Module 4

Implementing Storage Spaces and Data Deduplication

Module Overview

- Implementing Storage Spaces
- Managing Storage Spaces
- Implementing Data Deduplication

Lesson 1: Implementing Storage Spaces

- Enterprise storage needs
- What are Storage Spaces?
- Components and features of Storage Spaces
- Demonstration: Configuring Storage Spaces
- Changes to file and storage services in Windows Server 2016
- Storage Spaces usage scenarios
- Discussion: Comparing Storage Spaces to other storage solutions

Enterprise storage needs

In your storage planning, you should assess whether your storage solution needs to support capabilities such as:

- Mirror/parity support
- Data stripping
- Enclosure awareness
- Storage tiering
- Storage replication
- Data deduplication
- Data encryption
- Performance analysis

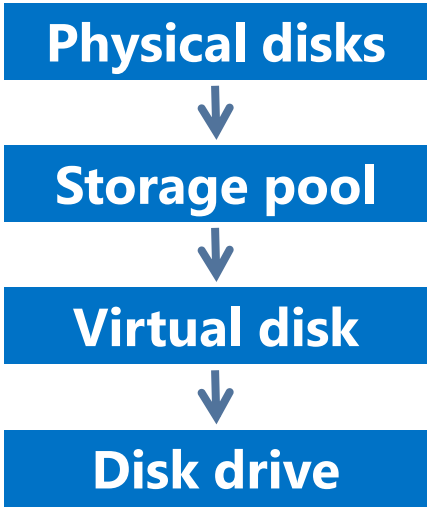
What are Storage Spaces?

Use Storage Spaces to:

- Add physical disks of any type and size to a storage pool
- Create highly-available virtual disks from the pool

To create a virtual disk, you need:

- One or more physical disks
- A storage pool that includes the disks
- Virtual disks (or storage spaces) that are created with disks from the storage pool
- Disk drives that are based on virtual drives



Components and features of Storage Spaces

Features to use Storage Spaces optimally in your environment

Feature	**Options**
Storage layout	Simple Two-way or three-way mirrors Parity
Disk sector size	512 or 512e 4 KB
Drive allocation	Data-store Manual Hot spare
Provisioning schemes	Thin provisioning space Fixed provisioning space
Stripe parameters	Number of columns Interleave

Demonstration: Configuring Storage Spaces

In this demonstration, you will learn how to:

- Create a storage pool
- Create a virtual disk and a volume

Storage Spaces usage scenarios

Storage Spaces features:

- Implement and easily manage scalable, reliable, and inexpensive storage
- Use inexpensive storage with or without external storage
- Combine multiple drives into storage pools that administrators can manage as a single entity
- Implement different types of storage in the same pool
- Grow storage pools as required
- Provision storage as required from existing storage pools
- Designate specific drives as hot spares

Storage Spaces usage scenarios

Limitations of Storage Spaces:

- Storage Spaces volumes are not supported on boot or system volumes
- The contents of a drive are lost when you introduce that drive into a storage pool
- You must have at least one drive in a simple storage pool
- Fault tolerant configurations have specific requirements:
  - A mirrored pool requires a minimum of two drives
  - Parity requires a minimum of three drives
  - Three-way mirroring requires a minimum of five drives
- All drives in a pool must use the same sector size
- Fibre Channel and iSCSI are not supported

Storage Spaces usage scenarios

When planning your Storage Spaces solution, consider:

