

# Microsoft Storage Directions and SMB23 futures

Thomas Pfenning, General Manager  
Jim Pinkerton, Partner Architect

# Short History of SMB >1

- **SMB 2**

- 20 Commands
- Durable handles
- Concurrency and compounding
- High bandwidth delay product
- Fairness with credits

- **SMB 2.1**

- Resilient handles
- Leasing
- Branch Cache
- Large MTU

- **SMB 3**

- Persistent handles
- Transparent failover
- Active-Active
- Multichannel
- Branch Cache update
- SMB Direct
- Directory Leases
- Remote VSS
- Witness Protocol
- Claims based authorization

# Windows Server 2012 Storage Platform

## Benefit

## Capability

## Features

Availability

- Maximize service/data availability

Lower Capital Expenditure

- Storage Virtualization
- New NTFS CHKDSK and Health Model
- Next generation scale, availability, and data integrity
- Storage Spaces
- ReFS
- Storage Efficiency

Lower Operational Expenditure

- Deduplication
- Ease of storage administration

Maximize Existing Hardware

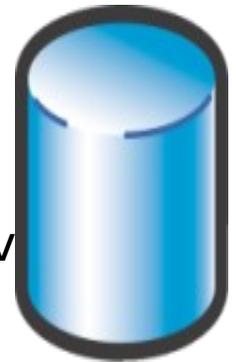
- Storage Management API with PowerShell & SMI-S integration
- Efficient capacity utilization
- Dynamic Data Movement
- Offloaded Data Transfer
- Thin Provisioning Integration

# CHKDSK – A Modern Approach

- **File system issues identified while the volume is online**
  - Pre-verification of corruption to ensure CHKDSK only runs when needed (eliminates false positives)
  - Online scanning determines repair steps that need to be taken
- **Minimized Downtime for Correction**
  - Online self-healing to immediately repair if possible
  - Volume can be taken offline at the administrator's discretion for minimal time for targeted repair of the issue (seconds)
- **Cluster integration for Continuous Availability**
  - Online spot-fixing of errors w/no down-time enabled w integration of Cluster Shared Volumes



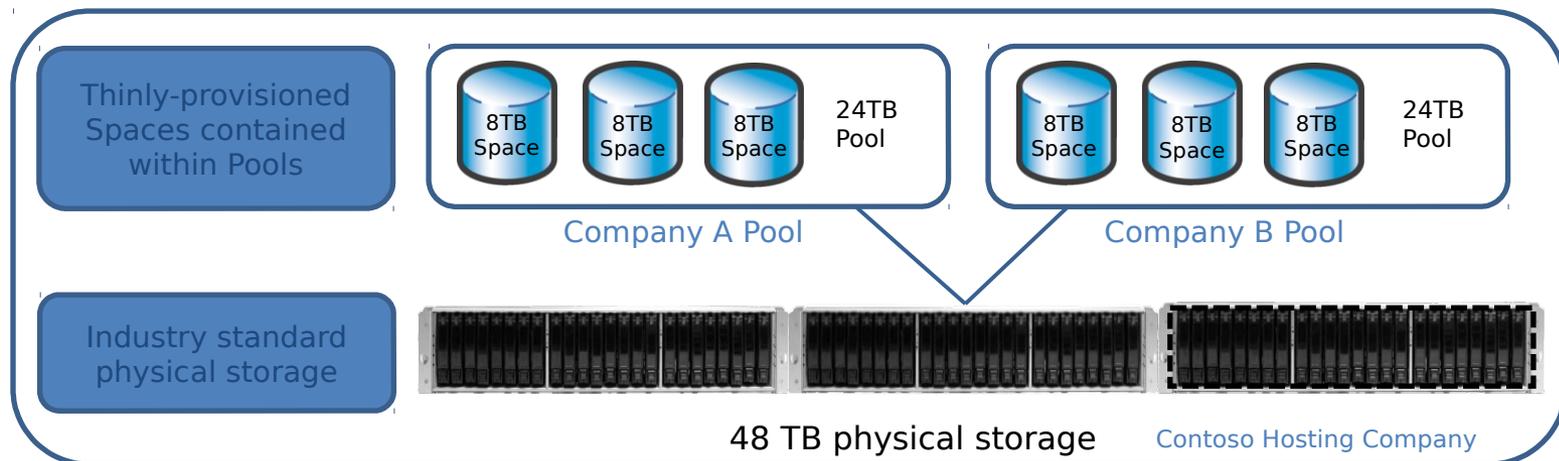
500 GB  
avg. size  
today



64 TB  
with Windows  
8

# Storage Spaces

- **Redundant storage**
  - Mirrored Spaces and Parity Spaces with integrated journaling
  - Cluster Support
- **Elastic capacity expansion**
  - Simply add more physical drives
- **On-demand provisioning with reclamation (trim/unmap) support**
  - For use by NTFS, Hyper-V, and other applications e.g. SQL
- **Managed via Windows Storage Management API**
  - Server Manager GUI and Powershell
- **Rapid Recovery - per pool hot spare support**
- **Application driven intelligent error correction**
- **Granular administrative control with tenant isolation**



