Technical Report

NDMP Cloud Extension Software

Best Practices Guide

Gerald Coon, NetApp
March 2015 | TR-4388

Abstract

This technical report outlines best practice considerations for using the NetApp® NDMP Cloud Extension software running in Amazon EC2.

TABLE OF CONTENTS

1 Introduction

To help NetApp customers select and deploy the most appropriate solutions for their IT infrastructure, NetApp provides up-to-date documentation describing its products. This technical report describes best practice considerations that you must take into account when using the NetApp NDMP Cloud Extension (NCE) software. NCE is a host-based utility that facilitates the interaction between the NetApp Data ONTAP® operating system and the Amazon Simple Storage Service (S3) cloud.

The report addresses considerations for each component of the NCE solution:

- An Amazon S3 resource that is either the source or destination of a Network Data Management Protocol (NDMP) data transfer
- NDMP services enabled on Data ONTAP (either Data ONTAP operating in 7-Mode, clustered Data ONTAP, or Cloud ONTAP®)
- A Linux® instance running the NCE software

1.1 Audience

The audience for this report is NetApp customers, partners, and employees who need to understand best practice considerations that are unique to the NCE software.

2 Amazon S3 Best Practices

This section covers considerations that pertain to Amazon Web Services (AWS).

2.1 Object-Based Storage

A bucket is a container for data objects that are stored in Amazon S3 cloud storage. Amazon S3 imposes restrictions on the number of buckets that you can create and on how the buckets are named. Keep the following limitations in mind when designing a solution that uses Amazon S3:

- Each AWS account can own up to 100 buckets.
- The buckets have very strict naming requirements, but these requirements are not enforced across all AWS regions. As a result, a bucket name might work in one region but not in another.

For detailed information about the restrictions that apply to Amazon S3 buckets, refer to Bucket Restrictions and Limitations.

2.2 AWS Connectivity Between Resources

AWS offers a wide range of networking configurations through its virtual private cloud (VPC) connectivity. Because companies normally divide out AWS resource responsibility between different groups, you must verify that VPC components are not configured in a way that impedes the network communication required for the NCE utility to function.

If you make any changes to security groups or network routes in the VPC, ensure that the communication channels between all resources are not affected. Your changes might prevent the NCE utility from completing connection tasks, and this issue can be very difficult to isolate and diagnose. Verify the following AWS networking configurations:

- Does the VPC subnet have a route to the Internet gateway if the subnet is external to the VPC resources that are involved?
- Does the VPC have an Internet gateway, or is the VPC routable to the Internet?
- Do the firewalls for Linux, Data ONTAP, and the VPC security group or the on-site WAN in a hybrid environment all allow both NDMP and the HTTPS protocol? NDMP uses port 10000 for the control

connection and negotiates the data connection port, but you can configure NCE with the `DATA_PORT_RANGE` option to limit the ports that can be used for data transfer.

Note: For more information about this NCE option, refer to the "NDMP Cloud Extension Software Installation and Administration Guide," which is available from the NDMP Cloud Extension Software (NCE) documentation library in the NetApp Support site.

2.3 AWS Credentials

When using the NCE software, keep the following information in mind to determine the best strategy for managing AWS credentials in your environment:

- You can use a configuration file with NCE to store your Amazon S3 credentials. This file must be updated if the account credentials change for any reason.
- By default, the `ndmpcopy` command, which performs the data transfer, deletes the configuration file after execution to maintain the credentials' security. You can retain the configuration file by running the `ndmpcopy` command with the `–nodeleteconf` option.
- The credentials that you specify in the configuration file must have read and write access to the Amazon S3 bucket that is the source or destination of the transfer. The NCE software cannot configure and enable access inside of Amazon S3 for these credentials.

For detailed information about AWS credential management, refer to Managing Access Permissions to Your Amazon S3 Resources.

3 Data ONTAP NDMP Best Practices

This section covers considerations that pertain to the Data ONTAP resource that is the source or destination of an `ndmpcopy` transfer.

3.1 NDMP Configuration

Data ONTAP NDMP capabilities are well documented in the "Data Protection Tape Backup and Recovery Guide" for your version of Data ONTAP. The guide is available from the Data ONTAP 8 documentation library in the NetApp Support site.

The following Data ONTAP settings are likely to affect NCE transfers:

- Clustered Data ONTAP has a firewall policy that can block NDMP access.
- Clustered Data ONTAP NDMP users are not the same as Data ONTAP users. You must create and enable an NDMP account and assign a password to it. To learn more about these tasks, consult the documentation for your version of Data ONTAP.

  Note: The Clustered Data ONTAP 8.3 NDMP Configuration Express Guide has simple and detailed instructions for how to create an NDMP account in the section titled "Configuring a Backup User for the Cluster."