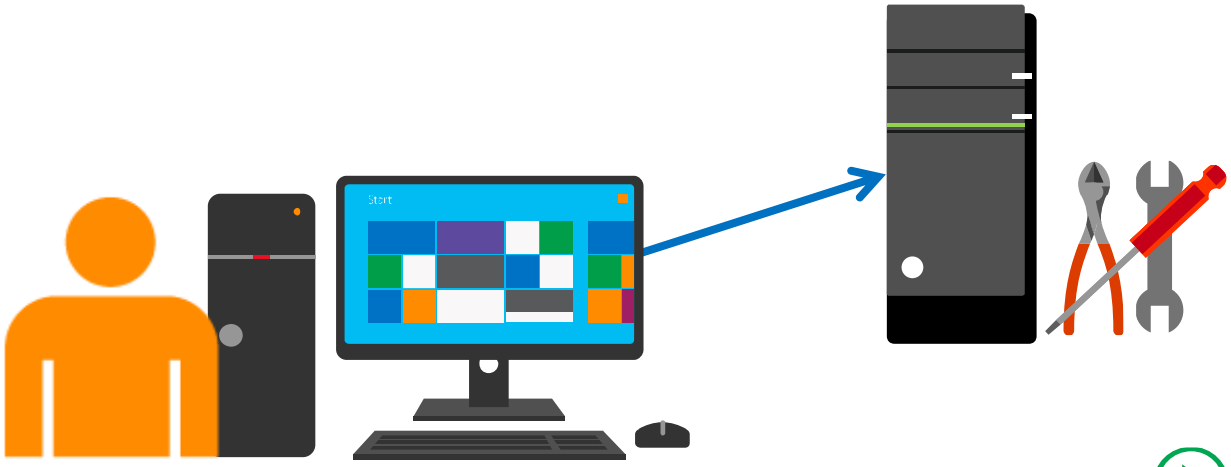
- Fault tolerance:
  - Provision virtual disks with mirrored and parity layout
- Performance:
  - Provision virtual disks with parity layout
  - Use disks of different types to provide tiered storage
- Reliability:
  - Use hot spare physical disks in case a physical disk fails
- Extensibility:
  - Add physical disks to a storage pool

Lesson 2: Managing Storage Spaces

- Managing Storage Spaces
- Managing disk failure with Storage Spaces
- Storage pool expansion
- Demonstration: Managing Storage Spaces by using Windows PowerShell
- Event logs and performance counters

Managing Storage Spaces

- You can manage Storage Spaces by using:
  - Server Manager
  - Windows PowerShell
  - Failover Cluster Manager
  - System Center Virtual Machine Manager
  - Windows Management Instrumentation
- Advanced management requires Windows PowerShell



Managing Storage Spaces

Windows PowerShell cmdlet	Description
Get-StoragePool	List storage pools
Repair-VirtualDisk	Repair a virtual disk
Get-PhysicalDisk | Where {$_.HealthStatus -ne "Healthy"}	List unhealthy physical disks
Reset-PhysicalDisk	Remove a physical disk from a storage pool
Get-VirtualDisk | Get-PhysicalDisk	List physical disks used for a virtual disk
Optimize-Volume	Optimizes a volume, performing such tasks on supported volumes and system SKUs as defragmentation, trim, slab consolidation, and storage tier processing

