

Dell EMC Integrated Data Protection Appliance DP4400

Scalable Data Protection for Microsoft Windows Server Software-Defined Solutions

September 2019

H17894

White Paper

Abstract

This white paper describes Dell EMC DP4400 data protection for Microsoft Windows Server Software-Defined (WSSD) Solutions. It describes DP4400 network topology best practices as well as backup and restore options. It also summarizes the functional testing that we performed to validate this data protection solution.

Dell EMC Solutions

Copyright

The information in this publication is provided as is. Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software that is described in this publication requires an applicable software license.

Copyright © 2019 Dell Inc. or its subsidiaries. All Rights Reserved. Dell Technologies, Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Intel, the Intel logo, the Intel Inside logo, and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries. Other trademarks may be trademarks of their respective owners. Published in the USA 09/19 White Paper H17894.

Dell Inc. believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

Contents

- Executive summary.....4
- IDPA DP4400 overview5
- Solution software stack6
- Dell EMC data protection components7
- Networking.....10
- Functional validation testing: S2D Ready Nodes and DP4400.....12
- Summary.....14

Executive summary

Business challenge

Organizations today are undergoing a digital transformation—examining how technology can improve all aspects of their operations in the face of increased complexity. Because IT is no longer a back-office function but rather one that is integral to the customer experience, the digital transformation requires IT transformation.

Data growth, application diversity, a growing user base, and resource constraints all contribute to the need for a modern IT infrastructure that enables organizations to do more with less. Many customers are modernizing their IT infrastructure by replacing legacy storage systems with hyperconverged infrastructure (HCI) solutions. Further, customers seek end-to-end solutions that go beyond optimizing resources to include backup/recovery and disaster-recovery data protection for their critical business assets.

Dell EMC Storage Spaces Direct Ready Nodes

Dell EMC Storage Spaces Direct (S2D) Ready Nodes are based on Microsoft's hyperconverged Windows Server 2016 cluster and Dell EMC PowerEdge servers. Microsoft has certified all the Dell EMC hardware to run virtualized workloads on premises. This paper focuses on Storage Spaces Direct Ready Nodes and how you can protect your data with the Dell EMC Integrated Data Protection Appliance (IDPA) DP4400. The DP4400 provides integrated data protection in small- and medium-size S2D Ready Node ecosystems. The DP4400 is simple to deploy and can be managed locally or remotely. For speed and efficiency, restores can be conducted locally.

Solution configurations enable maximum flexibility while including only validated and tested building blocks. Administrators can configure and tune solutions for performance levels, capacity, and availability as needed. Dell EMC provides sizing guidance, a best practices guide, and a detailed deployment guide.

Dell EMC S2D Ready Nodes extend the value proposition of software-defined storage. Dell EMC's engagement throughout the solution life cycle provides tangible benefits to customers. This end-to-end engagement starts with presales assistance with capacity and performance sizing, including solution optimization for specific workloads. It continues all the way through factory fulfillment and solution deployment. A full range of solution support offerings, ranging from 3-year, next-business-day support to mission-critical, 4-hour support worldwide, enables you to select support levels based on your business requirements.

Document purpose

This white paper provides an overview of IDPA DP4400 data protection for S2D Ready Nodes. It describes solution components and benefits, backup and restore options, and the functional testing that we performed to validate the solution.

Audience

This paper is for IT managers, administrators, and others who want to learn about the IDPA DP4400 data protection solution for small- and medium-size deployments of S2D Ready Nodes.

We value your feedback

Dell EMC and the authors of this document welcome your feedback on the solution and the solution documentation. Contact the Dell EMC Solutions team by [email](#) or provide your comments by completing our [documentation survey](#).

