available on PS Series technical documents and videos.

The following methodology may help you identify most PS Series storage performance issues:

- Review the overall environment.
- Compare the actual I/O load to the expected theoretical maximum.
- Determine the actual saturation level of the disks by comparing the observed I/O with the rule-of-thumb expectation.
- Verify that there is ample capacity available for load balancing, snapshots, and reserve space.

The overall goal is to find the bottleneck that is contributing to unacceptable responsiveness.

### 2.4.2 Network considerations

When setting up switches and host NICs, following best practices for PS Series storage typically eliminates the network as the source of a performance issue. Other issues will be apparent if the best practices are not followed.

TCP retransmits or other network issues may be avoided by following these practices:

- Isolate iSCSI traffic.
- Enable Flow Control.
- Ensure all switches are interconnected.

    – Non stacked: enable (LAGs) with rapid spanning tree
    – Stacked: use uplinks

- Enable Jumbo Frames (optional).

10Gbps iSCSI networks are capable of accommodating typical bandwidth requirements. The PS Series architecture effectively aggregates all the ports to address additional throughput needs.

### 2.4.3    Servers

The server should be eliminated as the source of a performance issue from a memory or processor perspective. Although spinning disk access is much slower than memory access speeds, other factors that are non-disk-related still need to be ruled out. For the highest availability and throughput, at least two NICs should be used. If Jumbo Frames are enabled on the switch ports, the NICs should also have Jumbo Frames of the same size enabled. In addition, the appropriate Host Integration Toolkits should be installed on the Servers attached to the PS Series SAN.

### 2.4.4    Tuning considerations

PS Series storage is a virtualized self-tuning system. The architecture is based on virtualizing all the disks and allocating volumes over those disks from within a page structure. This allows for appropriate data movement to equalize capacity and relieve performance constraints. To tune a PS Series system, a pool may need to be expanded by adding a new PS Series array, or volumes may need to be moved to other pools within a group. This is a fundamental design concept to the peer storage architecture where each member or peer will have its own disks, processor, memory, and iSCSI ports to support the capacity and workload in that enclosure. When a new PS Series array is added to an existing pool, all of the disks begin to have the data spread equally for all of the volumes in that pool. Those volumes then benefit from the additional capacity and the IOPS capability brought to that pool.

## 2.5    SAN HQ GUI navigation technique

In SAN HQ, many views are available and may be overwhelming with the amount of functionality available. This section describes a technique to help identify the problem quickly.

Below are some of the recommended views to help you determine the problem:

- **Combined Graphs**: Provides overall view of IOPS, MB/s, latency, and capacity trends. Can be used to observe possible issues, such as high latency, high-bandwidth needs, capacity availability, and resource counts such as iSCSI sessions.
- **Hardware / Firmware:** Provides overall hardware and software configuration. Is used to observe the number of pools, members, and RAID preferences.
- **Experimental Analysis:** A tool which should be used as a guide to provide the theoretical maximum IOPS based on the speed of drives, type of PS Series array, RAID preference, and read/write pattern. The assumption for this calculation is 8KB I/O size. Always review latency and actual disk IOPS and queue depth before making changes based on the **Experimental Analysis** tool.
- **Disks:** This shows IOPS at the individual disk level and typically indicates the source for high latency.

Table 1　　　General guidelines for IOPS per disk

| Disk type | Random IOPS per disk (8KB I/O size) |
|---|---|
| 7.2K NLSAS | 75-90 |
| 10K SAS | 130-185 |
| 15K SAS | 180-210 |
| SSD | Depends on I/O size Reads/Writes disk manufacture and model |

I/O per disk is dependent on many factors such as speed, interface, disk classification (enterprise or consumer), as well as the distribution across each individual drive. Although Table 1 is a good reference, care should be taken into consideration that the table is not used as an absolute. For instance, a 15K disk is capable of speeds over 300 IOPS with good response time when the data retrieved is not fragmented.

For more advanced analysis, use the following view:

- **I/O (all Volumes):** Show all volumes and identify the ones with the highest latency, IOPS, MB/s for a possible move to other resources, for instance, to another pool.

**Note:** Other views such as Network and Capacity are also necessary for additional details.

## 2.5.1　SAN HQ poll periods and I/O metrics averages

SAN HQ formats data in a way that can be read through various graphs and tables. The polling period is the sample that is in the visible view of the charts. To select an interval in the chart, simply click anywhere in the chart. The interval selected represents any averages which may appear in the tables under the charts or in summary areas next to the panel.
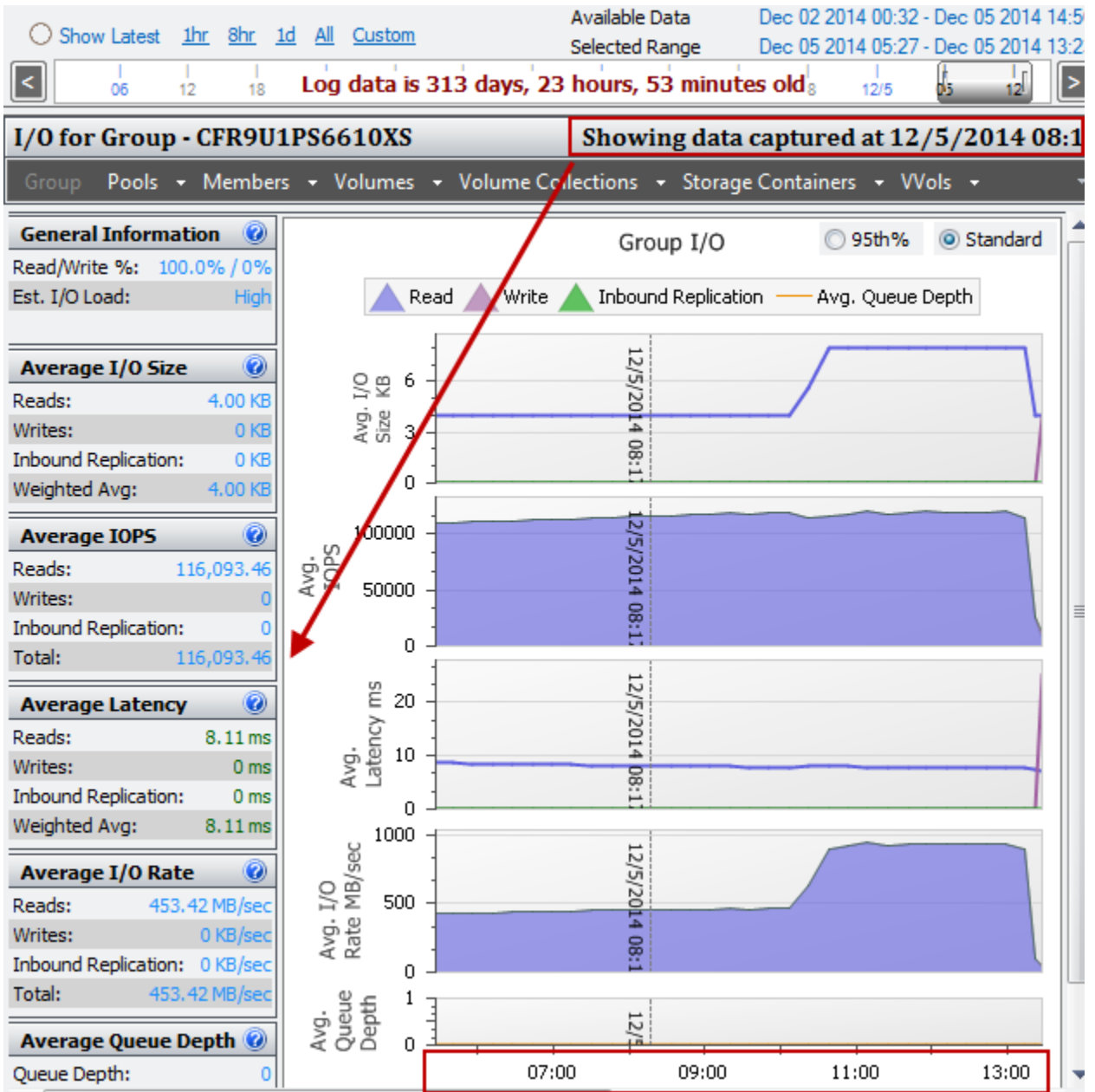
Figure 9    The sample period is from 06:00 to 13:00, while the summary (General Information) average is from the single interval selected (08:13)

To find sample ranges for these averages, drag the mouse while clicking on the chart. To view information about the interval period in a pop-up window, hover the mouse over any point in time on the graph. Both of these techniques are demonstrated in Figure 10.
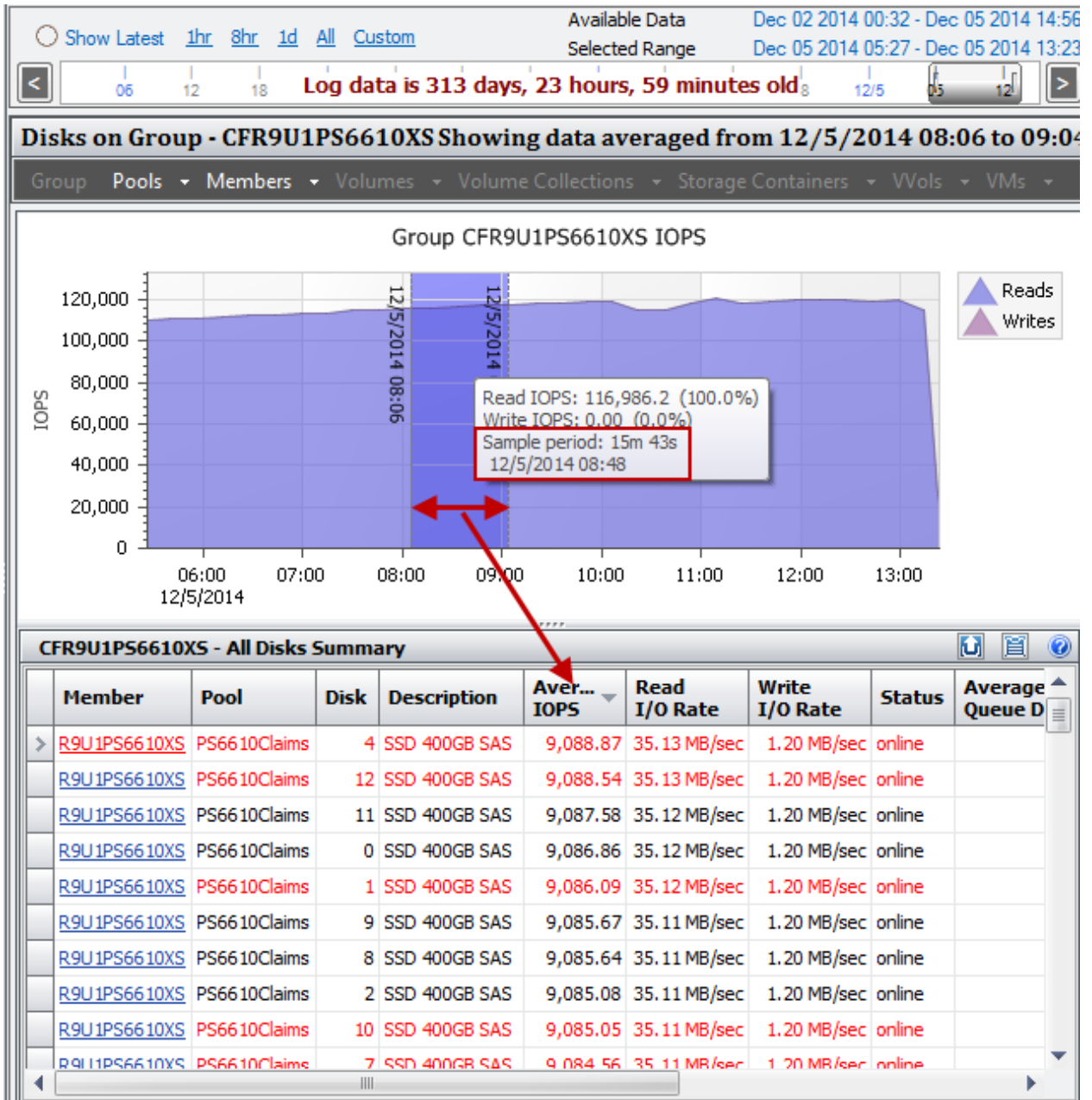
Figure 10     The interval span shows an average of all metrics over a span of an hour. Each interval in that span represents a sample of 15 minutes and 43 seconds.

The range between 08:06 and 09:04 is averaged for the data in the table below the chart. If no range is selected, only that interval's average is displayed in the table.

**Note:** It is important to understand the context for the averages of the polling periods as well as the sample intervals when reviewing SAN HQ.

# 3 Troubleshooting examples

Troubleshooting SAN issues using SAN HQ is easily demonstrated using examples. This section includes several issues with possible resolutions.

## 3.1 Example 1: Over utilized pool

This example demonstrates how to use SAN HQ to identify performance issues and navigates through the following SAN HQ panels:

- Combined Graphs
- Hardware / Firmware
- Experimental Analysis
- Disks

**Navigation tip:** In the left panel, click **SAN HQ Servers** > *Group* (**TMProd** in this example) > **Combined Graphs**.
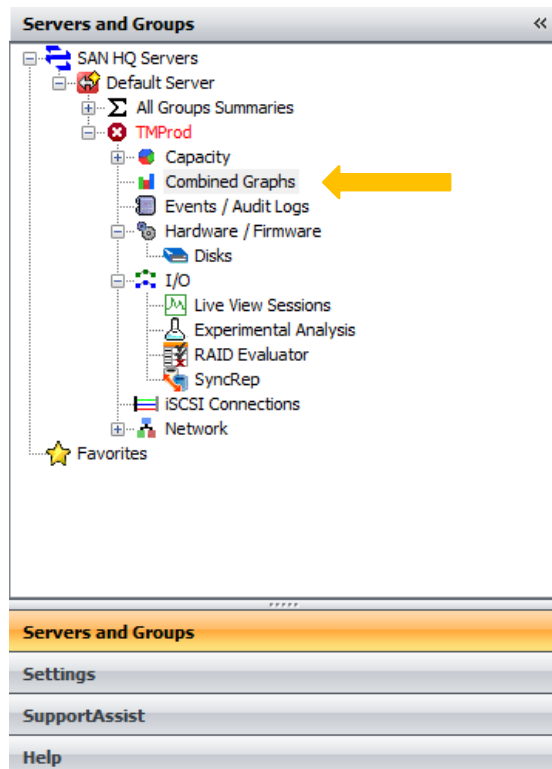


Figure 11    Navigation panel for SAN HQ combined graphs

### 3.1.1    Example 1: Before

Performance concerns are easily identified from the Combined Graphs dashboard where overall performance, capacity and trends are included in the display. For instance, if free capacity trends downward, then the focus on the volumes or snapshots consuming the most space. If latency were high in relation to the application tolerance, understanding which pools the members or volumes are taking the most resources from would be important.

