Windows Server 2016 Storage Spaces Direct Cluster

Marc Grote

Wer bin ich?

- Marc Grote
- Seit 1989 hauptberuflich ITler / Seit 1995 Selbststaendig
- MVP Forefront (2004-2014), MVP Hyper-V (2014), MVP Cloud and Datacenter (2015-2017), Microsoft MCT/MCSE Messaging/Security/Server/MCLC /MCITP*/MCTS*/MCSA*/MC* MCSE Private Cloud, Productivity, Cloud Platform and Infrastructure, Server Infrastructure, Exchange MCS Server Virtualization Hyper-V / System Center/ Azure MCITP Virtualization Administrator
- Buchautor und Autor fuer Fachzeitschriften
- Schwerpunkte:
  - Windows Server Clustering/Virtualisierung/PKI
  - System Center SCVMM/SCEP/DPM
  - Exchange Server seit Version 5.0
  - von *.Forefront reden wir nicht mehr ☹

Agenda

- Failover Cluster Neuerungen in Windows Server 2016
- Storage Spaces Direct Architektur
- Einrichtung eines S2D-Clusters
- Administration eines S2D-Clusters
- Storage-Erweiterung des S2D-Cluster
- S2D Troubleshooting

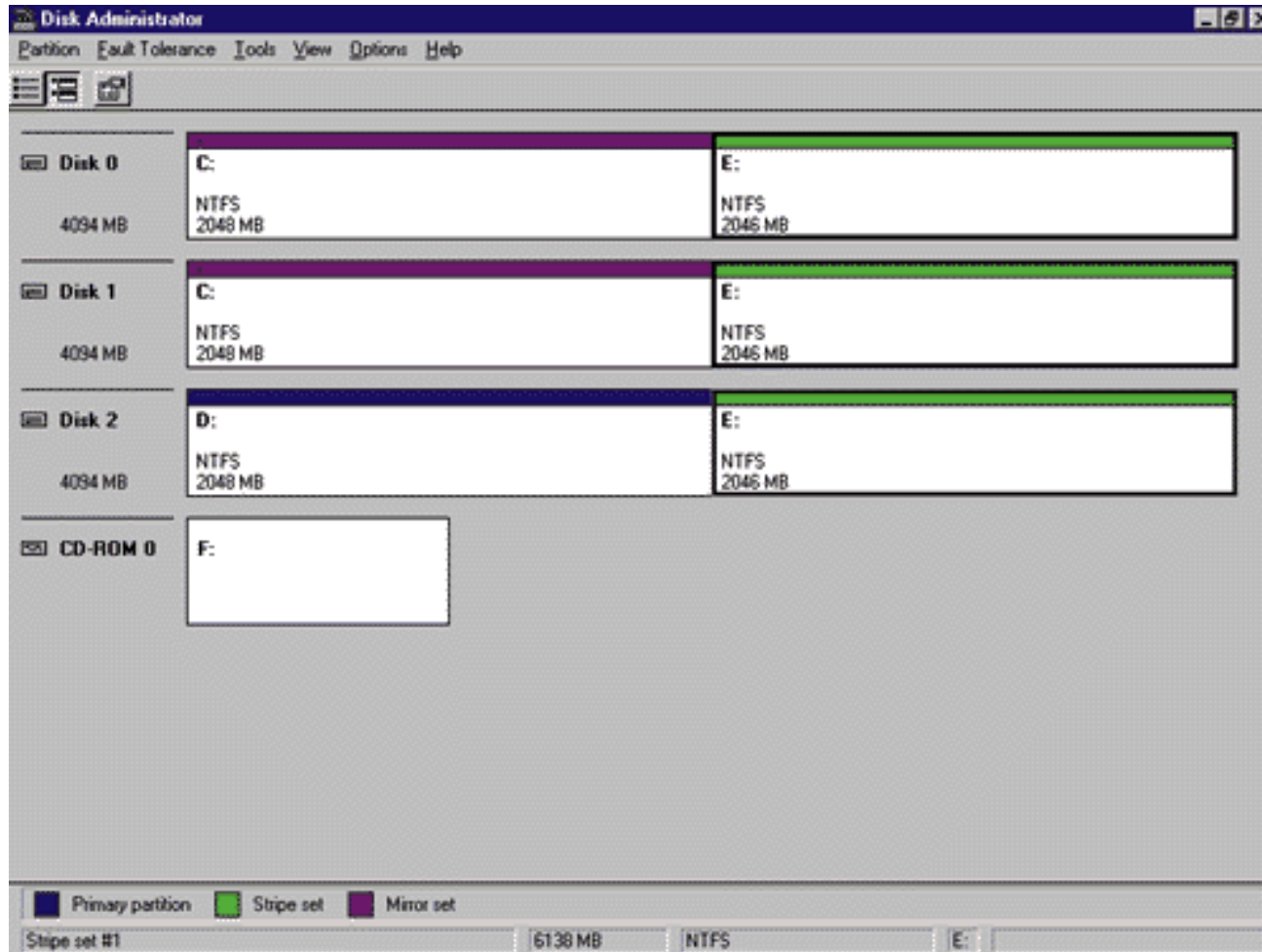# Failover Cluster Neuerungen in Windows Server 2016