Author: Sam Lucido

Contributors: Karthik Konaparathi, Ray Balli, Karen Johnson

IDPA DP4400 overview

Introduction

IDPA DP4400 is a converged appliance that provides integrated data protection with complete backup and centralized management capabilities in a small 2U Dell EMC PowerEdge chassis. The DP4400 solution includes Dell EMC Avamar backup and recovery software and the Dell EMC Data Domain deduplication storage system. The combination of Avamar software and the Data Domain system uses a multistreaming approach, with Data Domain Boost software accelerating backups. This data protection solution enables reduced network bandwidth usage, backup storage, and backup times.

Backup and recovery

Avamar backup and recovery software reduces network bandwidth and backup times by using variable-length subfile segments. Each subfile segment has a unique hash identifier, which the Avamar system uses to determine if a segment has been previously backed up. The use of a unique hash identifier enables Avamar to perform source-based deduplication. Subfile segments that have been backed up are not backed up again, saving network bandwidth and reducing backup times. Unique subfile segments that have not been backed up are compressed and sent to IDPA for protection.

Extensive performance testing by Enterprise Strategy Group (ESG) demonstrated that IDPA delivers significantly reduced backup times.¹ ESG compared an IDPA solution to Vendor A in progressively protecting a mixed-workload infrastructure that increased in scale by 100 virtual machines (VMs) with each test iteration. At the 400 VM count, IDPA showed a 64 percent faster backup time than Vendor A. The faster backup times were due in part to the significant source-side deduplication. In the lab review, the IDPA solution transferred 74 percent less data over the network than Vendor A.

IDPA DP4400 also provides target-based deduplication by integrating the Data Domain deduplication storage system. The system performs up to 99 percent deduplication processing in CPU and RAM. Because performance is not based on disk, the Data Domain system can accelerate and protect more backups in parallel. The deduplication occurs inline, which means that deduplication analysis is done in real time and only unique data is stored to disk. In a cost-efficiency analysis, ESG found that storage capacity cost over a 3-year period was 57 to 81 percent less with IDPA solutions than with comparable data protection solutions.²

¹ [ESG Lab Review: Efficiently Protect Virtual Environments with Integrated Data Protection Appliances from Dell EMC](#)

² [ESG White Paper: Economic Value of Data Domain and Integrated Data Protection Appliances \(IDPA\)](#)

Fast backups are critical for S2D Ready Node ecosystems, but many businesses mandate multisite recovery. With IDPA DP4400, administrators can schedule jobs that copy backups to other data protection solutions in the IDPA family, including other Data Domain systems.

Restoring a VM is simple and fast, enabled by a few clicks in IDPA System Manager. Advanced restore capabilities include restoring backups to a new VM, determining if the VM should be powered on or off, and choosing to restore backups to a new host or cluster. You can also restore individual files from a VM.

IDPA DP4400 configuration and scaling

The IDPA DP4400 solution is a fully integrated, customer-installable appliance. It comes from the factory with much of the configuration already completed. Customers can quickly implement data protection by performing some physical, network, and license setup tasks. For example, the administrator must download the license file for Data Domain Virtual Edition, Avamar Virtual Edition, and Data Protection Advisor from Dell EMC Software Licensing Central.

DP4400 is simple to deploy and upgrade. It grows in place from 8 TB to 24 TB in 4 TB increments and from 24 TB to 96 TB in 12 TB increments. This scaling capability enables organizations to start small and add capacity as their S2D Ready Node infrastructure grows. Expanding the solution requires license keys only. Expansion does not require additional hardware or downtime, so the IT organization can quickly react to rapidly expanding protection requirements.

Solution software stack

IDPA DP4400 is a fully integrated solution that comes with life cycle management and currently works with Windows Server 2016. The software component stack includes the following components:

- Appliance Configuration Manager
- Avamar Virtual Edition
- Cloud Disaster Recovery
- Cloud Long-Term Retention
- Data Domain Virtual Edition
- Data Protection Advisor
- Data Protection Search
- IDPA System Manager
- VMware vSphere

This is a controlled software stack that is maintained by Dell EMC. Dell EMC tests patches and updates and provides them to customers, eliminating the complexities of having to upgrade individual components.