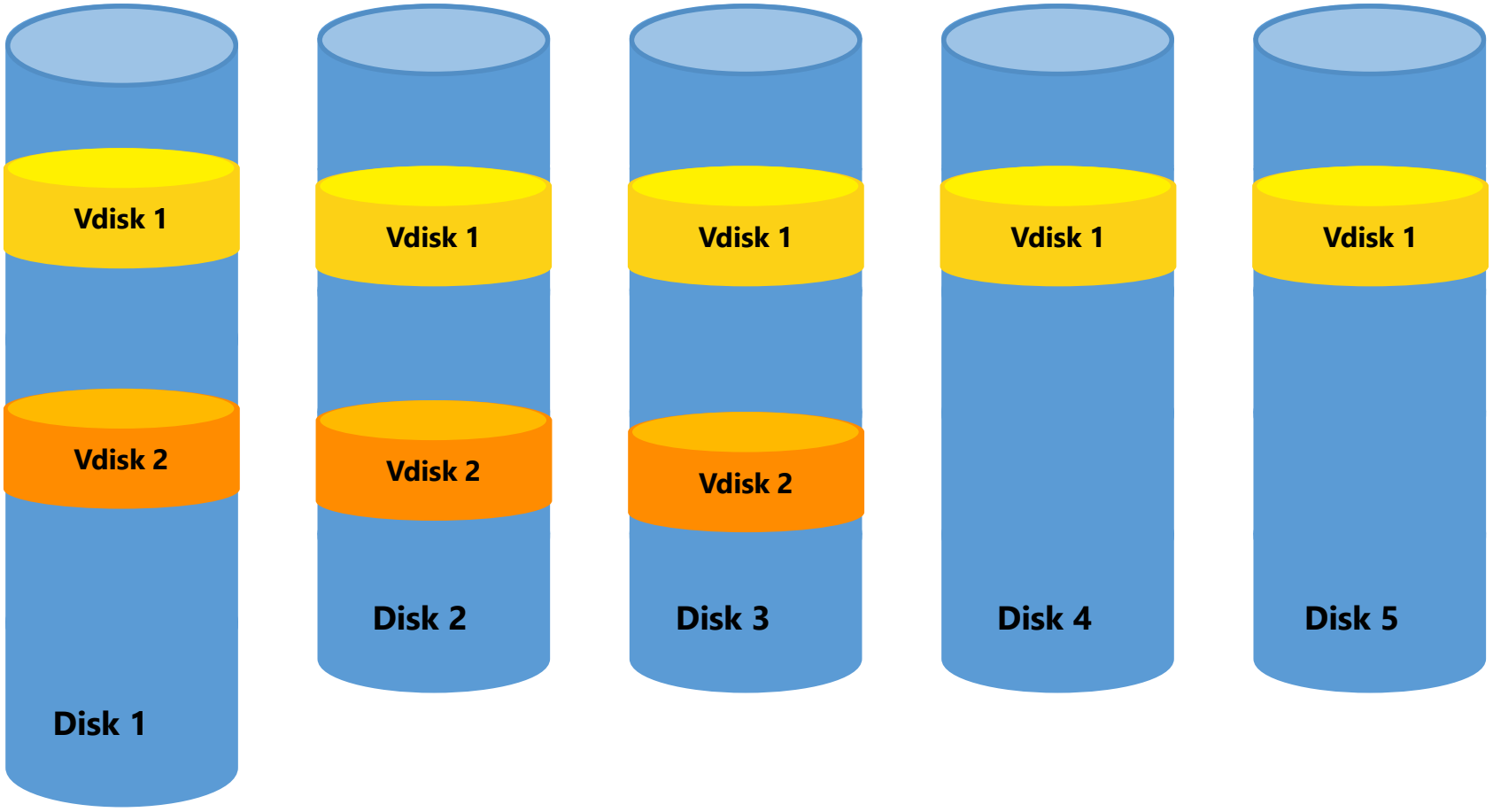
Managing disk failure with Storage Spaces

To enhance disk fault tolerance:

- Design a complete, fault-tolerant storage solution
- Deploy a highly available storage pool
- Verify hardware and firmware components
- Replace failed disks immediately
- Retain some unallocated space
- Be prepared for multiple disk failures
- Provide fault tolerance at the enclosure level

