



Technical Report

NetApp HCI for Oracle

Bobby Oommen, NetApp

April 2019 | TR-4768

Abstract

This paper introduces NetApp® HCI and describes how you can consume infrastructure linearly by starting small and then scaling compute and storage resources independently. NetApp HCI, powered by NetApp Element® software, enables you to build a private cloud solution and provides a simple and flexible architecture that helps elevate the value of IT.

TABLE OF CONTENTS

1	Introduction	4
1.1	Configuration.....	4
2	The Big Picture: VMware Private Cloud on NetApp HCI	4
3	Application Use Cases	5
3.1	Oracle Considerations	6
3.2	Performance Tiering with QoS.....	6
3.3	Data Protection	6
4	Oracle Licensing with NetApp HCI	7
4.1	CPU Optimization	7
4.2	Core Right-Sizing.....	7
5	Storage Configuration	8
6	Operating System Configuration	8
6.1	Add Packages.....	9
6.2	Update Kernel Parameters.....	9
6.3	Optimize Network Performance	9
6.4	Set User Limits.....	9
6.5	Set Permissions	9
7	Oracle Configuration	10
7.1	Install Prerequisite Packages.....	10
7.2	Configure Oracle ASM.....	10
7.3	Configure ASM Library.....	10
8	Microsoft Windows Configuration	10
9	Summary	10
	Where to Find Additional Information	11
	Acknowledgments	11
	Version History	11

LIST OF TABLES

Table 1	HCI Storage and Compute node technical specifications.....	4
Table 2	Compute configurations.....	7

LIST OF FIGURES

Figure 1) VMware private cloud on NetApp HCI.....	5
Figure 2) QoS policies in action.....	6

1 Introduction

One of the biggest challenges that you might have as a database administrator is how to manage multiple environments. As your data center footprint increases, it becomes difficult to gain insight into the workload characteristics of all the running applications. IT administrators are in a transition state, in which they would like to control and manage their compute and storage resources easily in a flexible consumption model. Flexibility and simplicity are key requirements when businesses like yours are planning to adopt new architectures. Reducing the TCO is an even bigger challenge, because the overall IT budget must align with the financial goal of your enterprise.

NetApp® HCI helps you build a private cloud strategy with a solution that can transform and empower your organization, so you can move faster, drive operational efficiency, and reduce costs. You can deploy NetApp HCI in minutes, eliminating the complex management of existing three-tier architectures.

NetApp HCI, powered by NetApp Element® software, is a hybrid cloud infrastructure for customers who want to break free from first-generation hyperconverged infrastructure limitations. These customers are looking for independent scaling of compute and storage resources while avoiding costly and inefficient overprovisioning. With NetApp HCI, you can run multiple application workloads that are distributed across all available storage nodes and get consistent performance while sharing resources across your entire data center.

For more information about NetApp HCI, see the [NetApp HCI product page](#).

1.1 Configuration

NetApp HCI is available with multiple configuration options, enabling you to right-size your environment. You can start with compute nodes that are defined as small, medium, or large. These resources reside as nodes inside a chassis. NetApp HCI also offers nodes with more resources for more intensive performance requirements in a 2-rack-unit (2RU) form factor.

As shown in Table 1, NetApp HCI H410C compute nodes scale from 8 cores to 40 cores and from 256GB to 1TB of RAM. They also offer the flexibility of mixing and matching any compute or storage node independently. For example, if your setup has two compute nodes and four storage nodes and you want to add more compute, you can do it without adding any storage. You simply add a compute node into the chassis. Having this scalability and flexibility enables you to build an efficient and agile cloud in your data center. For the data traffic side, each storage node comes with two 10/25GbE ports, and each compute nodes comes with four 10/25GbE ports.

Table 1) NetApp HCI compute node technical specifications.