# The Resilient File System (ReFS)

- **Data Integrity**

- All metadata corruption is always detected. Optionally, user data corruption is also detected.
- ReFS utilizes data redundancy offered by Storage Spaces to auto-correct metadata corruptions. Optionally, user data is similarly auto-corrected

- **Availability**

- ReFS maximizes dataset (volume) availability
- If corruption occurs and a redundant copy is unavailable, the corrupt portion of the namespace is removed even as non-corrupt portions remain unaffected and online

- **Scalability**

- ReFS efficiently scales to PB+ datasets comprising very large files or directories

- **Architectural Evolution**

- Architecture enables efficient evolution for new storage devices, new data types, and new access patterns.

- **ReFS compatibility**

- Support for failover clustering but no Clustered Shared Volumes
- NTFS features that ReFS does not support
  - Reparse Index
  - Hard Links
  - Named Streams
  - Extended Attributes (EA's)
  - Object ID's
  - Short Names
  - Fast MFT-like Enumeration
  - NTFS Compression
  - NTFS Encryption
  - NTFS Quotas
  - TxF
  - Deduplication

# SMB 3 Cluster-in-a-box Systems

HP X5000



Wistron (ODM)



Quanta (ODM)



Windows Server Support

- Released on Windows Storage Server 2008 R2
- Technology demonstration on Windows Server "8"

- Designed for Windows Server "8"

- Designed for Windows Server "8"

Market

Midrange NAS Server

- NAS server, Hyper-V appliance, SMB or Branch
- Private cloud performance NAS server

- NAS server, Hyper-V appliance, SMB or Branch
- Private cloud performance NAS server

Size

3U rack

3U rack

2U rack

Disks

- 36 x 2.5"
- 16 x 3.5"

- 24 x 2.5"

12 x 3.5"

Blades

2

2

2

CPU

2x Intel

2x SB EP

2x SB EP

Memory (per blade)

Varies by SKU

12 DDR3 DIMMs

16 DDR3 DIMMs

PCIe expansion slots (per blade)

1 x4, Gen 2

- 2 x16, Gen 3
- 1FH/HL, 1 HH/HL

1 x8, LP, Gen 3

Storage Controller

HP Cascade

- SAS Controller (Spaces)
- LSI High Availability MegaRAID

- SAS Controller (Spaces)
- LSI High Availability MegaRAID

External SAS (per blade)

- 4 x4 SAS
- Up to 8 Attached JBODs. 224 2.5", 108 3.5" (292 with MDS600) or mix

1 x4 SAS

1 x4 SAS

External network (per blade)

2x HP Flex-10 LOM  
4x 1 GbE mezz card

4x 1GbE LOM

2x RDMA-capable 10GbE or IB

Management Controller

HP iLO (integrated Lights-Out)

Emulex Pilot 3 iBMC

BMC

**“We don’t like their sound. Groups of guitars are on the way out.”**

**Decca Recording Co., executive, turning down the Beatles in 1962**

**"It is difficult to make predictions, especially about the future"**

**Mark Twain**

# Moving to a Cloud Design Point

- Discs, server, cluster, rack, row, data centers
- **Cluster of clusters**
  - Geo-replication
  - Coherent and non-coherent copies
- Parallel I/O
  - Hadoop is fastest growing File System
  - Sharding

# The Namespace challenge

- **Partitioned name space**
  - Rebalance name space when running out of capacity
  - Retire a file server from the name space
- **Tiering**
  - File vs. Directory level redirects
- **Hybrid Cloud scenarios**
  - As a tier and within the Cloud

# Functionality challenge

- Synchronization
  - File level vs. Chunk level
  - Filtering
- Storage Mobility
- E-2-E Deduplication
- Disaster Recovery
- Archival

# Device Challenge

- Explosion of connected mobile devices
  - Sync vs. browse
  - REST vs. connection oriented message blocks
  - Cloud vs. On Premise
  - Authentication
  - Authorization
  - Privacy