



OPENBENCH LABS

Storage Service Management

Analysis: **Manage SAN Storage as a Service** Via Path-Aware Performance Data



Analysis: Manage SAN Storage as a Service Via Path-Aware Performance Data

Author: Jack Fegreus, Ph.D.
Chief Technology Officer
openBench Labs
<http://www.openBench.com>
September 27, 2007

Jack Fegreus is Chief Technology Officer at openBench Labs, which consults with a number of independent publications. He currently serves as CTO of Strategic Communications, Editorial Director of Open magazine and contributes to InfoStor and Virtualization Strategy. He has served as Editor in Chief of Data Storage, BackOffice CTO, Client/Server Today, and Digital Review. Previously Jack served as a consultant to Demax Software and was IT Director at Riley Stoker Corp. Jack holds a Ph.D. in Mathematics and worked on the application of computers to symbolic logic.

Table of Contents

Executive Summary	04
Assessment Scenario	06
Storage Service Value Proposition	13

Executive Summary

“Using NetApp’s SANscreen Application Insight, SAN administrators can look beyond device-centric traffic metrics and correlate overall SAN traffic with quality of service issues.”

THE PATH FROM ASSET TO SERVICE

When it is time to act on an emerging opportunity or react to a competitive threat, today’s real-time enterprise depends on getting the right information, to the right people, in time to act and create value or mitigate risk. That goal has set off an explosion in the volume of data maintained on storage area networks (SANs).

openBench Labs Test Briefing: NetApp’s SANscreen® Application Insight 2.0

- 1) **Automate Discovery and Collection of Real-Time Data:** Application Insight discovers and collects SAN load metrics without agents across all switch ports, which allows storage administrators to rapidly identify and respond to critical performance issues which impacts the ability for IT to deliver storage services.
- 2) **Improve Service Quality:** Via SANscreen Service Insight, Application Insight leverages the service information discovered, and combines it with real time performance information to align SANs for maximum application performance at the lowest possible cost. Application Insight provides the ability to flag service compliance failures such as lack of redundancy and provides root-cause analysis of host-port congestion and a view into traffic balance across ports.
- 3) **Manage SAN Storage Utilization and Capacity:** Application Insight discovers traffic loads by host, array and switch port, which allows storage administrators to resolve bottlenecks, optimize storage tiers, uncover orphaned resources, and maximize SAN ROI.

To deal with that expanding volume of data, IT has focused on managing storage devices as individual assets. At best, this has meant investments in storage resource management (SRM) software solutions. More often, however, storage administrators simply rely on a combination of out-of-date spreadsheets, SRM data and data generated from home-grown scripts to represent the current

configuration state of storage resources.

The problem for IT is that no matter how well storage is managed as an isolated asset, only a limited value can be derived from the device. Asset properties and the state of storage devices are not sufficient to address the challenges to delivering IT support as a business service. In fact, most IT organizations lack a clear set of links that tie storage resources, with applications, and business value and that creates a gulf between IT and corporate executives.

Corporate executives think in terms of business processes. When it comes to the services that they need to support those processes, they expect IT to address issues of availability, business continuity, performance, and security. This requires IT to automate datacenter processes for building, maintaining, optimizing and auditing storage networks, That means IT must be able to create

policies and procedures that can effectively support a service level agreement (SLA) for storage-IT's most costly asset.

To define and support an SLA for a business process, IT must understand all of the interdependencies among storage devices, hosts, and SAN switches for each application that's a part of the process. The lack of such an overall understanding will negatively impact IT's ability to deliver processes for building, maintaining, optimizing and auditing storage networks. Unfortunately, the task of developing a service-centric process all too often burdens IT with costly labor-intensive tasks that require detailed application, datacenter, and business process knowledge. SANscreen® Application Insight, however, offers IT a simpler solution to optimizing storage as a service.

Application Insight uniquely empowers IT to align a SAN for maximum application performance at the lowest possible cost, by providing an automated solution for discovery, service monitoring and capacity management analysis. A key competitive differentiator is the construct of an access path.

An access path consists of physically connected resources, which represents a relationship between a particular application on a server, and its data on a storage device. This construct of a path is at the heart of SANscreen Service Insight, which is prerequisite for installing Application Insight. Using a service intelligence engine and a SQL-based configuration management database, Service Insight automatically discovers all of the interrelationships among physical resources on a SAN in order to provide IT with the critical end-to-end access paths needed to support business processes via an IT storage service.