NetApp HCI H410C Compute Nodes					
	RU	Solid-State Drive Size and Quantity	RU	Cores per Node	RAM per Node
Small	1RU, half-width	6 x 480GB	1RU, half-width	8–16	256GB
Medium		6 x 960GB		16–28	512GB
Large		6 x 1.92GB		28–40	768GB–1TB

2 The Big Picture: VMware Private Cloud on NetApp HCI

Moving to the private cloud is a journey. Most organizations want to start small and grow their environment over time. NetApp has developed, in collaboration with VMware, two documents that describe how to deploy VMware private cloud on NetApp HCI. These documents are [VMware Validated](#)

[Design \(VVD\)](#) and the [NetApp Verified Architecture \(NVA\)](#) for VMware private cloud on NetApp HCI. These documents provide guidelines, best practices, and deployment documentation for a VMware private cloud, using the VMware vRealize Suite of products and VMware NSX, with NetApp HCI.

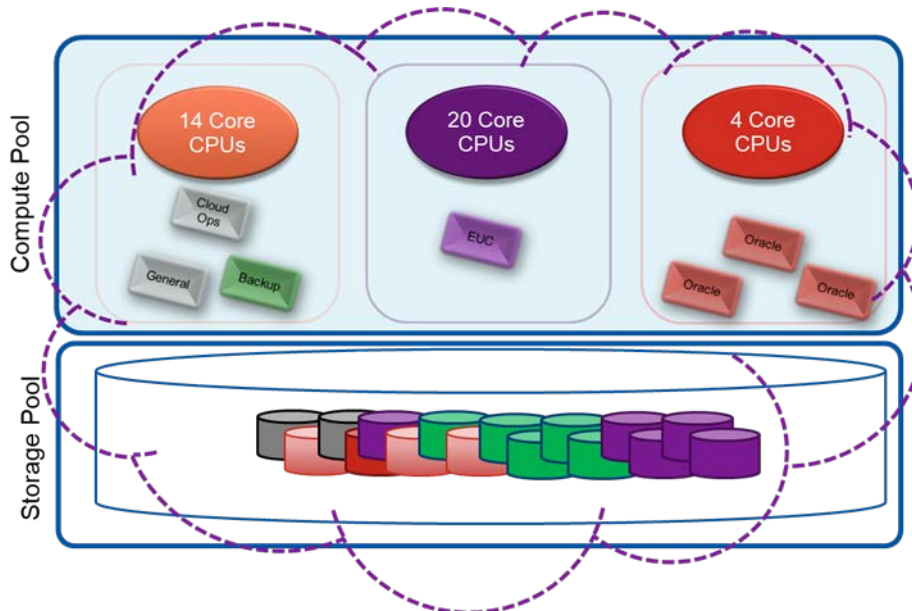
A typical goal when using a private cloud might be to automate parts of the environment. With vRealize Automation self-service, you can provision virtual machines (VMs) from a predefined catalog and quickly provision applications. For example, Oracle VMs can be delivered as an OVA file (template) from VMware and then deployed in the vSphere environment. The goal is a solution that offers microsegmentation of resources, provides extensive IT automation, and enables you to integrate your application workloads in your private cloud with hyperscalers through the NetApp Data Fabric.

Another objective when deploying and operating in a private cloud is increased agility with the ability to quickly deploy more applications or workloads. With NetApp HCI, you can easily deploy and support any workload or use case. Figure 1 shows how the compute and storage resources are laid out for a superior level of resource-sharing capabilities, which is similar to hyperscalers such as Amazon Web Services (AWS), Azure, and Google Cloud.

With this infrastructure, you can deploy the application or the workload to run on the desired CPUs and move that application or workload as your needs evolve. The storage that holds the data does not have to move and can be grown independently from the compute environment. The NetApp Data Fabric also enables you to connect your private cloud ecosystem on NetApp HCI to hyperscalers to deliver a hybrid multicloud experience.

For more information about the NetApp Data Fabric, see [What Is Hybrid Multicloud Experience?](#)

Figure 1) VMware private cloud on NetApp HCI.



