

» Tech OnTap Home

October 2006

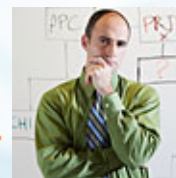
HIGHLIGHTS

- » **SAN Boot for VMware ESX 3.0.0**
- » **Configuration Tips for Swatch/Kiwi**
- » **Optimizing Oracle® 10g ASM**
- » **Visit NetApp at Oracle OpenWorld**

Should You Be an Early Adopter?

Two simple questions to determine whether you should go bleeding edge or join the quill pen crowd.

[More »](#)



BLOGGING WITH DAVE

Dave Hitz, NetApp Founder and EVP

"I just came across an internal e-mail that I sent to all of NetApp in 1996...these goals laid the foundation for what we have become."

» [Dave's Blog](#)

DRILL DOWN

- » **SnapVault® for NetBackup™ Live Chat**
Quiz product experts about disk-to-disk backup with Symantec® NetBackup 6.0.
- » **SnapVault Best Practices Guide (pdf)**
A step-by-step guide to designing intelligent backup solutions with minimal space consumption.
- » **NetApp Continues iSCSI Leadership**
NetApp storage enabled for 10GbE.
- » **Are You an Innovator?**
Win a NetApp Innovation Award for leadership or community contributions.



ADMIN RESOURCES

- » **Featured Tools: Swatch/Kiwi Syslog**
An inside look at how NetApp Global Services pros use simple monitoring tools. Includes syslog.conf/configuration edits.

TIPS FROM THE TRENCHES

Meeting the Four "Must Have" Storage Requirements of the Enterprise Grid

Maury Tiller, database storage architect, NetApp

For the grid to work, it needs uniform, reliable multiprotocol storage that understands the applications it supports. Virtualization, RAID technologies, Oracle's HARD initiative, and NetApp help address the requirements of the grid.

[More »](#)

NetApp SE Lab Report: SAN Boot with VMware ESX 3.0.0

Nick Triantos, global SAN SE, NetApp

One user's experiences upgrading to ESX 3.0.0, plus three suggestions for booting VMware servers from SAN:

- Enable the BIOS on only 1 HBA
- Modify the Execution Throttle/Queue Depth
- Modify the PortDownRetryCount parameter

[More »](#)

ENGINEERING TALK

How SnapManager® for Oracle Empowers Database Administrators

Alvin Richards, database and grid architect, NetApp

NetApp SnapManager for Oracle (SMO) software is tightly integrated with Oracle tools such as RMAN, RAC, and ASM.

Get the details on how SMO decreases capacity requirements for Oracle backups and clones, accelerates backup and restore, and optimizes ASM deployments.

[More »](#)

**VISIT NETAPP AT
ORACLE OPENWORLD**

Should You Be on the Cutting Edge?

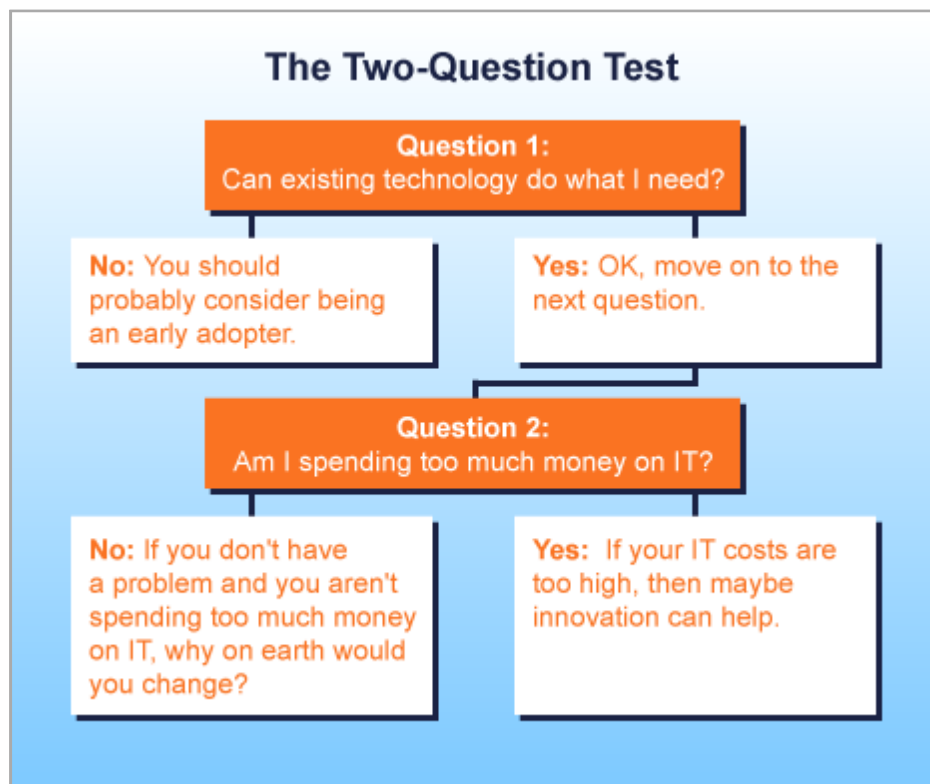
TechTalk Podcast: Hear NetApp founder Dave Hitz and Gartner analyst Bob Passmore discuss when it makes sense to be an early adopter.

At the right is the classic "technology adoption curve." The far left represents people on the bleeding edge, while the far right represents the "quill pen" crowd.



Startups typically tell customers that the correct place to be is on the very left side of the curve, a super-early adopter. However, more mature vendors recognize that being the first is not always the right approach. There are many good reasons for not wanting to change and for being a late adopter: change costs money, it requires new expertise, it introduces risk, and it is stressful.

The good news is there's an easy way to determine whether your organization or product should be an early or late technology adopter: the Two-Question Test.



Many companies become early adopters because they run into problems that exceed the limits of existing, mainstream technologies. Rendering a hundred thousand Orcs on a battlefield, for example, can be a challenge. Similarly, when the SEC changed the rules on how data must be protected, many organizations discovered that their existing technology offered inadequate protection, turning them into early adopters of disk-based compliance solutions.

The other logical reason to become an early adopter is because your IT costs are too high. Keep in mind, however, that usually the more important question is actually what percentage of total costs is IT related. They may seem high but at the same time represent a small portion of the overall budget.

RELATED INFORMATION

- [Examining the Future](#)
- [Data Center Proven Webcast Series](#)
- [Being an Early Adopter \(podcast\)](#)

Data Center Proven Webcast Series

In this three-part TechTalk series, listen to customers, analysts, and NetApp experts discuss best practices in data center design and management using scalable, high-performing, and cost-effective NetApp storage systems.

