Storage Spaces

Karan Mehra
Microsoft Corporation
Agenda

- Overview
- Fault Resiliency
- Flexible Utilization
- Server Availability
- Tailored Management
- Further Reading
- Q & A
Overview

- Storage Pools
  - Units of aggregation, administration and isolation
- Storage Spaces
  - Virtual disks
  - Provide fault resiliency and flexible utilization
- Industry standard commodity storage
  - SAS, USB and SATA disks
  - SAS enclosures
Overview

Diverse workloads

Storage Pools with Spaces

Industry standard commodity storage

Pool A
- Finance
- User Docs
- Mirror
- Parity
- SSDs
- 15K RPM HDDs

Pool B
- Archive
- Parity
- 7200 RPM HDDs
Fault Resiliency

- Simple
- Mirror
  - Rapid crash recovery
  - Read stream detection
  - Enclosure awareness
- Parity
  - Crash resiliency journal
  - Write-back cache
Fault Resiliency

Drive goes missing

Pool A

Mirror

Hot Spare
Fault Resiliency

Drive experiences intermittent errors
Fault Resiliency

Drive reports media errors or fails

Pool A

Mirror

Hot Spare
Fault Resiliency

- **Interfaces**
  - Read individual copies
    - FSCTL_MARK_HANDLE
    - CopyNumber
  - Repair specific copies
    - IOCTL_STORAGE_MANAGE_DATA_SET_ATTRIBUTES
    - DeviceDsmAction_Repair
  - Verify consistency of copies
    - IOCTL_STORAGE_MANAGE_DATA_SET_ATTRIBUTES
    - DeviceDsmAction_Scrub

- **Use case**
  - Data Integrity Scan task
Flexible Utilization

- On-demand provisioning
  - Thresholds
  - Alerts
- User-configurable capacity allotment
  - Manual selection
- Efficient capacity utilization
  - Weighted round robin
- Elastic capacity expansion
  - Add drives as needed
Flexible Utilization

Space

Slab

Threshold

Extent

Disk 0
Disk 1
Disk 2
Disk 3
Disk 4

Alert
Flexible Utilization

- **Interfaces**
  - Query allocated ranges
    - IOCTL_STORAGE_MANAGE_DATA_SET_ATTRIBUTES
    - DeviceDsmAction_Allocation
  - Trim ranges
    - IOCTL_STORAGE_MANAGE_DATA_SET_ATTRIBUTES
    - DeviceDsmAction_Trim

- **Use case**
  - Storage optimizer
Server Availability

- **High availability**
  - Integration with Failover Clustering
  - Spaces within a pool are independently accessible from different nodes
  - On-disk state machine to drive cross-node workflows

- **Scale out**
  - Integration with Cluster Shared Volumes
Server Availability

Unified Cluster Shared Volume Namespace

Node 1
- Pool A
  - Mirror

Node 2
- Pool A
  - Mirror

PBODs

SQL
Server Availability

- Interfaces
  - New Disk Policy
    - MSFT_StorageSetting
  - Attributes and methods on storage pools
    - MSFT_StoragePool
    - IsReadOnly
    - Attach
    - Detach
  - Attributes on storage spaces
    - MSFT_VirtualDisk
    - IsManualAttach

- Use case
  - Microsoft Cluster Service
Tailored Management

- Familiar object model
  - Storage Pool
  - Physical Disk
  - Virtual Disk

- Operational simplicity
  - Intelligent defaults

- Granular administrative control
  - ACL based security model
Tailored Management

File Server Manager → Storage Management API
System Center
ISV/IHV Application

Storage Management API
PowerShell and WMI

SMI-S Proxy Service

SMI-S HW Provider
SMI-S HW Provider
IHV SM Provider
Spaces SM Provider

Storage Array
Storage Array
Storage Array
JBODs
Tailored Management

- Interfaces
  - Query and manage objects
  - Indications
  - Associations

- Use case
  - PowerShell
Further Reading

- **Related sessions**
  - Tue at 11:20am: Windows File and Storage Directions
  - Wed at 4:20pm: ReFS - Next Generation File System for Windows

- **BUILD 2011 conference**
  - [Session 474](#): Platform storage evolved
  - [Session 446](#): Designing systems for continuous availability and scalability

- **MSDN**
  - [Blog](#): Virtualizing storage for scale, resiliency, and efficiency
  - [Blog](#): Windows Storage Team
  - [Doc](#): Windows Storage Management API