Below is an example **Combined Graphs** panel showing latency exceeding 25 ms. Also important to understand is the actual client experience in order to help associate application delays with the latency in SAN HQ. For instance, if a client is complaining about responsiveness around noon, then that time period should be the focus of the analysis.
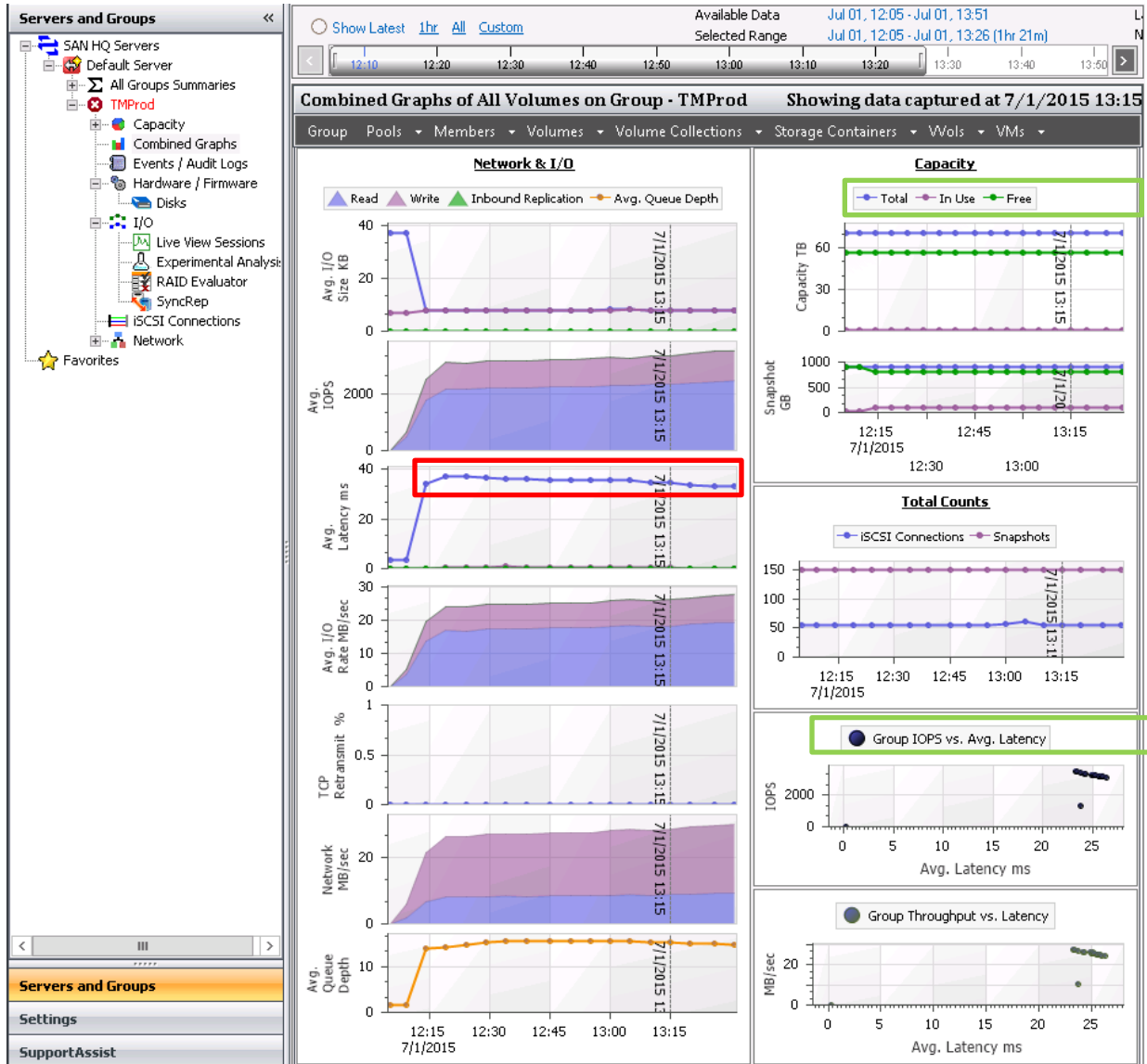


Figure 12    Combined graphs view. Notice latency over 25ms.

**Observations**: Latency is above 25ms. Total IOPS are shy of 3000 from the host, capacity and total iSCSI sessions are ok.

### 3.1.1.1 Hardware and firmware details

For a more thorough understanding of the array configuration, the **Hardware/Firmware** panel provides details. In this example, the **Hardware/Firmware** panel shows that two members each in their own pool. The default pool is in a RAID 6 configuration and contains a PS6210E with 7.2K 4TB NLSAS drives. The FASTPOOL15K is a new member added to the group and contains a single PS6210XV with 15K 146GB drives in a RAID 6 configuration.
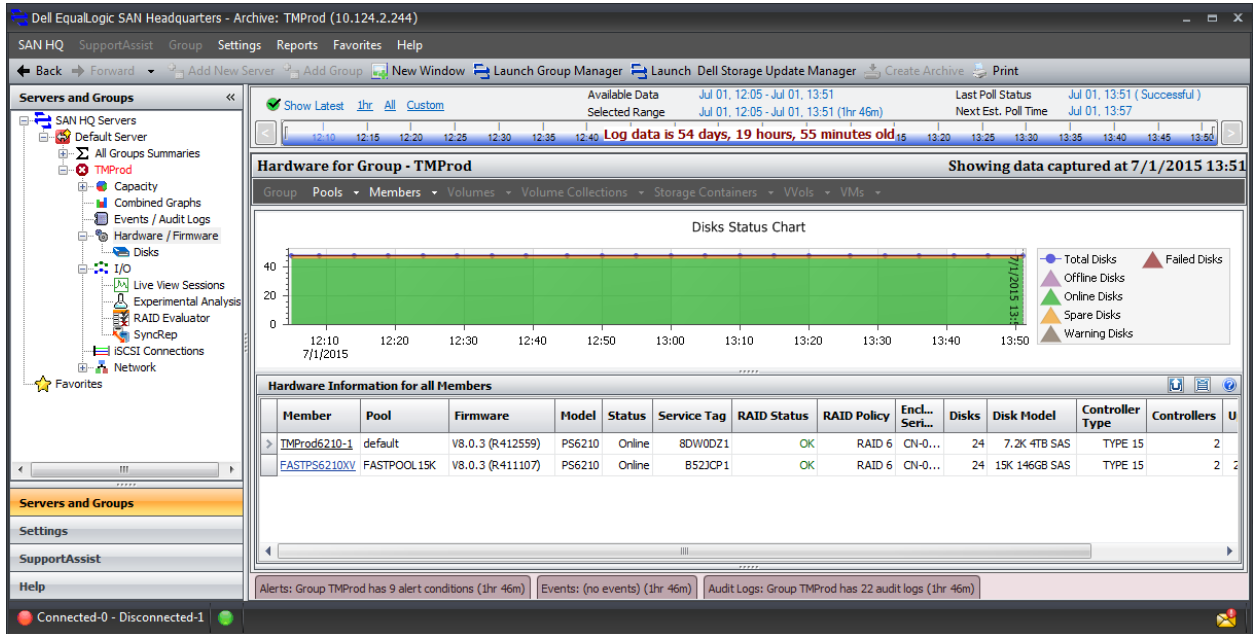


Figure 13    Hardware/Firmware shows two pools each with a single member

The default pool contains the volumes in this example and the administrator is aware that only the default pool is actively in use. The new array and pool were created for future use.

Note that all of the volumes are using the slower drives and additional resources from the new pool are available.

### 3.1.1.2 Experimental analysis

The **Experimental Analysis** panel shows the theoritical maximium IOPS based on groups, pools or members chosen. For instance, when the group is selected, all pools I/O potentials are included in the display. This may be useful for indicating the overall performance capability of the group in the case of multiple pools. However, in this case the UI is showing that the system is over the estimated maximum IOPS of the PS Series group even with the presence of the unused pool.
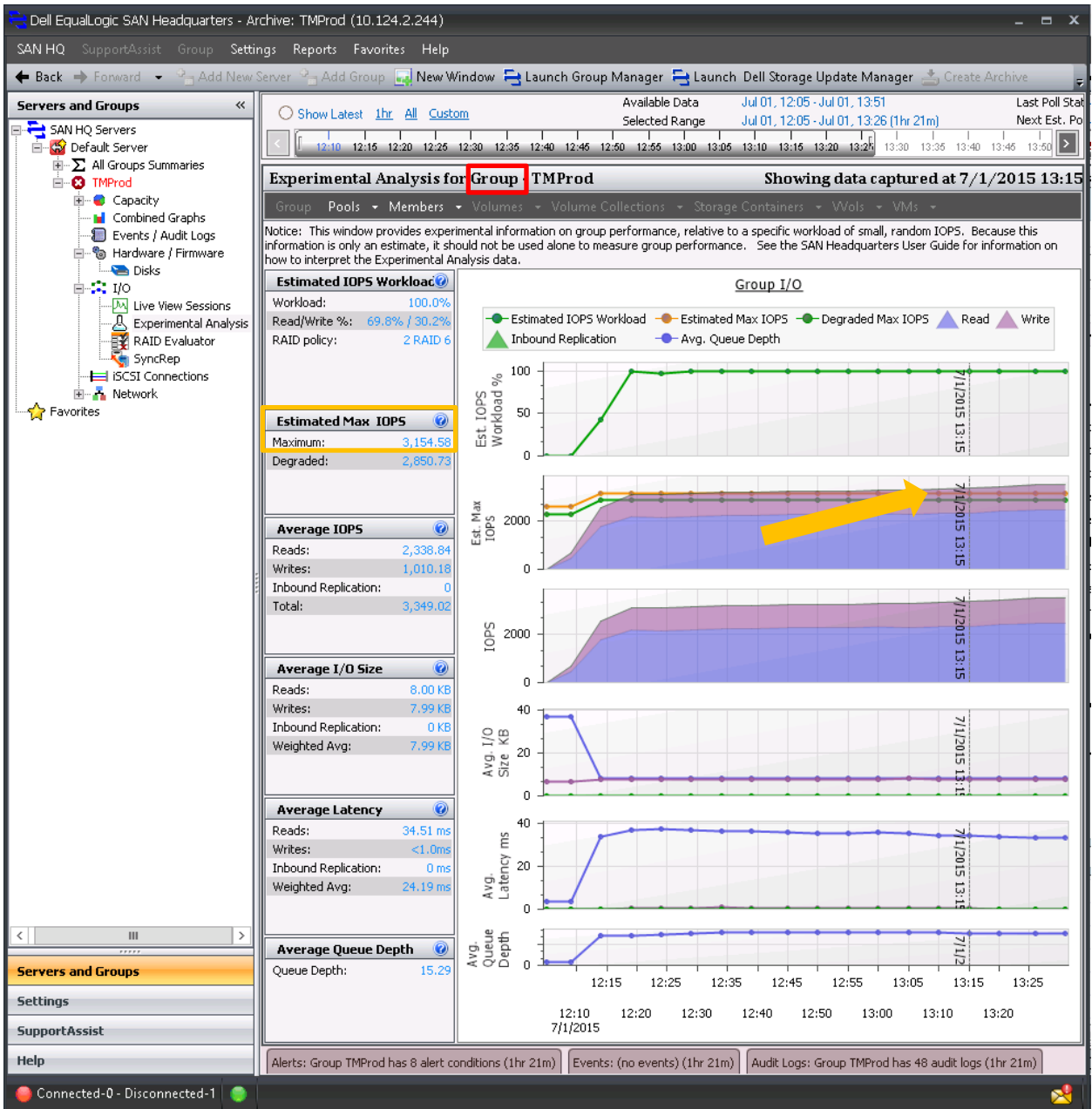
Figure 14    Experimental analysis of the group. The average IOPS is slightly over the estimated max IOPs.

The **Experimental Analysis** defaults to the Group I/O view. The **Hardware/Firmware** displayed two pools that have to be selected individually to view the current I/O and the estimated maximum IOPS. When the group is selected, IOPS are only about 6% over the estimated maximum. If the I/O load could be more evenly spread out, the latency may be reduced for certain volumes. First, analyze the utilization levels of the two pools.

In the **Experimental Analysis** panel, select the default pool to further expose the issues.
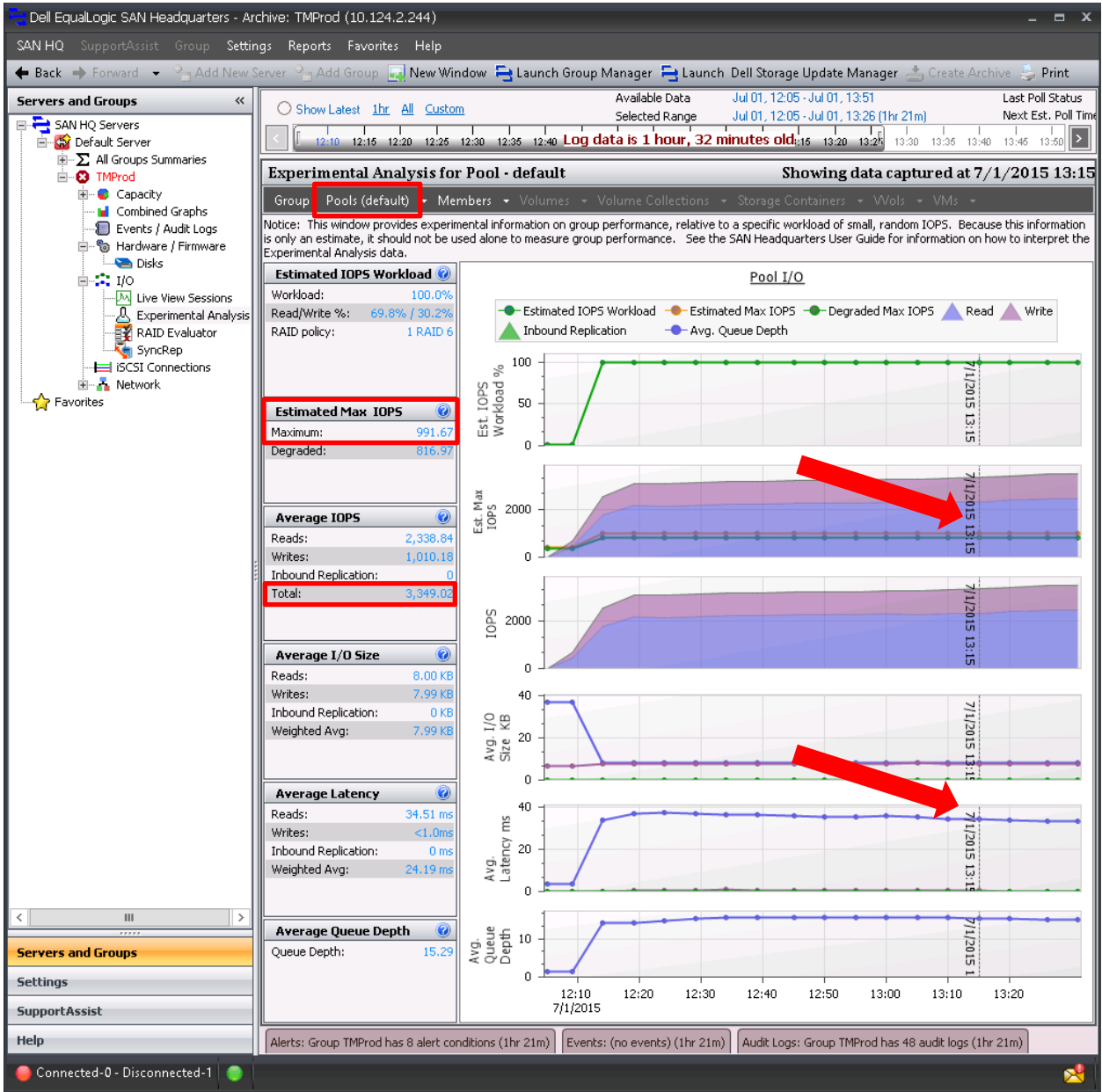


Figure 15    Experimental analysis of the default pool

The default pool shows that the current average IOPS far exceed the estimated maximum IOPS by 3.4 times. Latency is approximately 35ms for reads when observing this single pool. In this example, only the default pool has volumes allocated.

