

Resource: Using NFS as a data store.

Introduction

This document was created from a thread title [Using NFS as a data store](#). with the initial question: Can we have some discussion re:Using NFS as a datastore in terms of

performance, best practice etc?

Another reason for this document is that some VCPs are experiencing resistance in their organization against using NFS when the business case is clear and the technology viable.

Intended Audience

VCPs and Storage Professionals

Outline

- Introduction
- VMworld presentation
- Best practices

Author(s)

The VIOPS community

Using NFS as a data store

There has been loads on interest in running VI3 on NFS datastores ever since it was added as a storage protocol option in release 3. There have been several papers presented at VMworld in 2006 and 2007. One recent joint project comparing VM performance on VI3 with different protocols was recently published on the NetApp external web site as Technical Reference #3697 (<http://www.netapp.com/us/library/technical-reports/tr-3697.html>).

The bottom line is that performance is only one of several concerns with regards to which protocol should be chosen for shared storage pools for VI3. Ease of management, attainment of High Availability and your IT staff being familiar with a given protocol should also be factored in as well.

Resource: Using NFS as a data store.

To address best practices, there will be a paper presented at VMworld 2008 two weeks from now that is titled Joint best practices for running VI3 on IP storage. I am a co-speaker along with two NetApp experts. That presentation will be available after the conf and will outline some general guidelines for optimal configuration for HA, Performance and management optimization. If you are attending VMworld, the talk is TA 2784 and will be presented at 1:30 on Tuesday Sept 16th.

The bottom line is that many are using NFS datastores and it is performing very well. Even though some might frown, it is often more a result of NAS being out of their comfort zone more than it being a problem.

As a quick summary of best practices, I will share the following content from the before mentioned presentation:

- Use dedicated switches or VLANs for NFS datastores
- Use fast NICs (1 GbE or higher)
- Avoid over subscription of links and switches
- There is no need to place swap space on non IP storage
- Select no for the NIC teaming failback option
- If you ncrease max NFS mounts from 8 to 32 make sure to increase heap as well (see NetApp Technical reference 3428 for more details)
- Mount datastores the same way on all hosts
- Same host (hostname/FQDN/IP), export and datastore name
- Make sure NFS settings are persistent
- ONTAP CLI: use exportfs -p

Commentary from Tony Lamont (Sep 12 2008)

Hi Steve, the following NetApp Technical Support Bulletin dated September 11th explains the situation nicely but also raises a few questions of its own: "TSB-0805-02: Virtual Machines using NFS storage pauses when using VMware Snapshots." A snip from the bulletin says:"

Summary:

Note: This is an update to the TSB issued on this subject on May 5, 2008 and contains changes to the workaround and solution sections as a result of new information provided by VMware.

Virtual machines utilizing NFS data stores can experience an extended period where I/Os are suspended when using VMware ESX Server snapshots (VMsnaps).

Resource: Using NFS as a data store.

This behavior has been identified as bug SR 195302591 by VMware (NetApp Bug ID 264618), and NetApp has published a best practice recommendation in TR-3428. This TSB is intended to provide some additional information about potential issues & further expands on best practice recommendations.

Problem Description:

VMware ESX Server uses a locking mechanism to prevent an ESX Server host from simultaneously accessing the data files being used by another ESX Server host. When using NFS, ESX Server uses lock files to prevent multiple virtual machines from accessing the same set of .vmdk files. This results in extended periods of suspended I/O for these virtual machines.

Please note that this TSB will be updated after more information from VMware is available.

Workaround:

NetApp believes that there are several best practice steps that help to dramatically reduce the exposure to this ESX Server issue with NFS data stores.

If this suspension of I/O is experienced, NetApp's current best practice recommendations include all of the following:

1. Deploy the ESX Service Console ports on virtual switches with redundant NICs and links.
2. Deploy each Service Console port on the same virtual switch as any VMkernel port.
3. Deploy the ESX VMkernel Network Redundancy with redundant links.
4. Ensure that the Isolation response option is set to Power Off for each virtual machine using NFS storage.

NOTE: The Isolation response option default setting is Power Off in ESX Server for all protocols, and the recommended setting from both VMware and NetApp. However, in some FC configurations (not limited to NetApp FC storage) this setting may have been changed to Power On. Therefore, close attention should be paid to this setting in ESX Server configurations that may have initially used FC data stores and then used with NFS data stores.

5. Spanning tree protocol should be disabled on any network port using VMkernel connections.

Resource: Using NFS as a data store.

Important Note: The previous version of this TSB recommended changing the default setting of NFS.LockDisable from 0 to 1. This is incorrect. DO NOT change the value to 1. If the value had previously changed to 1, reset it to its default of 0."

Commentary from ukbrown (25 Sep 2008)

Had the same issue as you, as this was the reply from VMware. It looks like the patch below will enable as good performance as if locking were disabled.

VMware support reply below

You have some problems of consistency because I think you turn on the NFS.LockDisable to 1 means that you disable locks when accessing a disk in order to improve the performance. Just check it on "advanced setting" on your virtual center.

It was recommended to do that to improve the performance with NetApp but you risk of data inconsistency like in your case.

Just take a look to this document p.13 at the end of the page.

<http://www.netapp.com/us/library/technical-reports/tr-3428.html>

If you want to remediate to that just put 0 instead of 1 to NFS.LockDisable and put the patch for the performance (ESX350-200808401BG patch apply to ESX 3.5 update 1 or 2 only).

Update from Paul Manning (24 Nov 2008)

The current best practice for NFS is to not separate the VM swap space from the VMhome directory on a NFS datastore. The reason for the original recommendation was just good old fashioned conservatism. The thought was that if the IP traffic slowed the response time for swap space access that it could have a significant impact on the performance of the VM. However, the performance concern was not an issue that made this step of placing swap on another storage device a needed step.

Further, it turns out that it is a more simple solution to address the concern of performance degradation is to not over subscribe the memory of the VM. However the performance of the access to NFS storage compared to FC storage for swap space is not considered to be a significant enough delta that would make the separation of VM swap from the VMhome worth it.

So even though the KB article 1004082 has reference to separating them, it is no longer considered best practice. And that KB article is slated to be updated to be consistent

Resource: Using NFS as a data store.

with our current best practice of keeping swap in the VMhome Directory for VMs on NFS datastores.

Update from Vaughn @ NetApp (26 Nov 2008)

[Vaughn @ NetApp Blog Article](#)

Vaughn has provided a good update around the NetApp whitepaper - I know there's been some great collaboration between customers, VMware and NetApp and one of the end results is an update to the NetApp best practices doc. Vaughn explains it best in his blog entry above, but also here's the link to the NetApp paper:

[VMware on NetApp Best Practices Technical Report, TR-3428](#)