Hewlett Packard Enterprise

Best Practices for VMware VAAI Plug-in Deployment for XP7 Storage

Contents

Executive Summary 2
VAAI Overview 2
Best Practices and Recommendations 2
- Plug-in Installation 2
- Performance Considerations for Block Zeroing 3
- XCOPY Cloning Operations and Cache Configuration 3
- Back-end Storage Optimizations for XCOPY Cloning 3
- VAAI Interaction with other XP Program Products 3

Technical white paper

Executive Summary

This white paper gives a brief overview and explains the best practices for using the HPE Storage VMware® API for Array Integration (VAAI) plugin for the HPE XP7 Storage. This plug-in enables the offloading of some VMware vSphere operations to the storage array.

VAAI Overview

VAAI enablement in its current release of vSphere server consists of two parts:

- A VMware server-resident software drivers or "plug-ins" supplied by disk array vendors
- A VAAI enabled storage firmware/OS supplied by disk array vendors

These two components allow VMware vSphere hosts to exercise the capabilities defined by the VAAI primitives. Currently, the plug-in offered by HPE for use with XP7 Storage is capable of the following VAAI primitives:

Block zeroing/Block Initialization (WRITE SAME)
A common operation on virtual disks is to initialize large extends of the disk with zeroes to help isolate VMs and promote security. When performed by vSphere servers, this initialization consumes host CPU cycles, DMA buffers and HBA queue. By off-loading the processing to the disk arrays, the resource consumption in the ESX host can be reduced.

Full copy/Block Cloning (XCOPY)
Copy processing, such as creating a clone of a virtual machine is off-loaded to the disk array.

Because the vSphere host does not have to read/write data when copying it, the vSphere host workload is be reduced, as well as the traffic between the vSphere host and the disk array, resulting in improved copy performance.

Hardware assisted locking (Atomic Test and Set—ATS)
The ATS primitive removes the need for SCSI reservation for the purpose of disk-locking. ATS enables exclusive locking at the storage block (sector) level instead of the traditional LUN level locking performed by SCSI-2 reservations. SCSI reserve command contention can be reduced and an improvement of VMFS scalability can be realized.

Thin Provisioning Page Reclaim (UNMAP)
Once a virtual machine file has been deleted from a vSphere server, the underlying XP7 thin-provisioned storage pool can be made to reclaim the user storage space the virtual machine file was stored on and make it available once again as free space within this storage pool.

Out of Space Condition (VMSTUN)
This primitive will pause or "stun" a VM when the capacity of its underlying storage is about to be exceeded.

Quota-Exceeded Behavior
This primitive allows VAAI to provide alerts when the underlying storage exceeds pre-set thresholds. The thresholds are set via the array-specific management tools.

TP LUN Reporting
Primitive that allows VAAI to report whether or not a given LUN is thin-provisioned without the need to utilize vendor-specific management tools.

Best Practices and Recommendations

Plug-in Installation

The installation and verification of the XP7 plug-in is covered in the plug-in's user guide. It is however required to enable the XP7 Host Option Mode 0x54 prior to the installation of HPE's plug-in so that this step isn't simply overlooked by mistake after the plug-in installation is finished.

Performance Considerations for Block Zeroing

While the nature of the block zeroing primitive functionality as it is implemented on XP7 Storage is nearly cache-independent in nature, it is nevertheless a back-end disk- accessing process that must compete with any other unrelated disk access against any shared portion of allocated storage. It is recommended as a "best practice" to create any new VMDK with this functionality realizing that although the block-zeroing portion this action has been optimized, this may still result in performance attenuations arising from the array block-zeroing process competing with other parallel operations against that same portion of storage. **It is further suggested to execute such tasks during periods of time when reduced amounts of I/O is being done to the target data stores.**

XCOPY Cloning Operations and Cache Configuration

For XP7 Storage, the amount of array cache available to any cloning operation or collection of operations has a non-trivial effect on cloning performance. A performance benefit may be realized by allocating an amount of cache for cloning **up to the size of the VM, or the sum of the sizes of the VMs being cloned concurrently.** Given that array cache is a valuable and often limited resource in most array configurations, it is suggested as a "best practice" that using Cache Logical Partitions (CLPRs) to allocate quantities of cache that can be dedicated to the processing of VM cloning of larger VMs (> > 3 GB) when these operations are of a high priority or are time-sensitive.

Back-end Storage Optimizations for XCOPY Cloning

While the XCOPY cloning primitive has the potential to relieve a host of the VM cloning work, it does not exempt the user from design considerations that, if ignored, can negate any benefit the VAAI plug-in might be able to produce. These considerations are mainly in the areas of the host-connect or array "front-end" and the physical HDD-based storage, or array "back-end". The XCOPY cloning operation will show the strongest relative performance benefit when the hosts connected to the array "front-end" are bandwidth or processing-limited compared to the storage processing resources and HDDs allocated to these hosts from the array "back-end". For example, a host with a 2 Gb HBA will not be able to complete a VM cloning operation as fast as if it were to use the VAAI XCOPY primitive where the storage allocated to that same host is capable of 400 MB/sec or greater. The converse of the above situation is however not true.

The disk access patterns for a VAAI Plug-in cloning (XCOPY) operation are similar to that of an optimized tape backup or restore operation. Because of this, it is particularly necessary to avoid overloading the provisioned storage affected by these cloning operations by paying attention to the density and frequency of their execution. The XP7 Thin Provisioning program product offers the ability to combine several parity groups into a single storage pool that has a combined performance potential that is greater than any of its member parity groups. The use of Thin Provisioning has the potential of greatly mitigating any possible back-end bottlenecks a user of the VAAI XCOPY Cloning primitive might experience. It is therefore recommended as a "best practice" that Thin Provisioning be used whenever possible.

It is also important to note that a single instance of a VAAI Plug-in is capable of initiating up to 4 cloning operations concurrently. If the user attempts more than 4 on the same plug-in instance, the first four are serviced immediately and the remaining requests are queued for later processing.

VAAI Interaction with other XP Program Products

VAAI Primitives are supported with Business Copy, Fast Snap, Continuous Access, and External Storage as follows: XCOPY is supported with Continuous Access (Synchronous and Journal), Business Copy, and Fast Snap Primary Volumes. ATS is supported with Continuous Access (Synchronous and Journal), Business Copy, External Storage, and Fast Snap Primary Volumes. WRITE SAME is supported on Normal with Thin Provisioning.

Learn more at

hpe.com/us/en/storage/enterprise-xp.html

Sign up for updates

4AA5-1301ENW, January 2017, Rev.2