Application Insight extends the device attributes associated with an access path with real-time SAN traffic data. In particular, Application Insight maps the traffic data measured by SAN switch ports to hosts, ISLs, and arrays. Application Insight's unique ability to understand an application's path to storage combined with device load information results in unparalleled visibility into the cost and efficiency of delivering storage services to an application. The result is lower capital costs through improved resource utilization, improved traffic balance, identification of orphaned resources, validation of tiered-storage strategies, and improved application performance. Using NetApp's SANscreen Application Insight, storage administrators can look beyond device-centric traffic metrics and correlate overall SAN traffic with quality of service issues.

The ability to group, aggregate, and visualize a massive amount of SAN traffic data—by default, up to 10,000 FC switch ports are scanned every 10 seconds—allows SANscreen Application Insight to be thought of as an online analytical processing (OLAP) tool for predictive analysis. By slicing and dicing traffic data from varying path perspectives, administrators quickly correlate traffic problems, such as host congestion and multipath availability, with application performance, and service level policies. This enables rapid resolution of

performance bottlenecks and proactive optimization of host, array, switch, and fabric traffic distribution and utilization. More importantly, storage administrators can ensure that business applications receive the right level of performance from networked storage resources.

IT SERVICE MANAGEMENT

SANscreen Service Insight is the baseline product for NetApp SANscreen suite. The SANscreen suite products leverage and build on Service Insight to automate datacenter processes for building, maintaining, optimizing and auditing storage networks.

For IT to provide a storage service, the first step is to establish visibility into the end-to-end SAN infrastructure. SANscreen Service Insight allows IT teams to gain visibility into an accurate repository of the global storage infrastructure, including host-to-storage access paths, storage arrays, switch devices, tapes and hosts, as well as configuration changes. Service Insight updates this information in real time, reporting back the current status of all SAN objects (e.g., zoning, mapping, data paths, storage availability, port use, port status, etc.).

Assessment Scenario

“Application Insight is more than a monitoring tool to resolve known issues: Building on the service intelligence engine of SANscreen Service Insight, real-time traffic data can be used to discover the root causes of problems before users are impacted.”

As SANs continue to grow, SAN managers face increasing pressure to maintain accurate SAN documentation in complex dynamic environments, within strict time and resource constraints. Service Insight collects and stores SAN infrastructure data—including every change made in the storage environment—in a SQL database. This provides the historic data required for forensic analyses, audits and the automation of storage-related datacenter processes.

When the data is extended to include the collection and analysis of real-time SAN traffic data via Application Insight, storage administrators have both the transactional data and the critical analysis perspectives needed to align SAN devices within a service context. Through SANscreen Application Insight, IT acquires the information and tools needed to support a business-process SLA. Using SANscreen Application Insight, IT can deliver a particular level of application performance, prevent process disruption, and minimize the cost of service. What's more, IT can ring maximal benefits out of a working SAN, and extend the time horizon for capital investments to leverage Shugart's Law—the price per bit of storage halves every 18 months.

To enable IT to take on the role of a storage-service provider, Application Insight collects performance and error metrics every 10 seconds from up to 10,000 SAN switch ports. Using lossless data compression, all of that traffic data is stored in SANscreen's database for seven days. After seven days, the data is annualized and stored for viewing in summary histograms.

At any time, a storage administrator can group and analyze that traffic data over customizable time periods to conduct the necessary analysis to resolve a critical performance bottleneck, proactively increase device utilization, discover traffic trends, identify orphaned hosts, validate strategies for storage tiers or plan a change in the storage infrastructure. At all times, the goal is to reduce capital expenditures, save time and reduce labor costs. To that end, Application Insight adds three new options to the SANscreen Client GUI: Load Analyzer, Hosts, and Monitor.