- Virtual Machine Cluster Resiliency
  - Quarantine
  - Isolated
  - Konfiguration per PS (Get-Cluster ….)
- Cluster Rolling Upgrade
- Virtual Machine Load Balancing
- Cloud Witness in Microsoft Azure
- Host Guardian Service (VM Isolation fuer Tenants)

Failover Cluster Neuerungen in Windows Server 2016

- Start Order Prioritaet fuer VM im Cluster
- Storage Replica
- Storage Resiliency
- Site Aware Cluster
  - Failover Affinity, Storage Affinity, Cross-Site Heartbeating
- Verbessertes Cluster Log
  - TimeZone, VerboseDiagnostic EventViewer
- Active Memory Dump

Storage Spaces Direct

Quelle: http://windowsitpro.com/site-files/windowsitpro.com/files/archive/windowsitpro.com/content/content/2825/screen_03.gif
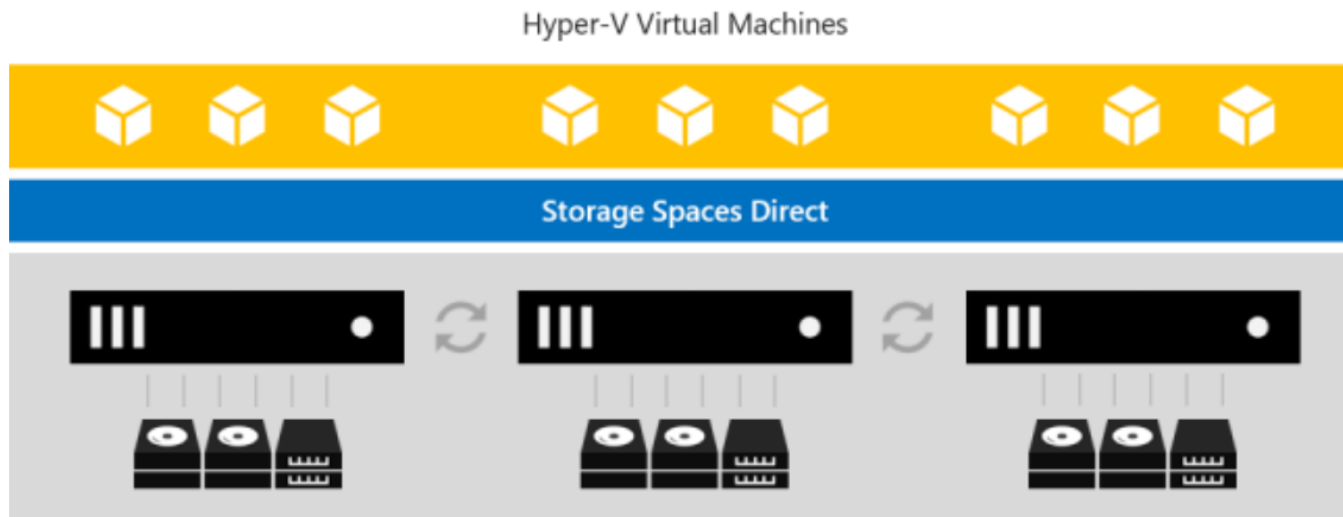
Storage Spaces Direct

