VMware® vSphere® VMFS-5 Upgrade Considerations

TECHNICAL MARKETING DOCUMENTATION

vmware®

Table of Contents

Introduction

VMware vSphere 5.0 introduced a new version of the flagship vSphere Virtual Machine File System (VMFS) known as VMFS-5. VMFS-5 offers a variety of new features, including:

• Larger single extent volume (64TB)

• Larger Virtual Machine Disks (VMDKs): 2TB – 512 bytes with a new unified 1MB block size

• More and smaller sub-blocks (8KB) to reduce the amount of stranded/unused space

• Improvements in performance and scalability via the implementation of the vSphere vStorage API for Array Integration (VAAI) primitive ATS (Atomic Test & Set) across all datastore operations.

vSphere 5.0 supports both VMFS versions 3 and 5. Therefore, it is not necessary to upgrade your VMFS volumes. However, customers can move to VMFS-5 to benefit from these features. A complete set of VMFS-5 enhancements can be found in the What's New in vSphere 5.0 Storage white paper.

VMFS-5 Enhancements

The following is a complete list of enhancements made in VMFS-5.

New Unified 1MB File Block Size

Earlier versions of VMFS used 1, 2, 4 or 8MB file blocks. These larger blocks were needed to create large files (>256GB). These different file blocks sizes are no longer needed to create large files on VMFS-5. Very large files can now be created on VMFS-5 using the new unified 1MB file blocks. Earlier versions of VMFS will still have to use larger file blocks to create large files.

Large Single Extent Volumes

In earlier versions of VMFS, the largest single extent was 2TB - 512 bytes. An extent is a partition on which one can place a VMFS. To create a 64TB VMFS-5, one needed to create 32 x 2TB extents/partitions and join them together. With VMFS-5, this limit for a single extent/partition has been increased to 64TB. This significantly reduces the management overhead when using very large VMFS volumes.

Smaller Sub-Blocks

VMFS-5 introduces smaller sub-blocks. Sub-blocks are now 8KB rather than 64KB as used in the earlier versions. With VMFS-5, small files (< 8KB, but > 1KB) in size will consume only 8KB rather than 64KB. This will reduce the amount of disk space stranded by small files. Also, there are many more sub-blocks in VMFS-5 than there were in VMFS-3 (32,000 on VMFS-5 compared to approximately 4,000 on VMFS-3).

Small File Support

VMFS-5 introduces support for very small files. For files less than or equal to 1KB, VMFS-5 uses the file descriptor location in the metadata for storage rather than file blocks. When these files grow beyond 1KB, they will then start to use the new 8KB sub-blocks. This will again reduce the amount of disk space stranded by very small files.

Increased File Count

VMFS-5 introduces support for greater than 120,000 files, a four-fold increase when compared to the number of files supported on VMFS-3, which was approximately 30,000.

ATS Enhancement

The Atomic Test & Set (ATS) Hardware Acceleration primitive is now used throughout VMFS-5 for file locking. ATS is a part of VAAI (vSphere Storage APIs for Array Integration). This enhancement improves the file locking performance over earlier versions of VMFS.

GPT

To handle much larger partition sizes, the GUID Partition Table (GPT) format will now be used to create VMFS-5 volumes. Historically, Master Boot Record (MBR) was used; however, this was limited to a maximum partition size of approximately 2TB. GPT overcomes this limitation and allows for much larger partitions to handle single extents up to 64TB (using VMFS-5).

New Starting Sector

VMFS-5 partitions will now have a starting sector of 2048. This is different from VMFS-3 which had a starting sector of 128. Moving to a starting sector of 2048 helps avoid alignment issues.

Considerations – Upgrade or Create New

While a VMFS-3 which is upgraded to VMFS-5 provides you with most of the capabilities as a newly created VMFS-5, there are some differences. Both upgraded and newly created VMFS-5 support single-extent volumes up to 64TB and both support VMDK sizes of ~2TB, no matter what the VMFS file-block size is. However additional differences, although minor, should be considered when making a decision whether to upgrade to VMFS-5 or create new VMFS-5 volumes.