Application Insight is, however, much more than a monitoring tool to resolve known issues: Building on the service intelligence engine of SANscreen Service Insight, real-time traffic data can be used to discover the root causes of problems before users are impacted. For that reason, openBench Labs set out to collect and analyze fabric data from a proactive storage-service management perspective, rather than reactive asset maintenance perspective. To that end, we needed to resolve three important questions:

- 1 Could we uncover hidden issues that could lead to potentially severe application availability problems?
- 2 Could we quickly and intuitively resolve those problems?
- 3 Could we pinpoint potential capital savings based on the levels of resource utilization?

To test that scenario, openBench Labs installed the two software components of SANscreen Service Insight along with the Application Insight 2.0 module on a dual-processor Intel® Xeon-based server running Windows Server® 2003 SP2.

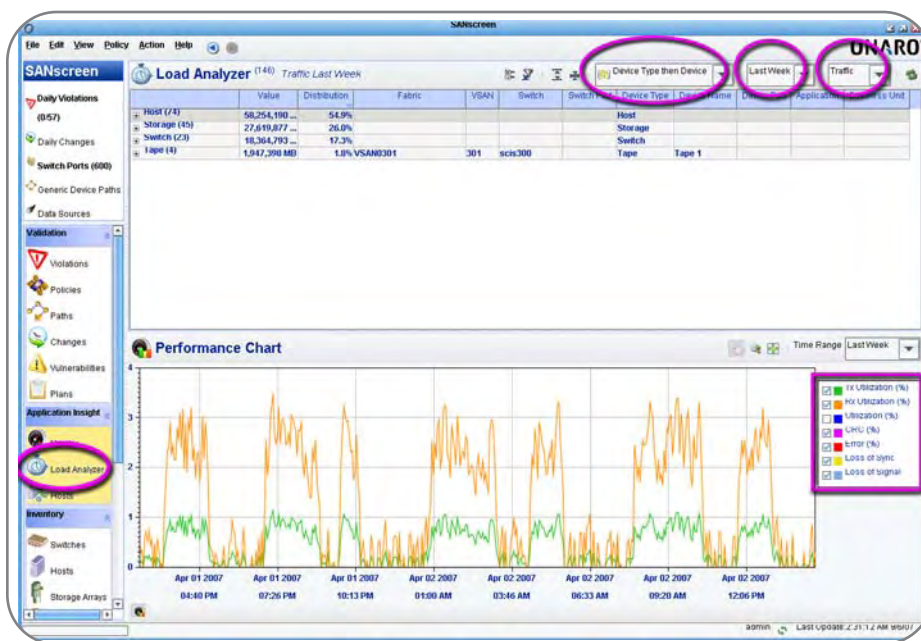
CAPACITY MANAGEMENT

Through the Load Analyzer, a storage administrator can perform sophisticated capacity management functions by slicing, dicing and visualizing all of the SAN traffic data collected from each switch port. That traffic can be aggregated and grouped by fabric then switch, device type then device, business unit then application then device, application then device, switch then device, business unit and project. Since data travels along a path, it has also direction—received or transmitted—so all aggregations can be applied to just received or transmitted traffic, as well as total traffic. Traffic is also a function of time so all of this analysis can be done over standard intervals—last hour, last day, last week—or a fully customizable time horizon.

Application Insight represents the volume of that traffic in terms of a raw

count, a distribution percentage based on the devices attached to a switch port, or a utilization percentage based on the physical properties of a specific path. Traffic is related to either application data or errors. Error-related data are further broken down by cause: loss of synchronization, loss of signal, CRC and general error rate. As with data, the volume of error is represented by either a raw count or an error rate percentage.

To simplify the visualization of traffic patterns within a single, performance chart, all data counts are presented on a percentage basis over a particular time horizon. Application traffic is represented as a percentage of path utilization; total errors are represented as a percent of total traffic; and each type of error is displayed as a percentage of the number of overall errors. This allows a storage administrator to identify any correlation between errors and traffic utilization. A quick resolution of any such correlation will then leave more bandwidth available for application data traffic.