Best Practices from Industry Experts

- [Part I: Building an Effective Data Center](#)
- [Part II: Effective Data/Storage Manageability](#) (includes transcripts!)
- [Part III: Effective Business Application Management](#) (includes transcripts!)

The Future of Storage



Like any technology, storage adoption generally moves in waves. New technologies emerge in response to the limitations and challenges of

existing technology, while other technologies mature and move into decline.

In June, NetApp founder and EVP Dave Hitz outlined his predictions for the next 20+ years of storage technology evolution.

Don't miss [Examining the Future](#).

Questions Every IT Organization Should Tackle

- Example 1: Suppose you are an oil company and you just spent \$5 billion on a refinery. And now you are trying to figure out whether to spend \$1 million or \$10 million on IT to support it. Without knowing anything else about the problem, the answer is clear. Spend the \$10 million. Just don't ever let that refinery go down.
- Example 2: In contrast, think about the Oracle® On Demand data center. What percentage of the cost of delivering that service is IT related? Maybe 100%? So it makes sense that Oracle would be an early adopter and have the world's biggest virtual mainframe.

But—if you don't have a problem and you aren't spending too much money on IT, why on earth would you change?

Once you've determined where you are on the technology adoption curve, we invite you to use the NetApp vision paper "[Exploring the Future of Data Management](#)" as a starting point for discussions with other members of your IT organization about your long-term approach to storage and data management.

Comment on this article

- What is the total volume of your organization's data storage today?
- At what rate is your organization's data growing?
- How much of your IT budget is developed to human processes?
- As what rate is the human component of your budget growing year over year?
- What is the total cost of downtime for your organization?
- What does your data growth curve look like? Is it steady or highly volatile?



MAURY TILLER

Database storage architect, Network Appliance

For the past six years, Maury Tiller has been responsible for the successful implementation of NetApp storage solutions in database and application-centric environments. Over that time, he has worked with dozens of IT teams to design and implement high-performance database storage environments that reduce planned and unplanned downtime, increase availability, and simplify backup and recovery. Prior to joining NetApp, Maury spent eight years as a database consultant/architect for various database vendors.

Meeting the Four "Must Have" Storage Requirements of the Enterprise Grid

By Maury Tiller

Oracle OpenWorld blasts off October 22nd in San Francisco. A major theme of the show will be how the grid has the potential to completely revolutionize IT. The grid vision, of course, is to enable companies to treat their compute resources as a utility rather than an asset.

As a former database consultant, I find this shift exciting, because it allows database systems to dynamically allocate CPUs and servers (nodes) to an application on demand. The workload associated with most database environments is extremely sporadic. A system can run almost idle for the majority of the time and at any given point (such as end of the month reporting) experience extreme workload. In the grid, compute resources can be sent to the application when the application needs it. When the application no longer needs the compute resources, they get reallocated elsewhere.

Not surprisingly, the role of storage changes in a grid environment. These infrastructures have specialized requirements in four key areas:

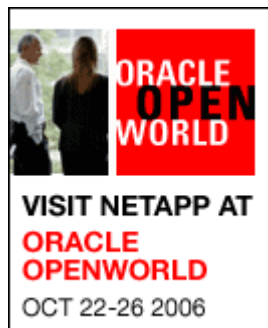
- Grid storage must appear uniform
- Grid storage must provide flexible connectivity
- Grid storage must be unbreakable
- Grid storage must be tightly integrated with application environments

In this article I'll examine each of these requirements and share examples from my experiences with NetApp to illustrate how these requirements can be met.

Requirement #1: Grid Storage Must Appear Uniform

In a grid, all servers are viewed as nodes and the management of those nodes is identical regardless of individual hardware. Management of one group of servers should be the same as managing another, even when the server groups have differing configurations.

The same applies to grid storage. Every individual storage system must appear the same to all nodes. Storage allocated on high-speed Fibre Channel disks should look the same to grid server nodes as storage allocated on more economical SATA disk drives (except of course from a performance standpoint). Furthermore, each individual storage system should have the same set of features and be manageable from a single interface in a central location.



RELATED INFORMATION

- [Fueling the Enterprise Grid](#)
- [Improving Database Agility](#)
- [Oracle IT's Implementation of Grid](#)
- [Architecting Storage for Resiliency](#)

Fueling the Enterprise Grid

A variety of trends are impacting the move toward a fully realized application virtualization model, including:

- Increasing pressure for true 24x7 operations and nondisruptive upgrades
- Emerging global namespace solutions and products supporting file-level migration in tiered storage environments
- Accelerating adoption of diskless boot architectures

Learn more. Read [Pursuing the Vision of Enterprise IT](#).

Increasing Database Agility for Test/Dev and QA

Replicating a large database for development, training, testing, or other purposes can be one of the most time-consuming tasks that a DBA can undertake. You have to carefully plan your methodology, provision enough storage to accommodate the copy, and then create a consistent replica of the data.

Learn how the use of NetApp technologies, including NetApp FlexClone and SnapMirror®, simplifies

There are two ways to achieve this. You can either deploy only storage systems that have the same management interface and feature set (which pretty much means storage from a single vendor), or you can use storage virtualization to accommodate the underlying management and feature differences between different storage systems and present a uniform interface to grid nodes.

The first approach (uniform storage) is the most straightforward for new grid deployments in which you aren't worried about leveraging existing storage assets. Even then, however, most companies want the option of utilizing different tiers of storage for different purposes to maximize economy. With many storage vendors this poses problems, since different classes of storage have different interfaces and feature sets. NetApp is the only major storage vendor to have solved this problem by providing a uniform management interface and feature set across its entry-level systems, mid-tier platforms, and high-end storage systems. For instance, a database running on a low-end NetApp FAS270 can utilize the same feature set as a database running on a high-performance FAS6070, and both systems can use either SATA drives, FC drives, or mixed SATA and FC drives.

A second approach to this requirement is to use storage virtualization to hide the complexity of underlying storage systems, thereby presenting a common interface to grid nodes and a single management interface for storage administrators. There is a mountain of literature describing the best ways to do this, and there are a variety of products on the market that can be used to achieve varying degrees of virtualization for heterogeneous storage. The NetApp solution involves V-Series storage appliances that virtualize storage arrays from a variety of leading storage vendors and provides access via NAS, SAN, or IP SAN while delivering the full NetApp feature set.

Regardless of the approach you choose, make sure your storage vendor offers a uniform management interface that can be used to efficiently provision, monitor, manage and protect data across entire your entire grid storage environment. To see an example of how NetApp customers can manage NetApp and virtualized storage environments, check out Tech OnTap's [guided tour of NetApp Operations Manager](#).