No Uniform Block Size

VMFS-5 upgraded from VMFS-3 continues to use the previous file-block size, which may be larger than the unified 1MB file-block size. This can lead to stranded/unused disk space when there are lots of small files on the datastore.

No New Sub-Block Size

VMFS-5 upgraded from VMFS-3 continues to use 64KB sub-blocks and not the new 8KB sub-blocks. This can also lead to stranded/unused disk space. The upgraded VMFS-5 also continues to use the original number of sub-blocks from the VMFS-3.

No Increase to the Maximum Number of Files per Datastore

VMFS-5 upgraded from VMFS-3 continues to have a file limit of 30,720 rather than new file limit of > 100,000 for newly created VMFS-5. This has an impact on the scalability of the file system.

Uses MBR

VMFS-5 upgraded from VMFS-3 continues to use MBR (Master Boot Record) partition type; when the VMFS-5 volume has grown beyond 2TB, it automatically and seamlessly switches from MBR to GPT (GUID Partition Table) with no impact on the running virtual machines.

Starts on Sector 128

VMFS-5 upgraded from VMFS-3 continues to have its partition starting on sector 128. Newly created VMFS-5 partitions will have their partition starting at sector 2048.

Here is a comparison of the metadata from a newly created VMFS-5 volume and a VMFS-5 volume that was upgraded from VMFS-3. The version, max files and max sub-blocks are highlighted. As you can see the version is the same, but there are differences in the amount of resources available in each file system.

```
~ # vmkfstools -Pv 10 /vmfs/volumes/newly-created-vmfs5/
VMFS-5.54 file system spanning 1 partitions.
File system label (if any): newly-created-vmfs5
Mode: public
Capacity 3298534883328 (3145728 file blocks * 1048576), 3297500987392 (3144742 blocks)
avail
Volume Creation Time: Tue Jun 14 14:35:53 2011
Files (max/free): 130000/129992
Ptr Blocks (max/free): 64512/64496
Sub Blocks (max/free): 32000/32000
Secondary Ptr Blocks (max/free): 256/256
File Blocks (overcommit/used/overcommit %): 0/986/0
Ptr Blocks  (overcommit/used/overcommit %): 0/16/0
Sub Blocks  (overcommit/used/overcommit %): 0/0/0
UUID: 4df771c9-f6419df2-81bc-0019b9f1ecf6
Partitions spanned (on "lvm"):
        naa.60a98000572d54724a34642d71325763:1
DISKLIB-LIB   : Getting VAAI support status for /vmfs/volumes/newly-created-vmfs5/
Is Native Snapshot Capable: NO
~ #
```

Figure 1. Newly Created VMFS-5

```
~ # vmkfstools -Pv 10 /vmfs/volumes/upgrade-testvol
VMFS-5.54 file system spanning 1 partitions.
File system label (if any): upgrade-testvol
Mode: public
Capacity 3298534883328 (3145728 file blocks * 1048576), 3297916223488 (3145138 blocks)
avail
Volume Creation Time: Mon Jun 13 13:03:04 2011
Files (max/free): 30720/30713
Ptr Blocks (max/free): 64512/64496
Sub Blocks (max/free): 3968/3968
Secondary Ptr Blocks (max/free): 256/256
File Blocks (overcommit/used/overcommit %): 0/590/0
Ptr Blocks  (overcommit/used/overcommit %): 0/16/0
Sub Blocks  (overcommit/used/overcommit %): 0/0/0
UUID: 4df60a88-8eaa51ea-3108-0019b9f1ecf6
Partitions spanned (on "lvm"):
        naa.60a98000572d54724a34642d71325763:1
DISKLIB-LIB   : Getting VAAI support status for /vmfs/volumes/upgrade-testvol
Is Native Snapshot Capable: NO
~ #
```