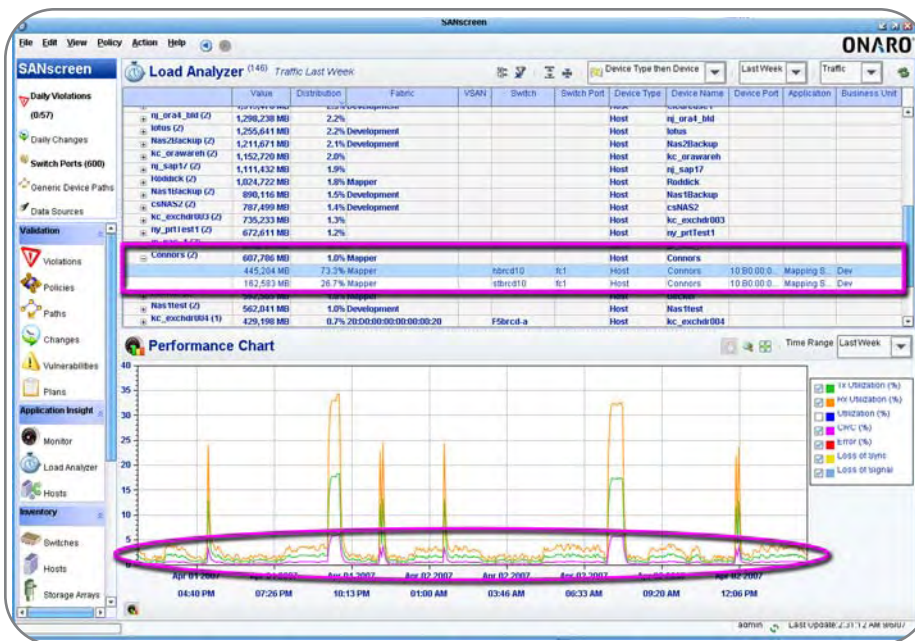


Moreover, host-to-storage paths are not the only SAN paths for which traffic balance is critically important. To create an extended SAN fabric with multiple switches, it is necessary to connect the switches via high-speed inter-switch links (ISLs), which are often created by bundling

We opened the Load Analyzer display, with a week's data grouped by device type, and charted data utilization and error rates. At this level of abstraction, roughly 55% of all the data was distributed to host devices and the breakdown of traffic- reads versus writes-was fully consistent with classic heuristics as reads made up 75% of traffic.

several standard FC switch ports. That configuration can be costly in terms of the allocation of switch ports and becomes very costly when the flow of data is not balanced, utilization is low and the ISL operates inefficiently.

A degraded ISL will negatively impact all traffic routed through that link. On a well-balanced SAN fabric, traffic will be dynamically routed over changing path combinations. As a result, serious application performance issues may be masked as random events to a storage administrator who lacks the tools needed to assess performance in an end-to-end service context. Worse yet, that impact of a degraded ISL will cascade down onto any Service Level Agreement (SLA) associated with an application that might access data through that link.



In addition to resolving load balance issues, the Load Analyzer SAN traffic distribution and utilization statistics make it easy for a storage administrator to identify both busy, under utilized resources and orphaned resources. Once identified, applications utilizing these resources can be

When openBench Labs drilled down on specific hosts, we readily identified the server Connors, which was lightly loaded with only 1% of the host traffic distribution, had a 73%/27% traffic load imbalance over paths. Looking at the performance of the heavily used path, we found a data framing problem-forcing data to be retransmitted and the path utilization rate to spike at around 30%.

analyzed as candidates for migration to higher-or lower-performing equipment or for consolidation, especially to virtual machines. As a result, IT can delay and perhaps avoid capital expenditures through increased host and array utilization, rigorously planned storage tiers, and the reclamation of unused devices.

SERVICE-ORIENTED TRAFFIC MONITORING

With regards to the path topology, the distribution of SAN traffic is closely related to the notion of risk. If a large number of storage array volumes were to be zoned, mapped, and masked through the same switch port, a failure at that port would cascade down to each volume. Making matters worse, hosts that use multi-path software for redundancy and load balancing are often configured incorrectly, which increases the risk of failure.

Validating multi-path configurations manually, however, is a very labor intensive, expensive, and frequently ineffective task. Nonetheless, when IT is faced with such tasks as adding applications or consolidating servers, issues such as availability and business continuity are very important. This makes the question, "How busy is a host?" a critically important issue, even for storage administrators.