3 Application Use Cases

In addition to the previously mentioned benefits, NetApp HCI is optimal for various use cases. NetApp HCI is architected to deliver exceptional value for the following use cases:

- Private cloud deployments
- Virtual desktop infrastructure (VDI) for end-user computing (EUC) environments
- Workload consolidation and enterprise applications

For applications, NetApp HCI provides several benefits, including:

- Oracle considerations
- Performance tiering with Element software quality of service (QoS)
- Data protection

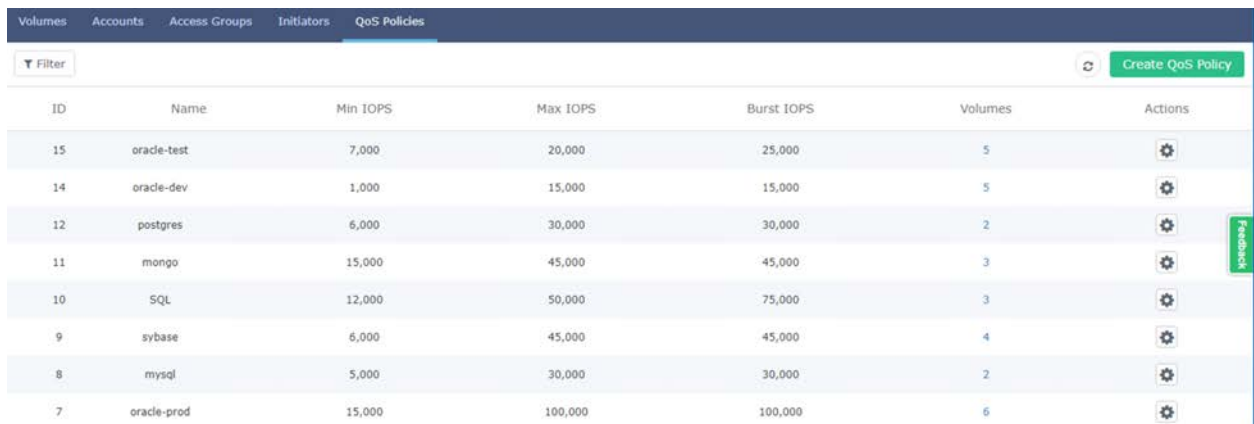
3.1 Oracle Considerations

NetApp HCI can support a wide range of database application use cases, such as database consolidation, dynamic resource allocation for development and testing environments, and integration of cloud infrastructure. There is no restriction on running single Oracle or Oracle Real Application Clusters (RAC) instances on NetApp HCI. However, you should also research other NetApp [products](#) if you want to achieve consistent submillisecond and high-bandwidth I/O performance.

3.2 Performance Tiering with QoS

The per-volume QoS controls of storage nodes help individual databases achieve sufficient I/O throughput without being affected by other databases that are running on the same NetApp HCI setup. As shown in Figure 2, you can achieve higher database density with the shared storage infrastructure by having several database instances.

Figure 2) QoS policies in action.



ID	Name	Min IOPS	Max IOPS	Burst IOPS	Volumes	Actions
15	oracle-test	7,000	20,000	25,000	5	
14	oracle-dev	1,000	15,000	15,000	5	
12	postgres	6,000	30,000	30,000	2	
11	mongo	15,000	45,000	45,000	3	
10	SQL	12,000	50,000	75,000	3	
9	sybase	6,000	45,000	45,000	4	
8	mysql	5,000	30,000	30,000	2	
7	oracle-prod	15,000	100,000	100,000	6	

VMware administrators have full control of each storage volume on which the database resides. These administrators can also perform all maintenance operations, including setting the QoS for each database copy, through the NetApp Element Plug-In for VMware vCenter Server. Administrators also can use REST APIs to achieve full automation and make storage management simpler and easier.

3.3 Data Protection

Storage snapshot copies provide a point-in-time view of the contents of an active file system or storage volume. You can use application-consistent snapshot copies for rapid recovery of corrupted datasets and to create space-efficient copies of datasets for development and testing. NetApp HCI supports having multiple volumes for the database and offers the same consistency point during the group snapshot copy. NetApp HCI is Data Fabric ready, increasing agility by expanding application resources across a wide range of cloud environments.