Dell EMC data protection components

Appliance Configuration Manager

Centralized management that simplifies protection of virtualized applications is at the core of this data protection solution. The Appliance Configuration Manager (ACM) provides a graphical, web-based interface for configuring, monitoring, and upgrading the solution. ACM provides comprehensive dashboards that include information about system capacity, backup and replication activities, health, and alerts. Administrators can manage all system backup activity from the ACM web interface.

Avamar Virtual Edition

Avamar Virtual Edition is a single-node Avamar server that enables full backups of:

- Virtualized and physical environments
- Enterprise applications
- NAS systems
- Remote offices
- Desktops and laptops

Avamar Virtual Edition offers a great deal of flexibility for protecting physical, vSphere, and Hyper-V environments. Depending on the application requirements, an administrator can choose a guest backup or an image-level backup, both of which offer extensive protection capabilities. For virtualized applications, Avamar Virtual Edition offers plug-ins for Microsoft Hyper-V that simplify protecting applications, meeting SLAs, and restoring data.

The Avamar server runs as a VM on the DP4400 solution and seamlessly integrates with Data Domain Virtual Edition. For physical client backup and restore, install Avamar client software on each server. For VM backup and restore for Hyper-V, choose between guest-operating-system and image-level options, as described in the following sections.

Avamar Virtual Edition for Microsoft Hyper-V

Avamar Virtual Edition provides the following two methods for backing up and restoring VMs that are running on Hyper-V:

- **Guest-operating-system backup and restore**—Requires installation of Avamar client software on each VM, enabling administrators to independently customize the backup of each VM. The Avamar plug-in enables the protection of data from multiple applications, including Microsoft Exchange, Microsoft SharePoint, IBM Db2, Oracle Database, and Microsoft SQL Server.
- **Image-level backup and restore**—Requires Avamar Client for Windows and the Avamar plug-in for Hyper-V Volume Shadow Copy Service (VSS) on the Hyper-V management operating system (parent partition). The host-based, image-level approach enables backups and restores of an entire image or specific drives of Microsoft Windows or Linux VMs. Using image-level backups, administrators can perform full backups of the VMs or the host component. The host component contains the role-based security configuration for Hyper-V. Unlike guest backups where backups occur in the VM, the image-level backups run from the Hyper-V management operating system.

The Avamar plug-in for Hyper-V VSS uses Microsoft’s volume snapshot service technology to perform backups. With this technology, administrators can create backup copies or snapshots of files or volumes, even while the files or volumes are in use. The VSS writer enables quiescing of the Windows operating system and applications to ensure data consistency during snapshot creation.

The following table describes some of the key differences between the two backup and restore options:

Table 1. Comparison of backup and restore methods for VMs running on Hyper-V

Criterion	Guest-operating-system backup and restore	Image-level backup and restore
Avamar software installation	Avamar file system client and optional application plug-in on each VM.	Avamar Client for Windows and the Avamar plug-in for Hyper-V VSS on only the management operating system.
Deduplication	Data within each VM.	Data at the image level.
Application consistency	Yes.	<ul style="list-style-type: none"> Supported if the guest operating system is Windows Hyper-V and integration components are installed in the guest operating system. Not supported if the guest operating system does not support VSS or if the guest operating system is Microsoft Windows and Hyper-V integration components are not installed in the guest operating system.
Application-aware backup and restore	Yes, through Avamar plug-ins for applications.	No.
VM status for backup	Must be running.	Does not need to be running.
Backup of unused file system space	No.	Yes.
Disaster recovery	Two-step recovery: <ol style="list-style-type: none"> 1. Load a known good operating-system image on the VM. 2. Restore backup data from Avamar software. 	One-step restore of backup data from Avamar. However, backups might be a “crash-consistent” snapshot of the full VM image, which might not reliably support a full system restore without data loss.