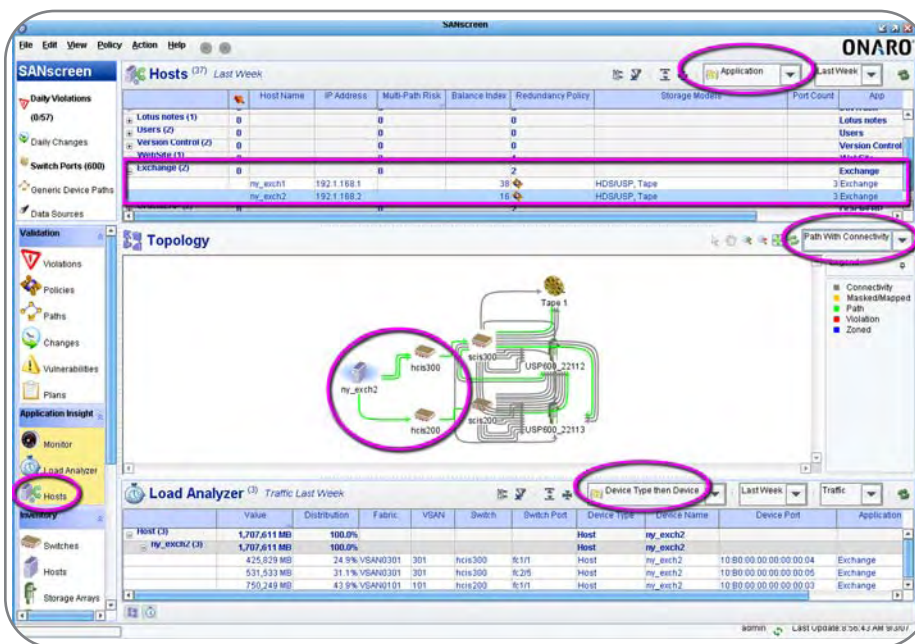
As it's name suggests, the Hosts display provides a more host-centric perspective on traffic data. Data associated with multipathing risk, load balancing, and resolving congestion can be initially grouped and aggregated by application, business unit, or operating system.

Within the Hosts display, Application Insight adds new data that includes a Port Count, which is the number of switch ports through which a host sends and

receives data. There are also structural data entries that contain useful SLA information, such as the name of application running on a server, the priority of that application, and whether the server has a redundancy policy in place. More importantly, the Hosts display includes a calculated Balance Index, which is an important measure of risk.

Application Insight calculates the standard deviation from the average traffic over all of a host's redundant paths using data collected from the switch ports. If redundant data paths on a host all have equal distribution, then the paths are said to be in balance and those paths will have a balance index of 0. A high Balance Index is a good indication that a problem is causing the server traffic to be out of balance across switch ports.

Within the Host display, a Load Analyzer detail view with traffic data can be opened. As a component of the Hosts display, this Load Analyzer detail view is specific to switch ports that are directly connected to the host's HBAs. Like the main Load Analyzer display, traffic data associated with the selected host can be grouped by fabric then switch, device type then device, business unit then application then device, application then device, switch then device, business unit and project.



A detailed Topology visualization can also be opened within the Hosts display. This Topology detail view uses icons and graphical elements to visualize the SAN as seen from the host. All physical connections appear as lines connecting the icons, which represent switches