- Lokal angeschlossener Festplattenspeicher von mehreren Servern werden zu einem gemeinsamen Speicher gebuendelt

- Verschiedene Ebenen der Fehlertoleranz moeglich

- Austausch der Paritaets/Spiegel-Informationen erfolgt ueber das Server Message Block-Protokoll (SMB 3.1.1)

- Kombination von Datentraegertypen moeglich (NVMe, SSD, HDD) für Cache, Storage Tiering

- Speicher wird per Storages Spaces und Storage Pools zu Speicherung von virtuellen Maschinen in einem CSV ueber NTFS/ReFS bereitgestellt.

- Windows Server 2016 Datacenter Lizenz notwendig

Storage Spaces Direct

Hyper-Converged

One cluster for compute and storage. The hyper-converged deployment option runs Hyper-V virtual machines or SQL Server databases directly on the servers providing the storage, storing their files on the local volumes. This eliminates the need to configure file server access and permissions, and reduces hardware costs for small-to-medium business or remote office/branch office deployments. See Hyper-converged solution using Storage Spaces Direct.
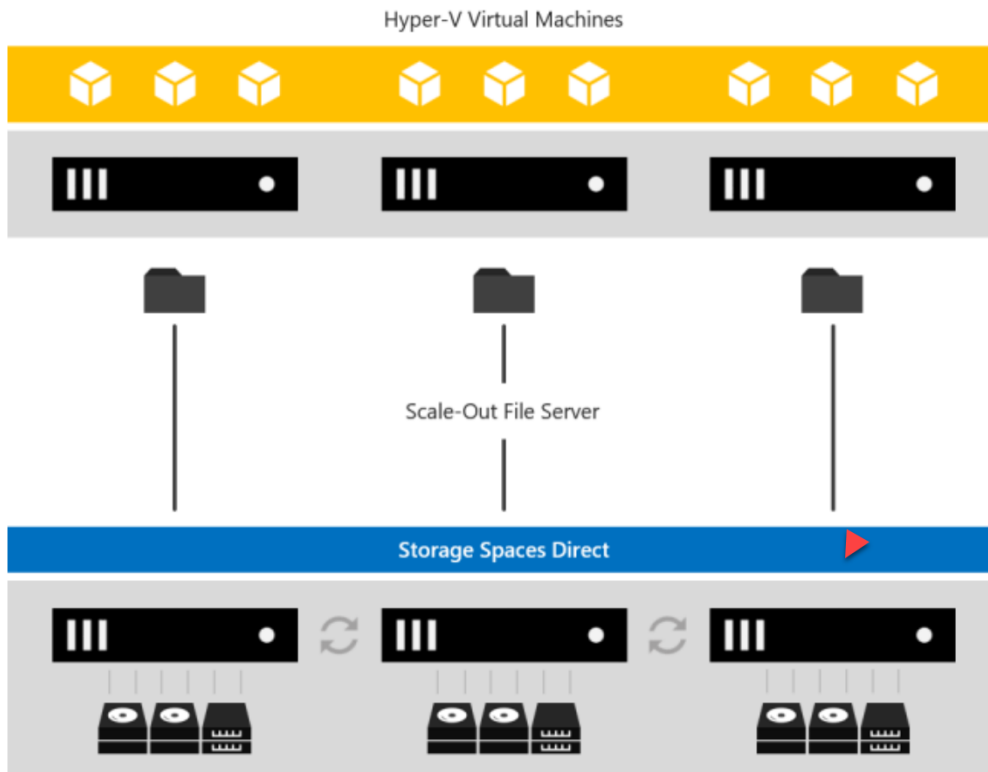
Quelle: https://technet.microsoft.com/en-us/windows-server-docs/storage/storage-spaces/storage-spaces-direct-overview
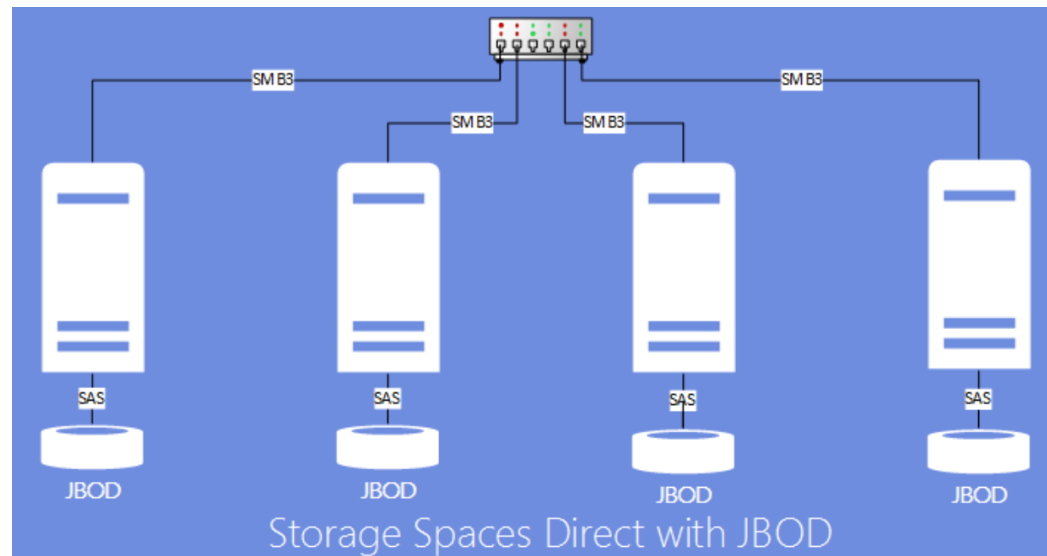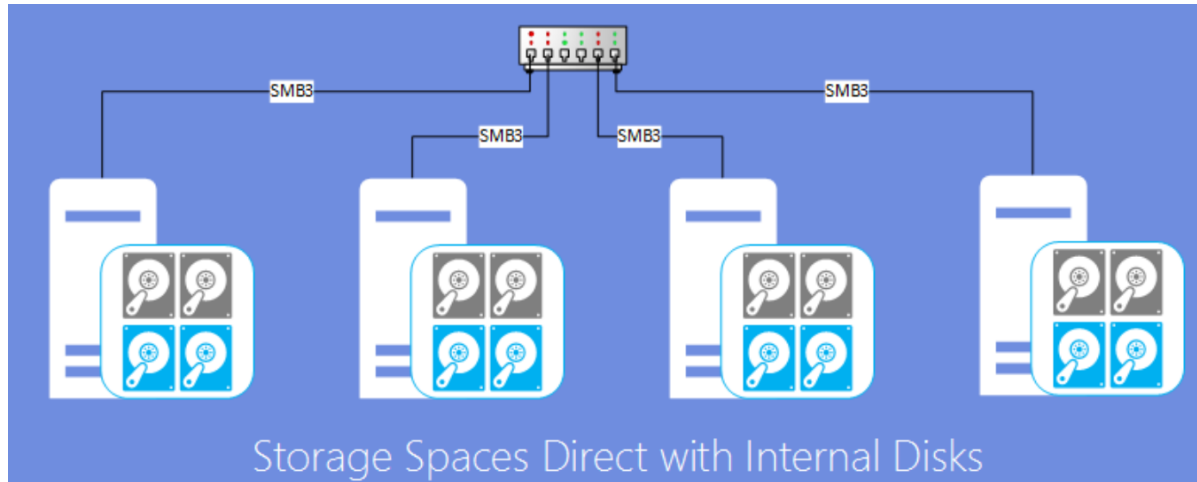
Storage Spaces Direct

Converged

Storage and compute in separate clusters. The converged deployment option, also known as 'disaggregated', layers a Scale-out File Server (SoFS) atop Storage Spaces Direct to provide network-attached storage over SMB3 file shares. This allows for scaling compute/workload independently from the storage cluster, essential for larger-scale deployments such as Hyper-V IaaS (Infrastructure as a Service) for service providers and enterprises.

Hyper-V Virtual Machines

Scale-Out File Server

Storage Spaces Direct

Storage Spaces Direct

Storage Spaces Direct with Internal Disks

Storage Spaces Direct with JBOD

Source: https://technet.microsoft.com/en-us/library/mt126109.aspx

Warum ReFS?