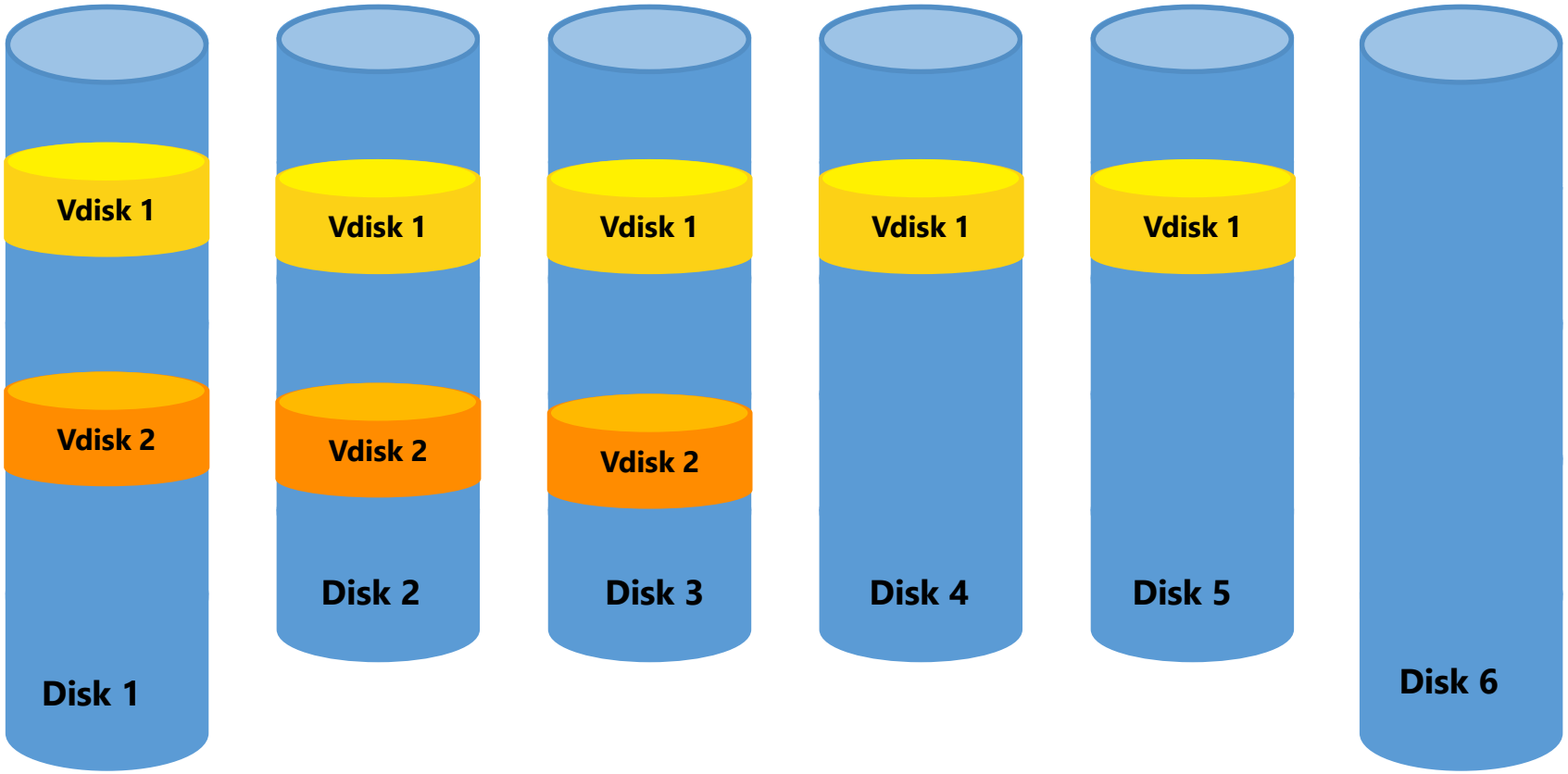
One of the main benefits of using Storage Spaces is **the ability to expand your storage pool** by adding additional storage.

Storage pool expansion



If you attempt to extend vdisk1, the maximum available space for that disk has already been used, even though more space is available within the pool on disk 6. This is because the layout required by vdisk1—due to the options chosen at creation (such as mirroring and parity)—needs five disks.
Therefore, to expand vdisk1, you would need to add four additional disks.

