

# Rockin' and Rollin' with SMB3

Ned Pyle, Microsoft

April 5, 2017

# Today's Presenters



**Ned Pyle**  
**SMB Program Manager**  
**Microsoft**



**John Kim**  
**Chair, SNIA-ESF**  
**Mellanox**

# SNIA at a glance



**160**  
unique member  
companies



**3,500**  
active contributing  
members



**50,000**  
IT end users & storage  
pros worldwide

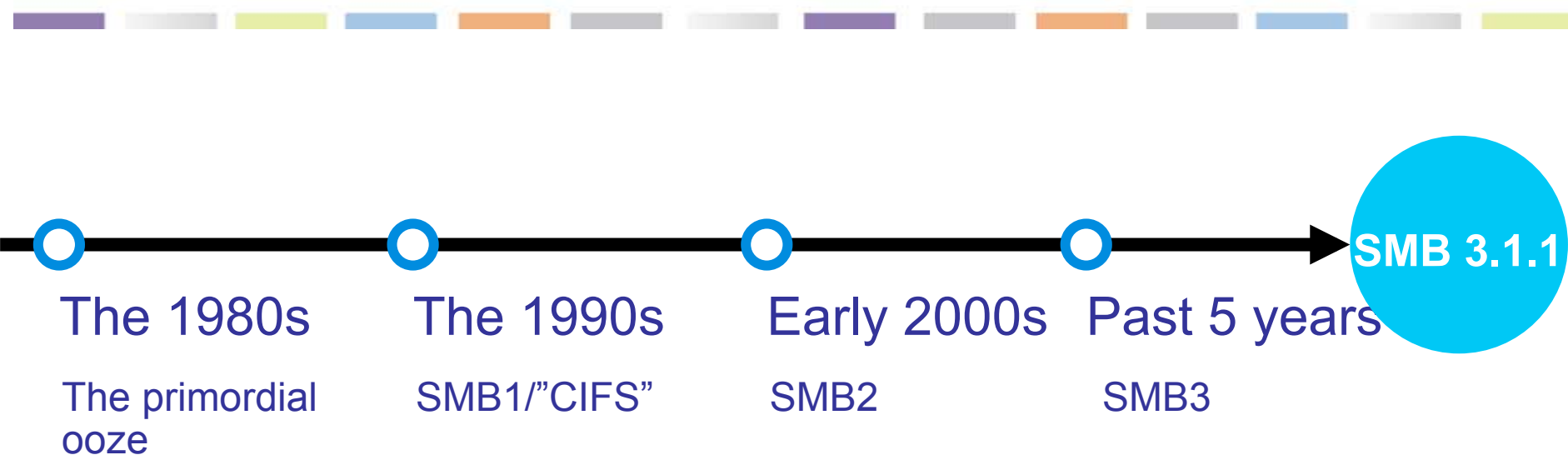
Learn more: [snia.org/technical](https://snia.org/technical)



- ◆ The material contained in this presentation is copyrighted by the SNIA unless otherwise noted.
- ◆ Member companies and individual members may use this material in presentations and literature under the following conditions:
  - ◆ Any slide or slides used must be reproduced in their entirety without modification
  - ◆ The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- ◆ This presentation is a project of the SNIA.
- ◆ Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- ◆ The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

**NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.**

# The evolution of SMB



DOS, Windows, LANMan(!)

Ubiquitously abused

Slow unsafe \$%^#

Will come back to this



SMB 1

## Windows Vista+ / Windows Server 2008+ User-optimized

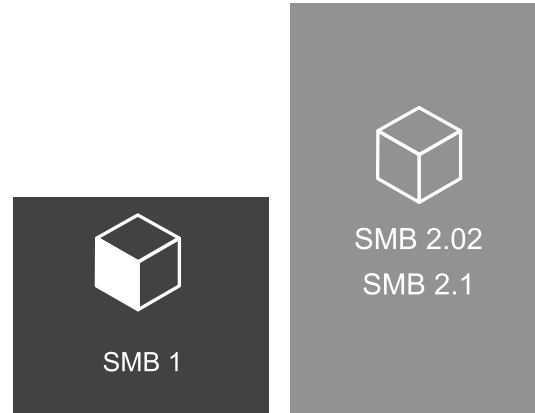
Request compounds, large reads and writes

Folder & file property caching

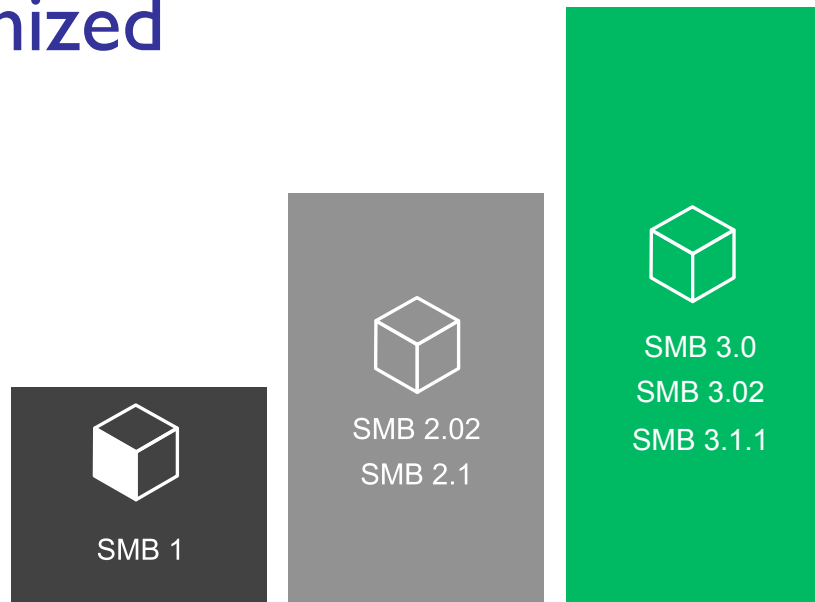
Durable handles

Improved message signing - HMAC SHA-256

Large MTU support



Windows 8+ / Windows Server 2012+  
Datacenter application-optimized  
Software-defined fabric  
Modern user