The FASTPOOL15K is also a RAID 6 PS Series member however, it currently has no I/O as expected and shown below.
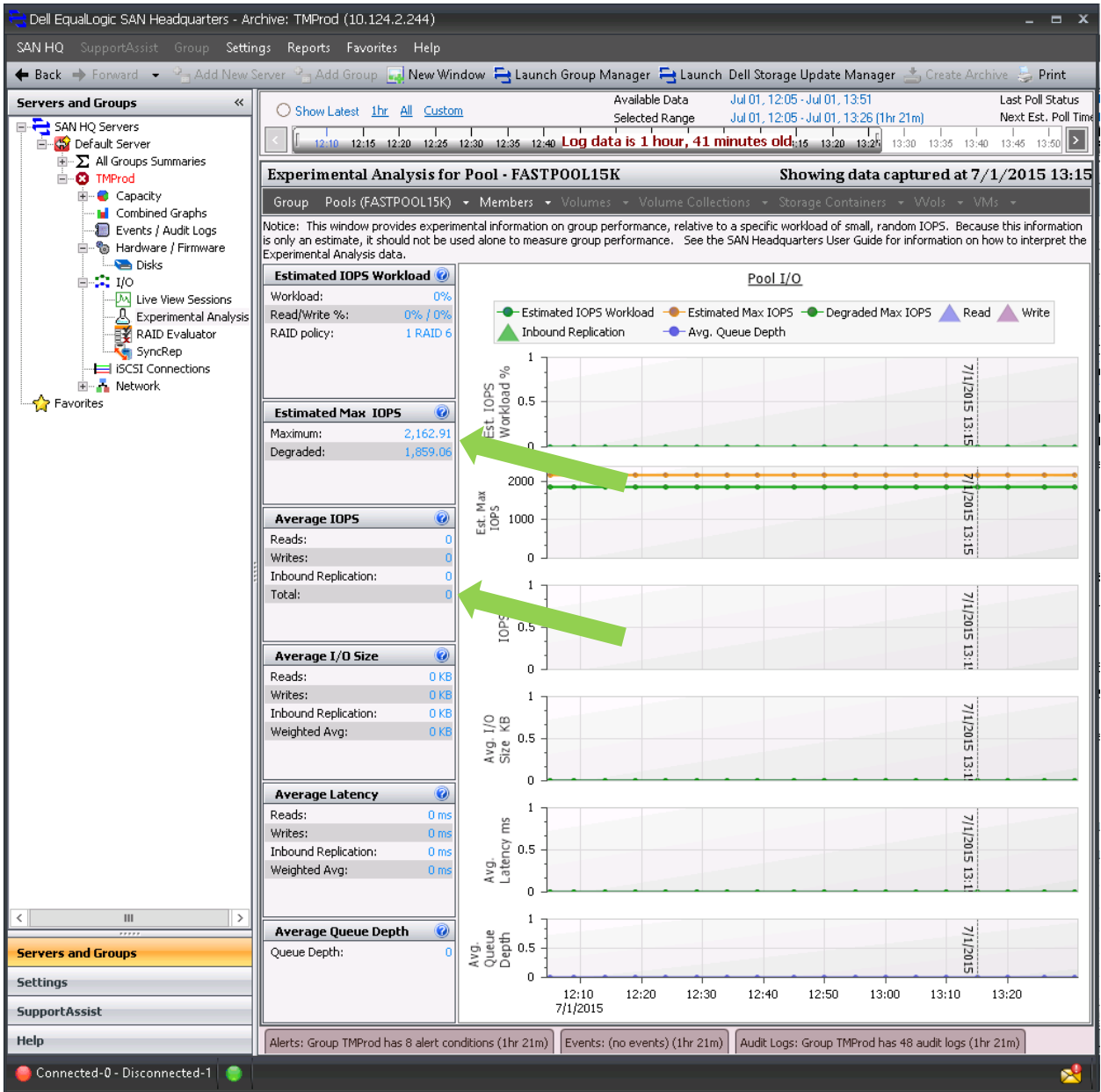
Figure 16     FASTPOOL15K Experimental Analysis panel. Currently no IOPs to this pool.

When the **Experimental Analysis** panel displays the estimated maximum IOPS at the group, both pools are be included in the averages for the performance metrics. The average includes the potential of both pools regardless of any actual I/O occurring. This is important to understand when observing the **Experimental Analysis** tool. These observations show resources that are available in the FASTPOOL15K and the default pool is oversaturated from a theoretical perspective. The next step is to confirm the saturation by observing the IOPS to each disk.

### 3.1.1.3    Disks

The Disk panel shows the individual IOPS for each disk, providing clues to the I/O profile. Under normal operations, the expected I/O rate from each disk is displayed. Disks that are over the expected IOPS, have high latency and queue depth reveal an over utilization. Table 2 shows that the expected IOPS for each disk in the NLSAS drives is approximately 85 IOPS/disk. The table data below the graph below shows that the IOPS/disk have exceeded this rule of thumb.
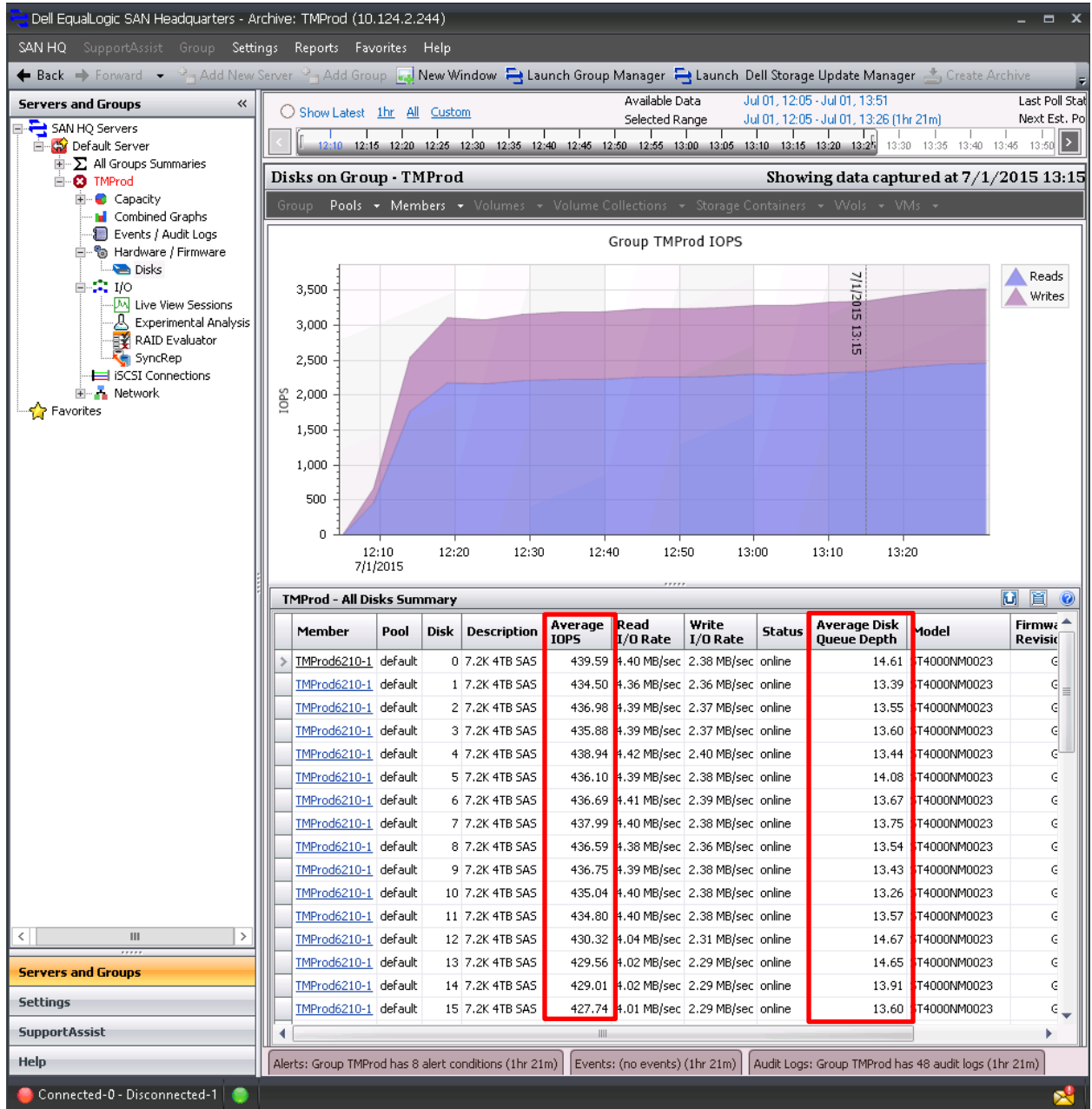


Figure 17        7.2K RPM Disks show over 430 IOPS for each disk

The disks in Figure 17 are 7.2K 4TB NLSAS disk, which typically are considered 100% utilized at 85 IOPS/disk. The Queue Depth on each disk is also indicating that disks may be waiting on results. A typical best practice is to keep the average disk queue depth below 10.

**Note:** The table is only showing the averages for the selected time (7/1/2015 at 13:15). The graph will represent the host IOPS while the table data shows the metrics to each individual disk.

### 3.1.1.4    Conclusions for Example 1 – Over utilized pool

This default pool is over utilized based on the latency, observations from the **Experimental Analysis** and the fact that the IOPS/Disk value exceeds the rule of thumb threshold. From these observations, it can be concluded that the capabilities of these drives have been exceeded.

Two options can be explored for correcting the situation:

- Move volumes to the unused FASTPOOLXV
- Merge pools

## 3.1.2    Example 1: Analysis of volumes to move

Since the example has one unused pool (FASTPOOLXV). One method to correct the over utilized pool is to move volumes from the default pool to FASTPOOLXV. To determine the best candidates for moving, first view all the volumes in the **I/O** panel by selecting **Group** in the menu, then **Pool** > **default** > **all volumes in pool default**. Sort the default pool IOPS in descending order. For this example, the four volumes at the top will be moved to the FASTPOOL15K
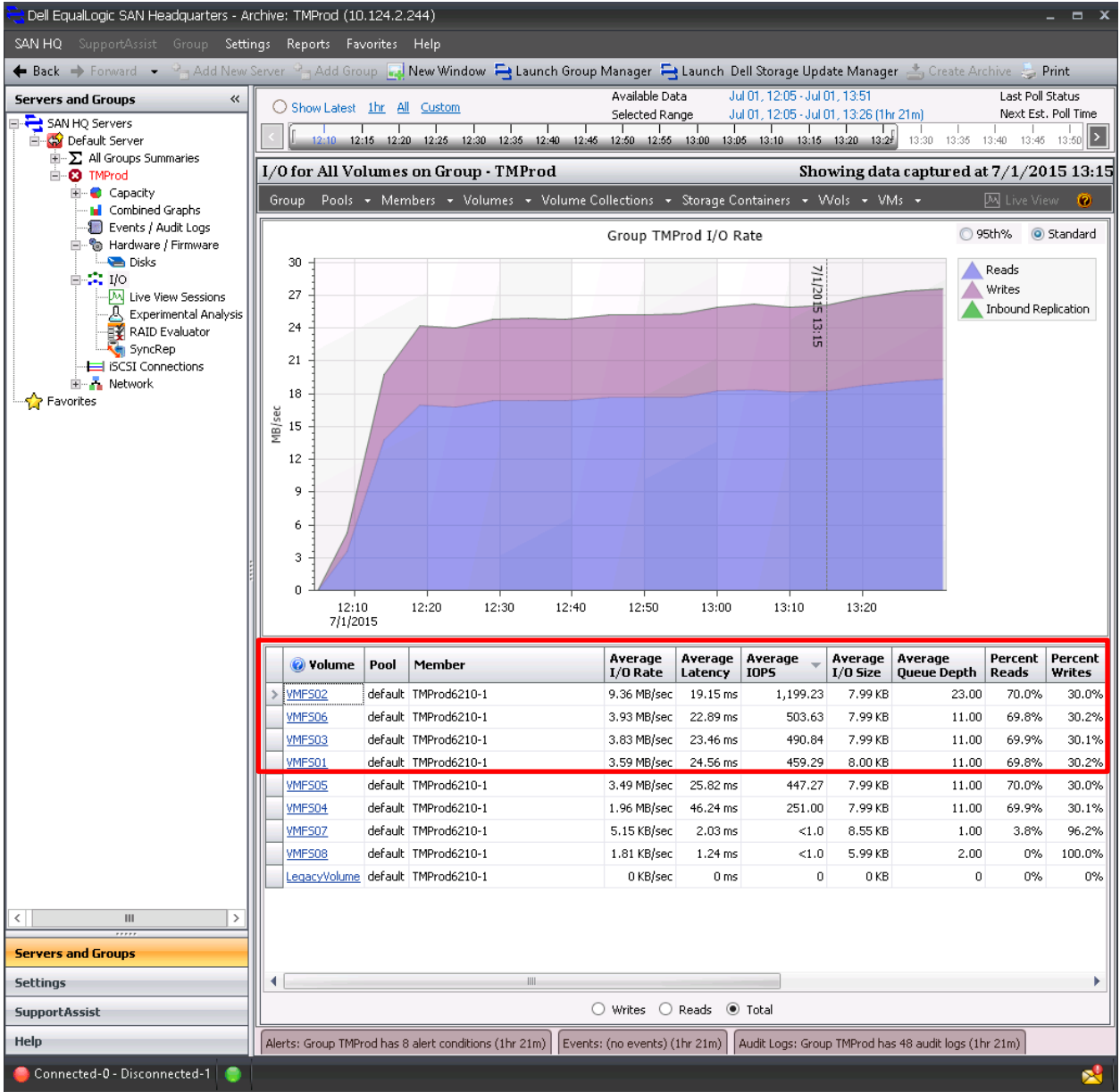
Figure 18    All volumes on Pool. Notice the table data showing highest IOPS by volume.

FASTPOOLXV contains a single PS6210XV in a RAID 6 configuration. Since this pool has a single member with 15K 146GB SAS drives, available capacity should be considered before moving.

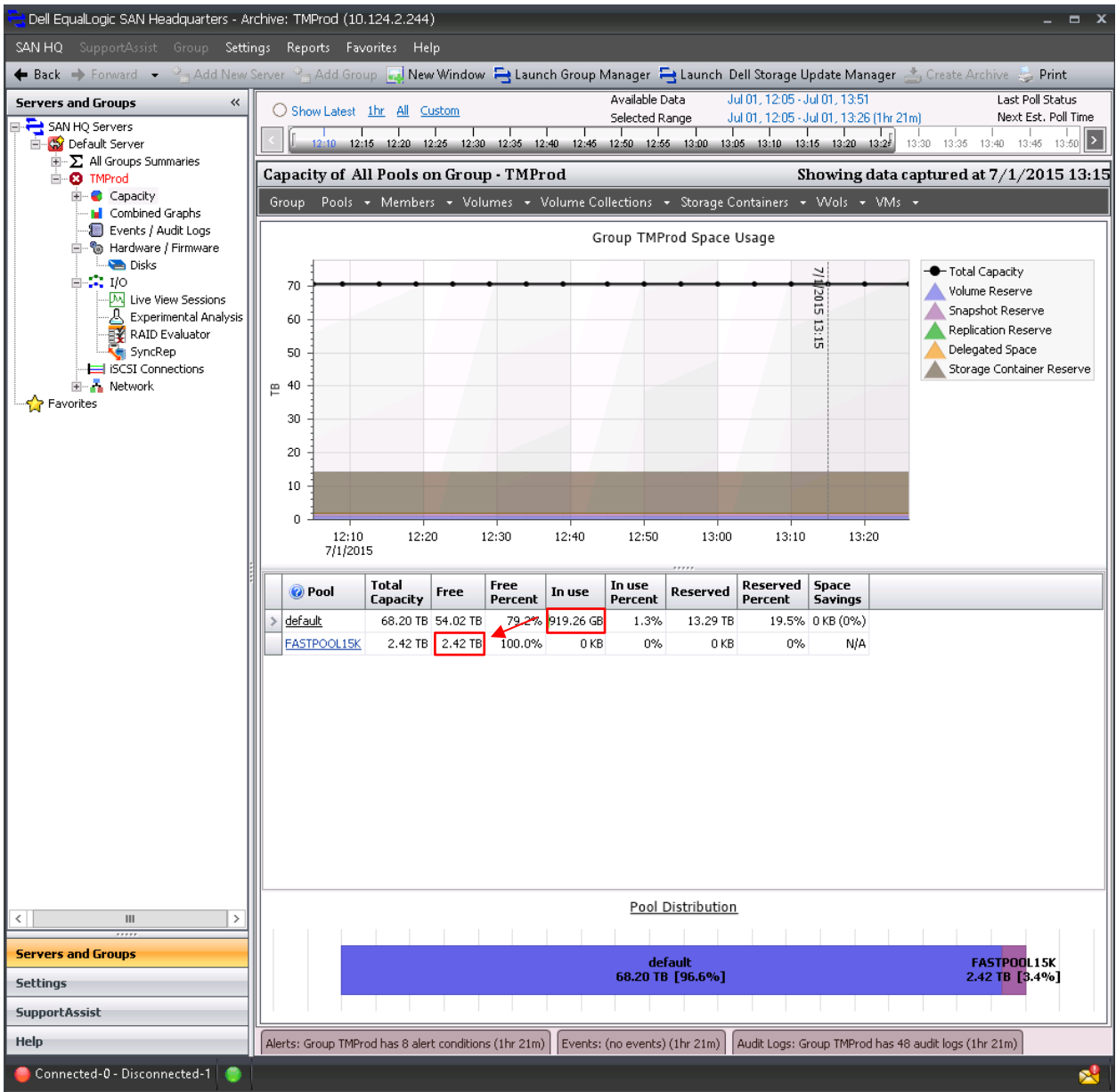The SAN HQ Capacity panel displays the overall capacity of the default pool as indicated in Figure 19.