- ReFS v3 (Resilient File System, Robustes Dateisystem)
- Integrity Streams – (Pruefsummen fuer Metadaten / optional fuer Dateidaten)
- S2D Integration – Echtzeit Tier-Optimierung und Online Reparatur von Datenkorruption
- Beschaedigte Daten werden aus dem „Namespace" Online entfernt
- Proaktive Fehlerkorrektur durch integrierten Datenintegritaetsscanner
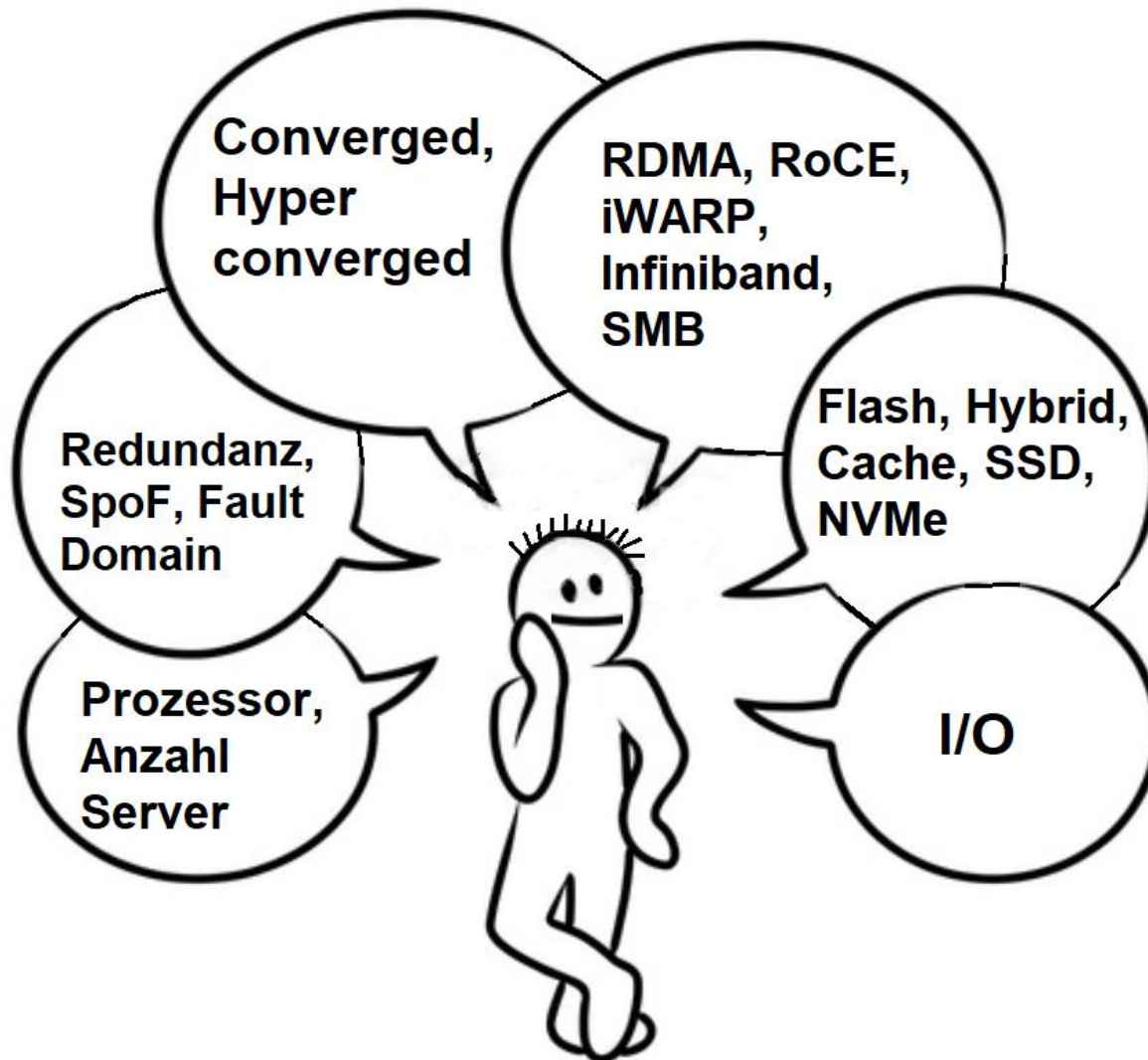
Warum ReFS?