Not CIFS. Never, ever CIFS

| Security  | Scale   | Performance          |
|---|---|----------------------|
| Encryption<br>Pre-auth integrity<br>Guest block | Transparent failover<br>Scale-out File Server | RDMA<br>Multichannel |

# Security

## End-to-end SMB encryption

Privacy

AES-128-GCM & AES-128-CCM

## SMB Signing updated

Integrity

AES-CMAC

Requires\* AES-NI CPUs

Per share or server

Introduced in Win10, KB3000483

## MitM prevention

Require Mutual Auth, Signing/Encryption

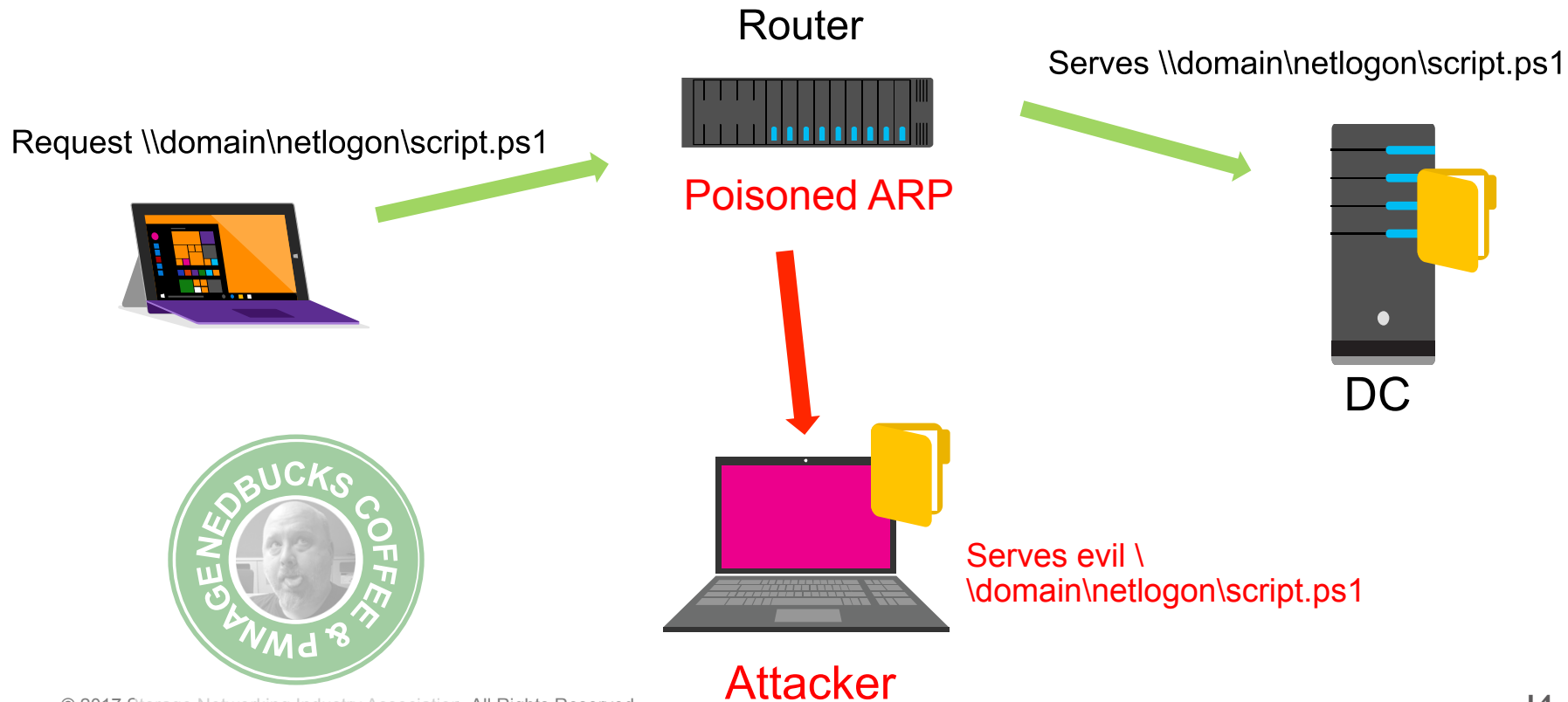
Client decides

## Specify UNC paths to protect

## Win10 defaults for AD DS

\\\*\SYSVOL

\\\*\NETLOGON



## MitM protection

Session Setup signed

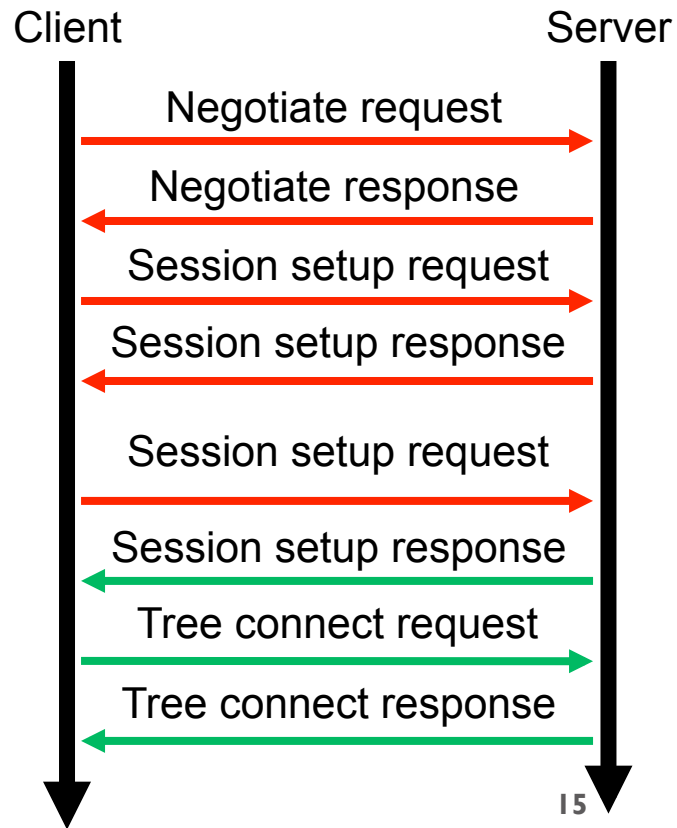
Tree Connect encrypted

No SMB capabilities downgrade possible

## Replaces “Secure negotiate”

## Cannot be disabled

## Update WAN accelerators



## Stop Guest usage on failed auth

## Bad behavior allowed by default

Because SOHO NAS

## Group Policy

“Enable Insecure Guest Logons”