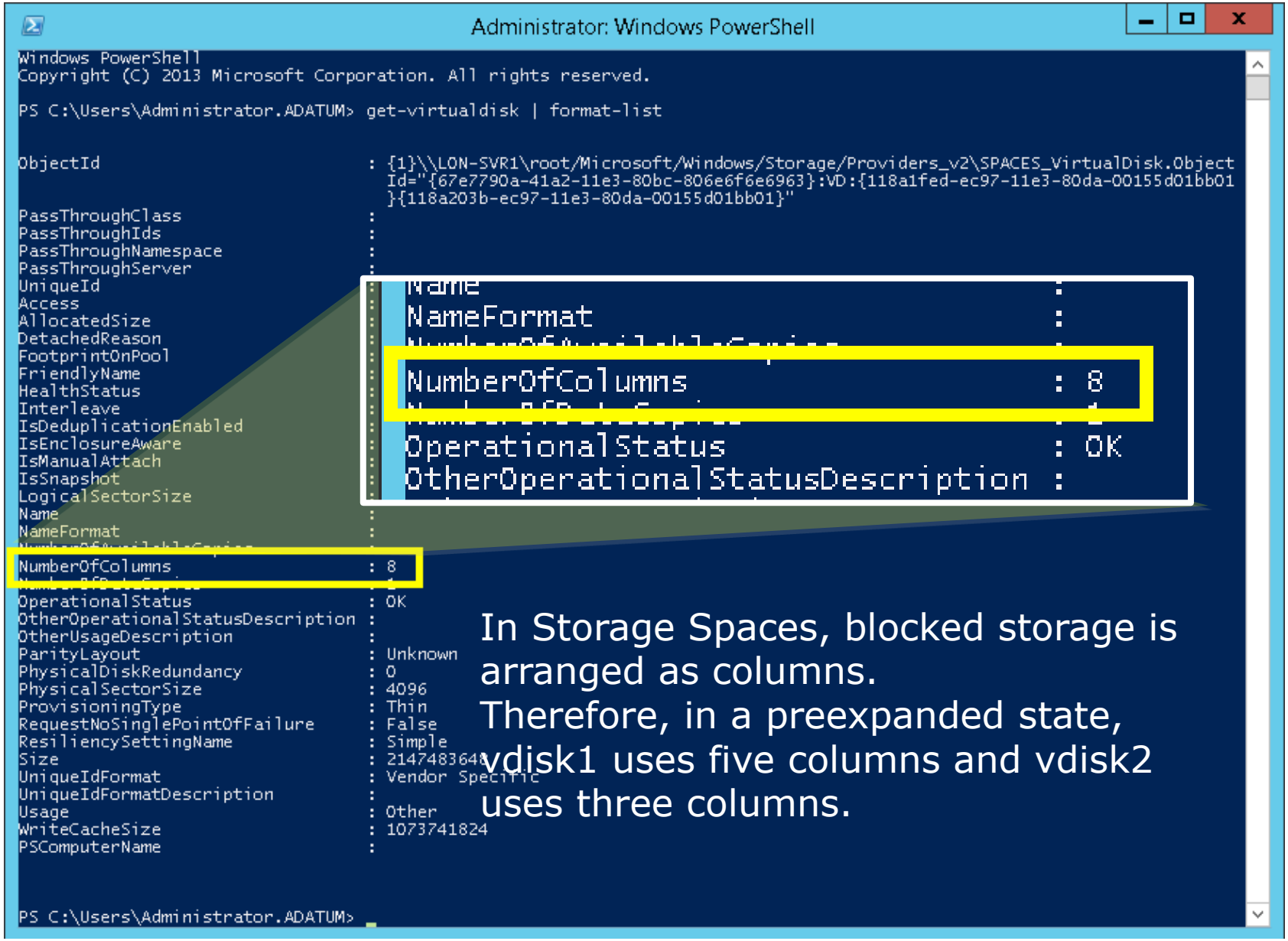
Storage pool expansion



However, if you attempt to extend vdisk2, you can do so because that disk is currently distributed across three devices and there is available space across those three devices to extend it.

Vdisk2 might just be a virtual disk that used two-way mirroring. This means that data on disk1 is duplicated on disk2 and disk3. If you wish to expand a virtual disk with two-way mirroring, it has to have the appropriate number of columns available to accommodate the needs of the virtual disk

Storage pool expansion



In Storage Spaces, blocked storage is arranged as columns.
Therefore, in a preexpanded state, vdisk1 uses five columns and vdisk2 uses three columns.

Demonstration: Managing Storage Spaces by using Windows PowerShell

In this demonstration, you will learn how to use Windows PowerShell to:

- View the properties of a storage pool
- Add physical disks to a storage pool

Lab A: Implementing Storage Spaces

- Exercise 1: Creating a Storage Space
- Exercise 2: Enabling and configuring storage tiering

Logon Information

Virtual machines: **20740C-LON-DC1**
20740C-LON-SVR1
User name: **Adatum\Administrator**
Password: **Pa55w.rd**

Estimated Time: 40 minutes

Lab Scenario

Adatum corporation has purchased a number of hard disk drives and SSDs and you have been tasked with creating a storage solution that can utilize these new devices to the fullest. With mixed requirements in Adatum for data access and redundancy, you must ensure that you have a redundancy solution for critical data that does not require fast disk read and write access. You also must create a solution for data that does require fast read and write access.

You decide to use Storage Spaces and storage tiering to meet the requirements.

Lab Review

- At a minimum, how many disks must you add to a storage pool to create a three-way mirrored virtual disk?
- You have a USB-attached disk, four SAS disks, and one SATA disk that are attached to a Windows Server 2012 server. You want to provide a single volume to your users that they can use for file storage. What would you use?