Figure 19 Capacity for FASTPOOL15K shows 2.42TB of free space

Since these are small volumes (100GB each), the volumes can be safely moved to the destination pool. The Group Manager GUI prevents the move if space is not available. For this example, analyzing the individual volume capacity is not necessary since the entire space in use of the default pool is well below the free space in the FASTPOOL15 pool.

### 3.1.2.1 Moving the volumes with EqualLogic Group Manager

Volumes may be moved without interruption to other existing pools within the group. This allows applications to benefit from the additional resources quickly. IOPS are sustained during the move by prioritizing the application IOPS over the internal group operations (such as moving volumes).
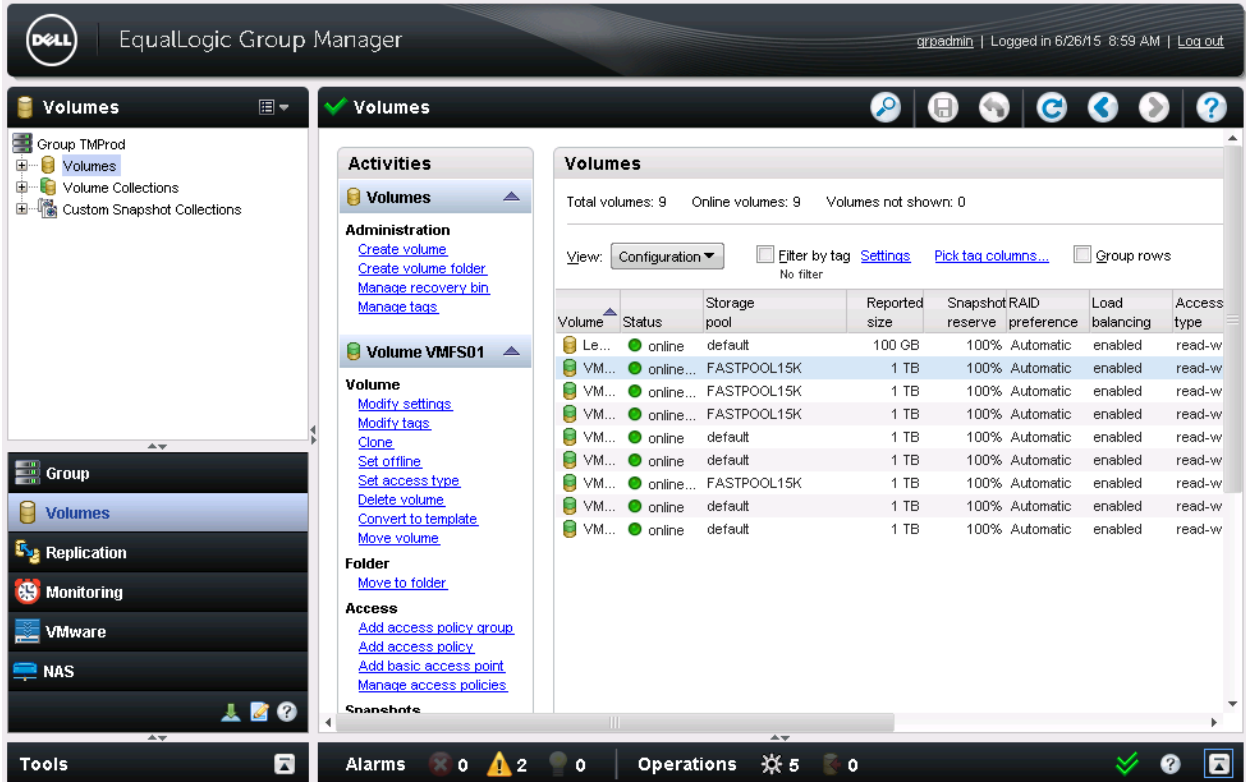
Figure 20    Volumes split between the default and the FASTPOOL15K in the Group Manager GUI

### 3.1.3    Example 1: After volumes are moved

After the volumes completed the move, the same performance test was run again to show the improvement. The following graphs from the I/O panel show the improved latency as well as additional IOPS achieved by splitting the volumes into the two pools.
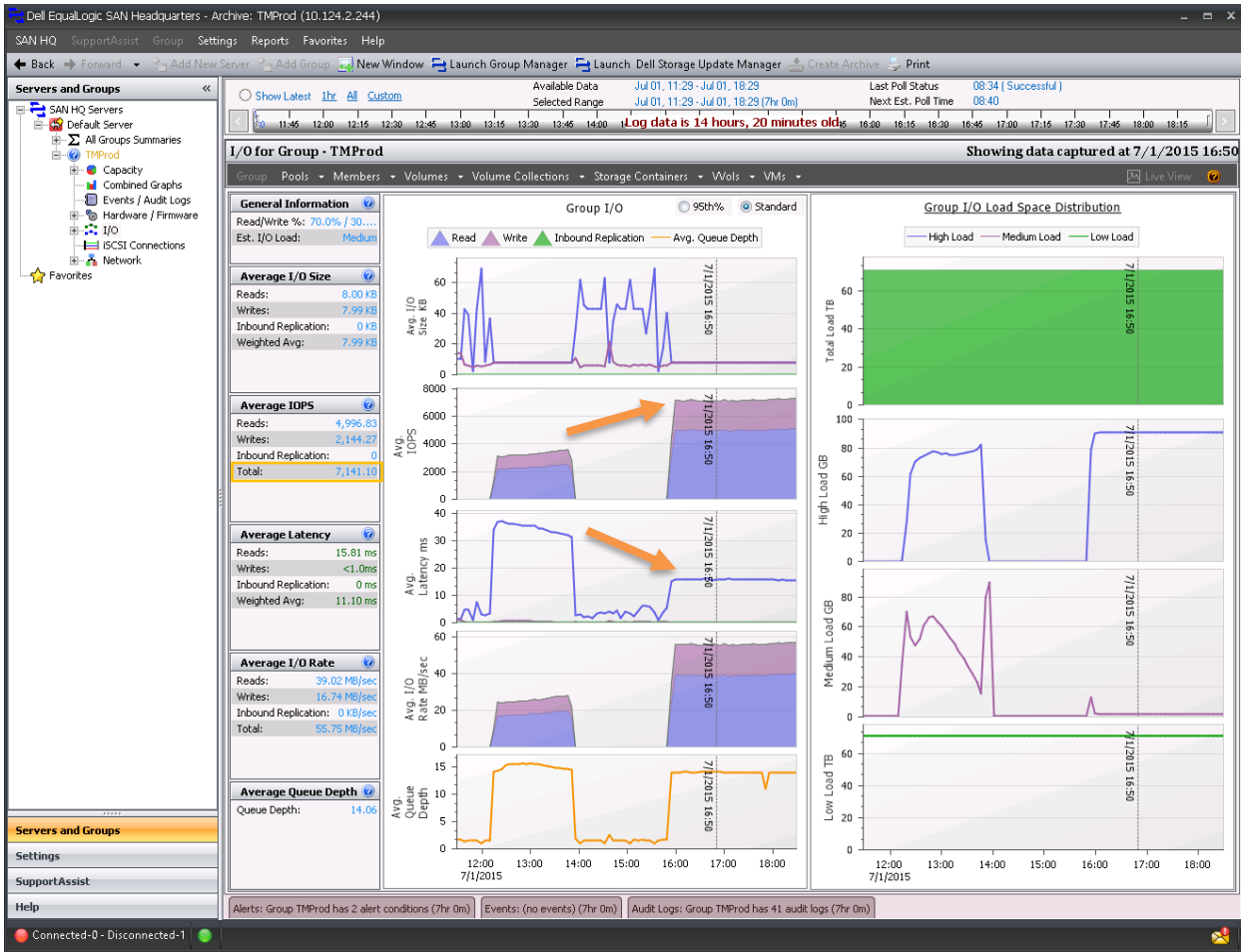
Figure 21    Group view from the I/O panel showing the results of moving the volumes

The arrows in Figure 21 indicate the improvement made after moving the volumes to the new pool. The increase in IOPS indicates that the pool had pent-up demand which was relieved by moving the volumes to the pool with faster drives. In addition, a lower latency was achieved even with the higher I/O load.
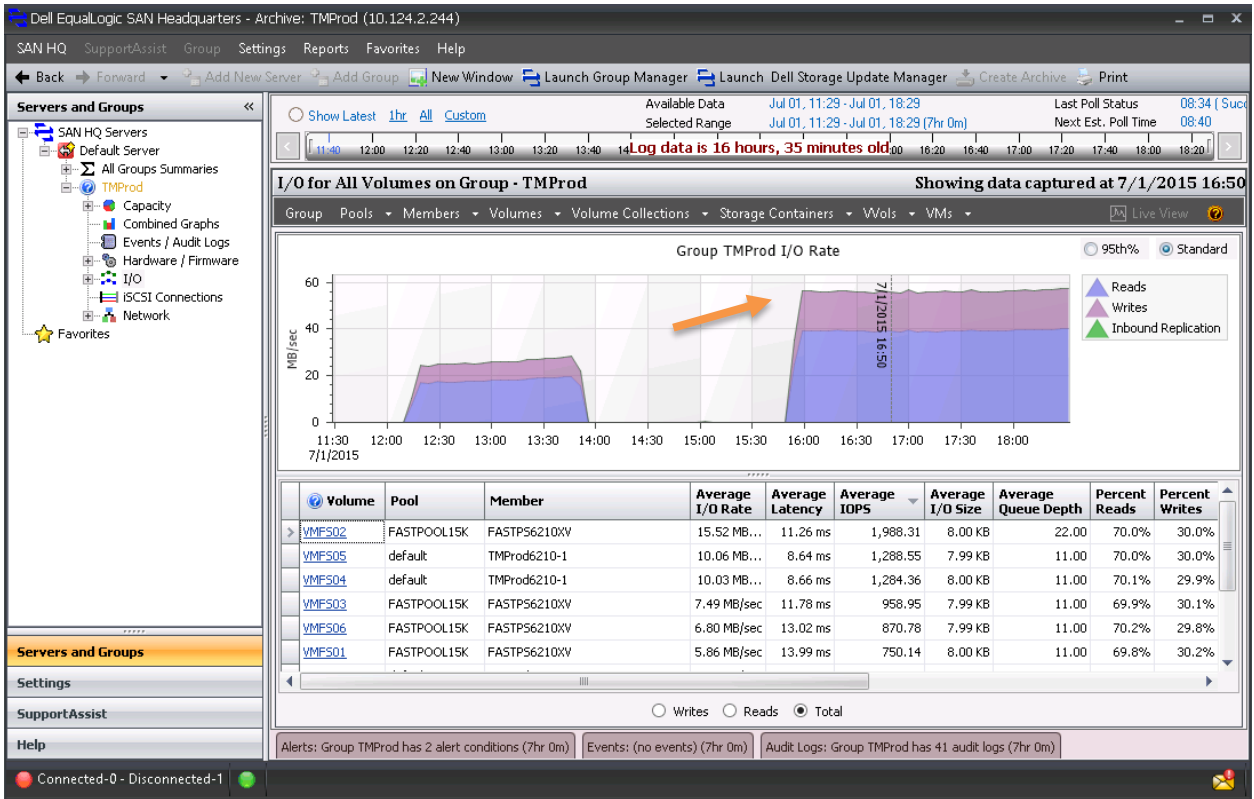
Figure 22     Volume IOPS distribution after the move showing more I/O activity to the FASTPOOL15K

### 3.1.3.1     Iterative process of moving volumes

Manually moving volumes to another pool is a technique which allows for appropriate placement of workloads based on I/O characteristics or business importance. This methodology may be an iterative process and to fully even the load, several volume moves may need to be considered to balance the workload to the desired goal. An example of a second move of the volumes based on their IOPS profile is shown below. Here, we moved volumes that needed more IOPS to the FASTPOOL15K and ones with less of an I/O need to the default pool. The final result is that vmfs05, vmfs04 were moved to FASTPS6210XV and vmfs01, vmfs06 were moved to the default pool.