key=“HKLM\Software\Policies\Microsoft\Windows\LanmanWorkstation”

DWORD=“AllowInsecureGuestAuth”

## If SMB1 enabled, meaningless





SMB3 security doesn't  
matter if SMB1 is enabled

I'm sorry Mr.  
Johnson, but the  
test is positive.

Please read this  
pamphlet...



Deprecated years ago

Removable since WS2012 R2/Win8.1

Disable-able since Vista/2008

Gone in WS2016 Nano

Uninstalled by default in certain upcoming releases

# Why might SMB1 still be in use?

## XP & WS2003 under custom support agreement

Yeah right

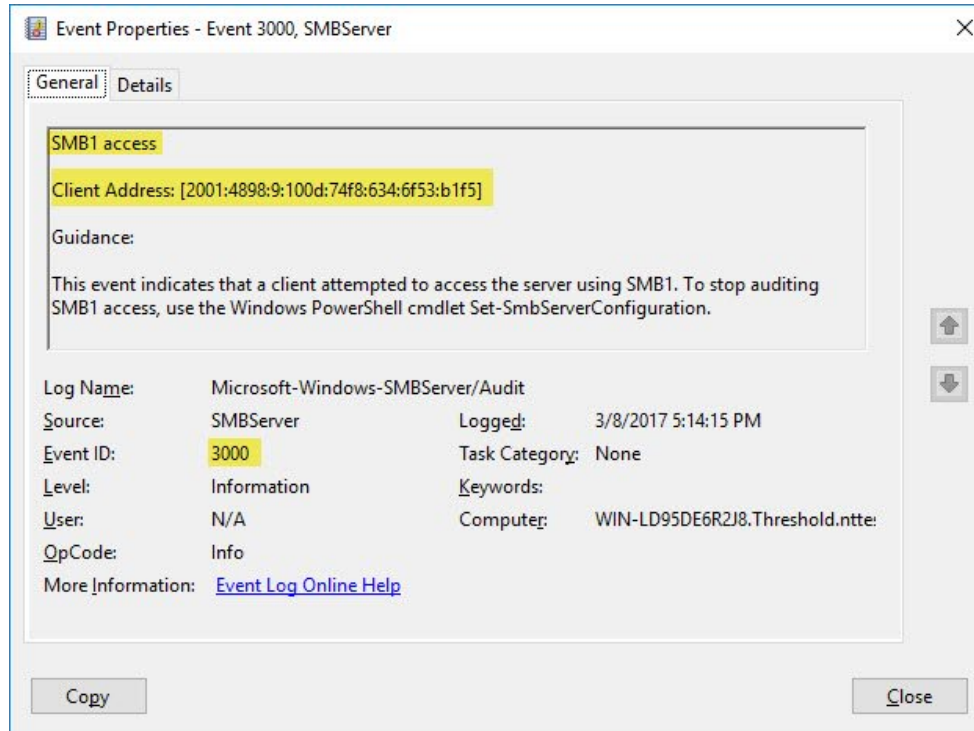
## Antique firmware printers

“Scan to share”

## Evil NAS devices

## Older Linux

# Set-SmbServerConfiguration -AuditSmb1Access \$true



Zap it

KB2696547



# Question break I



# Scale



# SMB Transparent Failover

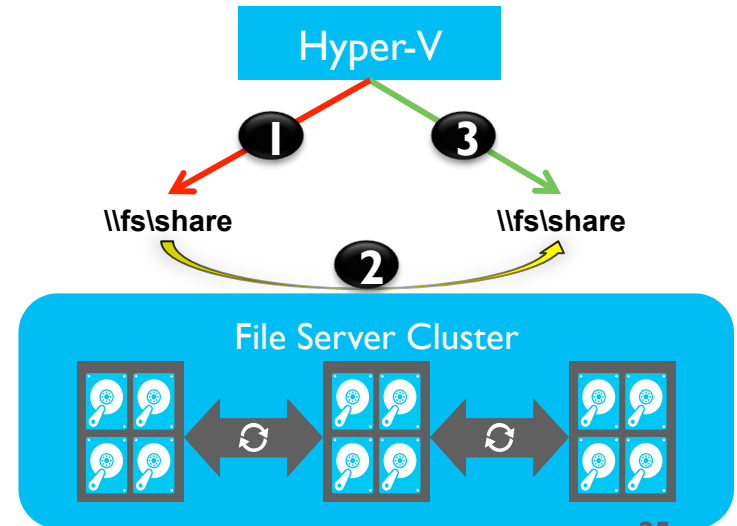
## Failover transparent to application

Zero downtime – small IO delay during failover

Planned & unplanned failovers

Clustered

- 1 Normal operation
- 2 Failover share - connections & handles lost temporary stall of IO
- 3 Connections & handles auto-recovered Application IO continues with no errors