- Clustergroesse 4K und 64K. 4K empfohlen, 64K nur bei grossen und sequentiellen Datenstroemen
- VM Optimierung
  - Snapshot Zusammenfassung (Metadata Update)
  - Erstellung von VHDX-Dateien fester Groesse
- NTFS nur verwenden, wenn ReFS nicht supported ist

Einschraenkungen von ReFS

Funktion	ReFS	NTFS
File system compression	Nein	Ja
File system encryption	Nein	Ja
Data Deduplication	Nein	Ja
Transactions	Nein	Ja
Hard links	Nein	Ja
Object IDs	Nein	Ja
Short names	Nein	Ja
Extended attributes	Nein	Ja
Disk quotas	Nein	Ja
Bootable	Nein	Ja
Supported on removable media	Nein	Ja
NTFS storage tiers	Nein	Ja

Source: https://docs.microsoft.com/en-us/windows-server/storage/refs/refs-overview

Planung ist das halbe Leben

Hardware-Anforderungen

- 2-16 Server (ideal) vom gleichen Modell und Hersteller
- Intel Nehalem Prozessor oder aktueller
- Max. 26 HD pro Server, 416 HD gesamt
- Maximum 1 PB RAW Kapazitaet pro Storage Pool, max. 100 TB Speicherplatz pro Server empfohlen
- 4 GB RAM fuer 1 TB Cache Diskspeicher
- 10 Gbit/s NIC fuer Cluster Kommunikation
- Mindestens 2 NIC fuer Redundanz und Performance
- NIC RDMA, iWARP oder RoCE faehig
- Alle Server gleiche HD (Typen, Groesse, Firmware)

S2D-Einrichtung

- Test-Cluster -Node S2D-NODE1.DOMAIN.TLD, S2D-NODE2.DOMAIN.TLD -Include "Storage Spaces Direct", Inventar, Netzwerk, "Systemkonfiguration"

- New-Cluster -Name S2D-CIM -Node S2D-NODE1.DOMAIN.TLD, S2D-NODE2.DOMAIN.TLD -NoStorage -StaticAddress 192.168.16.100

- Set-ClusterQuorum -NodeAndFileShareMajority "\\FS1.DOMAIN.TLD\S2D-Witness"

S2D-Einrichtung

- Enable-ClusterStorageSpacesDirect -PoolFriendlyName S2D-CIM -CacheState Disabled -SkipEligibilityChecks -Autoconfig:0 (in VM!)

- New-StoragePool -StorageSubSystemFriendlyName *Cluster* -FriendlyName S2D-CIM -ProvisioningTypeDefault Fixed -PhysicalDisk (Get-PhysicalDisk | ? CanPool -eq $true)

- Get-StorageSubsystem *cluster* | Get-PhysicalDisk | Where MediaType -eq "Unspecified" | Set-PhysicalDisk -MediaType HDD

S2D-Einrichtung

- Get-PhysicalDisk | where {$_.size -lt 10GB} | Set-PhysicalDisk -MediaType SSD

- New-StorageTier -StoragePoolFriendlyName S2D-CIM -MediaType SSD -FriendlyName Performance -Verbose

- New-StorageTier -StoragePoolFriendlyName S2D-CIM -MediaType HDD -FriendlyName Capacity -Verbose

- 1..3 | forEach {New-Volume -Size 10GB -FriendlyName "Volume $_ " -FileSystem CSVFS_ReFS -StoragePoolFriendlyName S2D-CIM}

S2D Ausfallsicherheit

Two-way mirror

Two-way mirroring writes two copies of everything. Its storage efficiency is 50% – to write 1 TB of data, you need at least 2 TB of physical storage capacity. Likewise, you need at least two hardware 'fault domains' – with Storage Spaces Direct, that means two servers.

> ⚡ **Warning**
>
> If you have more than two servers, we recommend using three-way mirorring instead.