Table 2     IOPS analysis for volume placement

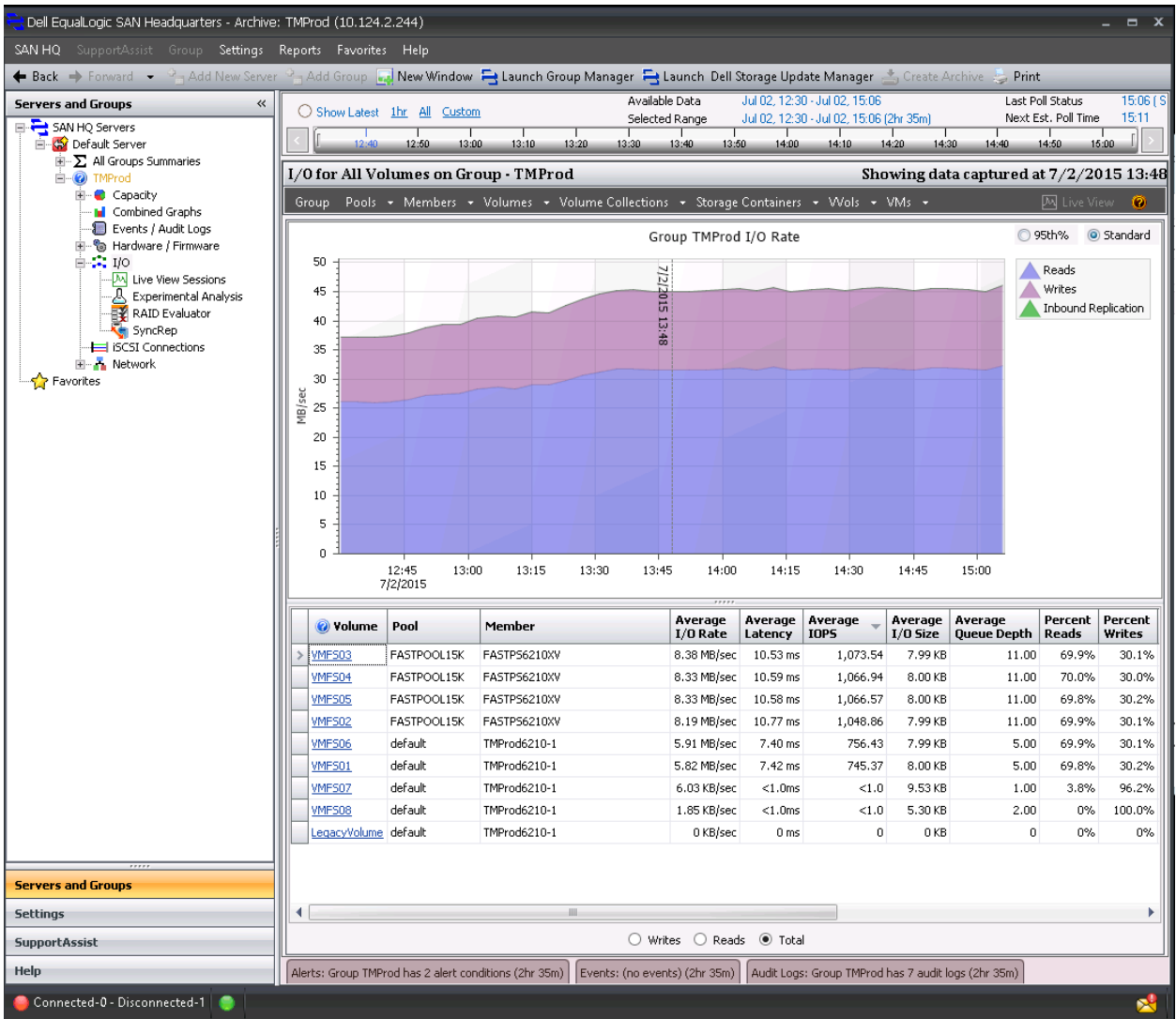| Volume | IOPS | First Move | Second Move |
| --- | --- | --- | --- |
| Vmfs02 | 1988 | FASTPOOL15K | NO |
| Vmfs05 | 1288 | Default | FASTPOOL15K |
| Vmfs04 | 1284 | Default | FASTPOOL15K |
| Vmfs03 | 958 | FASTPOOL15K | NO |
| Vmfs06 | 870 | FASTPOOL15K | Default |
| Vmfs01 | 750 | FASTPOOL15K | Default |

**Figure 23** After the second volume move, the IOPS favor the faster pool.

### 3.1.3.2 Example 1 Move: Conclusion

This example demonstrated a method of associating the faster member to the higher workload. First, a few volumes were moved and the activity monitored. As a result, the decision was made to move the volumes again according to their new workload profile.

## 3.1.4 Example 1: After merging pools

Merging the two pools simply allows the PS Series to virtualize all of the volumes across the aggregate spindles. The PS Series will move data appropriately based on capacity between the resulting members. Typically, this is the simplest method.

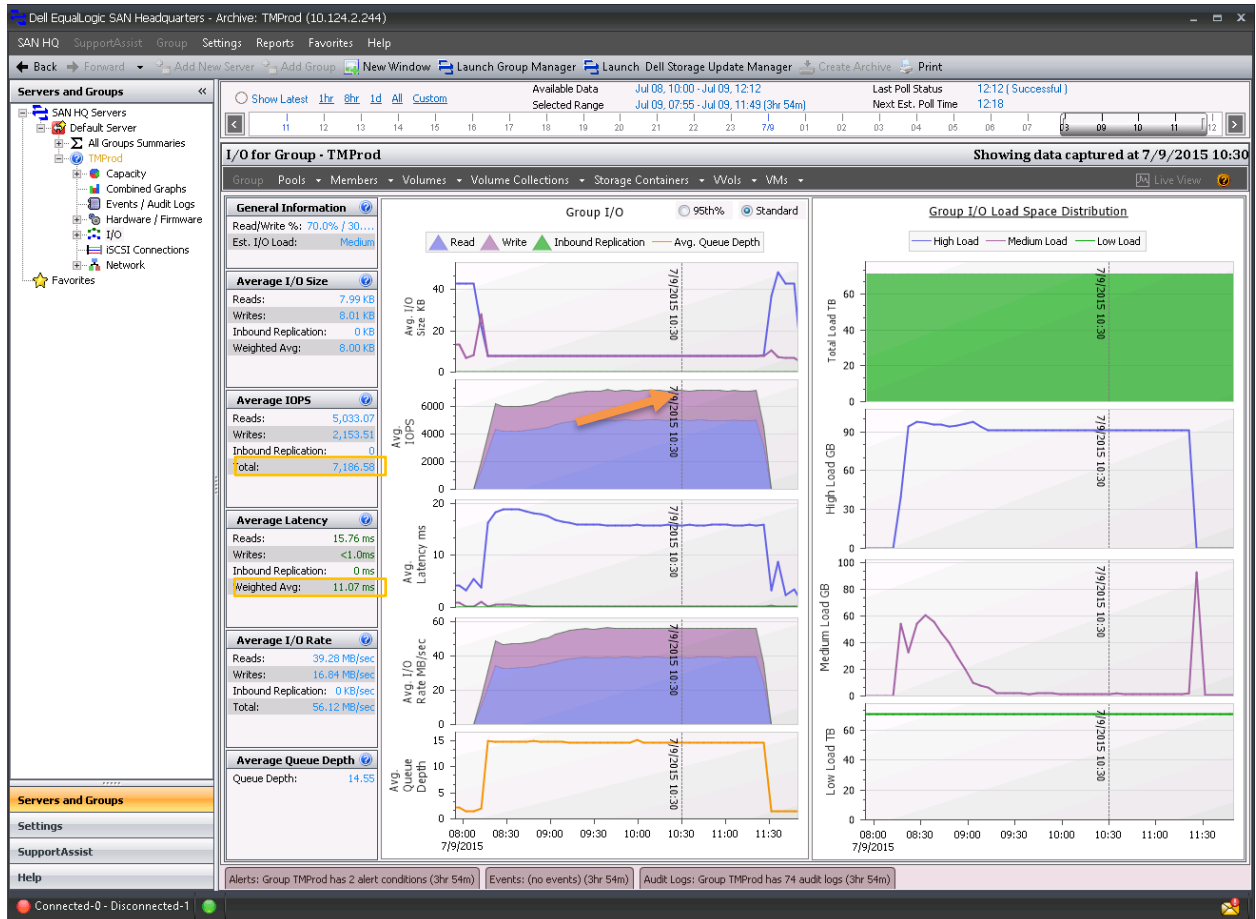Notice the results are similar to moving the volumes to separate pools.



Figure 24    Results after merging the two pools. Latency is below 20ms and IOPS increased to over 7000.

### 3.1.4.1 Example 1 - Over utilized pool analysis: Conclusion

This example touched on several key concepts for performance analysis and resolution. First we analyzed the array and determined the disks in the pool are over utilized.  Then we used some common techniques to help alleviate the problem. The first method involved moving volumes while the second method simply takes advantage of the PS architecture by simply merging two separate pools together.

Manually moving the volumes to a new pool is a more deliberate method which allows the administrator to control where the volumes are placed. This method may be appropriate to meet business objectives. For instance, a marketing department that funded the additional array wants the isolation guaranteed. The second method of merging members

into a single pool simplifies administration and allows the array to optimize for performance and capacity. Both methods have their merits and typically achieve very similar performance results.

## 3.2 Example 2: Performance planning

As a storage administrator, planning for new workloads and correcting bottlenecks are frequent tasks. This next example will demonstrate how a solution may be designed to adequately plan and correct bottlenecks. In this situation, the first example of the over utilized pool is used to determine the best solution for fixing the performance issues. First, the **Experimental Analysis** panel is used to indicate how many more IOPS are needed.
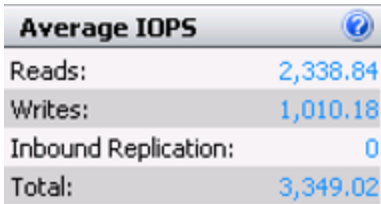
### 3.2.1.1 Calculate the IOPS shortage

Using the previous examples **Experimental Analysis** display before the fix, notice the estimated maximum IOPs for the default pool.



Figure 25      Estimated maximum IOPS for the PS6210E with RAID 6

Also, notice the total average IOPS that represents all of the I/O arriving at the PS Series group.



Figure 26      All IOPS for the PS Series group

To estimate the additional IOPS needed, subtract the total average IOPS (3349) from the estimated maximum IOPS (991) to find the difference (-2358). So our goal is to accommodate the shortage of 2358 IOPs. These IOPs are considered the total front-end IOPS needed.

*Additional front-end IOPS needed = Total average IOPS – Estimated maximum IOPS*

The delta we need to support is at least 2358 IOPS coming from the hosts at the same average, 70% reads and 30% writes, with an 8KB I/O size. I/O size is important since 8KB is the typical size that the rule of thumb methods use for sizing. Next, determine the back-end IOPS (total disks IOPS).

The formula below shows how back-end IOPS are approximated.

*Total disk IOPS = reads + (writes x write penalty)*

The "write penalty" is the additional IOPS needed to protect the disks with RAID. Table 3 shows the values for the available RAID policies for PS Series arrays.

Table 3        Write Penalty based on RAID

| RAID Policy | Write Penalty |
|-------------|---------------|
| RAID 10     | 2             |
| RAID 50     | 4             |
| RAID 6      | 6             |

**Note:** RAID Policy is defined on the PS Series member. The Pool may have multiple members at different RAID policies.

For this example, a RAID 6 solution will be considered which has a write penalty of 6. Plugging in .7 to represent the reads and .3 to represent the writes we get the non-reduced formula below:

*Total disk IOPS = (2358 x 0.7) + ((2358 x 0.3) x 6)*
                        *= 4480*

The number of disks needed to support this workload will need to be determined. Pick 15K disks from the rule of thumb in Table 1 (210 IOPS per disk). The last step is to divide the total disk IOPS needed by the rule of thumb IOPS per disk.

*Total disks needed = Total disk IOPS/Rule of thumb IOPS per disk*

**Note**: The rule of thumb IOPS per disk value comes from a best practice formula that is based on disk seek latency, drive speed, drive type and form factor.

*Total Disks Needed = 4480/210*
                        *= 22*

Because we are using the PS6210XV (15K disks) in a RAID 6 (23 effective disks), this solution is appropriate.

However, consider the pent-up demand effects described below.

### 3.2.1.2    Pent-up demand

The IOPS per disk are within reason for 15K drives (210 IOPS). However, additional IOPS may be pending due to the relatively high latency. Pent-up demand is difficult to predict and often requires further detailed analysis to find an accurate model.

Actual results may vary. As seen in the first example, the load increased to near 7000 IOPS (vs. 3400 IOPS) while maintaining a reasonable response time. Review the results in Figure 24.

### 3.2.1.3　Example 2 - Performance planning: Conclusion

A new PS6210XV will have 24x15K disks and should be added to the group. The new member would handle the current IOPS. However, the pent-up demand may indicate an additional member may need to be added to the pool. This sizing methodology should be useful to help properly design a solution to meet the desired performance requirements.

## 3.3　Example 3: SQL application load example

For an application-relevant example, consider a Microsoft SQL load process. In this case, the application is simply inserting data into new tables across multiple databases.

The **Experimental Analysis** panel below shows the load on the system. Notice that the estimated maximum IOPS exceed the total average IOPS during this time line.

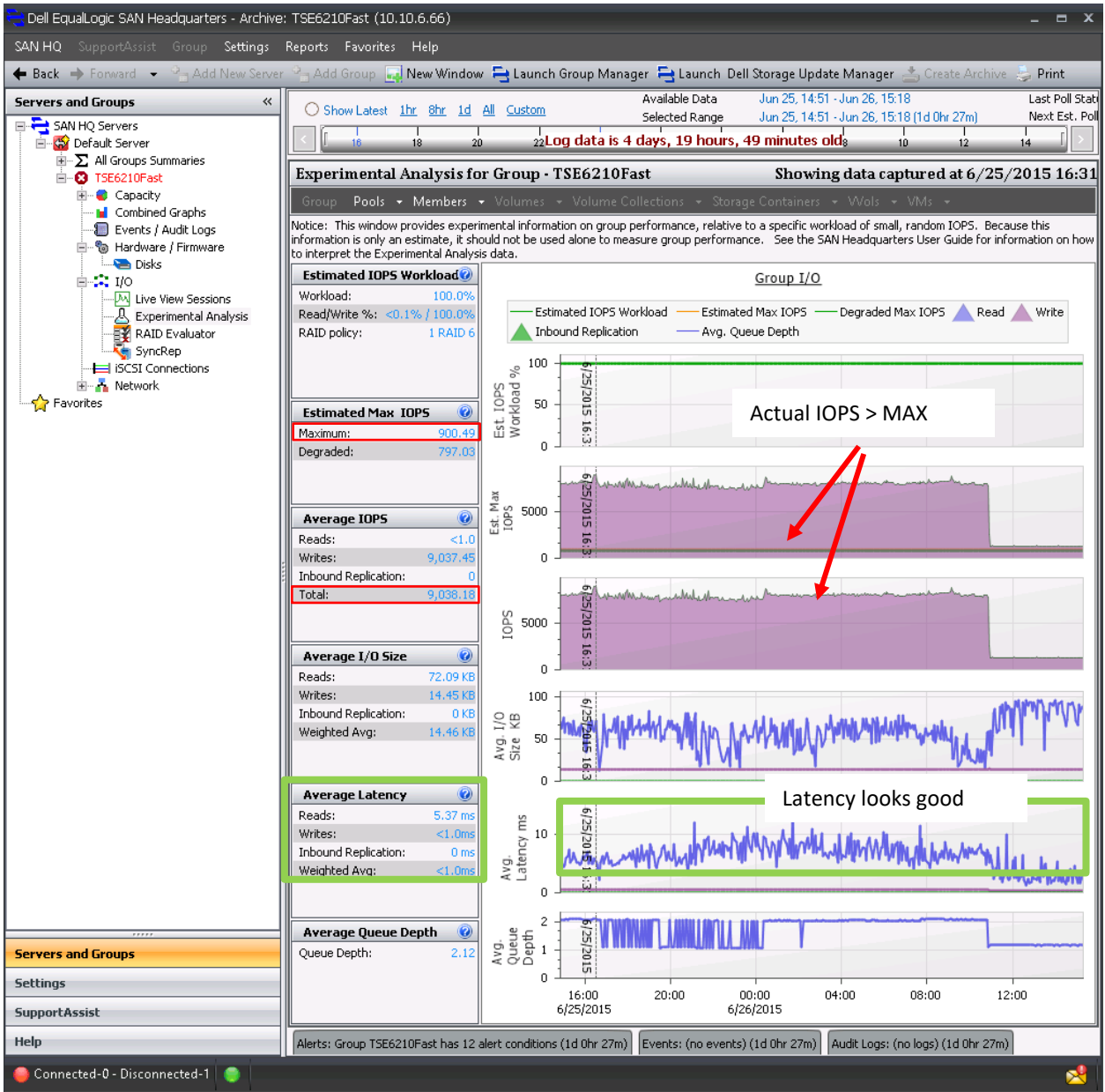Next, observe the results to determine if there is a performance issue.

Figure 27　　SQL Insert commands loading tables show IOPS exceed the estimated maximum IOPS, however performing with low latency

From the **Experimental Analysis** panel, the overall I/O load appears to indicate that this PS Series group is exceeding its capabilities. However, latency is below 15ms. For this part of the application 15ms latency is acceptable.