## Server app storage

Example: Hyper-V and SQL Server

Increase available bandwidth by adding nodes

CSV

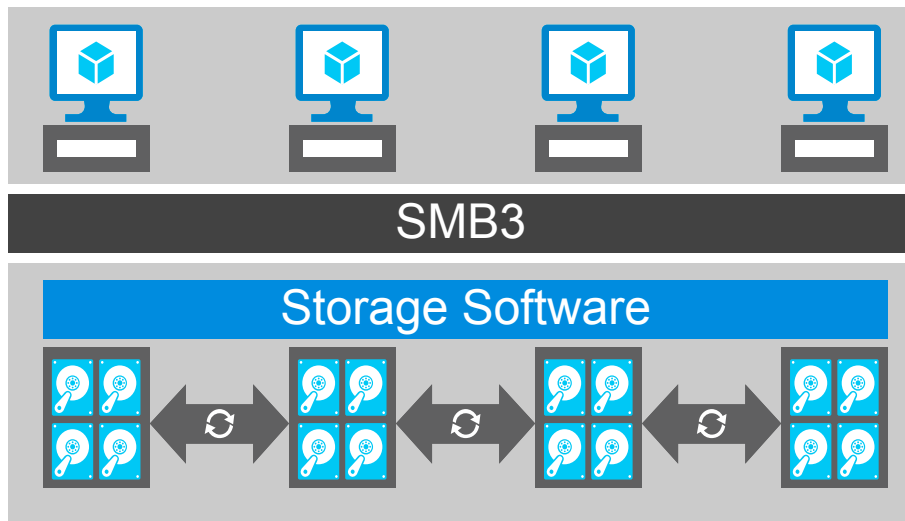
## Key capabilities

Active/Active file shares

Fault tolerance with zero downtime

Fast failure recovery

Forced unit access/write through





# Performance

# SMB Multichannel

## Full Throughput

Bandwidth aggregation

## Automatic Failover

## Automatic Configuration

SMB detects & uses multiple paths

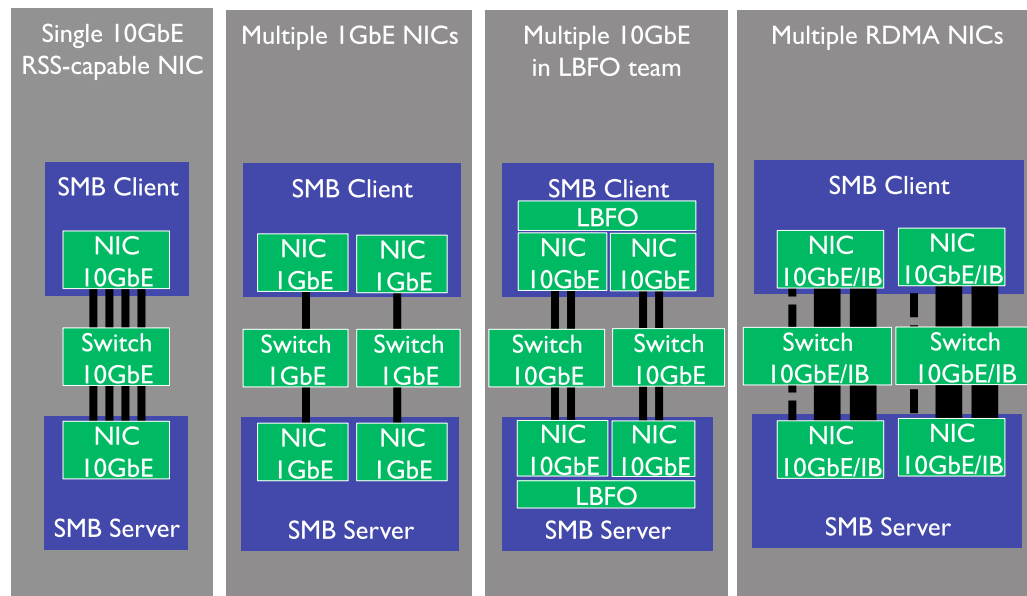
## Requires one:

One+ RSS-enabled NIC

Multiple network adapters

Teamed NICs

One+ RDMA NICs



# SMB Direct (SMB over RDMA)

High throughput

Low latency

Minimal CPU

Load balancing

Automatic failover

Bandwidth aggregation

Required hardware

iWARP

RoCE v1 & V2

InfiniBand

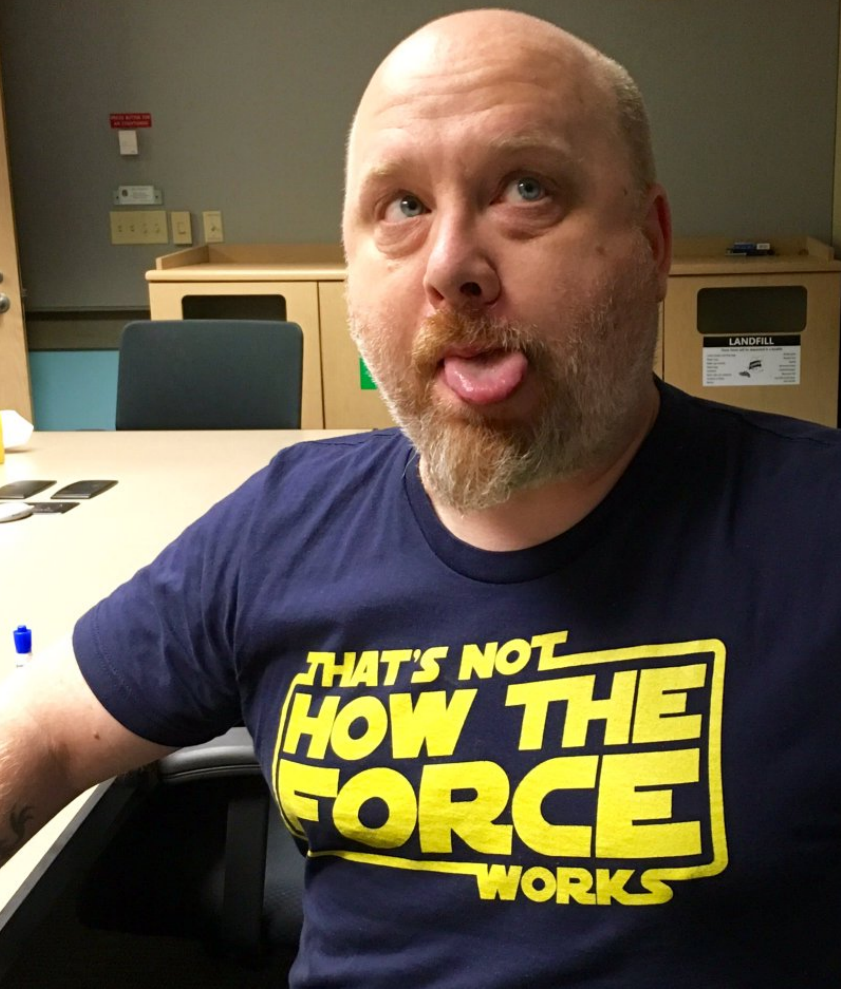
| NIC                   | Throughput  |
|-----------------------|-------------|
| 1Gb Ethernet          | ~0.1 GB/sec |
| 10Gb Ethernet         | ~1.1 GB/sec |
| 40Gb Ethernet         | ~4.5 GB/sec |
| 32Gb InfiniBand (QDR) | ~3.8 GB/sec |
| 56Gb InfiniBand (FDR) | ~6.5 GB/sec |