Requirement #2: Grid Storage Must Provide Flexible Connectivity

Grid infrastructure is dynamic and fast moving. Storage has to be provisioned quickly and presented to different nodes, running different applications potentially on different operating systems. Some applications may require shared CIFS or NFS access while other applications need block-based access via iSCSI or Fibre Channel. This creates a significant challenge for grid storage systems and an additional concern for storage administrators.

Most vendors can support at least two of the four major storage protocols. Network Appliance is the only major storage vendor that provides native support for CIFS, NFS, iSCSI, and Fibre Channel protocols on its storage systems without the need for a front-end gateway / router (which typically adds latency and additional management burdens).

During a recent customer conversation I realized that—in addition to enabling support for multiple applications and requirements—flexible connectivity can also simplify the move to a grid architecture.

The customer I was talking with currently has a large UNIX® server for his Oracle® database and plans to eventually move to a set of Linux® nodes in an Oracle RAC cluster. Since the UNIX server is currently generating all database I/O, Fibre Channel connectivity is a must. However, after the RAC migration, I/O will be spread across the individual nodes all working together. The I/O requirement for each individual node will then drop to the point that it can be satisfied via NFS, which is cheaper and easier to manage. Since a single NetApp storage system can support both types of connectivity, the customer did not have to purchase an additional storage system just to accommodate the new connectivity model. And, as the customer's grid grows, he will have the flexibility to adapt to new connectivity requirements without making wholesale storage infrastructure changes.

Requirement #3: Grid Storage Must Be Unbreakable

Back when I was a database consultant, the only time I remember hearing people talk about storage was when it was broken. In a grid environment in which racks full of server nodes connect to racks of storage and capacity is dynamically allocated and de-allocated as necessary, the consequences of storage failure are even more

the creation of local and remote database replicas. This can streamline the database application development, test, and deployment process to improve business agility.

Get the details. [Read Reducing Time-to-Deployment](#).

NetApp vs. EMC and HP: IT Decision-Maker Perspectives

After interviewing storage managers and DBAs at more than 20 European and North American enterprises, Mercer Management Consulting identified key differentiators between database environments using NetApp, EMC, and HP storage.

Findings include:

- In a typical FC SAN environment, NetApp solutions cost less than EMC CLARiiON (32%), HP EVA (42%), and EMC Symmetrix (62%).
- NetApp environments require less primary disk capacity (52%) and data protection disk capacity (20%) than EMC and HP environments.
- Database storage administrators using NetApp storage can recover twice as fast from a database crash or corruption.

Get the details. [Read the full report](#).

noticeable. As a result of a single disk failure, a hapless storage administrator can quickly get the whole company's attention. Therefore, it's critical that every single storage system in the grid is unbreakable.

Potential single points of failure are not acceptable with storage systems in this category. Even conventional RAID technologies such as RAID 5 may be too risky, while RAID10 and 0+1 may be prohibitively expensive. Disk capacities have grown but resiliency (failure rate) has stayed roughly the same. Larger disks and larger storage volumes mean longer RAID reconstruction times, increasing the chance that another disk will go bad (or merely return a new media error) during reconstruction. Obviously if that happens you're looking at unplanned downtime, significant restore efforts, and potentially significant data loss.

Intelligent RAID technologies such as the NetApp RAID-DP™ (Double Parity — a SNIA RAID 6 standard) solution can eliminate the possibility that a double-disk-failure will bring down a critical database or application with minimal incremental expense and negligible performance impact. Up to two disks per RAID-DP RAID group can fail at the same time without causing any downtime or loss of data, and this is only one of several unique disk resiliency features available from NetApp. The bottom line is that RAID-DP offers better protection from a double-disk drive failure than RAID 0+1 while yielding the storage efficiency typically seen only in the less reliable and slower RAID 5 configurations.

To further increase database resiliency, Oracle has started a new initiative called Hardware Assisted Resilient Data (HARD). Oracle buried a checksum and a signature into each of its blocks. The idea is that when the page is written to the storage device, the storage device reads the checksum and signature and verifies the page's consistency, eliminating the chance of silent data corruption in the data path. NetApp is the only storage vendor to offer this protection across all primary and secondary Oracle-compatible storage platforms.

Requirement #4: Grid Storage Must Be Tightly Integrated with Application Environments

While rock solid storage is important, it's not enough. You need to maximize the availability of every application or database relying on the grid. This means that standard maintenance activities like backing up data, cloning databases, or expanding storage can't disrupt end users and—if the unthinkable happens—restores must be lightning-fast. This means that storage must be tightly integrated with each environment to enhance its individual operation.

Network Appliance storage offers a flexible set of features and functionalities that can be leveraged by any application to meet unique requirements. For instance, an Oracle grid utilizing NetApp storage can take advantage of any or all of the set of underlying NetApp features:

- FlexVol™ allows database storage volumes to be sized, resized, managed, and moved independent of the underlying physical storage. Volumes are automatically spread across available spindles so that the hot spots that frequently plague conventional storage volumes are eliminated.
- Snapshot™ technology allows a database to capture frequent point-in-time data copies that serve as checkpoints for re-starting. Because of their ubiquity and near-zero performance impact, NetApp snapshots also uniquely satisfy the "near-CDP" category of the emerging Continuous Data Protection industry. Snapshot copies can be used to protect the database data files and the application files that reside outside the database. Many customers are now using Snapshot copies to protect ORACLE_HOME and APPL_TOP from errors that can occur when installing a new patch or an upgrade to the application software.
- SnapRestore® can almost instantaneously revert a database to any saved Snapshot so that operations can be re-started if a problem occurs with the application.
- With FlexClone, a clone of a storage volume can be created instantaneously without consuming additional storage space until changes are made. For instance, an Oracle DBA can create a clone of an entire database (even one stored in an ASM disk group) in a matter of minutes and bring up that database under a new database id (ORACLE_SID) using minimal incremental storage.

For more information on managing storage for Oracle database environments, don't miss the article on [SnapManager® for Oracle](#) also in this month's issue.

backup/restore and cloning capabilities. These can include synchronous and async replication, disaster recovery, encryption, workload prioritization, etc. Ideally, it should be possible to add these services to existing grid storage incrementally as they become necessary. Grid-based applications can then utilize the underlying services that they require to achieve business objectives. It probably won't surprise you to learn that NetApp meets this requirement perfectly.