S2D Ausfallsicherheit

Three-way mirror

Three-way mirroring writes three copies of everything. Its storage efficiency is 33.3% – to write 1 TB of data, you need at least 3 TB of physical storage capacity. Likewise, you need at least three hardware fault domains – with Storage Spaces Direct, that means three servers.
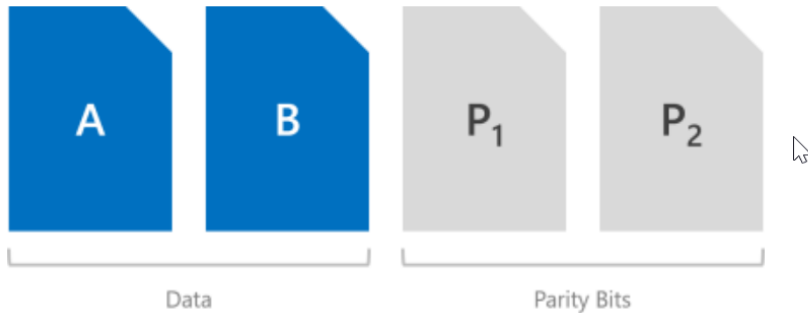
Three-way mirroring can safely tolerate at least two hardware problems (drive or server) at a time. For example, if you're rebooting one server when suddenly another drive or server fails, all data remains safe and continuously accessible.

Data   Copies

S2D Ausfallsicherheit

Dual parity

Dual parity implements Reed-Solomon error-correcting codes to keep two bitwise parity symbols, thereby providing the same fault tolerance as three-way mirroring (i.e. up to two failures at once), but with better storage efficiency. It most closely resembles RAID-6. To use dual parity, you need at least four hardware fault domains – with Storage Spaces Direct, that means four servers. At that scale, the storage efficiency is 50% – to store 2 TB of data, you need 4 TB of physical storage capacity.

Data   Parity Bits

The storage efficiency of dual parity increases the more hardware fault domains you have, from 50% up to 80%. For example, at seven (with Storage Spaces Direct, that means seven servers) the efficiency jumps to 66.7% – to store 4 TB of data, you need just 6 TB of physical storage capacity.

Data   Parity Bits

S2D Ausfallsicherheit

Resiliency types

Resiliency	Failure tolerance	Storage efficiency
Two-way mirror	1	50.0%
Three-way mirror	2	33.3%
Dual parity	2	50.0% - 80.0%
Mixed	2	33.3% - 80.0%

Minimum scale requirements

Resiliency	Minimum required fault domains
Two-way mirror	2
Three-way mirror	3
Dual parity	4
Mixed	4

S2D Storage Erweiterung

Virtual Disk auflisten

- Get-VirtualDisk

Size und Max Size der vDisk anzeigen

- Get-VirtualDisk CSV-01 | Get-Disk | Get-Partition | Get-Volume

Anzeige ob Storage Tiers vorhanden sind

- Get-VirtualDisk CSV-01 | Get-StorageTier

Name des Storage Tier anzeigen

- Get-VirtualDisk CSV-01 | Get-StorageTier | Select FriendlyName

Resizen des Storage Tier

- Get-StorageTier CSV-01 | Resize-StorageTier -Size 10GB

S2D Storage Erweiterung

Resize der Partition ...

Virtual Disks auflisten

- $VirtualDisk = Get-VirtualDisk CSV-01

Anzeige der Partitionen

- $Partition = $VirtualDisk | Get-Disk | Get-Partition | Where PartitionNumber -Eq 2

Vergroessern auf Maximum

- $Partition | Resize-Partition -Size ($Partition | Get-PartitionSupportedSize).SizeMax

S2D Troubleshooting

- Datentraeger pruefen
  - http://cosmosdarwin.com/Show-PrettyVolume.ps1
  - Server Manager / Failover Cluster Console
- Auto Tiering Statistik: https://www.hyper-v-server.de/management/microsoft-storage-spaces-tiering-statikstiken-unter-windows-server-2012-r2-und-windows-server-2016/
- Failover Cluster Konsole / Validation Test / Ereignisanzeige

S2D Troubleshooting

- Test Cluster Health: https://github.com/Microsoft/diskspd/blob/master/Frameworks/VMFleet/test-clusterhealth.ps1