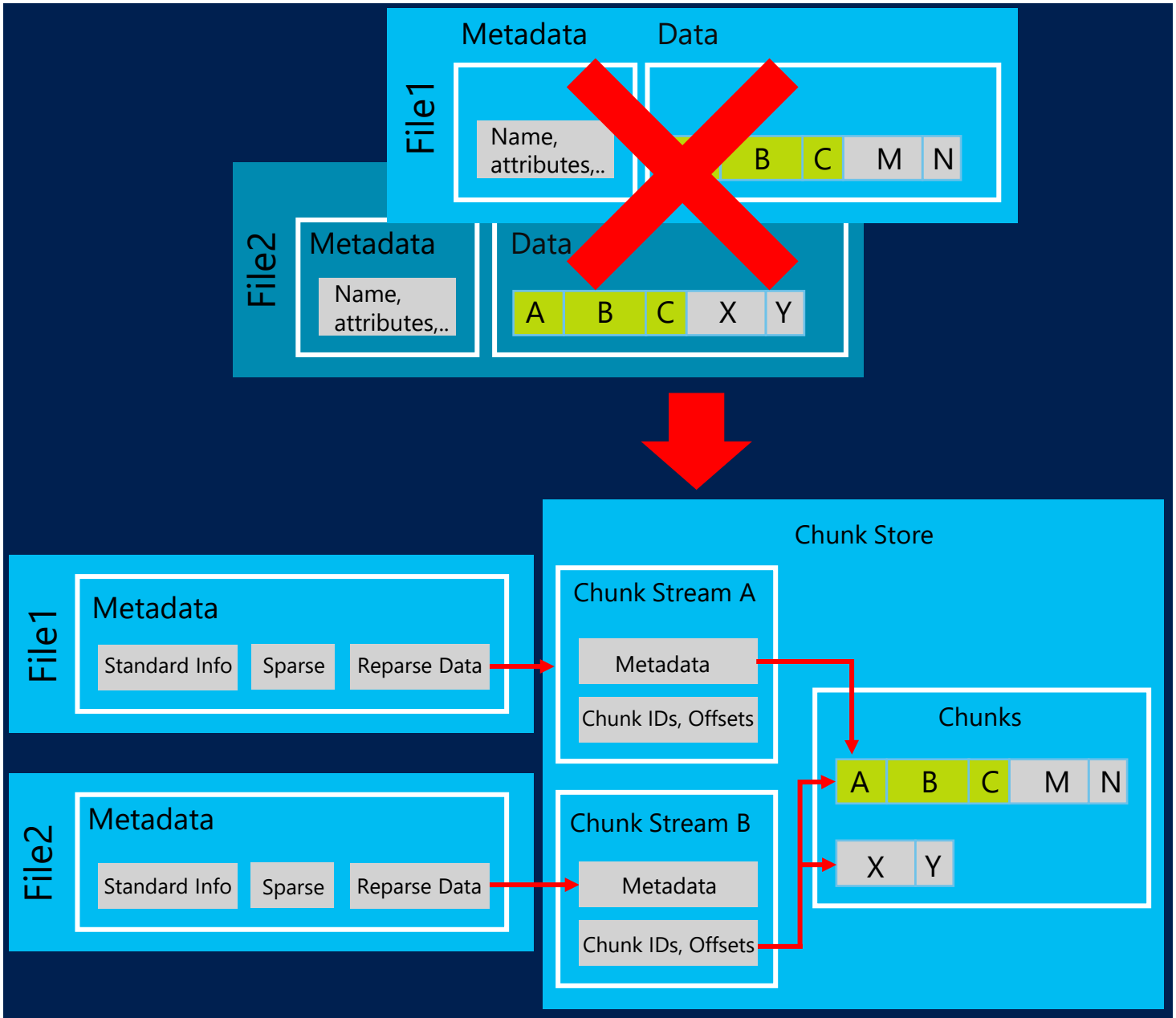
Lesson 3: Implementing Data Deduplication

- What is Data Deduplication?
- Data Deduplication components
- Deploying Data Deduplication
- Demonstration: Implementing Data Deduplication
- Usage scenarios for Data Deduplication
- Monitoring and maintaining Data Deduplication
- Backup and restore considerations with Data Deduplication

What is Data Deduplication?

- Data Deduplication:
  - Identifies and removes duplications within data without compromising the data's integrity or fidelity
  - Has the goal to store more data on less space
- When you enable Data Deduplication **on a volume**, a background task runs with low-priority that:
  - Segments data into small, variably-sized chunks
  - Identifies duplicate chunks
  - Replaces redundant copies with a reference
  - Compresses chunks

What is Data Deduplication?