Achieving Success with the Grid

The grid may have the potential to revolutionize IT, but it can't be successful if you don't get your storage infrastructure right. While I certainly wouldn't claim that NetApp is the *only* vendor capable of providing storage for a grid environment, I do believe that NetApp storage addresses the specific requirements and demands of the grid better than any other solution on the market.

But don't take my word for it. The best proof point comes from Oracle directly.

It is not by accident that NetApp storage fits so nicely in a grid environment, but rather the result of dedicated joint development initiatives between the two companies. Oracle10g™ Grid Control, for example, has been out for almost a year and is designed to manage and monitor all aspects of the Oracle Grid. Today it can monitor storage from only one vendor: Network Appliance. This is because only NetApp delivers the storage capabilities that grids require:

- A uniform interface that spans the product line and virtualization technologies to incorporate heterogeneous storage into your grid infrastructure
- All NetApp storage supports both block- and file-oriented connectivity options including NFS, CIFS, FC SAN, and IP SAN (iSCSI)
- Proven reliability with no single points of failure
- A complete set of storage services that dramatically increase database availability by offloading routine tasks from compute nodes while enabling business-critical capabilities such as (logical and physical) disaster recovery, compliance, and encryption

You can learn more on this topic at Oracle OpenWorld in San Francisco. Don't miss the below session and stop by Booth 1614 for demos on quick and cost-effective cloning, workload prioritization, and more.

Strategies for Reducing Complexity and Cost:

A Practical Approach to Grid Computing

Tuesday, October 24, 2006, 1:45–2:45 p.m.

The ongoing coexistence of newer and older technologies has made it more complex and more difficult for organizations to optimize and run their infrastructure efficiently. This joint session by NetApp and Oracle will explore virtualization technologies and a practical approach for moving to the grid.

- [See a complete list of NetApp speakers at OpenWorld.](#)
- [Sign up by October 20 for a free OpenWorld Visitor Plus Pass.](#)

Comment on this article



NICK TRIANTOS

Global SAN Systems Engineer, NetApp

A member of the elite Global Systems Engineering Group, Nick helps top NetApp global enterprise customers solve their toughest technical challenges. Nick has been in systems or support engineering roles for nearly 15 years, including positions at HP as an account support engineer (Server Group) and presales technical consultant (Storage Group). Nick maintains a [blog](#) and has previously authored Tech OnTap articles on choosing [iSCSI or FC SAN](#) and [tips for simplifying SANs](#).

NetApp SE Lab Report: SAN Boot with VMware ESX 3.0.0

By Nick Triantos

Enterprises with large numbers of servers are increasingly turning to diskless servers that boot from SAN (FC or IP) to reduce costs, consolidate storage, and streamline provisioning. Although SAN boot is not new, the introduction of the blade server has helped accelerate its adoption. Blade servers provide greater manageability, reduced hardware costs, and simpler cable management in addition to power, cooling, and real-estate savings.

One of the most popular platforms that lends itself to booting from SAN is VMware ESX Server. More and more enterprises are deploying VMware ESX Server to consolidate hundreds of physical servers into a few diskless blades in a single-blade chassis. Others have elected to deploy VMware using racks of standalone 1U Intel-based servers.

Given that most servers today ship with internal SATA drives, which are not supported to host the ESX Server image, SAN boot becomes an attractive choice. Storage-based snapshots, cloning, and replication techniques provide the added advantage of being able to quickly recover a corrupted ESX image from a clone or a snapshot and restore it on the same physical server or replicate it to a remote site for DR purposes.

Like most technology vendors, Network Appliance has labs in the vast majority of its sales offices. These labs are primarily used for technology demonstrations, and they utilize a variety of operating systems and third-party software. The NetApp Dallas office also has a systems engineering lab, where NetApp SEs can delve deep into specific technologies. This lab has included a VMware ESX Server environment since the 2.5.1 release, and we have been booting ESX Server from a fabric since the 2.5.3 release.

In August I decided to upgrade to the new ESX 3.0.0 release to get a feel for the changes, and I'm definitely impressed with the results.

With ESX 3.0.0, VMware has made significant advancements in supporting boot from SAN. The multiple requirements from the previous release have eased. Based on my experience, the setup process is quick and easy and—at least for testing purposes—the environment has been working flawlessly.

Requirements	ESX server 2.5.x	ESX server 3.0.0
Config HBA	Supported	Supported
Enable HBA	Not supported	Supported
VMFS volume for Boot Image	Supported	Supported
Raw Device Mapping for Boot image	Not Supported	Supported and Recommended for Array Based Snapshots
Booting from lowest number LUN	Required	Required
Booting from lowest Target	Required	Not Required
If booting from an Active/Passive array the lowest numbered path must be Active	Required	Not Required
The HBA used for Booting can not be dedicated to the VMkernel	Required	Not Required
The boot LUN must be visible from lowest number HBA	Required	Required
The boot LUN must be visible from the lowest number Storage Processor	Required	Not Required
HBA at lowest number Bus and PCI Slot	Required	Not Required
Bootable Fabric	Required	Required
Direct Connect to external Storage	Not Supported	Not Supported

A comparison of SAN Boot Requirements for ESX Server 2.5.x and 3.0.x.
(click to enlarge)

RELATED INFORMATION

- [Love VMware, Hate Backups?](#)
- [NetApp and VMware ESX Server: Going Virtual from Server to Storage \(pdf\)](#)
- [VMware SAN Configuration Guide \(pdf\)](#)
- [Tips: Reducing SAN Complexity](#)

Love VMware But Hate Backups?

IT teams typically report that server virtualization reduces the number of servers, network ports, floor space, maintenance, and electricity required. Once the honeymoon is over, however, many of these same IT teams find themselves trapped in a backup nightmare.

In a recent Tech OnTap article, NetApp systems engineer and VMware expert Vaughn Stewart explains how to simplify backups for virtual infrastructures.

Get the details. [Read the article.](#)

Best Practices for Effective Business Application Management

Auto-parts manufacturer Magna International consolidated 65 servers into 4 physical servers using VMware and NetApp storage.

In a recent Webcast, Magna IT systems analyst Chris Green describes his experiences with NetApp and VMware and explains how the

Setting Up ESX 3.0.0 for SAN Boot

Setting up VMware for SAN boot with NetApp storage was a breeze. The whole process didn't take more than 20 or 25 minutes from the time I provisioned the boot LUN to the time the ESX image installation had completed.

The following table shows our setup.