The actual pressure on the disks should also be verified by clicking **Hardware / Firmware** > **Disks** as shown below. For the time interval selected, the table data shows each disk is well below the rule of thumb maximum of 185 IOPS per disk.
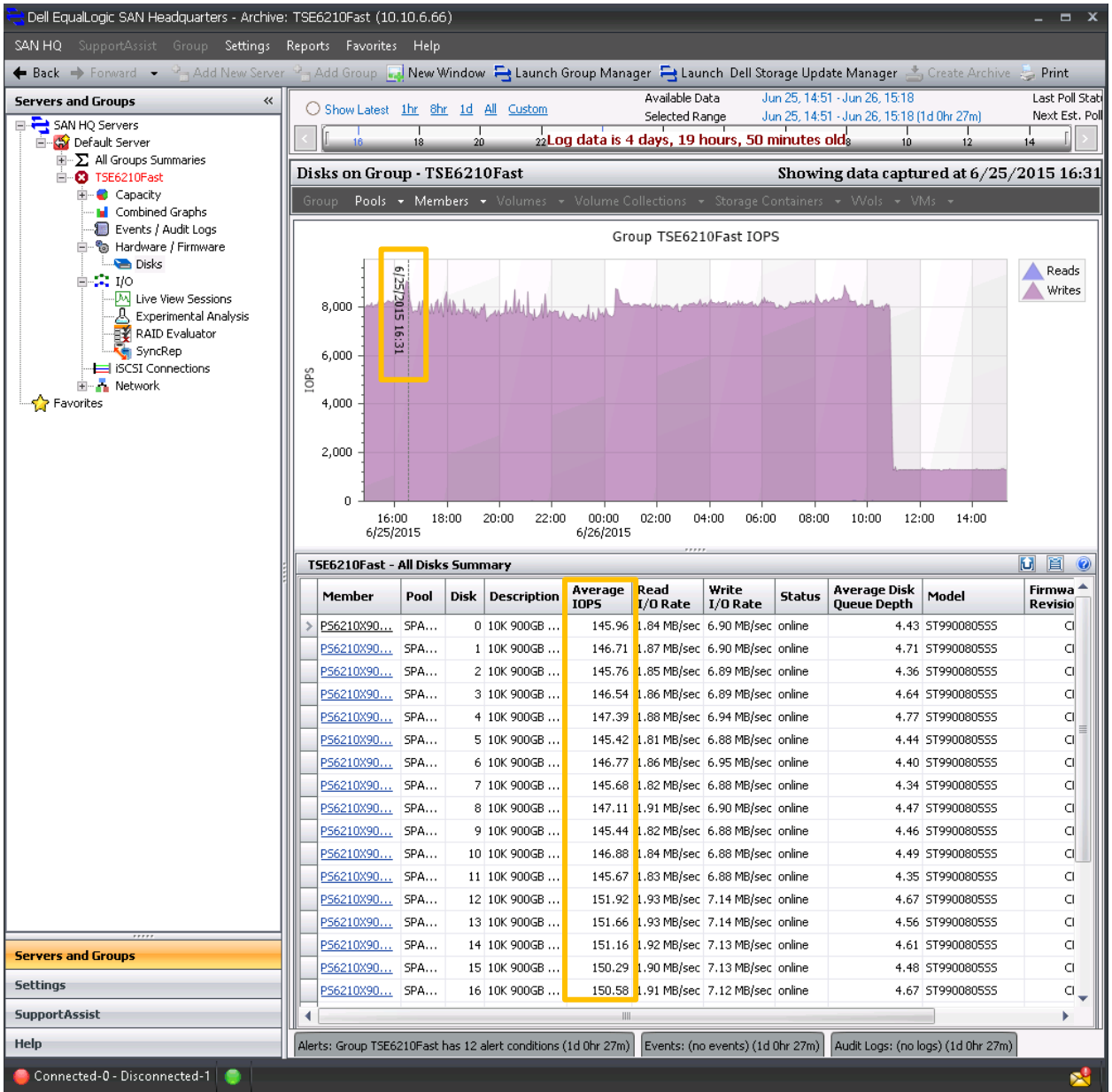
Figure 28    Disk table data

### 3.3.1    Example 3 - SQL application load: Conclusion

Although the **Experimental Analysis** seems to indicate that the maximum capabilities of this array are being exceeded, the actual workload latency and disk IOPS are within acceptable performance criteria. One of the goals of this example is to show the importance of understanding that the estimated maximum IOPS should be used as a tool and validated by reviewing latency and actual disk IOPS measurements.

## 3.4 Example 4: Benchmark of Exchange

This example is from a Microsoft Exchange Solution Review Program (ESRP). From an analysis perspective, we want to see how well the PS Series array handles a large mailbox configuration. This test typically pushes the limits of the resources to show maximum efficiencies for the underlying solution.

The **Combined Graphs** show very distinct I/O patterns. One part of the I/O is significantly lower than the other. Latency also appears very high for most of the sample.
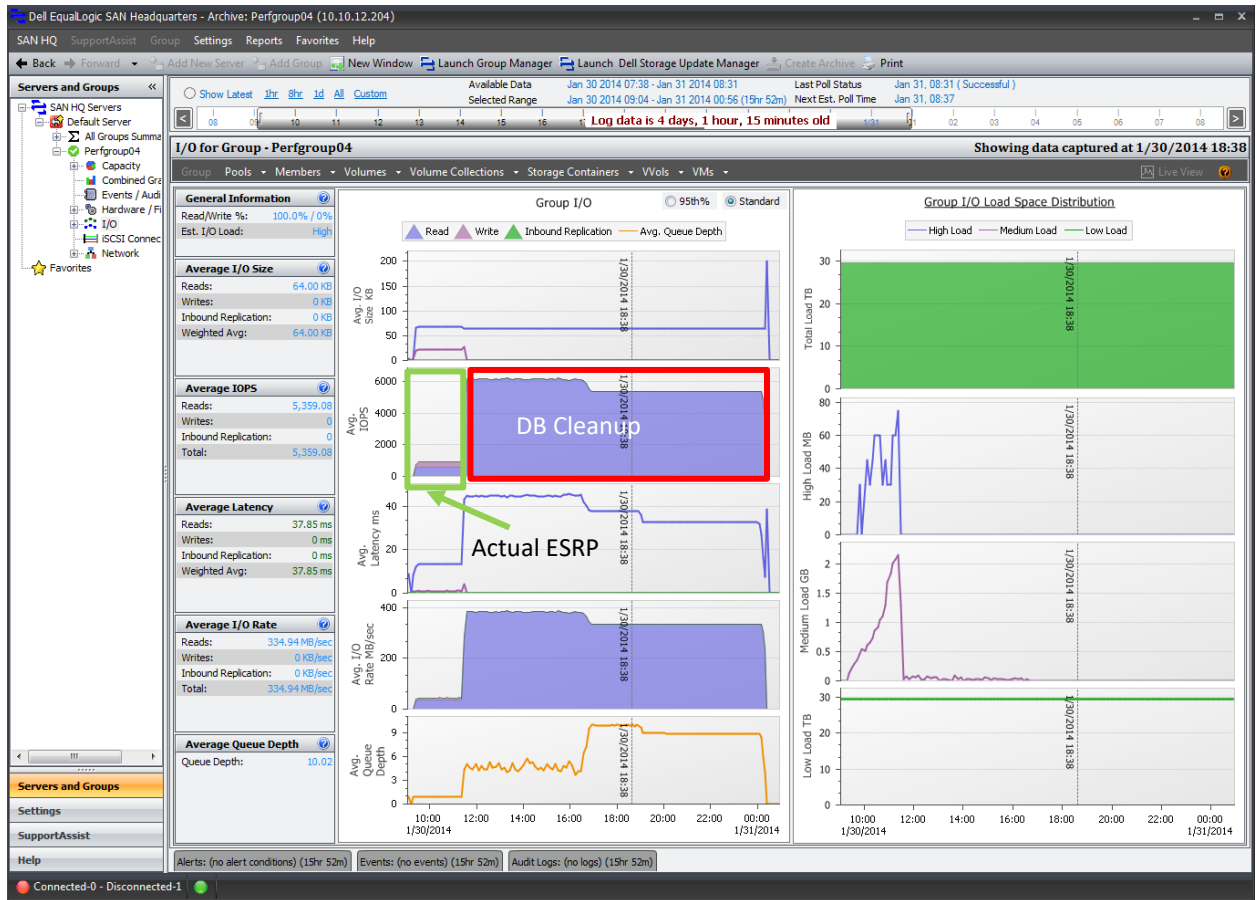


Figure 29     Initial view of the complete ESRP as shown in SAN HQ

However, from what is known about the ESRP mailbox rating is that the first part of the test and is the actual application execution. The second part of this sample is showing IOPS during the database cleanup and is not part of the ESRP certification criteria.

Focus on the relevant sample periods to be sure context is understood while reviewing the SAN HQ data. For the purpose of this example, the time sample is selected to show only the actual ESRP workload test. This is accomplished by shrinking the time range bar to include only the 9:00 AM to 11:00 AM period.
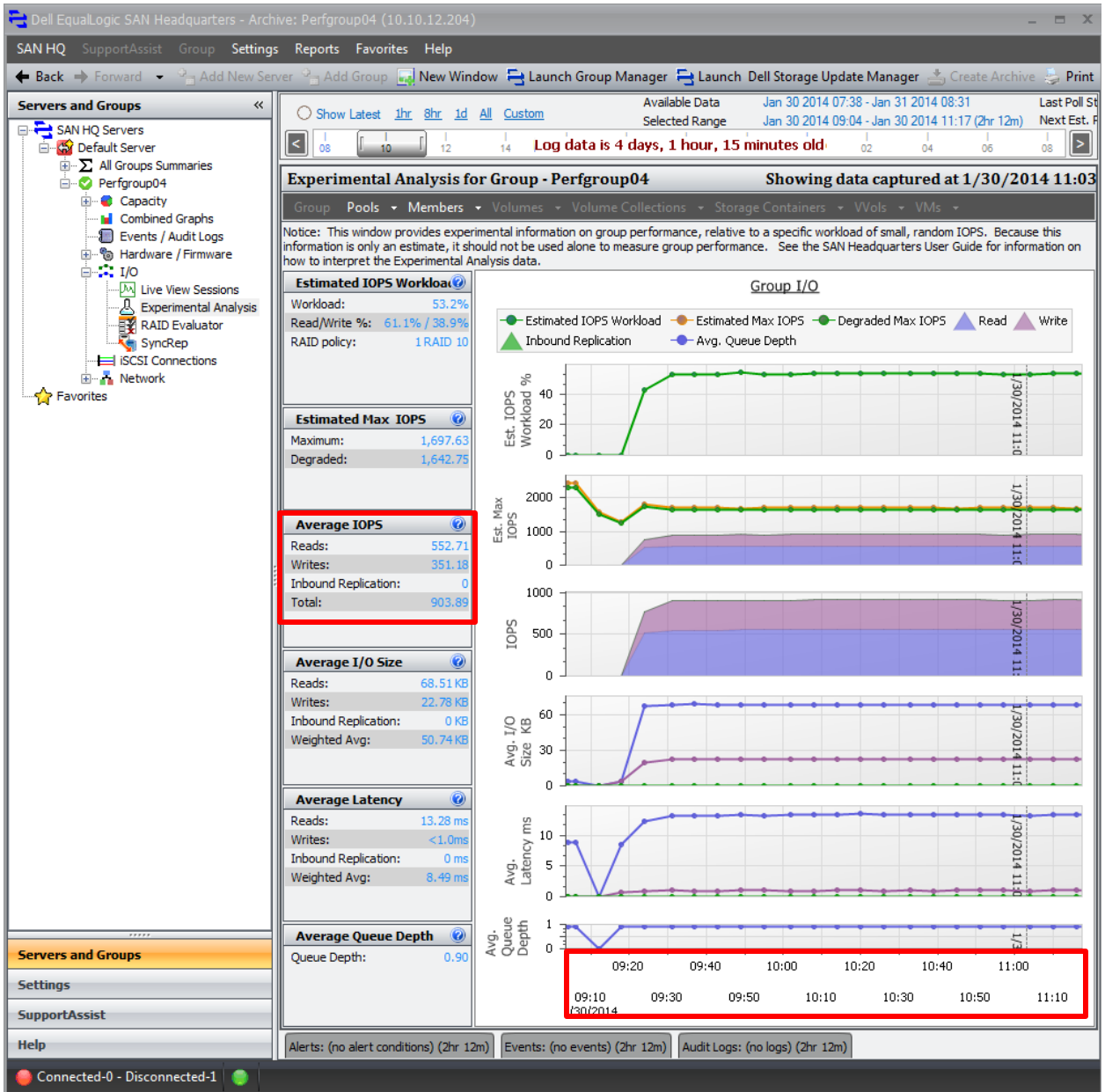
Figure 30    **Experimental Analysis** shows latency below 14ms, steady IOPS of 900 and plenty of headroom from estimated maximum IOPS

The adjusted view provides a more accurate overview of the I/O load, and indicates a successful test. Also important is the length of an interval period considered for the averages. The interval for each sample in this example is over six minutes. If more granularity is needed, SAN HQ provides this with the Live View tool demonstrated in the next example.

## 3.5 Example 5: Live View

**Live View** allows for up to a 10-minute sample at one-second intervals. More precise sample intervals helps in difficult troubleshooting scenarios, planning for new applications or just establishing a baseline profile of an existing application. **Live View** sessions may also be saved for later analysis.

Using the Microsoft Exchange ESRP example to demonstrate the benefits of **Live View,** Figure 30 shows that the IOPS are around 900 and latency averages are less than 9ms for both reads and writes.



Figure 31      I/O total for the ESRP tests

**Live View** for a single two-minute sample exposes a more dynamic I/O pattern at the one-second interval average.
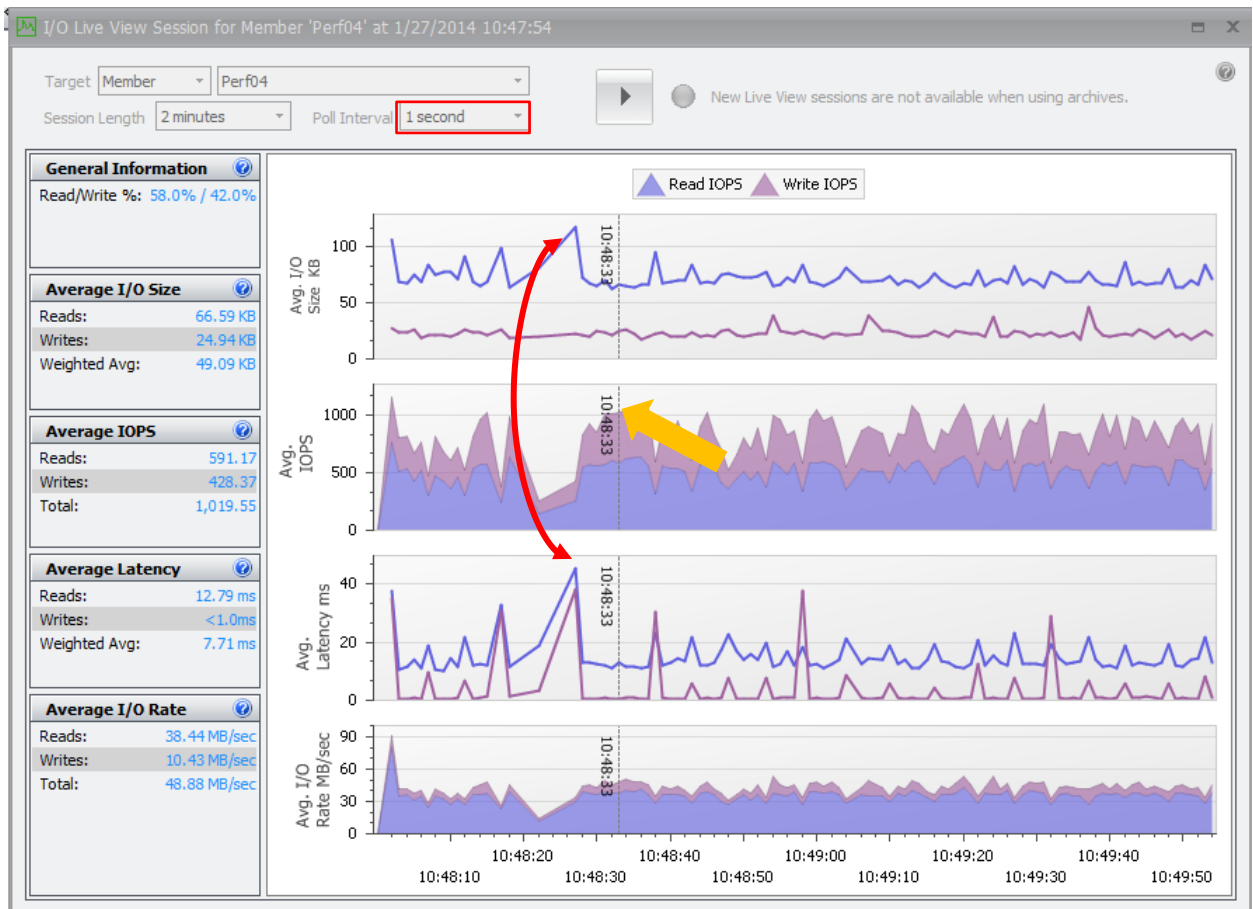


Figure 32      **Live View** during the ESRP test shows the one second poll interval of latency and I/O

This **Live View** sample shows I/O spikes of over 1000 and a few latency spikes. The latency spikes correspond to larger I/O sizes (as expected). In this case, **Live View** is providing a more detailed picture of the ESRP I/O profile. Although the spikey behavior when shorter intervals are measured is evident, this does not necessarily represent a problem. More accurately, these spikes should be considered when sizing to solutions close to the maximum capabilities of the array.