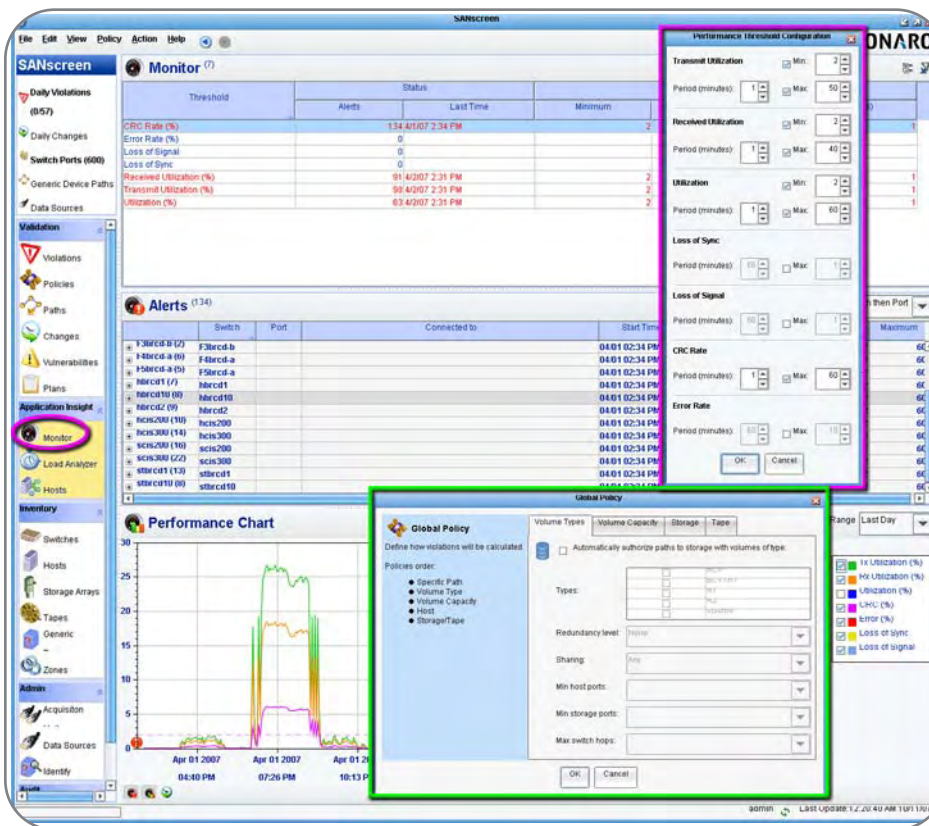
We opened the Hosts display, with top-line data grouped by application. Selecting MS Exchange as our application, we easily identified two servers running that application. Narrowing the focus to a specific server, we discovered a well-balanced traffic flow across the three switch ports to which the server was attached. Via the Topology display, we could also drill down on the two Hitachi Data Systems arrays and the tape library to which there were data paths connecting this server.

and storage devices.

For even more detailed analysis, physical connections can be further categorized by the way in which these were logically defined. Color keys identify connections as zoned, masked, mapped, simple connectivity, or a violation. By selecting a switch in the Topology display, the Load Analyzer Window in the

Host display will provide a traffic distribution analysis for all hosts connected to that switch. This enables a storage administrator to determine all of the hosts sharing a switch.

Finally, the third display introduced by SANscreen Application Insight is that of an alarm Monitor. A storage administrator uses the Monitor to set and edit error conditions along with upper/lower bounds on traffic utilization. These alarm thresholds for performance can be related to total port utilization or broken out by transmit utilization and received utilization. Alarms can also be set on the number of errors, such as signal or synchronization loss, CRC, and the rate of errors.



Whenever any such threshold is violated during a specified sampling time, an alert is automatically triggered. The alert will provide granular information concerning switch ports and the devices attached to the switch ports. In addition to being stored in the database, alerts can be emailed or sent to another systems management framework application via SNMP traps.

Within the Monitor display, openBench Labs found time-consuming tasks, such as setting alerts and creating SAN policies, were greatly simplified through the presentation of easy-to-use menu screens.