Server	IBM x346
CPU	2x Xeon 3.2GHz
Memory	8GB
FC HBA	2x QLA 2340
FC Switch	MDS 9120
External Storage	NetApp FAS3050c
Table 1. NetApp SE Lab Setup	

After the installation, it was time to create virtual machines (VMs) and install the guest operating system. I chose to install VMs on LUNs over iSCSI so I could get a feel for VMware's implementation. Configuring the iSCSI initiator was a breeze, and I was able to install the guest operating system with no issues. Given that ESX currently does not provide a multipathing mechanism for iSCSI LUNs, I chose to implement NIC teaming, which essentially serves the same purpose.

Suggested Edits to the Default ESX 3.0.0. Configuration

If you're interested in SAN boot with ESX 3.0.0, there are some things you need to consider. First, I highly recommend that before making any HBA purchasing decisions you contact your storage vendor and carefully review the VMware *I/O Compatibility Guide for ESX Server 3.0*. You will find that certain model HBAs are not supported for SAN booting.

Additionally, there are a number of tweaks and customizations that you can make to achieve higher performance and nondisruptive failovers if hardware failures occur. I recommend three simple changes to the default ESX Server 3.0.0 setting:

- Enable the BIOS on only *one* HBA.
- Modify the Execution Throttle/Queue Depth.
- Modify the PortDownRetryCount parameter.

I go into more detail on each point in the following sections. However, keep in mind that this advice hasn't been fully tested or approved by NetApp engineering, so I can't claim that this is the right answer for all environments.

Tip #1: Enable the BIOS on only *one* HBA.

You need to enable the BIOS on the second HBA *only* if you need to reboot the server while either the original HBA used for booting purposes, the cable, or the FC switch has failed. In this scenario, you would use QLogic Fast!UTIL to select the active HBA, enable the BIOS, scan the BUS to discover the boot LUN, and assign the WWPN and LUN ID to the active HBA. However, when both HBA connections are functional, only one needs to have its BIOS enabled.

Tip #2: Modify the Execution Throttle/Queue Depth.

The Execution Throttle/Queue Depth signifies the maximum number of outstanding commands that can execute on any one HBA port. The default for ESX 3.0.0 is 32, but

deployment helped improve server utilization, reduce deployment time, and achieve DR efficiency.

Other guest speakers include:

- Brian Byun, VP, Products and Alliances at VMware
- Rich Clifton, VP and GM, NetApp Networked Storage business unit

Watch the [TechTalk Webcast](#).



Screen Shot of the Current Lab Environment
(click to enlarge)

the best value for your environment depends on a couple of factors:

- Total number of LUNs exposed through the array target ports
- Array target port queue depth

The formula to determine the value is:

Queue Depth = Target Queue Depth / Total number of LUNs mapped from the array

This formula will guarantee that a simultaneous fast load on every LUN will not flood the Target Port resulting in QFULL conditions. A QFULL condition signifies a Target port's inability to process more I/O than is capable. In most operation systems, upon receiving a QFULL condition from the Target, the HBA driver will typically decrease a LUN's maximum queue depth to the minimum value, typically "1", thereby throttling I/O to the Target port. When the Target stops issuing QFULL conditions, the HBA driver will start gradually increasing the LUN queue depth value, thereby slowly increasing I/O to the Target port.

Here's an example of how the above formula can help you avoid a QFULL condition. If a Target Port has a queue depth of 1024 and 64 LUNs are exposed thru that port, then the Queue Depth on each host should be set to 16 outstanding I/Os per LUN. This is the safest approach and guarantees no QFULL conditions:

16 Outstanding I/Os per LUN x 64 LUNs = Target Port Queue Depth

But—be careful. If a separate queue depth calculation for each host is performed using the above formula, then you still have the potential for QFULL conditions.

Here's why. Let's expand the previous example and assume that we have a total of 64 LUNs and four ESX hosts, each of which has 16 LUNs mapped.

Performing the calculation for each ESX host separately would yield: Queue Depth = 1024 / 16 LUNs = 64 Outstanding I/Os per LUN. However, a simultaneous fast load on all 64 LUNs across four ESX servers would yield: 64 Outstanding I/Os per LUN x 64 LUNs = 4096 which is much greater than Queue Depth of the Physical Array Target Port. This is an undesirable condition that, under certain circumstances, will generate a QFULL and throttle I/O.

To Change the Queue Depth on a QLogic HBA

- a. Create a copy of /etc/vmware/esx.conf.
- b. Locate the following entry for each HBA:
`/device/002:02.0/name = "QLogic Corp QLA231x/2340 (rev 02)"`
`/device/002:02.0/options = ""`
- c. Modify as shown:
`/device/002:02.0/name = "QLogic Corp QLA231x/2340 (rev 02)"`
`/device/002:02.0/options = "ql2xmaxqdepth= xxx"`
Where xxx is the queue depth value.
- d. Reboot.

Tip #3: Modify the PortDownRetryCount parameter.

The PortDownRetryCount parameter value must be set to the value recommended by your storage vendor, using Fast!UTIL. This setting specifies the number of times the adapter's driver retries a command to a port that is returning Port Down status. This value for ESX server is $2 * n + 5$, where n is the value of PortDownRetryCount from the HBA BIOS.

You can change this value directly in the HBA, or you can do it after you've installed ESX by editing the /etc/vmware/esx.conf file. To edit the file, locate the "options=" entry under the HBA model you are using and make the following change.

To Change the PortDownRetryCount Parameter

- a. Create a copy of /etc/vmware/esx.conf.
- b. Locate the following entry for each HBA:

```
/device/002:02.0/options = ""
```

c. Modify as shown:

```
/device/002:02.0/name = "QLogic Corp QLA231x/2340 (rev 02)"
```

```
/device/002:02.0/options = "qlport_down_retry= xxx"
```

Where xxx is the value recommended by your storage vendor. The equivalent setting for Emulex HBAs is "lpfc_nodedev_tmo". The default is "30".

d. Reboot.

Overall Assessment

So far, my experience with SAN booting VMware ESX 3.0.0 has been nothing but positive. From a procedural perspective, the process is certainly much easier than with previous releases. In addition, I've found the ESX host's reliability during storage controller failover testing to be rock solid so far.

Visit **NetApp** at VMWorld and see a demo of how to combine storage and server virtualization.



Comment on this article



ALVIN RICHARDS

Database and Grid Architect, Network Appliance

Alvin Richards joined NetApp in 2004 after over 16 years at Oracle. His wealth of database knowledge and experience has allowed him to serve as the leader of a virtual startup organization within NetApp focused on developing applications that allow databases to take full advantage of NetApp technologies. Among other roles, Alvin has served as the development manager, engineering lead, architect, and evangelist for SnapManager® for Oracle®.