3.2 Data ONTAP Version

Some but not all versions of Data ONTAP have been tested with the NCE software, but the list of tested versions will grow. Before you install NCE or change your Data ONTAP version, consult the following documents to verify which Data ONTAP versions are supported with the NCE software:

- The release notes for your version of Data ONTAP
- The release notes for NCE
- The "NDMP Cloud Extension Software Installation and Administration Guide"

Note: The NCE documentation is available from the NDMP Cloud Extension Software (NCE) library in the NetApp Support site.

3.3 Node-Scoped NDMP Mode or SVM-Scoped NDMP Mode

In clustered Data ONTAP 8.2 and later, NDMP can be configured to control data transfers either at the node level or at the storage virtual machine (SVM; formerly called Vserver) level. The NDMP mode of operation that you choose affects the logical interfaces (LIFs), credentials, and resources that can be used to copy data from and to storage running Data ONTAP.

If your NCE setup is not working as expected, verify which NDMP mode of operation is enabled in Data ONTAP. For information about the command to use for this task, refer to this NetApp Support page.

Note: For detailed information about the NDMP modes of operation, refer to the "Tape Backup and Recovery Guide" for your version of Data ONTAP. The guide is available from the Data ONTAP 8 documentation library in the NetApp Support site.

3.4 Data ONTAP Resource Naming

You can use NCE to back up resources running either Data ONTAP 7-Mode or clustered Data ONTAP:

- In Data ONTAP 7-Mode, the NDMP backup path for the resource must be in the format `/vol/volume_name`.
- In clustered Data ONTAP, the NDMP backup path for the resource must be in the format `/vserver_name/volume_name`.

  Note: In the Data ONTAP CLI, Vserver is the term used to refer to an SVM.

Using the `/vol/volume_name` format with clustered Data ONTAP causes the `ndmpcopy` transfer to fail.

4 NCE Best Practices

This section covers considerations that pertain to the Linux instance running the NCE software.

4.1 Keeping Linux Up to Date

Providing detailed instructions for how to update your version of Linux is beyond the scope of this technical report, but keep the following information in mind:

- The NCE software has dependencies on Java®. Any software that manipulates Java settings or versions should be tested to determine its impact to the NCE solution before it is deployed in a production environment. Any version of JRE later than version 1.7 update 61 is expected to work with NCE, but NetApp recommends that you verify whether NCE will be affected before you install a new Java version.
- Linux has both firewall and advanced network configuration capabilities. Before you make any changes to your firewall or network configuration, carefully consider if the changes will affect the network communications required for NCE.
- If NCE will be running on an Amazon Linux EC2 instance, you can improve network performance by enabling enhanced networking interfaces on your instance. For information about how to configure this feature, refer to Enabling Enhanced Networking on Linux Instances in a VPC.

5 NCE Performance

This section covers considerations that pertain to the performance of the NCE software. The section covers only high-level performance expectations because the actual observed performance will depend on a number of factors:

- The other Data ONTAP workloads that are processed concurrently with the `ndmpcopy` transfer
- The type of data
- The available network bandwidth
- The performance of Amazon S3

The fastest and most accurate way to gather data on the performance of your specific configuration is to try out the NCE software. Documenting all performance problems that could occur with traditional data center storage, networking, hybrid cloud connections, Amazon EC2 resources, and Amazon S3 is beyond the scope of this technical report.

5.1 Instance Running NCE Software

The NCE's function is primarily administrative work, but you must provide adequate resources to the Linux instance that runs the NCE software.

The hardware requirements for NCE have been set at 4 CPUs with 15GB of RAM. The expectation is that the CPU capabilities of the resource running the NCE utility be equivalent to the capabilities of an Amazon EC2 instance that has this same specification.

5.2 Network Bandwidth Performance

Infrastructure for Data ONTAP, Amazon EC2 (optional), and the selected Amazon S3 destination resource all can influence the performance of an NDMP backup. In addition, the type of EC2 instance that you select controls how much bandwidth is available to the host.

The amount of available bandwidth directly affects the amount of time that is required to complete a data transfer. For more information about Amazon EC2 networking, refer to the column "Networking Performance" under the section "Instance Types Matrix" on the Amazon EC2 Instances page.

5.3 NCE Scaling

NCE has been released with the expectation that virtual machines and Amazon EC2 resources are inexpensive and easily deployed. Each `ndmpcopy` transfer requires about 300MB of host memory; therefore, you may exceed the host memory capacity when you invoke multiple instances of `ndmpcopy`.

If more host memory is required, NetApp recommends that you deploy more than one virtual machine or EC2 instance to facilitate as many `ndmpcopy` instances as needed.

Version History

Version	Date	Document Version History
Version 1.0	March 2015	Initial 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.