4 Oracle Licensing with NetApp HCI

4.1 CPU Optimization

The use of dedicated compute and dedicated storage servers is a key differentiator with NetApp HCI because it means that core-licensed products do not run on storage server CPUs. As a result, you can reduce the number of cores that must be licensed for products such as Oracle Database and can simplify software management and auditing.

The NetApp HCI architecture is also more efficient from a licensing and performance point of view because cores that are licensed for Oracle are not given the task of storage work. The CPU work that is required to present and to maintain a LUN, including features such as compression and deduplication, is isolated to a different system. It therefore does not create a “NetApp HCI tax” on the compute hypervisor, including its cores that were licensed for Oracle work.

4.2 Core Right-Sizing

The processing capabilities of CPUs have increased much faster than the processing demands of most database workloads. Sometimes databases are limited by CPU work, but it is generally a result of the processing limits of a single core and is not a limitation of the CPU. The result is an increasing number of idle cores on database servers that still must be licensed for the Oracle Database software. This underutilization of CPU resources is a waste of capital expenditure, not only in terms of licensing costs, but also in terms of the cost of the server itself, heat output, and so on.

The compute configurations in Table 2 are available with the NetApp HCI solution and, to optimize the licensing costs, you can deploy the minimum cores that are required to service a workload. The options range from 40 cores at 2.0GHz to just 8 cores at a substantially faster 3.6GHz.

Table 2) Compute configurations.

Name	RAM in Gigabytes	CPU Cores/Type	Speed
H410C-14020	384	2 Xeon Silver 4110, 8 core	2.1GHz
H410C-15020	512		
H410C-17020	768		
H410C-25020	512	2 Xeon Gold 5120, 14 core	2.2GHz
H410C-27020	768		
H410C-28020	1,000		
H410C-57020	768	2 Xeon Gold 6138, 20 core	2.0GHz
H410C-58020	1,000		
H410C-35020	512	2 Xeon Gold 5122, 4 core	3.6GHz
H410C-37020	768		

Oracle licensing requirements can be complicated, but as a rule, all cores in which the software product can run without deliberate user intervention must be properly licensed. Virtualization adds requirements because of the native clustering capabilities. There is no requirement to license all ESX clusters under a single VMware vCenter Server, although the ability to migrate workloads between clusters does increase the chance of user error that leads to a license violation. A more common approach is to fully license a particular ESX cluster for core-licensed Oracle products. This approach reduces the risk of user error in

VM placement while avoiding the licensing of unnecessary cores. All nodes are then available to provide high-availability (HA) services.

When it is not feasible to license all the cores, such as a large single ESX cluster, VMware affinity rules are often used to segment the cluster. For example, two out of six ESX nodes might be dedicated for Oracle work, and VMware affinity rules can prevent the software from running elsewhere. You must verify that new databases are properly restricted to the desired nodes. Finally, you can use [VMware policies and affinity rules](#) to limit the way that running database instances are tied to specific CPU cores. No controlling Oracle documentation currently exists for Oracle virtualization license policies. The recommended option is to reduce the number of cores that you deploy in a single server so that servers can be fully licensed wherever possible.

Next, by using the previously outlined guidelines, you explain your intended deployment of Oracle products on the NetApp HCI system to your Oracle account team, including the number of cores you are expected to license. If the Oracle account team expresses any concerns, ask them to clarify the policies with documented, binding authority that covers your scenario.

5 Storage Configuration

The configuration that is described in this document is for an Oracle database that uses an Oracle Automatic Storage Management (ASM) layout. You can also use a logical volume manager (LVM) configuration on Linux to configure the Oracle application. NetApp HCI supports presentation of the storage in a 4K sector size (native mode) and in a traditional 512-byte sector size (512e). Provided that there is no partition misalignment at the host level, there is no performance effect from choosing emulation mode.