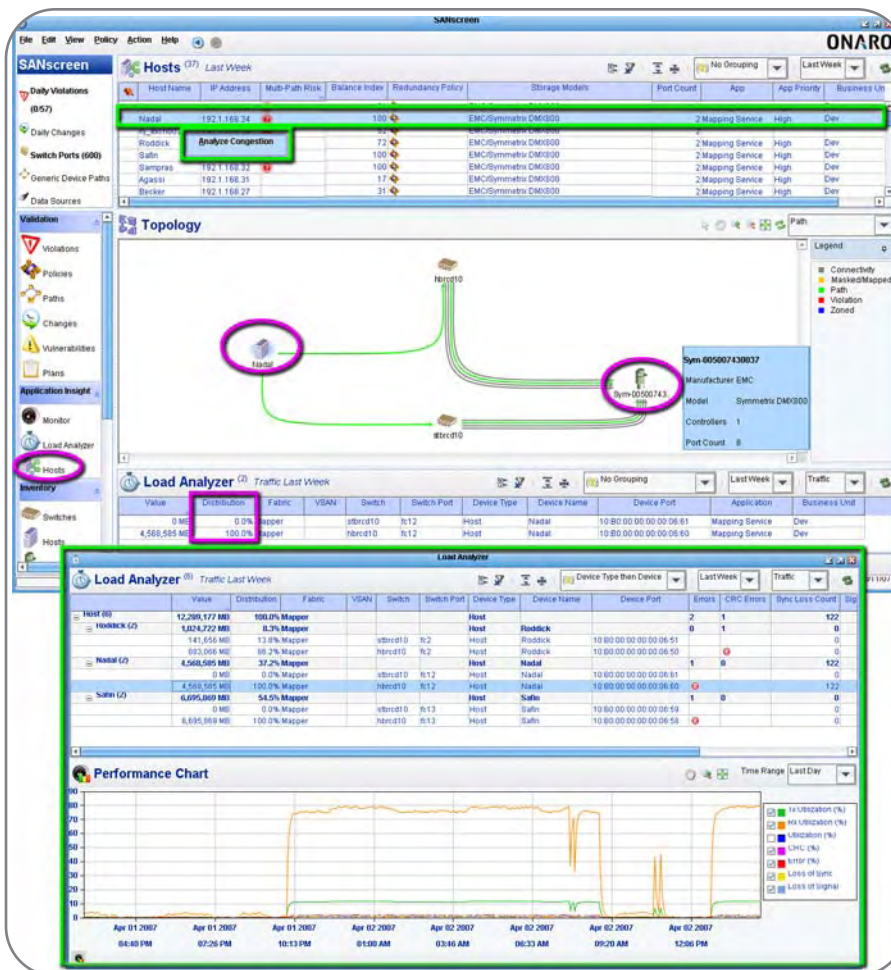
framework application via SNMP traps.

CUTTING THROUGH CHAOS AND CONGESTION

The ability to automatically discover changes in the SAN fabric and analyze those changes with respect to the flow of data over identifiable access paths from storage devices to switches to servers makes Application Insight an important tool for any storage administrator. What distinguishes Application Insight is that it is not just a tool to help resolve known issues and problems. The promise of Application Insight is that it enables IT to identify bottlenecks and empower IT to re-route data flows proactively to under-utilized devices and ports before end

users are stymied by glacial I/O throughput, error conditions or path failures leading to application brownouts or worse yet failures.

Among the key elements of any SAN fabric are the Fibre Channel ports of storage arrays. Through these ports, volumes are "exposed" to hosts. In theory, these storage ports should be some of the most carefully planned and balanced in a fabric. In practice, these ports more often than not must support the aggregate traffic of a random collection of logical volumes.



With the volume of storage needed for new and existing applications and regulatory requirements for data availability and security all growing exponentially, the number of logical storage devices that must be exposed through a fixed number of storage ports makes the manual tracking of these mappings a Herculean effort. Even more daunting is the task of balancing traffic for these devices. Via the integration of Application Insight with Service Insight, an administrator can locate congested ports, analyze the

Through the Host display, we began an analysis of the host Nadal, which was running the Mapping Service application.. We quickly identified a multi-path risk condition, which is flagged by a Balance Index of 100: Nadal's traffic is going through only 1 HBA, within an active/active environment. This is reinforced in the distribution statistics of the Load Analyzer Detail View.

By selecting Analyze Congestion on the host Nadal, we were able to see a bigger picture based on the end-to-end path configuration data discerned by SANscreen Service Insight. Through Analyze Congestion, we see that, Nadal is "competing" for the same storage port resources as servers Roddick and Safin. This could degrade the performance of the Mapping Service application and create an SLA issue.

cause of the congestion, and identify the SAN devices that are most likely the cause of the congestion. While a number of factors can cause congestion at a storage port, three issues are at the heart of most congestion problems:

- Too many hosts mapped to logical volumes exported through a port;
- A busy high-traffic application monopolizing the port;
- A malfunctioning switch port.

What makes this root-cause analysis an easy task to accomplish is SANscreen's construct of a data path, which is leveraged by the service intelligence engine and the database found in SANscreen Service Insight. Service Insight automatically discovers all of the interrelationships among physical resources on a SAN: Application Insight extends all the device attributes associated with an access path with real-time SAN traffic data.

STORAGE SERVICE VALUE PROPOSITION

“To bridge device-centric technology silos and deliver an actionable end-to-end storage service context, SANscreen Application Insight was built on the notion of an end-to-end data path that connects an application on a host to a data volume in a storage array.”