| HBA        | Throughput  |
|------------|-------------|
| 3Gb SAS x4 | ~1.1 GB/sec |
| 6Gb SAS x4 | ~2.2 GB/sec |
| 4Gb FC     | ~0.4 GB/sec |
| 8Gb FC     | ~0.8 GB/sec |
| 16Gb FC    | ~1.5 GB/sec |

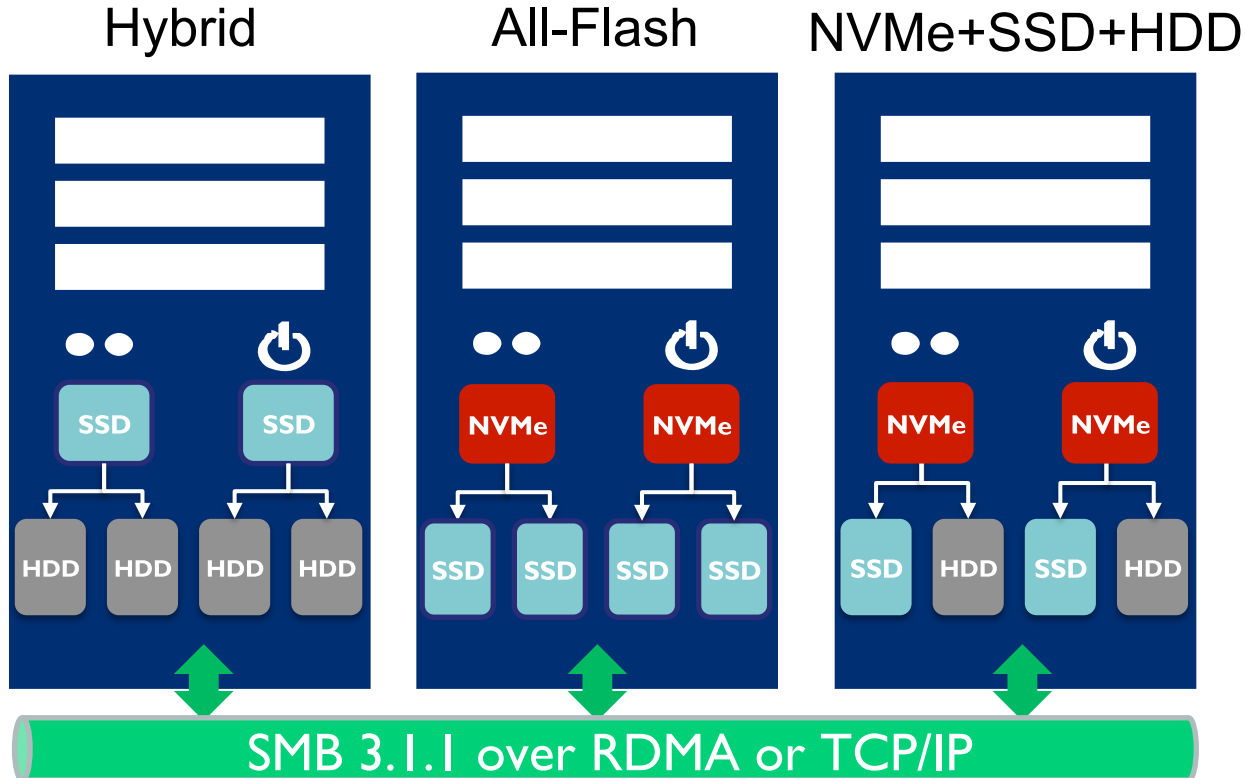


# Question break 2

# Real world examples

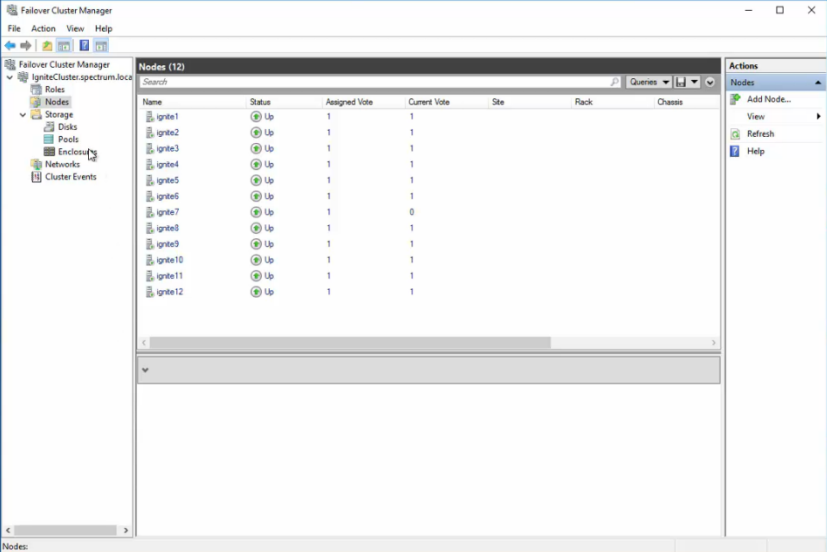


# Storage Spaces Direct (hyper-converged storage cluster)





# Storage Spaces Direct quasi-demo



The screenshot shows the Failover Cluster Manager interface. The left-hand navigation pane is expanded to show 'Nodes (12)'. The main pane displays a table of 12 nodes, all with a status of 'Up'. The table columns are Name, Status, Assigned Vote, Current Vote, Site, Rack, and Chassis. The nodes are named ignet1 through ignet12. The right-hand pane shows the 'Actions' menu with options: Add Node..., View, Refresh, and Help.

| Name    | Status | Assigned Vote | Current Vote | Site | Rack | Chassis |
|---------|--------|---------------|--------------|------|------|---------|
| ignet1  | Up     | 1             | 1            |      |      |         |
| ignet2  | Up     | 1             | 1            |      |      |         |
| ignet3  | Up     | 1             | 1            |      |      |         |
| ignet4  | Up     | 1             | 1            |      |      |         |
| ignet5  | Up     | 1             | 1            |      |      |         |
| ignet6  | Up     | 1             | 1            |      |      |         |
| ignet7  | Up     | 1             | 0            |      |      |         |
| ignet8  | Up     | 1             | 1            |      |      |         |
| ignet9  | Up     | 1             | 1            |      |      |         |
| ignet10 | Up     | 1             | 1            |      |      |         |
| ignet11 | Up     | 1             | 1            |      |      |         |
| ignet12 | Up     | 1             | 1            |      |      |         |

12 HPE Servers  
24 Mellanox ConnectX-4 100GbE NICs  
1 Mellanox Spectrum 100GbE switch

# Storage Spaces Direct quasi-demo

The screenshot displays the Failover Cluster Manager interface for a cluster named 'IgniteCluster:spectrum.local'. The 'Pools (1)' section shows 'Cluster Pool 1' in an 'Online' state with a health status of 'Healthy', owned by 'Ignite12', and an operational status of 'OK'. The free space is 55.7 TB and the used space is 49.0 TB.

The 'Cluster Pool 1' table lists the following physical disks:

| Name              | Health Status | Operational Status | Used Space | Capacity | Allocation | Bus Type | Enclosure Name                 | Slot Number |
|-------------------|---------------|--------------------|------------|----------|------------|----------|--------------------------------|-------------|
| PhysicalDisk-5057 | Healthy       | 14                 | 1.04 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC41F7B40C00 |             |
| PhysicalDisk-5001 | Healthy       | 14                 | 1.01 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC413F57E800 |             |
| PhysicalDisk-5028 | Healthy       | 14                 | 1.03 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC4151B99F00 |             |
| PhysicalDisk-5035 | Healthy       | 14                 | 1.03 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC4111060A00 |             |
| PhysicalDisk-5036 | Healthy       | 14                 | 1.01 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC41804D9B00 |             |