For the database layout, follow one of the storage configurations:

- iSCSI LUNs that are managed by the iSCSI initiator on the VM, not on the hypervisor
- Separate datastore and storage LUNs for Oracle data files, redo logs, and archival logs

To get the following benefits, you should avoid the use of datastores for Oracle files:

- **Transparency.** When a guest VM owns the file systems, it is easier for you to identify the source of the file systems for your data.
- **Manageability.** When a guest VM owns the file systems, manageability of provisioning, monitoring, and data protection across the environment becomes easier.
- **Portability.** When a VM owns its file systems, the process of moving an Oracle environment becomes simpler. File systems can easily be moved between virtualized and nonvirtualized guests and to other hypervisors.

Note: NetApp strongly recommends that you use QoS policies to set the IOPS characteristics when you provision the volumes for the Oracle database. The use of QoS policies enables better manageability of IOPS settings at the volume level.

6 Operating System Configuration

Before you proceed, follow the recommendations in [TR-4639: SolidFire for Linux Configuration Guide](#) to implement the OS. These recommendations apply to Red Hat Enterprise Linux 7.x, to distributions of Oracle Linux (OL), and CentOS. This report assumes that a default installation of a software development workstation exists. You can use alternate distributions if they are fully compatible with the Oracle grid infrastructure and with Oracle Database software.

6.1 Add Packages

After you install the base OS, you must update it to meet Oracle grid infrastructure and Oracle Database installation requirements. For further information, see the [Oracle documentation](#) and the release requirements, specifically the section “Install Prerequisite Packages.”

6.2 Update Kernel Parameters

Update the kernel parameters for your host operating system to the following values:

```
net.ipv4.ip_local_port_range = 9000 65500
net.core.rmem_default = 4194304
net.core.rmem_max = 16777216
net.core.wmem_default = 262144
net.core.wmem_max = 16777216
net.ipv4.ipfrag_high_thresh = 524288
net.ipv4.ipfrag_low_thresh = 393216
net.ipv4.tcp_rmem = 4096 524288 16777216
net.ipv4.tcp_wmem = 4096 524288 16777216
net.ipv4.tcp_window_scaling = 1
net.core.optmem_max = 524287
net.core.netdev_max_backlog = 2500
net.ipv4.tcp_mem = 16384 16384 16384
fs.aio-max-nr = 1048576
net.ipv4.tcp_no_metrics_save = 1
net.ipv4.tcp_moderate_rcvbuf = 0
vm.min_free_kbytes=262144
vm.swappiness=10
```

6.3 Optimize Network Performance

Consider the following guidelines for optimal network performance:

- Enable jumbo frames for all data network interfaces.
- The RAC interconnect should be configured with a different subnet than the RAC virtual IP address (VIP address) so that cluster interconnect traffic is fully isolated.
- Use separate virtual LANs (VLANs) for the management and data traffic.
- Add two virtual network interface cards (vNICs) for data traffic.

6.4 Set User Limits

Oracle recommends that you set the user limits to an optimal number because user limits affect application performance. For this solution, the following user limits were applied for both the `grid` and the `oracle` user. The same settings were applied on all RAC nodes. See the [Oracle documentation](#) for any updates to the following settings:

```
grid soft nproc 2047
grid hard nproc 16384
grid soft nofile 1024
grid hard nofile 65536
oracle soft nproc 2047
oracle hard nproc 16384
oracle soft nofile 1024
oracle hard nofile 65536
```

6.5 Set Permissions

Before the Oracle installation, you should set appropriate permissions so that Oracle can access the underlying devices on the NetApp HCI storage. You can set the device permissions on multipath devices by creating an `udev` rule file that allows the owner of the Oracle Database software to access the devices.