Cloud Disaster Recovery

Cloud Disaster Recovery (Cloud DR) enables recovery of one or more of your protected VMs to a cloud provider. Supported clouds include Amazon Web Services (AWS), AWS GovCloud, and Microsoft Azure. Administrators can create runbooks that define disaster plans to recover multiple VMs to the cloud. The Cloud DR runbook offers features such as

recovery orchestration definition, including network and security groups association, VM boot-order definition, and instance-type selection. The DP4400 solution manages Cloud DR and runbooks, so administrators can easily and efficiently implement and monitor cloud recovery.

Cloud DR also provides the option of creating a rapid-recovery copy of a protected VM to enable the fast recovery of the VM to the cloud. Creating a rapid-recovery copy starts with a rehydrating process that converts the VMDK files to the cloud provider's required format. The cloud administrator then needs to run only the recovered VM. The rapid-recovery copy feature speeds the process of moving applications to the cloud during a planned outage.

Many IT organizations must annually test and validate their disaster recovery plans. Cloud DR can recover a VM to the cloud as part of disaster recovery testing. It enables temporary access to a virtual cloud instance to retrieve data or verify that the recovery VM is working before running a failover.

Cloud DR also supports moving a recovered production application in the cloud back to the data center. From the DP4400 appliance, the administrator directs Cloud DR to copy the cloud application to a new on-premises VM. The administrator can monitor the failback process from the activity screen and verify that the production VM is again running within the data center.

Cloud Long-term Retention

The IDPA DP4400 solution includes Cloud Long-term Retention (Cloud LTR), which enables IT organizations to push backups to the cloud for long-term retention and protect up to 14.4 PB of data. The administrator can natively tier deduplicated backups to a private, public, or hybrid cloud with no need for a separate cloud gateway. Pushing backups to the cloud can provide a total-cost-of-ownership benefit to customers. For example, Amazon Glacier and Dell EMC Elastic Cloud Storage (ECS) can provide lower long-term retention costs than maintaining backups within the data center.

Data Domain Virtual Edition

Data Domain Virtual Edition is a software-defined data protection solution that is integrated into IDPA DP4400. Data that is sent to the DP4440 system is also deduplicated at the Data Domain level. This capability for target-side deduplication is a function of the Data Domain global deduplication engine. The global deduplication engine uses Stream-Informed Segment Layout (SISL), which assigns a unique hash identifier to each 8 KB (or larger) segment. Variable-length segmentation optimizes deduplication rates by breaking up incoming streams based on the data structure. The ability to break up data streams increases parallelization and deduplication efficiency.

The analysis of the incoming data streams occurs in RAM, which means that the Data Domain system can analyze large groups of data to maintain consistently fast backups. In addition to using memory, the Data Domain system uses a mature and proven deduplication algorithm that is designed to optimize performance and capacity. The key benefits are that similar VMs can have the same data and redundant data is deduplicated.

Real-time analysis of incoming data in the Data Domain system is referred to as inline deduplication. With inline deduplication, only unique data is sent to and stored on disk. Inline deduplication eliminates the need for a disk staging area and compute resources for post-process deduplication. The Data Domain Data Invulnerability Architecture protects the data

that is written to disk. All data that is read from and written to the Data Domain system is verified, which protects against data integrity issues. Continuous fault detection and self-healing ensure that data remains recoverable throughout its life cycle on the DP4400 solution.

Data Protection Advisor

As the number of applications grows, IT organizations look for ways to proactively monitor and analyze their data protection environment. Dell EMC Data Protection Advisor provides a single, comprehensive view of the data protection environment and activities. It provides automated monitoring and alerting across the backup and recovery infrastructure, replication technologies, storage platforms, enterprise applications, and virtual environment.

Data Protection Advisor includes robust reporting functionality that enables administrators to identify and diagnose issues, mitigate risks, and forecast future trends. Administrators can use the included report templates or create additional reports to meet the organization's requirements.