- VM Fleet: https://blogs.technet.microsoft.com/larryexchange/2016/08/17/leverage-vm-fleet-testing-the-performance-of-storage-space-direct/
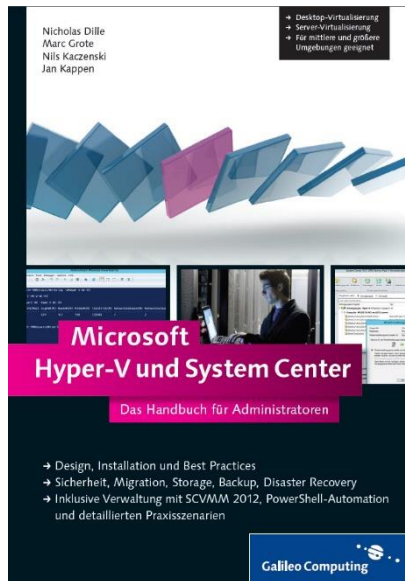
S2D mit VMM

- VMM kann einen neuen S2D-Cluster erstellen oder bestehende S2D-Cluster verwalten

- OS und Netzwerke konfigurieren

- Hosts unter VMM-Verwaltung bringen (Hyper-V Rolle und VMM-Agent werden installiert)

- S2D auf mit Baremetal Deployment erstellten Hosts und Nano Server (Nano Strategie?) nicht moeglich

- Nur eine Storage-Klassifizierung fuer Storagepool

- Dateiserver Rolle und Failover Clustering Feature werden installiert
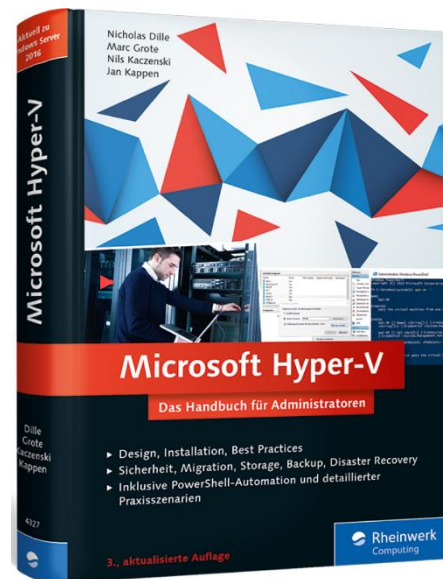
S2D mit VMM

- Storage Replica und Data Deduplication wird aktiviert

- Cluster Validierung wird durchgefuehrt

- Storage Spaces Direct werden aktiviert und ein Storage Pool angelegt

- Virtual Disk und Cluster Shared Volume (CSV) muss anschliessend angelegt werden

- Fuer ein PoC in einer Nested Virtualization sind zusaetzliche Schritte erforderlich bzw. Nacharbeiten
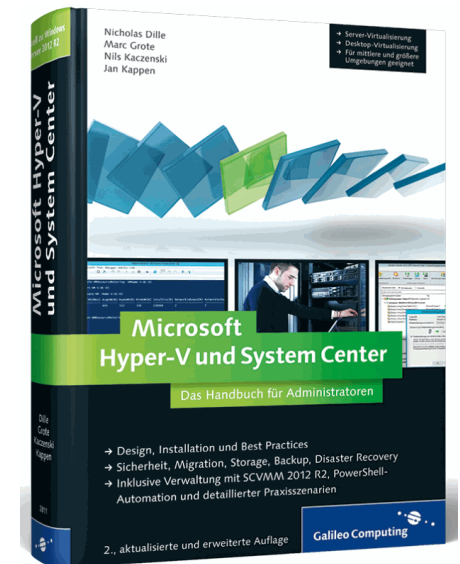
Die Buecher

Hyper-V und System Center 2012

Hyper-V und System Center 2012 R2

Hyper-V 2016

Kontakt

- E-Mail: marc.grote@it-consulting-grote.de
- Web: https://www.it-consulting-grote.de
- Blog: https://blog.it-consulting-grote.de
- XING: https://www.xing.com/profile/Marc_Grote2
- Mobile: +4917623380279