## 3.6    Example 6: Raid Evaluator

SAN HQ includes a reliability-modeling tool known as the **Raid Evaluator**. The **Raid Evaluator** provides expected relative reliability based on the current RAID Policy and the expected performance with different RAID policy choices.

In the ESRP test, the administrator wanted to gain usable capacity by changing to a RAID 6 configuration. In order to model this change the possible RAID Policy was changed to a RAID 6 from RAID 10 which is the current RAID Policy. This provides the additional capacity and a higher relative reliability score since now two parity disks are available in the event of a hard drive failure.
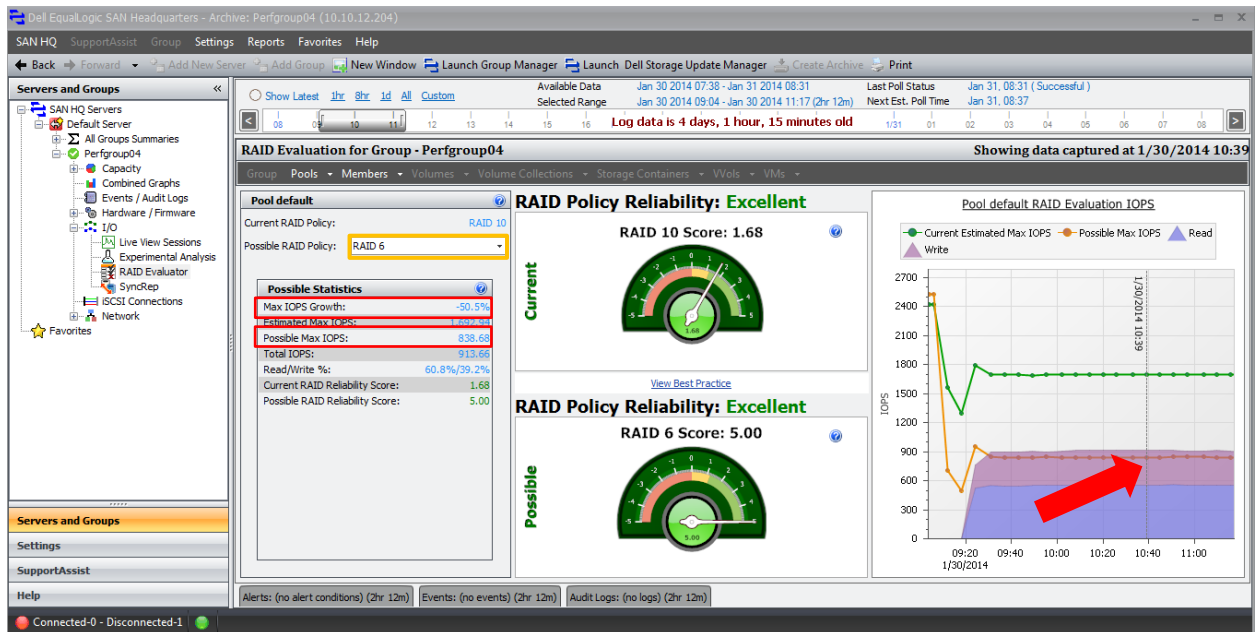


Figure 33      RAID Evaluator showing higher reliability; however **Possible Max IOPS** are lower

Notice the right side of the dashboard models the expected possible maximum IOPS from a RAID policy change. Since modifying the RAID policy from RAID 10 to RAID 6 would increase the number of writes to the disks, we would expect the estimated maximum IOPS to be lower. See Table 3 for the write penalties based on RAID policy.

The RAID evaluator models the differences that RAID choices have on performance. Since RAID 6 requires more I/O for writes the maximum IOPS growth is -50% for the time interval selected. For this example, changing the RAID policy may have an adverse effect on the performance.

**Note:** Verify the possible impact on performance before changing the RAID policy

The example below shows a situation where choosing the more reliable RAID policy has little impact on the current IOPS. The difference between the current IOPS and the possible maximum IOPS allows for headroom even considering the maximum IOPS for RAID 6 will be less than the RAID 50 policy.
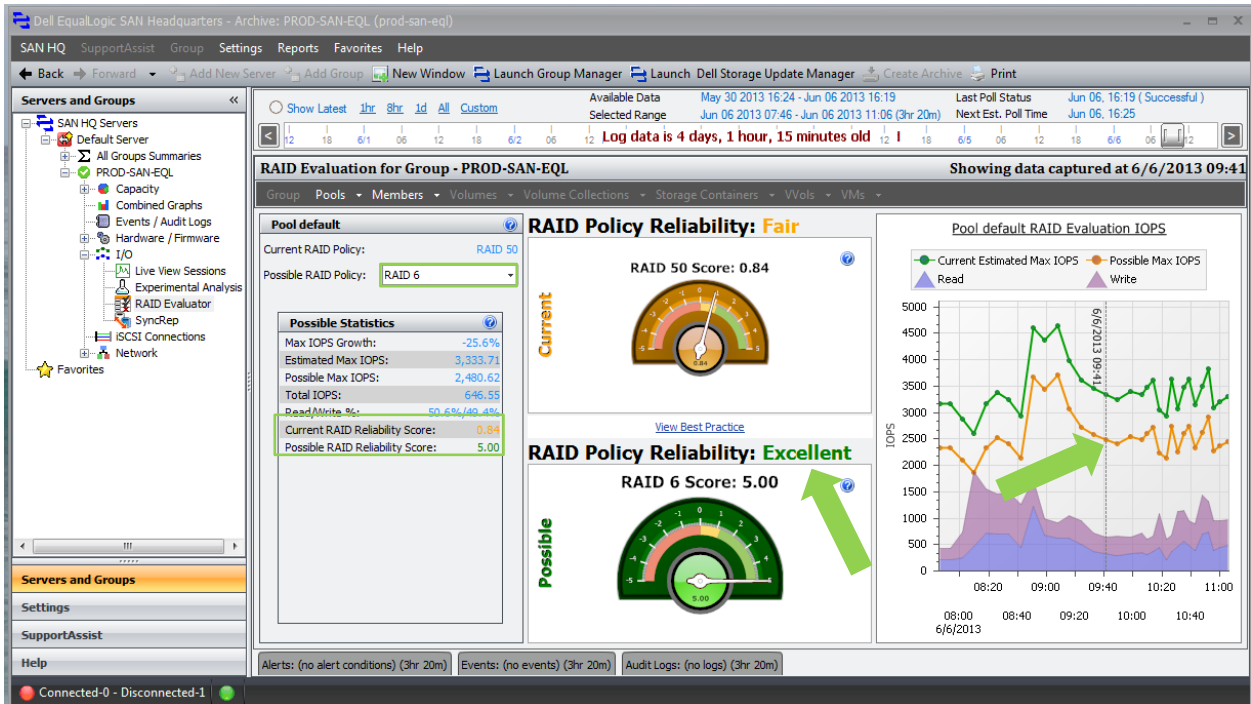


Figure 34    RAID Policy change improves the reliability and has little effect on the performance

## 3.7    Example 7: PS Series hybrid performance

The PS Series offers several models to meet the business, performance and budget needs. One high performing model is known as a hybrid which offers high performance SSDs in combination with large capacity NLSAS drives intended for less active data. The architecture uses an advanced page optimization methodology which exploits the faster drives to accommodate the performance needs while utilizing larger drives for additional capacity.

The example below is from the PS Series PS6610ES hybrid array that includes both SSDs and NLSAS. Notice in the table below the graph how the bulk of the IOPS are processed by the SSD drives.
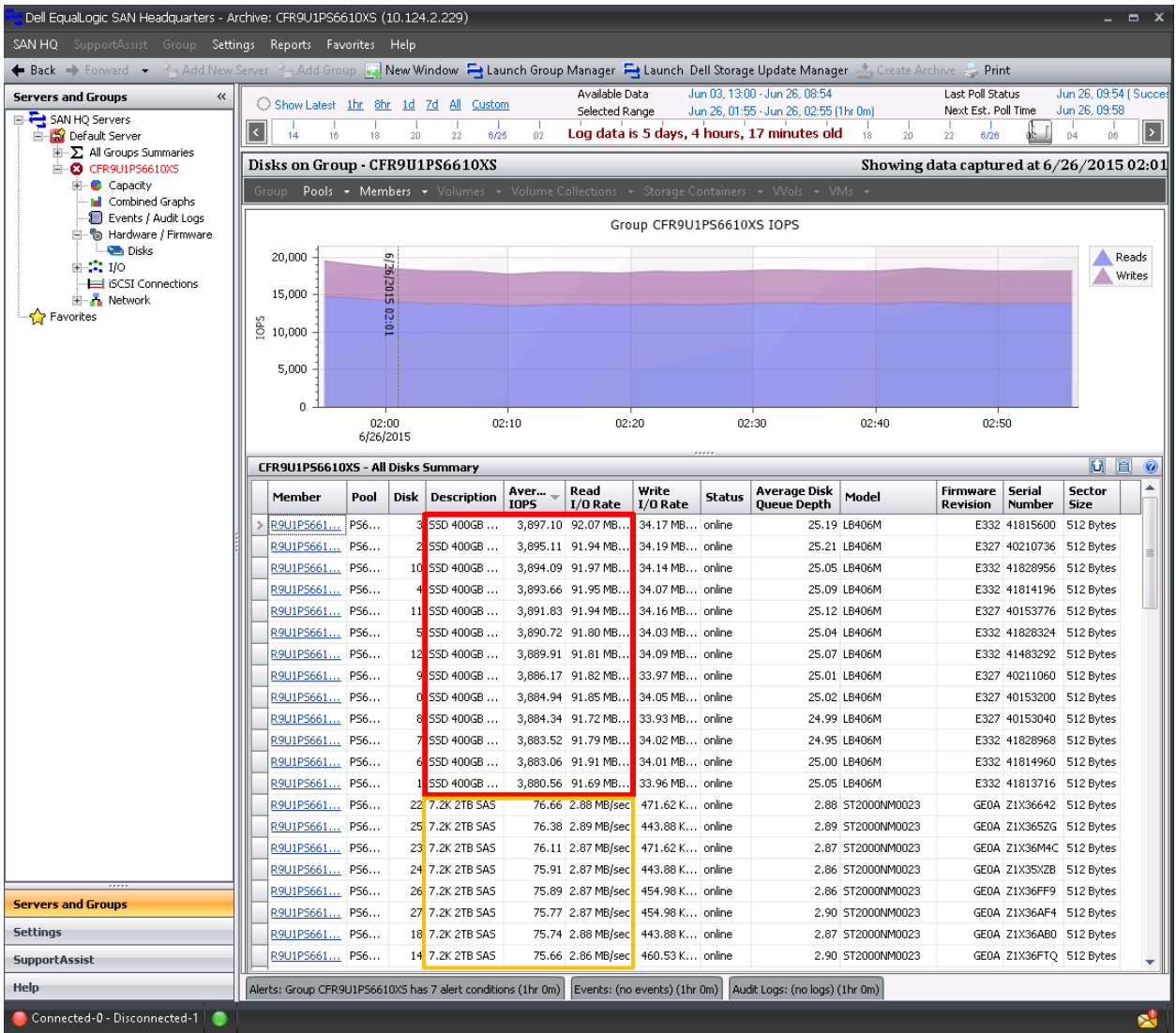
Figure 35   Disk performance on the hybrid showing the SSDs performing around < 4000 IOPS while the NLSAS is near 76 IOPS

The table below the chart shows the SSDs in excess of 3800 IOPS while the 7.2K SAS drives are near 75 IOPS per disk. SAN HQ can be used here to view the placement of data in relation to the I/O needs.

In the I/O panel, **Group I/O Load Space Distribution** data is represented according to its frequency of access to the hybrid.
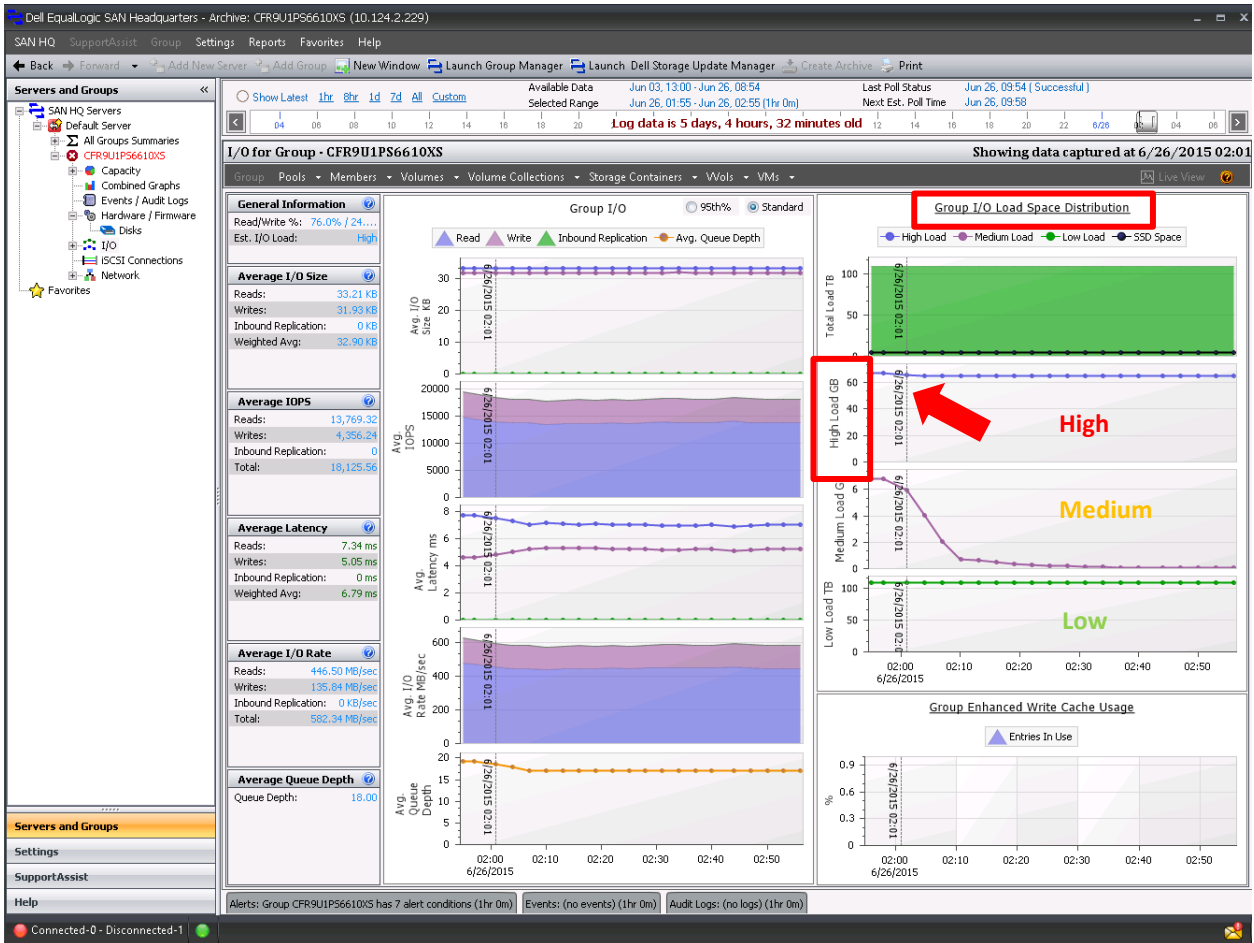
Figure 36    **Group I/O Load Space Distribution** indicates around 62GB are high load pages, 6GB are medium load pages and over 100TB are consider low load

The **I/O** panel displays the breakdown of high, medium and low load distribution. In this example, the highest load needs are about 60GB of space and that is accounted for on the SSD drives. However, medium and low loads may also exist on the SSDs; these may be moved to the 7.2K drives appropriately as determined by the Hybrid Advanced Page Load Balancer algorithm. More information about the PS Series load balancers may be found at PS Series technical documents and videos.