Figure 2. Upgraded VMFS-5

Online Upgrade

Upgrading a VMFS-3 to a VMFS-5 file system is a single-click operation. Once you have upgraded the host to VMware ESXi™ 5.0, go to the Configuration tab > Storage view. Select the VMFS-3 datastore, and above the *Datastore Details* window, an option **Upgrade to VMFS-5** will be displayed:
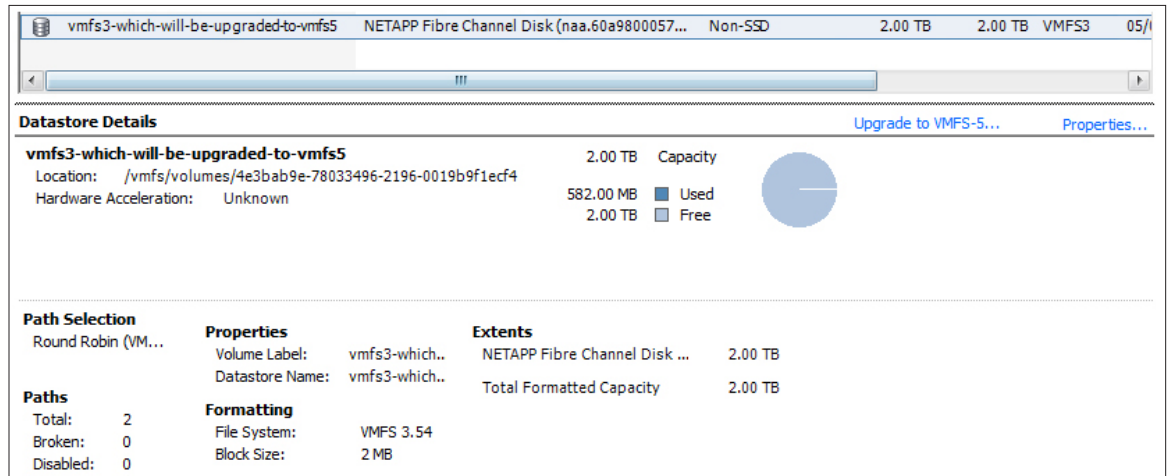

Figure 3. Upgrade to VMFS-5

The upgrade process is online and non-disruptive. Virtual machines can continue to run on the VMFS-3 datastore while it is being upgraded. Upgrading the VMFS file system version is a one-way operation. There is no option to reverse the upgrade once it is executed. Additionally, once a file system has been upgraded, it will no longer be accessible by older ESX/ESXi 4.x hosts, so you need to ensure that all hosts accessing the datastore are running ESXi 5.0. In fact, there are checks built into vSphere which will prevent you from upgrading to VMFS-5 if any of the hosts accessing the datastore are running a version of ESX/ESXi that is older than 5.0.

As with any upgrade, VMware recommends that a backup of your file system is made prior to upgrading your VMFS-3 file system to VMFS-5.

Once the VMFS-5 volume is in place, the size can be extended to 64TB, even if it is a single extent, and ~2TB Virtual Machine Disks (VMDKs) can be created, no matter what the underlying file-block size is. These features are available 'out of the box' without any additional configuration steps.

Conclusion

As you do not get the complete set of enhancements when upgrading a VMFS-3 to a VMFS-5, VMware recommends using newly created VMFS-5 volumes if you have the luxury of doing so. You can then seamlessly migrate the virtual machines from the original VMFS-3 to VMFS-5 using vSphere Storage vMotion. If you do not have the available space to create new VMFS-5 volumes, then upgrading VMFS-3 to VMFS-5 will still provide you with most of the benefits that come with a newly created VMFS-5.

About the Author:

Cormac Hogan is a Senior Technical Marketing Manager responsible for storage in the Cloud Infrastructure Product Marketing group at VMware. His focus is on core VMware vSphere storage technologies and virtual storage in general, including the VMware vSphere Storage Appliance. He was one of the first VMware employees at our EMEA HQ in Cork, Ireland, back in April 2005. He spent two years as the Technical Support Escalation Engineer for Storage before moving into a Support Readiness Training role, where he developed training materials and delivered training to Technical Support and our Support Partners. He has been in Technical Marketing since 2011.

• Follow Cormac's blogs at http://blogs.vmware.com/vsphere/storage

• Follow Cormac on Twitter: **@VMwareStorage**

vmware®