SANSCREEN APPLICATION INSIGHT QUICK ROI

- 1) **End-to-End Access-Path Management:** To support storage service management, the SANscreen suite of products provides IT visibility of the global storage infrastructure, including host-to-storage access paths, storage arrays, switch devices, tapes and hosts, as well as configuration changes.
- 2) **Proactive Management of SLA Issues:** For rapid resolution of QoS issues, SANscreen Service Insight and Application Insight automatically collect, organize and present both Access Path and SAN Load data in an application-centric context to isolate and diagnose problems.
- 3) **Service-centric Resource Optimization:** Resources that are imbalanced, underutilized, saturated, orphaned or do not comply with tiered-storage strategies can be readily identified. This enables IT to either defer capital expenditures, or remove latent quality risks, which impacts application availability.

"We see computers everywhere except in the productivity statistics." In eight words, Nobel Laureate economist, Robert Solow, summed up a problem confronting many enterprises. When CEOs ask, "What are our returns on IT investments," too often the response is a deafening silence.

Those CEOs have discovered what the Center for eBusiness at MIT's

Sloan School of Management dubs "The Productivity Paradox of Information Technology." To resolve this issue, top business management consultants are telling CEOs that IT must evolve organizationally from being a technology provider to being a service provider. In that process, the general consensus is that IT procedures must be made more rigorous, formal, and-perhaps most importantly-automated. Simply put, the notion of doing IT efficiently is now being augmented with the notion of doing IT correctly.

At the frontline of any change in IT is storage management. Fear of regulatory compliance violations, especially with respect to Sarbanes-Oxley, the Health Insurance Portability and Accountability Act, and the Federal Rules of Civil Procedure, which sets the discovery rules for email, has led IT to adopt defensive

policies to store and save all information by default. That strategy has sent the volume of managed stored data skyrocketing.

Fortunately, storage technology improvements continue to shrink the footprint, increase the performance, and mitigate the impact of disk drives on capital costs, which have risen much more modestly than the growth in stored data. Operationally, however, there are no mitigating factors that will passively lessen the cost of storage management. For storage operations, greater storage volume equates to greater complexity and IT's highest rising labor costs.

The general rule of thumb is that operating costs for managing storage on a per-gigabyte basis are three to ten times greater than the capital costs. That's because provisioning and management tasks associated with storage resources are highly labor-intensive and often burdened by the bureaucratic inefficiencies. For IT, the bottom line for controlling storage costs comes down to minimizing the time to manage while maximizing resource utilization. Full realization of that bottom line, however, has fallen victim to the rapid growth in storage experienced at datacenters over the past several years.

As IT decision makers scrambled to maintain the volume of storage required to support business initiatives, they often added resources from multiple vendors based primarily on the lowest price. At the same time, storage vendors seeking to build brand equity in a price-oriented market tied advanced management features to software that runs explicitly on their SAN devices, which effectively partitioned SAN management into device-centric technology silos.

That left IT focused on managing storage assets with a combination of SRM software, scripts and/or out-of-date spreadsheets. Meanwhile, the corporate executives, who are the clients of IT, think in terms of application-centric business processes. When it comes to the services they need to support a business process, they expect IT to address issues of availability, business continuity, performance, and security. Information about the properties of storage devices is not sufficient to support an SLA for a business process.

To bridge device-centric technology silos and deliver an actionable end-to-end storage service, SANscreen Application Insight was built on the notion of an end-to-end data path that connects an application on a host to a data volume in a storage array. At the heart of the SANscreen solution, an SQL-based database serves as the repository for both SAN load data and the interrelationships among physical resources, virtual resources, and service level policies. A core service intelligence engine automatically associates and maps all of this data to service-centric end-to-end data paths to assess the impact on the overall IT infrastructure.

Using SANscreen Service Insight with Application Insight IT successfully extends datacenter automation to storage with its real time, heterogeneous, service level views of the storage environment. Organizations using SANscreen manage

storage as a true end-to-end IT service. SANscreen's actionable service level information is consumable by both storage teams and non-technical storage users, effectively integrating storage into the entire IT service delivery chain. IT can provide a rigorously managed storage service that helps ensure that valued information is readily available, secure, and delivered at the lowest possible cost.