Networking

Introduction

Networking is a key component of backup over the company network. Considerations include network throughput and bandwidth. Throughput refers to the amount of data that is sent and received within a specific timeframe. Network administrators monitor throughput, usually in bits per second (b/s), because it indicates how many packets are arriving at the destination. In the case of the DP4400 solution, network throughput determines how fast data can be transferred, which affects backup and recovery times.

Network bandwidth refers to the amount of data that can be sent and received at one time. The greater the bandwidth the more data that can be sent and received. Today many networks are 1 Gb/s or 10 Gb/s, but, with the rapid pace of network innovation, faster networks are the future. Because bandwidth determines the amount of data that can be sent and received, it impacts how many backups or recoveries can run in parallel.

DP4400 has four 10 GbE network ports to connect to a customer's network. As part of the network setup, the DP4400 solution comes with the Network Validation Tool (NVT). The NVT enables customers to run multiple tests to validate the network configuration prior to the DP4400 installation. The goal of the NVT is to facilitate and accelerate the deployment of the DP4400 solution in the customer's network.

DP4400 in an S2D Ready Node converged topology

DP4400 supports only a flat network, meaning that the network is not segmented or separated to isolate traffic. The backup network is configured on the DP4400 10 GbE rack Network Daughter Cards (rNDCs) using Switch Embedded Teaming (SET). SET, which was introduced in Windows Server 2016, enables multiple network adapters to be joined within the vSwitch. Up to eight network adapters can be joined, but they must be identical—the same make, model, driver, and firmware. SET enables the network to support both host and VM application backups.

The use of dedicated network cards on the hosts and the DP4400 appliance means that Quality of Service (QoS) configuration is not needed. The following figure shows a network configuration example:

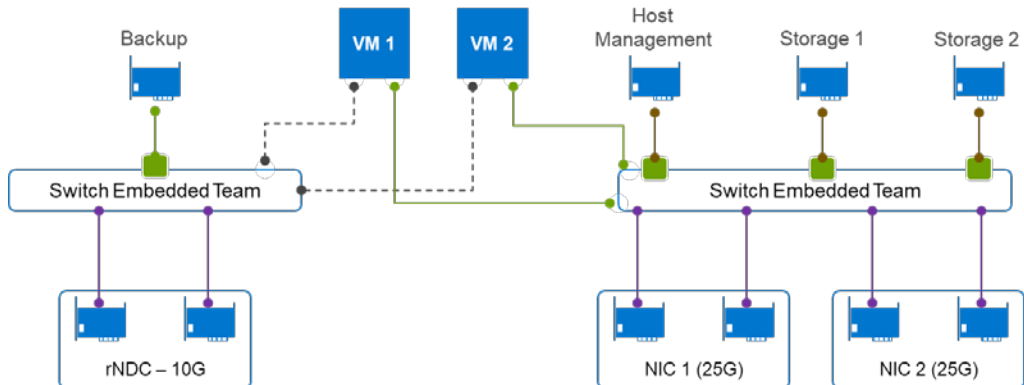


Figure 1. Example of network configuration with Switch Embedded Teaming

DP4400 in an S2D Ready Node nonconverged topology

The storage network is configured on physical adapters in a nonconverged S2D Ready Node environment while rNDCs are used for management and VM traffic. Because management and VM traffic do not need flow control, you can configure a backup network on the same SET team, as shown in the following diagram. You can configure bandwidth shaping to ensure appropriate QoS for egress traffic from the host.