7 Oracle Configuration

7.1 Install Prerequisite Packages

NetApp recommends that you run the `runcluvfy.sh` script that is part of the Oracle installation package under the `grid` directory. This script thoroughly verifies all the required packages for each RAC node. You can run the script from the first RAC node, and it validates all the other nodes that are part of the cluster. The script can be executed as follows and is available in the `installer` directory of the grid installation software:

```
./runcluvfy.sh stage -pre crsinst -n rac1,rac2 -fixup -verbose -asm -asmdev  
/dev/mapper/crs1,/dev/mapper/crs2,/dev/mapper/crs3
```

7.2 Configure Oracle ASM

The storage nodes automatically distribute and protect Oracle Database data across all nodes in the NetApp HCI cluster. For optimal distribution of data, use ASM to stripe data over all the storage volumes. This implementation helps confirm that all nodes can process Oracle writes and that Oracle Database I/O requests are uniformly distributed over all the nodes. As your capacity and performance requirements increase, you can dynamically create and add storage volumes to Oracle ASM disk groups.

If the overall throughput requirement of the database goes beyond 600MBps, NetApp recommends that you create more than one storage volume. You should also configure the ASM software to use a large allocation unit (64MB). The larger allocation unit reduces the amount of metadata and metadata updates that ASM requires to support a given amount of storage.

The storage nodes provide complete fault isolation through a comprehensive shared-nothing architecture. Therefore, ASM mirroring within a single cluster is not appropriate. For extreme fault tolerance, ASM mirroring can be used between clusters.

Note: ASMLib is a requirement on Linux when you use a 4KB native volume. Any volumes that you create with the default 512e are accessible by Oracle ASM.

7.3 Configure ASM Library

Oracle ASMLib simplifies storage deployment and management for Linux OS. Before you assign any LUN or volume to ASMLib, you must set the parameter `ORACLEASM_USE_LOGICAL_BLOCK_SIZE` to `true`.

8 Microsoft Windows Configuration

This section describes how to configure the Microsoft Windows OS on NetApp HCI. The Windows OS automatically aligns disk partitions to storage sector boundaries. No specific tuning of the partition table layout or device queue depth settings is required with Windows on NetApp HCI. To configure Microsoft Windows, complete the following tasks:

1. Enable jumbo frames on the network interface cards (NICs).
2. To identify storage devices, enable Microsoft Multipath I/O (MPIO) and configure the Microsoft device-specific module (DSM).
3. Configure at least four iSCSI sessions per storage device.

9 Summary

The NetApp HCI architecture is designed to run private cloud solutions where you need the flexibility to easily deploy and consume IT resources to support various applications and workloads. The VMware

private cloud on NetApp HCI serves multiple simultaneous applications while delivering an exceptional end-user experience. With its ability to scale compute and storage resources independently, and to connect to existing on-premises resources, the NetApp HCI solution empowers you to transform your existing environment to a private cloud operation. The VMware private cloud on NetApp HCI offers a hybrid multicloud experience to public cloud providers.

Where to Find Additional Information

To learn more about the information that is described in this document, review the following documents and websites:

- NetApp product documentation
<https://docs.netapp.com>
- TR-4639: Configuring SolidFire on Linux for Element Software
www.netapp.com/us/media/tr-4639.pdf
- NVA-1122: VMware Private Cloud on NetApp HCI
www.netapp.com/us/media/nva-1122-design.pdf
www.netapp.com/us/media/nva-1122-deploy.pdf
- TR-4748: Build Your NetApp Data Fabric
www.netapp.com/us/media/tr-4748.pdf
- TR-4643: Configuring SolidFire on Windows for Element Software
<https://www.netapp.com/us/media/tr-4643.pdf>

Acknowledgments

The author would like to thank following people for their help in writing this document:

- Jeff Steiner
- Chris Reno
- Michael Phelan
- Calvin Chinn

Version History

Version	Date		Document Version History
Version 1.0	April 2019		Initial document release.

Refer to the [Interoperability Matrix Tool \(IMT\)](#) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

Copyright Information

Copyright © 2019 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

Data contained herein pertains to a commercial item (as defined in FAR 2.101) and is proprietary to NetApp, Inc. The U.S. Government has a non-exclusive, non-transferrable, non-sublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b).

Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.