| PhysicalDisk-5027 | Healthy       | 14                 | 1.03 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC41DE29E000 |             |
| PhysicalDisk-5030 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC4190AF9900 |             |
| PhysicalDisk-5026 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC4111060A00 |             |
| PhysicalDisk-5010 | Healthy       | 14                 | 1.01 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC4151B99F00 |             |
| PhysicalDisk-5058 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC41F7B40C00 |             |
| PhysicalDisk-5045 | Healthy       | 14                 | 1.01 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC418080D000 |             |
| PhysicalDisk-5036 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC4151B99F00 |             |
| PhysicalDisk-5032 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC418080D000 |             |
| PhysicalDisk-5013 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC418080D000 |             |
| PhysicalDisk-5042 | Healthy       | 14                 | 1.03 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC41CC00EA00 |             |
| PhysicalDisk-5019 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC4151B99F00 |             |
| PhysicalDisk-5002 | Healthy       | 14                 | 1.02 TB    | 2.18 TB  | Automatic  | NVMe     | SES Enclosure 50CBAC413F57E800 |             |

The interface also shows a 'Summary' tab for 'Cluster Pool 1' and a 'Physical Disks' tab. The background of the screenshot is a Windows 10 desktop with the Start menu open.

Single Storage Pool

48 Micron 9100 NVMe 2.4TB

# Storage Spaces Direct quasi-demo

The screenshot displays the Failover Cluster Manager interface. The main pane shows a list of 336 roles, all of which are virtual machines (VMs) named 'vm-base-IGNITE1-19' through 'vm-base-IGNITE1-23'. The status of all roles is 'Running'. The 'Owner Node' for all roles is 'ignite1', and the 'Priority' is 'Medium'. The 'Information' column is currently empty.

| Name               | Status  | Type            | Owner Node | Priority | Information |
|--------------------|---------|-----------------|------------|----------|-------------|
| vm-base-IGNITE1-19 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-20 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-21 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-18 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-15 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-16 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-17 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-26 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-27 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-28 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-25 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-22 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-23 | Running | Virtual Machine | ignite1    | Medium   |             |

The detailed view for the selected role 'vm-base-IGNITE1-19' shows the following information:

- Virtual Machine: vm-base-IGNITE1-19**
- Status:** Running
- CPU Usage:** 0%
- Memory Demand:** 573 MB
- Integration Services:** 10.0.14393
- Computer Name:** WIN-7JHGSCDU5QJ
- Date Created:** 8/4/2016 9:35:40 PM
- Version:** 8.0
- Up Time:** 6:56:52
- Assigned Memory:** 1024 MB
- Heartbeat:** OK
- Operating System:** Windows Server 2016 Datacenter
- Operating System Version:** 10.0.10011
- Unable to determine monitored services:**