Empowering DBAs by Integrating Oracle and NetApp Technologies

By Alvin Richards

Backup and recovery for Oracle Databases is often a complex task, especially in Fibre Channel SAN environments, where there's a whole storage fabric to deal with.

Our primary design goal with SnapManager for Oracle (SMO) was to create a framework that would empower DBAs so that they could easily manage backup and other common storage-related tasks independently from storage or host admin teams. SMO accomplishes this by integrating closely with Oracle tools, including RMAN, RAC, and ASM, to merge the core value of NetApp technology with the task flows that DBAs routinely perform.

As part of my role at NetApp, I frequently meet with customers to explain the inner workings of this application. This article draws on those experiences to explain how SMO decreases capacity requirements for Oracle backups and clones and accelerates backup and data restoration processes, especially in customer environments that are using Oracle Automatic Storage Management (ASM).

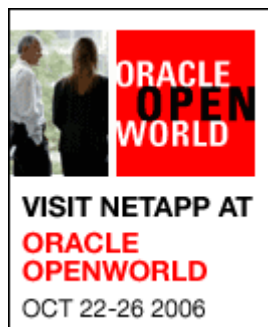
Making Oracle Data Management More Efficient

SMO allows DBAs to manage their own data while offloading data management work from Oracle servers to NetApp storage. From a central management system a DBA can, for instance, look at the state of all Oracle backups across multiple databases even if they are running on different servers and operating systems. From the standpoint of a storage administrator, SMO enables the DBA to perform common tasks of backup, restore, and clone without the needing the assistance of the storage administrator.

An SMO repository—similar in principle to an RMAN catalog—stores all the data that SMO needs. Profiles stored in this repository provide the credentials to complete Oracle operations, including:

- Oracle username and passwords (DES encrypted)
- Host and O/S account for each database instance
- RMAN repository connection info
- ASM connection info

With this information securely stored in the repository, a DBA can initiate operations without having to worry about accounts,



RELATED INFORMATION

- [SMO Product Overview](#)
- [SMO Best Practices Guide \(pdf\)](#)
- [Improving Agility for Test/Dev](#)
- [Visit NetApp at OpenWorld](#)

NetApp vs. EMC and HP: IT Decision-Maker Perspectives

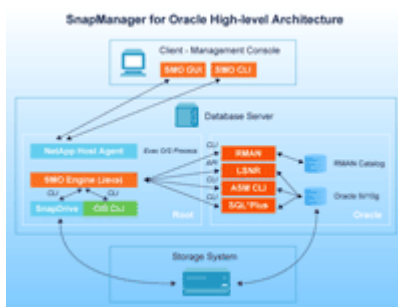
After interviewing storage managers and DBAs at more than 20 European and North American enterprises, Mercer Management Consulting identified key differentiators between database environments using NetApp, EMC, and HP storage. Findings include:

- In a typical FC SAN environment, NetApp solutions cost less than EMC CLARiiON (32%), HP EVA (42%), and EMC Symmetrix (62%).
- NetApp environments require less primary disk capacity (52%) and data protection disk capacity (20%) than EMC and HP environments.
- Database storage administrators using NetApp storage can recover twice as fast from a database crash or corruption.

Get the details. [Read the full report \(pdf\)](#)

Increasing Database Agility for Test/Dev and QA

Replicating a large database for



passwords, and so on. Through the SMO interface, a DBA can execute backups, restores, and clones that leverage proven NetApp technology to achieve extreme time and space efficiency.

[SnapManager for Oracle High-Level Architecture \(click to enlarge\)](#)

Here's how we accomplish this:

Snapshot™ Backup

Integrating Oracle backup capabilities with NetApp Snapshot copies greatly reduces the time it takes to perform a consistent backup. Because NetApp Snapshot copies only preserve changed blocks, they are extremely space efficient. SMO can optionally register the backup with RMAN to ensure not only that RMAN can take advantage of Snapshot backups, but also that the DBA still preserves the unique value of RMAN. This can be important for block-level recovery. Restores can be initiated from RMAN or SMO.

For flexibility, SMO provides the option of either full or partial backups.

- A full backup puts all tablespaces into "hot backup" mode at once, creates Snapshot copies of the associated volumes, returns the tablespaces to normal operation, and then performs a log switch, archives the logs, and makes a Snapshot copy of the archived logs. Full backups use the least Snapshot copies to get the job done, but may keep tablespaces in "backup mode" for a longer period.
- A partial backup does essentially the same thing, but iterates through the tablespaces one at a time. This ensures that each tablespace is in "backup mode" for the shortest possible time, but uses more Snapshot copies since a volume may be used by more than one tablespace.

The option you choose depends on the operating conditions of the application and business processes the DBA is supporting. Having the whole database in backup mode can be bad for a busy database since CPU utilization on the Oracle server increases 10 to 15%, and there is increased I/O since Oracle will write full blocks rather than partial blocks to the logs while in hot backup mode.

A key feature of SMO is that it maps the database contents (data files, archive logs, and so on) each time a backup is created. This means that if you add another data file or tablespace, change the archive log location, and so on, SMO will automatically take this into account each time a backup is performed. No longer do you have to worry whether you have updated your backup script to include that new tablespace or data file.

Restore

In a typical database restore scenario, you have to try one or more backups to get the database back to the exact point in time you require. If you're restoring from tape and you have to check multiple backups, it can take a very long time. SMO allows you to restore very quickly from data stored on disk, so you can quickly choose the backup you need and get up and running much faster.

The SMO restore algorithm uses NetApp Single File SnapRestore® (SFSR) capability to restore the requested parts of the database. If you're using NFS, the number of SFSRs that must be performed is equal to the number of files required to restore the database, and the SFSR operations are carried out in parallel to improve the performance. For an FC or IP SAN, SMO performs an SFSR on each LUN rather than each file, which greatly reduces the number of objects that need to be restored. If the LUN cannot be restored in its entirety, because the LUN contains files that are not pertinent to the database, SMO ensures that just the required files are restored in place into the LUN.

Cloning

One of the common tasks that DBAs ask for is to copy a gold master of a production database for development, testing, or training. DBAs typically spend a lot of time provisioning storage or getting storage provisioned and then copying the data. Using NetApp FlexClone™ technology, SMO can create an instantaneous copy of the database without using any additional physical space (until changes are made to the clone or original). Normally, if you copy a 500GB database, you're going to need 1TB

development, training, testing, or other purposes can be one of the most time-consuming tasks that a DBA can undertake. You have to carefully plan your methodology, provision enough storage to accommodate the copy, and then create a consistent replica of the data.