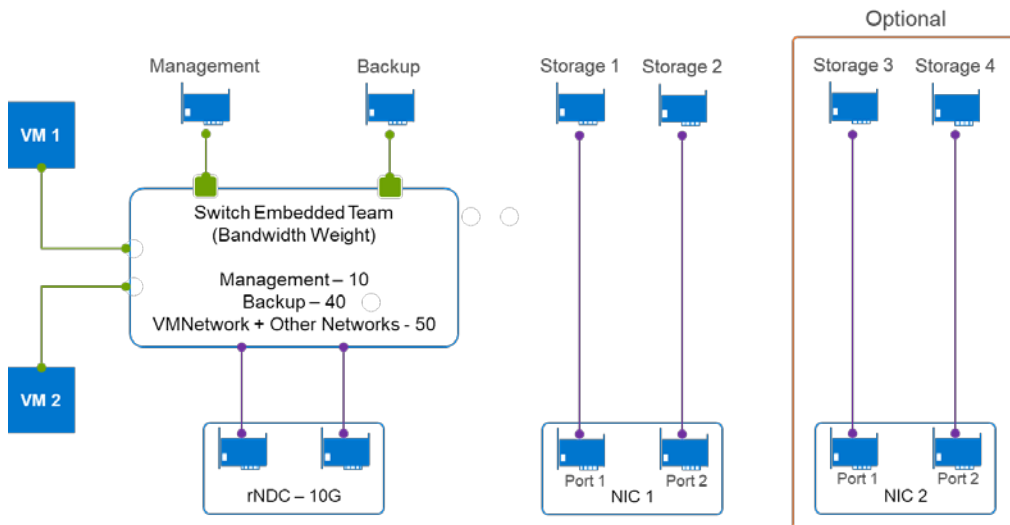


Figure 2. Example of backup network configuration with Switch Embedded Teaming

The following table shows the minimum bandwidth that is required for each network:

Network	Minimum bandwidth weight
Management	10
Backup	40
VM network/other (default minimum bandwidth)	50

Dell EMC offers ProDeploy Plus for your data protection needs. A recent study by [Principled Technologies](#) found that customers using ProDeploy services spent 74 percent less time on in-house administration and 85 percent less time on project planning. Dell EMC network experts can work with your teams to validate network requirements, reduce risk, and accelerate time-to-value.

Functional validation testing: S2D Ready Nodes and DP4400

Introduction

Using S2D Ready Nodes with Dell EMC servers with internal drives provides a software-defined storage solution at a fraction of the cost of traditional SAN or NAS arrays. Microsoft first released its software-defined storage solution, Storage Spaces Direct, with Windows Server 2016. In 2019, to reflect the integration of Storage Spaces Direct with Azure Stack, Microsoft changed the Storage Spaces Direct name to Azure Stack HCI. Azure Stack HCI is experiencing rapid growth in popularity, with more than 10,000 clusters worldwide running within the first 18 months of general availability.³

Note: At the time of this writing, DP4400 supports Windows Server 2016 only.

The rapid growth of the platform has challenged many customers to find a protection solution that enables the IT organization to start with a right-sized solution and then grow incrementally. Using Dell EMC's [Microsoft Storage Spaces Direct Ready Nodes](#) (now Azure Stack HCI Ready Nodes), our engineering lab ran a group of data protection tests to assess whether pairing the DP4400 appliance with Ready Nodes provides an optimal protection solution for VMs. Dell EMC S2D Ready Nodes offer the following benefits:

- **Confidence**—The nodes are preconfigured with components that are tested by Dell EMC and certified by Microsoft.
- **Convenience**—The nodes make S2D adoption convenient and reduce administrative work with fewer interfaces, fewer steps, and less need for specialized knowledge.
- **Customer support**—The nodes are supported by Dell EMC, which means that customers have only one vendor to call.

In our Dell EMC labs, we created a two-node S2D Ready Node cluster. We used the Microsoft Cluster Shared Volumes (CSV) file system, which enables multiple nodes to read and write to the same LUN (disk). CSV provides resiliency in the case of a node failure. The clustered roles can fail over quickly to a surviving node.

We also used Resilient Change Tracking (RCT), which increases backup reliability, scalability, and performance. RCT is a native Change Block Tracking (CBT) feature that was introduced in Windows Server 2016. RCT accelerates backups by providing a way to query a virtual hard disk for changes since the last backup. For every virtual hard disk, two files are created—an .rct file and an .mrt file. Backup solutions such as IDPA DP4400 with Avamar software can use these two files to determine the unique blocks to back up.