336 virtual machines

512KB random read workload

# Storage Spaces Direct quasi-demo

The screenshot shows the Hyper-V Roles console with 336 roles listed. Below it, a Windows PowerShell performance monitor window displays the following data:

|   | IOPS           | Reads          | Writes     | BW (MB/s)      | Read           | Write    | Read Lat (ms) | Write |
|---|----------------|----------------|------------|----------------|----------------|----------|---------------|-------|
|   | <b>308,388</b> | <b>307,922</b> | <b>466</b> | <b>161,115</b> | <b>161,111</b> | <b>4</b> |               |       |
| 0 | 26,739         | 26,690         | 48         | 13,813         | 13,813         |          | 4.098         | 0.621 |
|   | 23,693         | 23,646         | 47         | 12,398         | 12,397         |          | 4.598         | 0.551 |
| 1 | 26,866         | 26,818         | 48         | 14,061         | 14,061         |          | 4.042         | 0.465 |
| 2 | 27,983         | 27,941         | 43         | 14,589         | 14,588         |          | 3.855         | 0.468 |
|   | 24,111         | 24,076         | 35         | 12,596         | 12,596         |          | 4.528         | 0.795 |
|   | 24,731         | 24,697         | 34         | 12,935         | 12,935         |          | 4.405         | 0.520 |
|   | 27,165         | 27,136         | 29         | 14,227         | 14,227         |          | 3.996         | 0.491 |
|   | 24,409         | 24,382         | 28         | 12,783         | 12,783         |          | 4.438         | 0.524 |
|   | 25,960         | 25,907         | 53         | 13,583         | 13,583         |          |               |       |
|   | 24,829         | 24,795         | 34         | 13,000         | 13,000         |          |               |       |
|   | 26,137         | 26,106         | 31         | 13,640         | 13,640         |          |               |       |
|   | 25,764         | 25,728         | 36         | 13,489         | 13,489         |          |               |       |

A large teal box in the bottom right corner of the PowerShell window displays **160 GB/s**.

# Storage Spaces Direct quasi-demo

The screenshot shows the Hyper-V Roles console with 336 roles listed. Below it, a Windows PowerShell window displays performance metrics for a storage configuration. A callout box indicates a throughput of 'or 1.2+ Terabit/s'.

|   | IOPS           | Reads          | Writes     | BW (MB/s)      | Read           | Write    | Read Lat (ms) | Write |
|---|----------------|----------------|------------|----------------|----------------|----------|---------------|-------|
|   | <b>315,997</b> | <b>315,765</b> | <b>232</b> | <b>160,367</b> | <b>160,365</b> | <b>2</b> |               |       |
| 0 | 27,154         | 27,137         | 17         | 13,867         | 13,867         |          | 3.995         | 0.612 |
| 1 | 23,949         | 23,936         | 12         | 12,550         | 12,549         |          | 4.556         | 0.447 |
| 2 | 26,431         | 26,418         | 13         | 13,851         | 13,850         |          | 4.108         | 0.492 |
|   | 31,603         | 31,581         | 22         | 14,218         | 14,218         |          | 3.440         | 0.482 |
|   | 28,886         | 28,862         | 25         | 12,787         | 12,787         |          | 3.773         | 0.610 |
|   | 24,211         | 24,186         | 25         | 12,681         | 12,681         |          | 4.475         | 0.508 |
|   | 26,365         | 26,336         | 28         | 13,808         | 13,808         |          | 4.121         | 0.557 |
|   | 24,780         | 24,765         | 16         | 12,984         | 12,984         |          | 4.121         | 0.485 |
|   | 26,422         | 26,408         | 14         | 13,845         | 13,845         |          |               |       |
|   | 24,293         | 24,281         | 12         | 12,730         | 12,730         |          |               |       |
|   | 26,387         | 26,355         | 32         | 13,677         | 13,677         |          |               |       |
|   | 25,516         | 25,500         | 16         | 13,369         | 13,369         |          |               |       |

# Storage Spaces Direct quasi-demo

The screenshot displays the Failover Cluster Manager interface for a cluster named 'IgniteCluster:spectrum.local'. The Roles tab shows several virtual machines (VMs) running on the 'ignite1' node. The Task Manager window shows the CPU usage for the 'Intel(R) Xeon(R) CPU E5-2697 v3 @ 2.60GHz' is at 26%. A PowerShell window shows the results of an IOPS test, with a total of 254,961 IOPS across 10 nodes. A callout box highlights the 26% CPU utilization.

| Name               | Status  | Type            | Owner Node | Priority | Information |
|--------------------|---------|-----------------|------------|----------|-------------|
| vm-base-IGNITE1-19 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-20 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-21 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-18 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-15 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-16 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-17 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-26 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-27 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-28 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-25 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-22 | Running | Virtual Machine | ignite1    | Medium   |             |
| vm-base-IGNITE1-23 | Running | Virtual Machine | ignite1    | Medium   |             |

```
CSV FS IOPS
-----
total 254,961
ignite1 26,338
ignite10 24,408
ignite11 26,479
ignite12 27,090
ignite2 23,635
ignite3
ignite4
ignite5 25,024
ignite6 25,376
ignite7 23,650
ignite8 26,027
ignite9 26,935
```