Learn how the use of NetApp technologies, including NetApp FlexClone and SnapMirror®, simplifies the creation of local and remote database replicas. This can streamline the database application development, test, and deployment process to improve business agility.

Get the details. [Read Reducing Time-to-Deployment.](#)

of storage in total right away. FlexClone makes the whole process much more time and space efficient because instead of storing a new physical copy, FlexClone simply stores the changed blocks.

Tight Integration with ASM

One of the key advantages of SMO that differentiates it from any other vendor's management tools is its extremely close integration with Oracle Automatic Storage Management (ASM). ASM is designed to simplify the deployment of Oracle Databases by replacing external volume managers and file systems with Oracle virtualization technology. ASM provides mirror and striping and data load balancing. By understanding the data layout within ASM, SMO delivers the same high level of backup, restore, and cloning capabilities to sites using ASM.

NetApp had three main challenges to solve when it came to integrating SMO with ASM:

- Make consistent Snapshot copies of ASM disk groups even when they span volumes and storage systems
- Provide ability to restore a single file within an ASM disk group
- Clone ASM disk groups

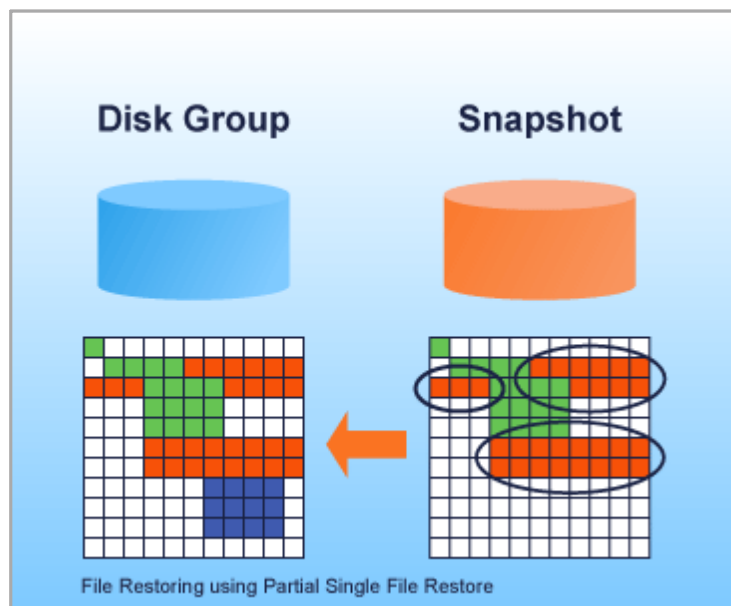
Consistency Groups

ASM has the ability to put disk groups on different volumes or even different storage systems for performance and scalability. In order for SMO to be able to create a consistent backup of a database that spans volumes or storage systems, it has to ensure that all Snapshot copies take place at a single point in time. Consistency groups, introduced in Data ONTAP® 7.2, make this possible.

Partial File SnapRestore

NetApp maps each ASM disk group as an individual LUN or file within the WAFL® (Write Anywhere File Layout) file system. WAFL is a storage virtualization technology that is a part of Data ONTAP. Since an ASM disk group can contain files from more than one database, it's obviously not sufficient in many cases to simply restore the entire disk group.

This challenge was solved using a new NetApp technology that was created for use with SMO: partial file SnapRestore (PFSR). PFSR allows SMO to restore a range of blocks from a Snapshot copy that represents a single Oracle file. This is possible because SMO understands the way Oracle lays out data in ASM. The following diagram shows how the semantic knowledge of the ASM file system allows SMO to restore just the blocks that present the red file using PFSR. The process is extremely efficient and takes place within the NetApp storage system. In laboratory tests, a table with 10 million rows can be restored in 10 seconds versus the two to three minutes required using native RMAN capabilities.



Disk Group Cloning

A general problem associated with block-oriented storage is that LUNs are typically stamped with a unique identifier. If you copy a LUN and introduce it back to the same host, the host will complain because it can't have two LUNs with the same identifier. The same situation holds with ASM. NetApp understands the ASM label and integrates with ASM to create a unique new label for disk groups cloned with FlexClone. This makes it possible for SMO to create a clone of an ASM disk group that can be introduced back to the same host, which allows database clones or backup verifications to be performed on the same host.

More Time for Value-Added Tasks

The key to the success of SMO is the close integration between NetApp and Oracle. Because SMO thoroughly understands Oracle data, the application can coordinate activities to ensure that backups are consistent and integrated with the RMAN catalog. During the development process, we invested a huge amount of our engineering effort to ensuring that we had the right methodologies to deal with possible failure scenarios so that we never leave a database in an unusable state. Even when you introduce ASM virtualization into the mix, NetApp understands the technology and does the right thing to give DBAs the capabilities they need and the results they expect.

With SMO, a DBA can get the basic and time-consuming tasks done quickly and space efficiently, freeing time to focus on high-value tasks such as tuning and capacity planning. The combination of SMO, core NetApp technology, and Oracle ensures that you can do more with less, more quickly.

You can catch Alvin speaking at Oracle OpenWorld in San Francisco on October 25.

Optimizing ASM Deployments for Resiliency, Data Protection, and Performance

Wednesday, October 25, 2006, 2:30-3:30 p.m.

As Oracle ASM and RAC are deployed more extensively, people want to understand the impact of using ASM in various storage environments. This presentation will discuss both general and NetApp specific approaches and strategies for deploying ASM and will consider the initial deployment, backup, restore, and cloning for ASM in standalone and RAC configurations, for both production and test systems.

- [See a complete list of NetApp speakers at OpenWorld.](#)
- [Sign up by October 20 for a free OpenWorld Visitor Plus Pass.](#)

Comment on this article



Adam Herger

Professional Services Engineer, Network Appliance Global Services (NGS)

Adam Herger is a NetApp Professional Services engineer currently engaged on a six-month customer project that includes migration of 100TB of storage to Data ONTAP® 7G, implementation of near-line storage, and definition and implementation of best practices for ongoing management. Adam has over 20 years of experience in a wide range of IT positions with an extensive background in networking including LAN, MAN, and WAN installation and management; system administration; and storage management. He holds a BS degree in math/computer science and an MBA.

Tools of the Month: Swatch and Kiwi Syslog Daemon

Every month Tech OnTap showcases a free tool that just might make your life a little easier. [Recommend a tool](#) and get a FREE NetApp jacket.

Author: Swatch was developed by Todd Atkins on sourceforge.net. Kiwi Syslog Daemon is freeware developed by Kiwi Enterprises.