³ [Storage Spaces Direct: 10,000 Clusters and Counting!](#)

From the DP4400 appliance, we installed the Avamar Client plug-in for Windows on each of the two nodes and the Avamar Client plug-in for Hyper-V VSS on the management node. With these two Avamar plug-ins, we ran image-level backups and restores of VMs in the cluster to validate that the backup and restore of the VMs would work as expected. We then ran additional tests to validate file-level and folder-level restores to the backup cluster and VM as well as to a new cluster and VM. Across all seven tests, the DP4400 solution successfully backed up and restored VMs, files, and folders. The backup and restore procedures in our experience were intuitive, which is important when the backup administrator must recover from an unexpected data loss.

The DP4400 plug-in for Hyper-V streamlined our backup and restore procedures. Both activities required only two to four mouse clicks. We monitored progress across all the tests to validate success. The ability to monitor progress, particularly with restores, is important because the business can be promptly notified when a restored VM is available.

Backup and restore tests and results

This section summarizes our solution backup and restore tests and test results.

Test 1: Full image-level backup

For the first test, we ran a full image-level backup on a VM. The DP4400 system processed the backup request, and the Avamar software protected the entire VM. The test was completed successfully with no issues.

Test 2: Incremental image-level backup

In this test, we built upon the full image-level backup test by performing an incremental backup. Avamar software used Windows Server RCT to accelerate backing up only the changed blocks. The test was completed quickly and successfully with no issues.

Test 3: Full image-level restore of a single VM to its original location

Using the full image-level backup, we restored the VM to its original node. We monitored progress in the activity screen and further validated success by logging in to the restored VM.

Test 4: Multiple image-level restores of VMs to their original location

In this test, we simulated a runbook in which multiple VMs had to be restored to the original cluster. Restoring multiple VMs worked as expected with no errors.

Test 5: Full image-level restore of a single VM to a new location

The goal of this test was to simulate the loss of a cluster and the recovery of a VM to a new cluster. We restored the single VM to the new cluster without any errors or additional complexity.

Test 6: File- and folder-level restore on the same cluster and in the same VM

In this test, we simulated the loss of files and folders within a VM. We restored the files and folders without any errors. This test shows the granular ability to restore files and folders from an image-level backup.

Test 7: File- and folder-level restore to a new cluster and VM

We ran this test to ensure that we could restore important files and folders to another VM on a different cluster. This final test was completed successfully. No errors occurred during the restore of the files and folders.

Summary

HCI enables organizations to respond quickly and effectively to new digital opportunities. It enables IT teams to use modern software-defined solutions that can reduce costs and simplify deployments and scalability. Dell EMC S2D Ready Nodes, based on Dell EMC PowerEdge servers, are validated and tested solutions that can be configured as needed for performance, capacity, and availability.

The IDPA DP4400 data protection solution for small- to medium-size S2D Ready Node ecosystems is a compact 2U appliance. It incorporates data protection software components that significantly reduce backup storage requirements, backup time, and network bandwidth usage. DP4400 is a comprehensive solution that includes, among other components, Avamar Virtual Edition backup and recovery software and the Data Domain Virtual Edition deduplication storage system. With a graphical, web-based interface, administrators can easily configure, monitor, and upgrade the solution, either locally or remotely. Organizations can begin with 8 TB and, through licensing alone, expand capacity up to 96 TB without adding hardware.

Dell EMC labs performed functional testing of the DP4400 solution with Dell EMC S2D Ready Nodes to validate the solution's capabilities and ease of use. The DP4400 system successfully performed image-level backups and restores as well as folder-level and file-level restores.

For more information about this solution, see [Dell EMC and Microsoft Azure: Storage and Data Protection for Today's Cloud-Ready Business](#). For more information about Avamar software for Hyper-V, see the [Dell EMC Avamar for Hyper-V User Guide](#).