Data Deduplication components

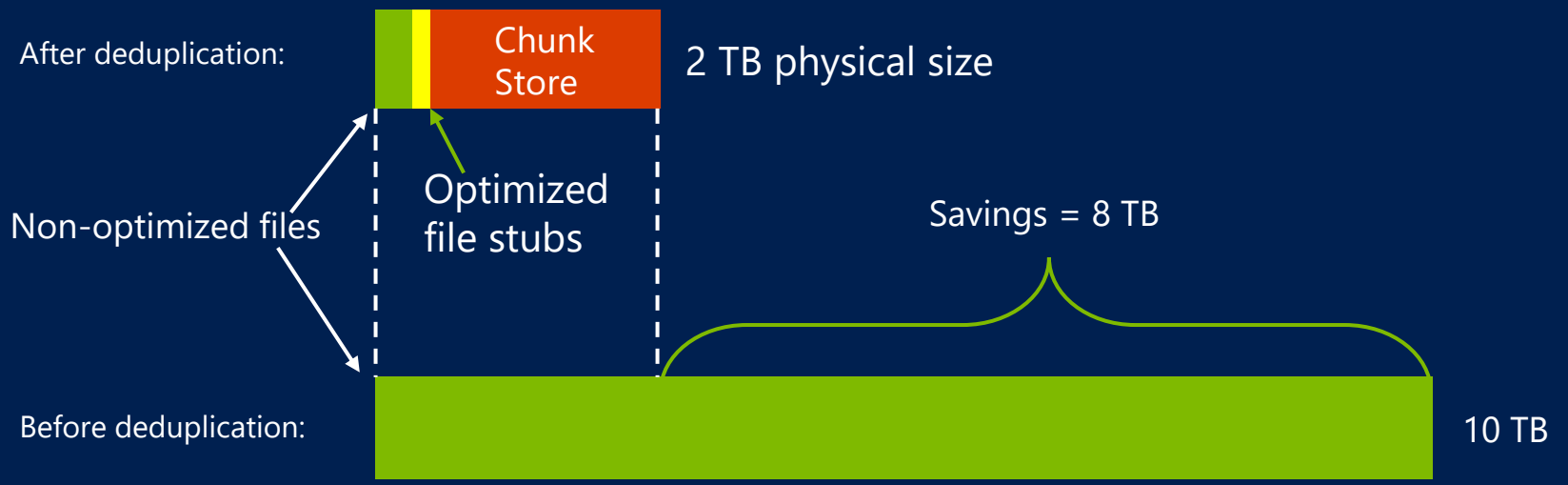
The Data Deduplication feature consists of several components:

- A filter driver, which monitors local or remote I/O
- The Data Deduplication role service, which controls the three available job types:
  - Optimization
  - Garbage Collection
  - Scrubbing