**CPU** Intel(R) Xeon(R) CPU E5-2697 v3 @ 2.60GHz  
26% 2.57 GHz

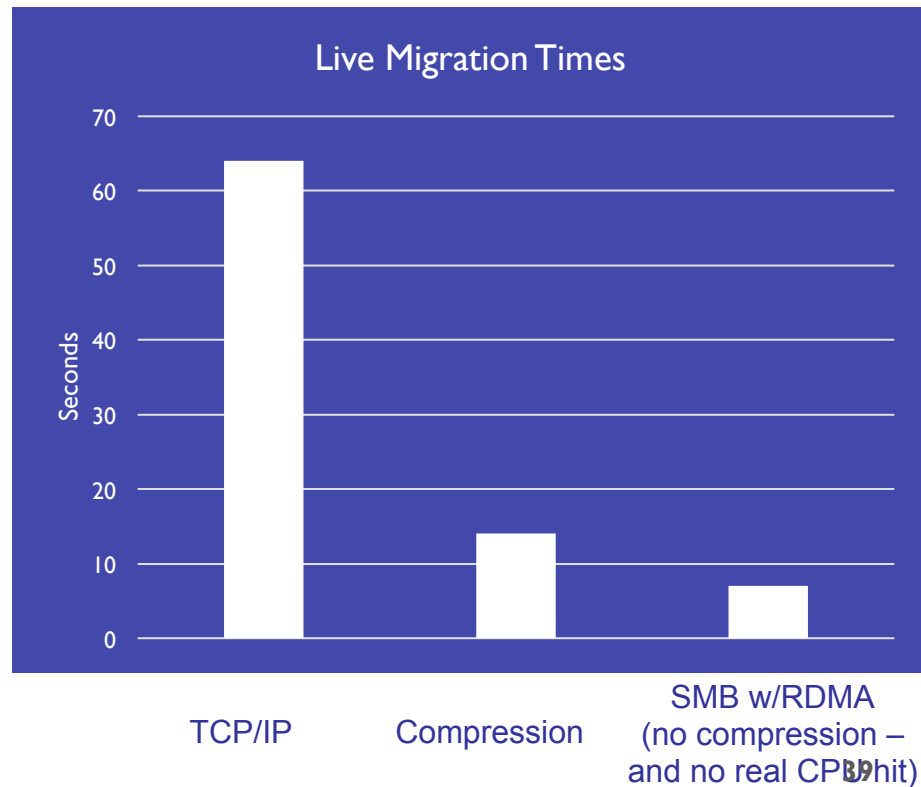
**write** Read Lat (ms) Write Lat  
3 1 4.073 0.622  
4.504 1.386  
4.099 2.038  
4.005 1.497  
4.588 1.046  
- -  
- -  
4.378 0.465  
4.282 1.458  
4.598 1.170  
4.168 0.935  
4.029 1.214

Only 25% CPU  
utilization

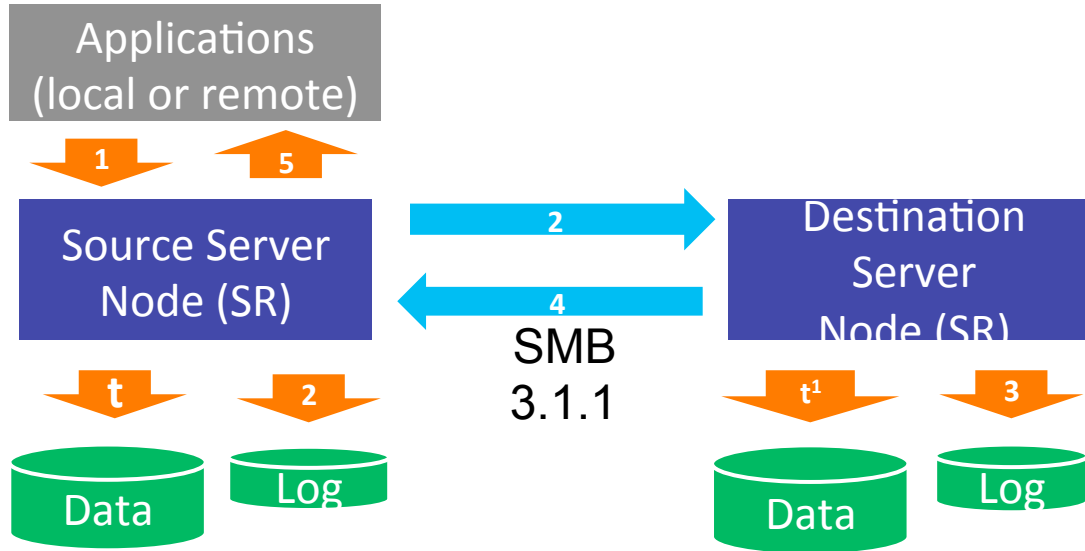
## SMB transport for Live Migration

RDMA

Multichannel over multiple NICs



# Storage Replica (block replica engine for DR)





# SMB3 Users & Vendors

Apple

EMC

NetApp

Ryussi

Tuxera

Visuality

Etc...

SMBI removal - [aka.ms/stopusingsmb1](https://aka.ms/stopusingsmb1)

SOFS best practices - [aka.ms/sofsyaynay](https://aka.ms/sofsyaynay)

SMB & Windows Server - [aka.ms/windowsserver](https://aka.ms/windowsserver)

SMB team blog - [aka.ms/serverstorage](https://aka.ms/serverstorage)

Good SMB blog - [aka.ms/josesmb3](https://aka.ms/josesmb3)

# 2017 SNIA SMB3 Plugfest

- Co-located at SNIA's Storage Developer Conference, September 11-14, Santa Clara, CA
- Enables vendors to bring their implementations of SMB3 to test, identify, and fix bugs in a collaborative setting with the goal of developing interoperable products
- 20+ companies and 80+ attendees are expected
- Evening Reception and Open House scheduled for September 11<sup>th</sup>

# 2016 SMB3 Plugfest

## SNIA SMB3 PLUGFEST PARTICIPANTS





# Final questions

# After This Webcast

- Please rate this webcast and provide us with feedback
- This webcast and a PDF of the slides will be posted to the SNIA Ethernet Storage Forum (ESF) website and available on-demand
- [www.snia.org/forums/esf/knowledge/webcasts](http://www.snia.org/forums/esf/knowledge/webcasts)
- A full Q&A from this webcast, including answers to questions we couldn't get to today, will be posted to the SNIA-ESF blog: [sniaesfblog.org](http://sniaesfblog.org)
- Follow us on Twitter @SNIAESF

**Thank you!**