What These Tools Do: Both tools monitor log files for specified patterns or messages and can generate a variety of notification actions when a match occurs. Swatch (simple watcher) was designed to monitor log files produced by UNIX's syslog facility and has evolved to monitor many types of logs. Kiwi Syslog provides similar capabilities specifically for Windows® systems.

How They Work: Swatch is a Perl-based tool that uses standard pattern matching to catch or exclude specific messages in almost any log file. Swatch monitors logs as they are written, carrying out specific actions as defined in the .swatchrc configuration file. Actions that Swatch can take include sending e-mail to a list of recipients, sending notifications to the console, or executing scripts or programs. If you need to monitor more than one log file, you can simply start multiple instances of the tool.

Kiwi Syslog receives, logs, displays, and forwards syslog messages from hosts such as routers, switches, UNIX hosts, and any other syslog-enabled device.

Why Adam Thinks They're Cool: These tools allow you to capture events and receive immediate notification. The great thing about them is that instead of configuring them to watch for specific things, you can configure them to ignore routine log messages. That way, you get notified when rare events occur that you aren't specifically watching for.

How Adam Uses Swatch and Kiwi Syslog Daemon: When a customer engages NGS, we immediately install a number of tools to monitor the customer's storage systems, such as Swatch (Linux® or UNIX®), Kiwi Syslog (Windows), and MRTG. (See the May Tech OnTap Tool of the Month featuring [MRTG for Filers](#).) Swatch and Kiwi Syslog allow us to receive e-mail notification whenever a significant event occurs. We configure these tools to ignore all the routine messages, so we are immediately notified if anything important occurs, no matter what it is. With many other tools, including SNMP-based monitoring, you configure the tool to watch for specific events.

Swatch Setup: After installing Swatch on a central administrative host, we edit the syslog.conf file on each NetApp storage system of interest to direct them to forward system messages to a central administrative host:

RELATED INFORMATION

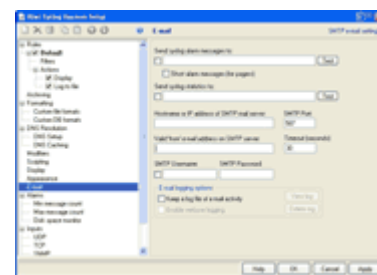
■ [Swatch website](#)

■ [Kiwi Syslog Daemon](#)

■ **Previously Featured Tools:**

- [Firefox](#)
- [Cacti](#)
- [MRTG for Filers](#)
- [Nagios](#)
- [SIO](#)
- [ToasterView](#)

Kiwi Syslog Daemon Screen:



```
# Syslog.conf
#
# Log messages of priority info or higher to the console and to
/etc/messages
*.info                                /dev/console
*.info                                /etc/messages
*.info                                @adminhost
# Edit and uncomment following line to log all messages of priority
# err or higher and all kernel messages to a remote host, e.g. adminhost
*.err;kern.*                          @adminhost
*.debug                               @adminhost
```

Next, we configure the .swatchrc configuration file to ignore the common syslog entries for both the administrative host and NetApp systems. Our standard configuration file is designed to ignore the common messages in a Linux system log as well as common NetApp messages that are not of interest:

```
#
# Swatch Configuration File
#
# /usr/local/bin/swatch
#
### Ignore this stuff from the Admin Host.
ignore=/sendmail|nntp|xntp|ntpd|faxspooler|telnetd|SIGCHLD|PAM|pam|ttloop|
sudo|application bug|Workaround activated.*|LOGIN|login|last message
repeated/
### Ignore this typical logging information from filers.
ignore=/raid.mirror.resync.*|kern.uptime.filer*|logged in|HTTP
Authentication|ok:info|is running NetApp Release|Cluster Notification mail
sent|rmc.driver:warning|mountd_main:warning|Permission denied|kernel: nfs:
server
g01|rsd_0:notice|config_async_0:notice|license.db.migrate.vol.readonly:wa
rning|scrub|Premature
EOF|CfTimeDaemon|sshd_0:info|e0:warning|ispfc_main:error|has nothing new
for|contains no new data/
### Grab everything else that falls through the above ignore list(s).
watchfor /.*/
echo
mail addresses=recipient1\@mailhost:recipient2\@mailhost
throttle=60:00,use=regex
```

With this configuration, Swatch will catch a variety of events that occur infrequently, including:

- Volume usage exceeding threshold (volume full, inode full)
- Failed disk drive
- Filer panic
- Snapshot™ copies filling up snapreserve space
- Snapshot failures
- NVRAM battery low
- CPU savage

To complete the installation, we also configure init.d (on Linux or UNIX admin hosts) to ensure that Swatch gets started immediately any time the administrative host is rebooted.

Kiwi Syslog Daemon Setup: When you install Kiwi Syslog, you have the option of installing it as an application or a service. You'll want to install it as a service so it's always on. After Kiwi Syslog is installed, you'll configure your NetApp systems to forward syslog messages using the same method as described earlier for Swatch.

Rather than configure rules using a configuration file, you specify filtering rules using the Kiwi Syslog graphical user interface. Kiwi Syslog has an "Exclude" function that provides the same functionality as the "ignore" rule we used for Swatch. You should specifically filter out the same strings as we did in the .swatchrc file shown earlier. You must put each string in double quotation marks. Here's a list of the strings to exclude:

- raid.mirror.resync
- kern.uptime.filer
- logged in
- HTTP Authentication
- ok:info
- is running NetApp Release

- Cluster Notification mail sent
- rmc.driver:warning
- mountd_main: warning
- Permission denied
- kernel: nfs: server g01
- rshd_0:notice
- config_async_0:notice
- license.db.migrate.vol.readonly:warning
- scrub
- Premature EOF
- CfTimeDaemon
- sshd_0:info
- e0:warning
- ispf_main:error
- has nothing new for
- contains no new data

Note there is no need to filter out the typical admin host stuff because it won't appear.

Once rules are configured, you can then set up Kiwi Syslog to notify you via e-mail when an event occurs using its graphical interface.

Caveat: With either utility, your administrative host will need to be able to send mail outside the data center. Be careful to not send company or sensitive data via these tools to a personal email account. Sometimes, that syslog messages can have something you'd rather not get out, but that is usually pretty low risk.

For Swatch, you must have Perl 5 installed with the following Perl modules:

- time::hires
- date::calc
- date::format

In addition to the freeware version, there is a commercial version of Kiwi Syslog that might offer greater functionality.

[Download Swatch](#)

[Swatch Discussion forums](#)

[Download Kiwi Syslog Daemon](#)

Want to recommend a tool? If we feature it, you get a free FREE NetApp jacket.

🚩 **Nominate a tool**