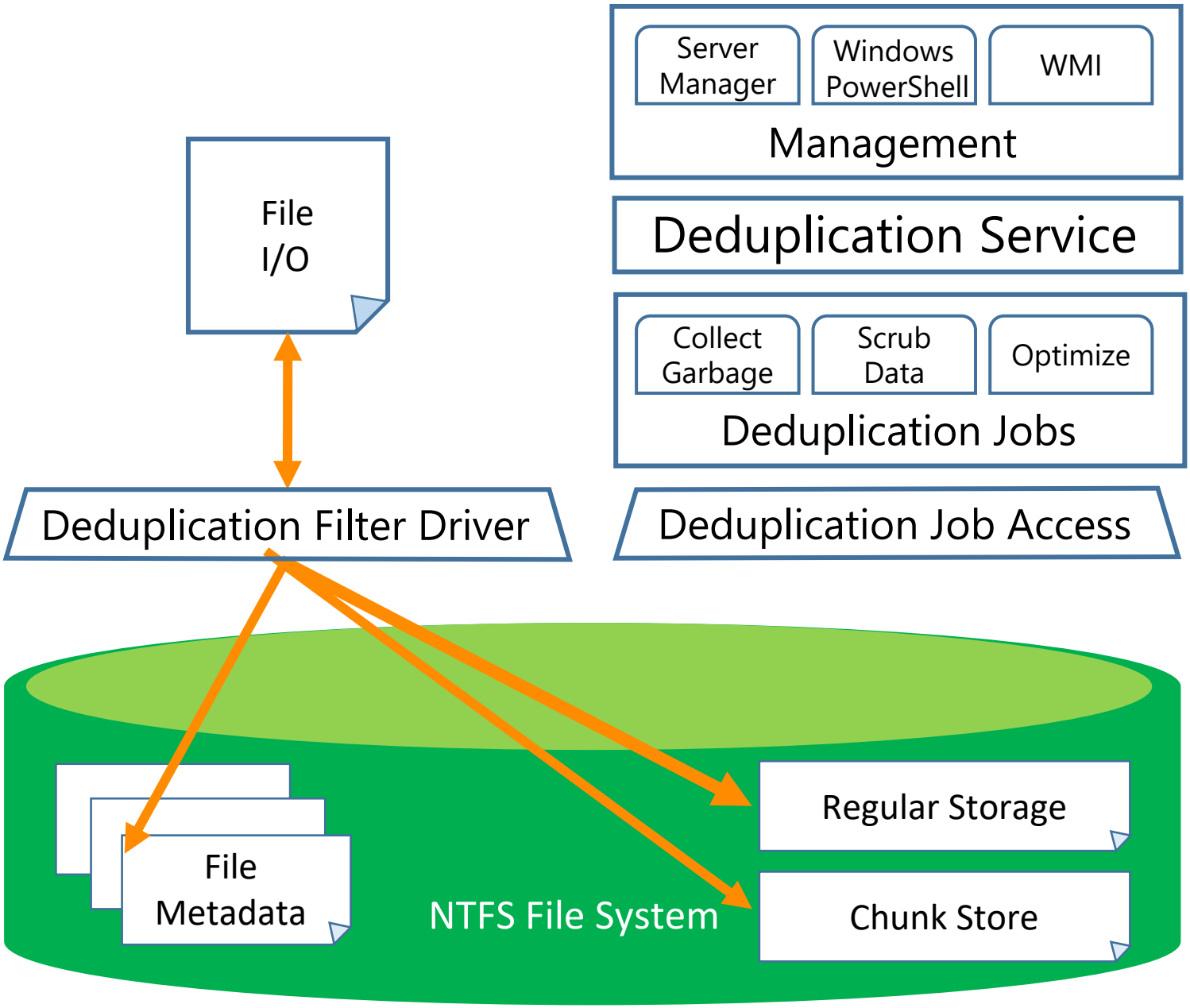
Data Deduplication components

Data Deduplication disk space savings

A view of optimized data (80% savings)



Data Deduplication components



Deploying Data Deduplication

Prior to installing and configuring Data Deduplication in your environment, you need to plan your deployment using the following steps:

- Target deployments
- Determine which volumes are candidates for deduplication
- Evaluate savings with the Deduplication Evaluation Tool
- Plan the rollout, scalability, and deduplication policies

Deploying Data Deduplication

After completing your planning, you need to use the following steps to deploy Data Deduplication to a server in your environment:

- **Install** Data Deduplication components on the server
- **Enable** Data Deduplication
- **Configure** Data Deduplication jobs
- Configure Data Deduplication schedules

Demonstration: Implementing Data Deduplication

In this demonstration, you will learn how to:

- Install the Data Deduplication role service
- Enable Data Deduplication
- Check the status of Data Deduplication

Usage scenarios for Data Deduplication

Consider using Data Deduplication for the following

File shares

Software deployment shares

.vhd libraries

Lab B: Implementing Data Deduplication

- Exercise 1: Installing Data Deduplication
- Exercise 2: Configuring Data Deduplication

Logon Information

Virtual machines: **20740C-LON-DC1**
20740C-LON-SVR1
User name: **Adatum\Administrator**
Password: **Pa55w.rd**

Estimated Time: 40 minutes

Lab Scenario

After you have tested the storage redundancy and performance options, you decide that it also would be beneficial to maximize the available disk space that you have, especially on generic file servers. You decide to test Data Deduplication solutions to maximize storage availability for users.

New: After you have tested the storage redundancy and performance options, you now decide that it would also be beneficial to maximize the available disk space that you have, especially around virtual machine storage which is in ever increasing demand. You decide to test out Data Deduplication solutions to maximize storage availability for virtual machines.

Lab Review

- Your manager is worried about the impact that using data deduplication will have on the write performance of your file servers' volumes. Is this concern valid?