## 3.8 Example 8: VMware vSphere Virtual Volumes

Starting with ESX 6.0, a new feature is available with VMware® vSphere® known as Virtual Volumes, or vVols. This feature is supported on PS Series firmware version 8 and later and SAN Headquarters 3.1 and later. vVols are different from the traditional PS Series volumes and have their own views to show capacity and performance.

Information about SAN HQ vVols:

- Volumes are not associated with vVols they are a separate object type.
- I/O graphs show all the I/O to the Group, Pool or member that has vVols.
- Containers have vVols and their associated files. Several file types are associated with a vVol including – config (.), data (.vmdk) and swap (.vswp).
- Most application I/O is associated with the data (.vmdk) files within a vVol.
- The minimum interval for vVols in **Live View** is 35 seconds.

vVol capacity may be viewed with SAN HQ in several ways:

- Within the group **Capacity** panel, select **Storage Container**, **VVols**, or **VMs** to view the capacity attributes.
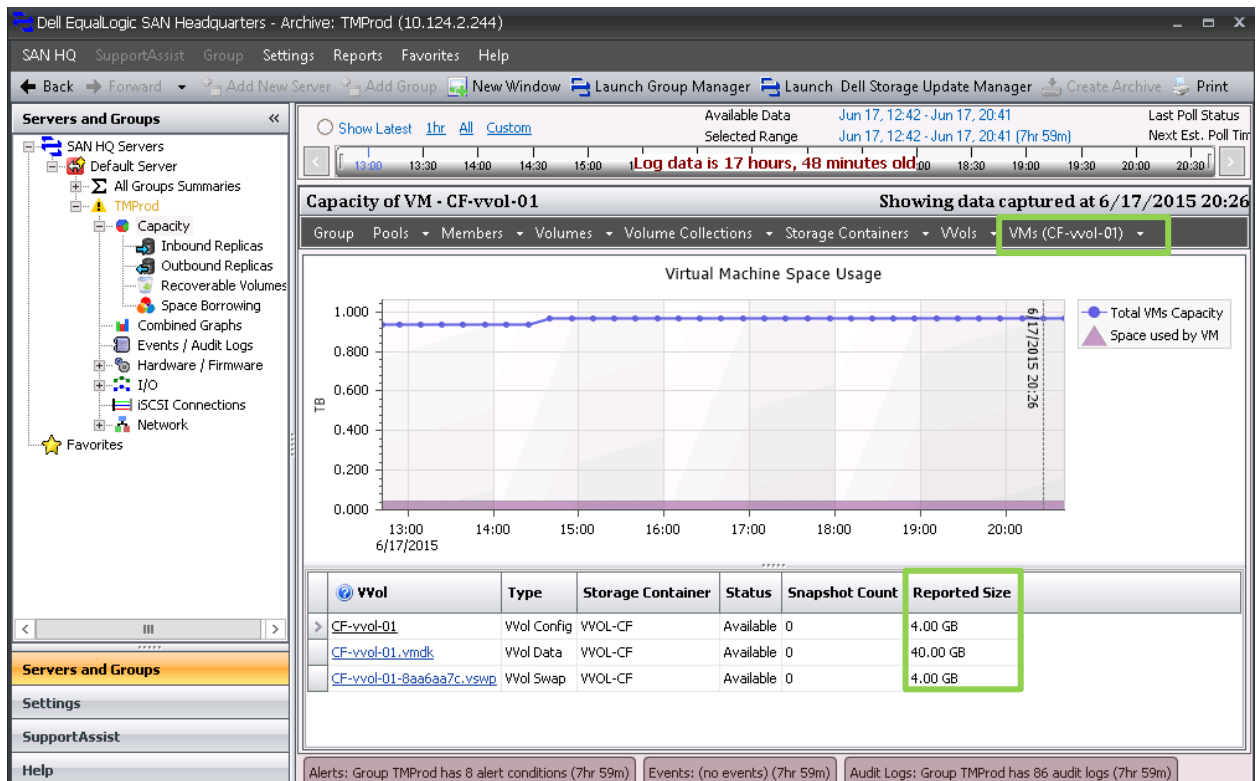


Figure 37        Capacity of all vVol types within a vVol virtual machine (VMs)

- From **All Group Summaries** navigate to **Volume Capacity Summary** (either select the volume type in the search or view the vVol).
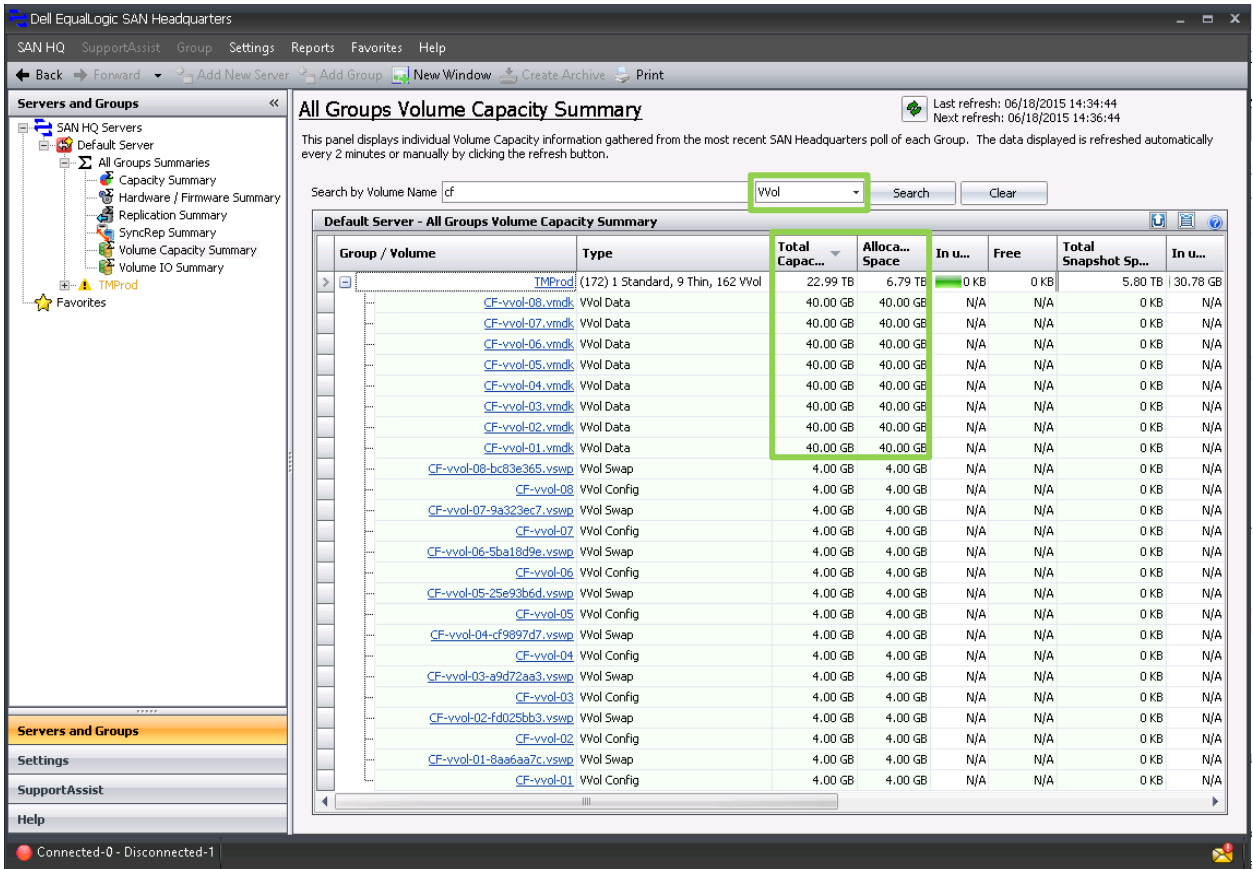
Figure 38    vVol capacity shown in the **All Groups Summaries** > **Volume Capacity Summary**

- View vVol performance by selecting **Group**, **Pool, Member**, **Storage Container**, **VMs** or **VVols** menu options. In Figure 39, performance metrics are demonstrated for the storage container "test" (which is on this array).
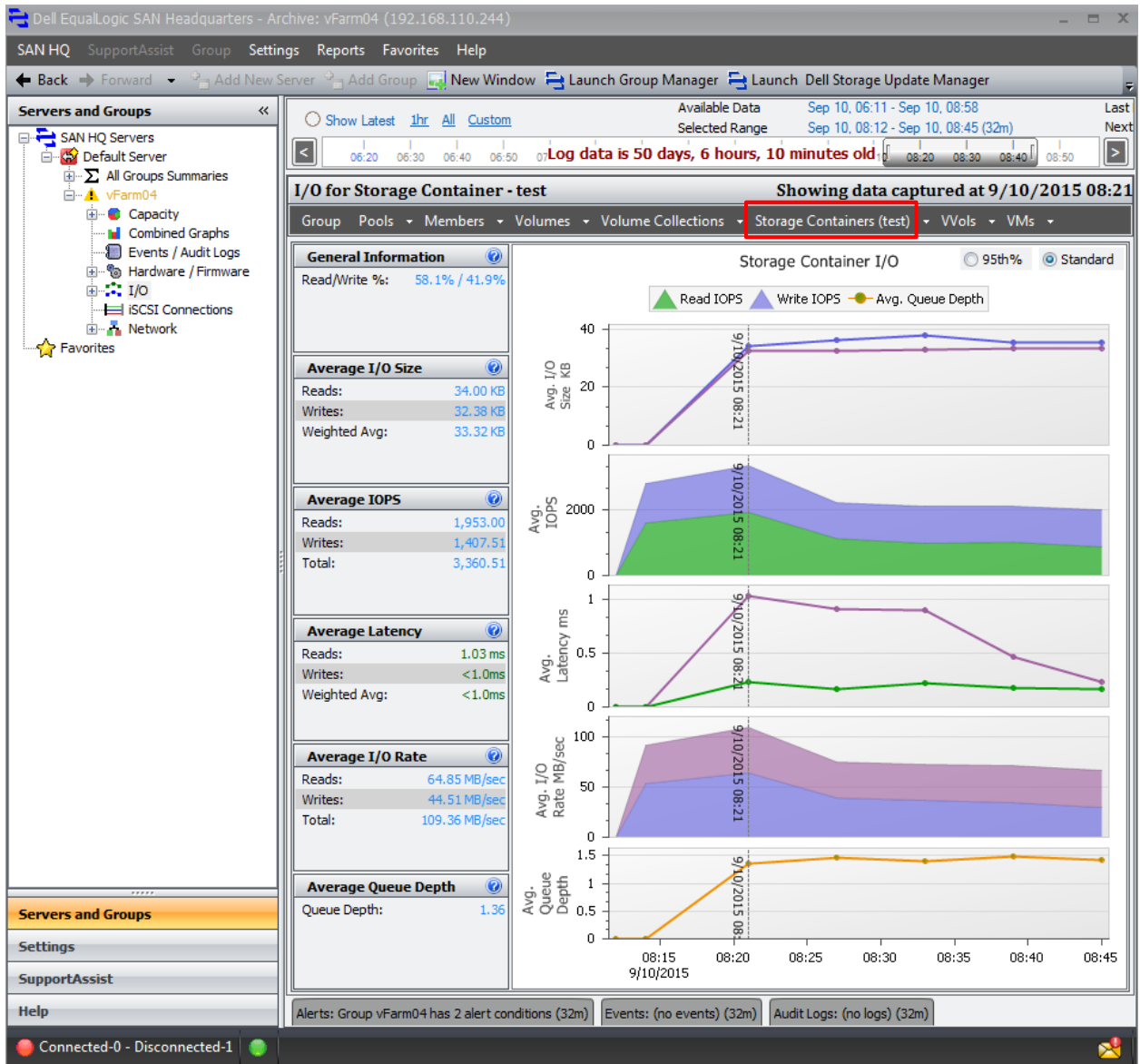


Figure 39    I/O panel showing all metrics for the storage container

# 4 Summary

SAN HQ acts like an in-flight data recorder for your PS Series group that is a powerful monitoring and analysis tool. It provides SAN administrators with valuable insight into the health of their storage environment. The easy-to-use graphical interface provides information on PS Series group capacity, I/O performance, network data, member hardware and configuration, and volume data. With the ability to show trends and export metrics for further reporting and analysis, SAN HQ is a key component in the constant daily battle that administrators face to do more with fewer resources.

# A   SAN Headquarters tips and techniques

Several practices are helpful for day to day monitoring of the PS Series environment with SAN HQ. The following tips are for reference.

## A.1   Creating SAN HQ archives

Over time, SAN HQ automatically compresses the time for each interval to save log space. Archives are one way to capture the more granular interval periods. For instance creating an archive each month to capture the busiest month-end processing allows the data to be reviewed offline using the un-compressed interval. This may also be automated through a command (SANHQClient.exe) to provide an archive that can be scheduled through Windows Task Scheduler. See the *Dell EqualLogic SAN Headquarters Installation and User's Guide* for more information.

## A.2   Enable SAN HQ SupportAssist

SupportAssist allows Dell support personnel to provide intelligent, proactive support for the PS Series storage. In addition, SAN HQ archives may be loaded directly to SupportAssist for further analysis when needed.

## A.3   Create reports to review performance, capacity and configuration trends.

SAN HQ has many reports available to show the average performance, capacity, alerts, configuration and advanced software status such as replication. These reports may be run manually through the GUI or using the command-line interface (SANHQClient.exe) and scheduled through Windows Task Scheduler. See the *Dell EqualLogic SAN Headquarters Installation and User's Guide* for more information.

# B Additional resources

Dell.com/support is focused on meeting your needs with proven services and support.

Referenced or recommended Dell publications:

- TR1092 EqualLogic Hybrid Array Load Balancer:
- Dell EqualLogic SAN Headquarters Installation and User's Guide:
- Dell TechCenter for PS Series storage
- Dell EqualLogic Configuration Guide: