

Azure Stack HCI

Monitoring and Troubleshooting – High Level



Hyperconverged Architecture

Hyperconverged Troubleshooting!

Azure Stack HCI consists of the following technologies each with their own lengthy troubleshooting methodologies.

Technologies are:

- Hyper-V
- Failover Clustering
- Storage Spaces Direct
- Windows Admin Center
- Networking
- RDMA
- Software Defined Networking

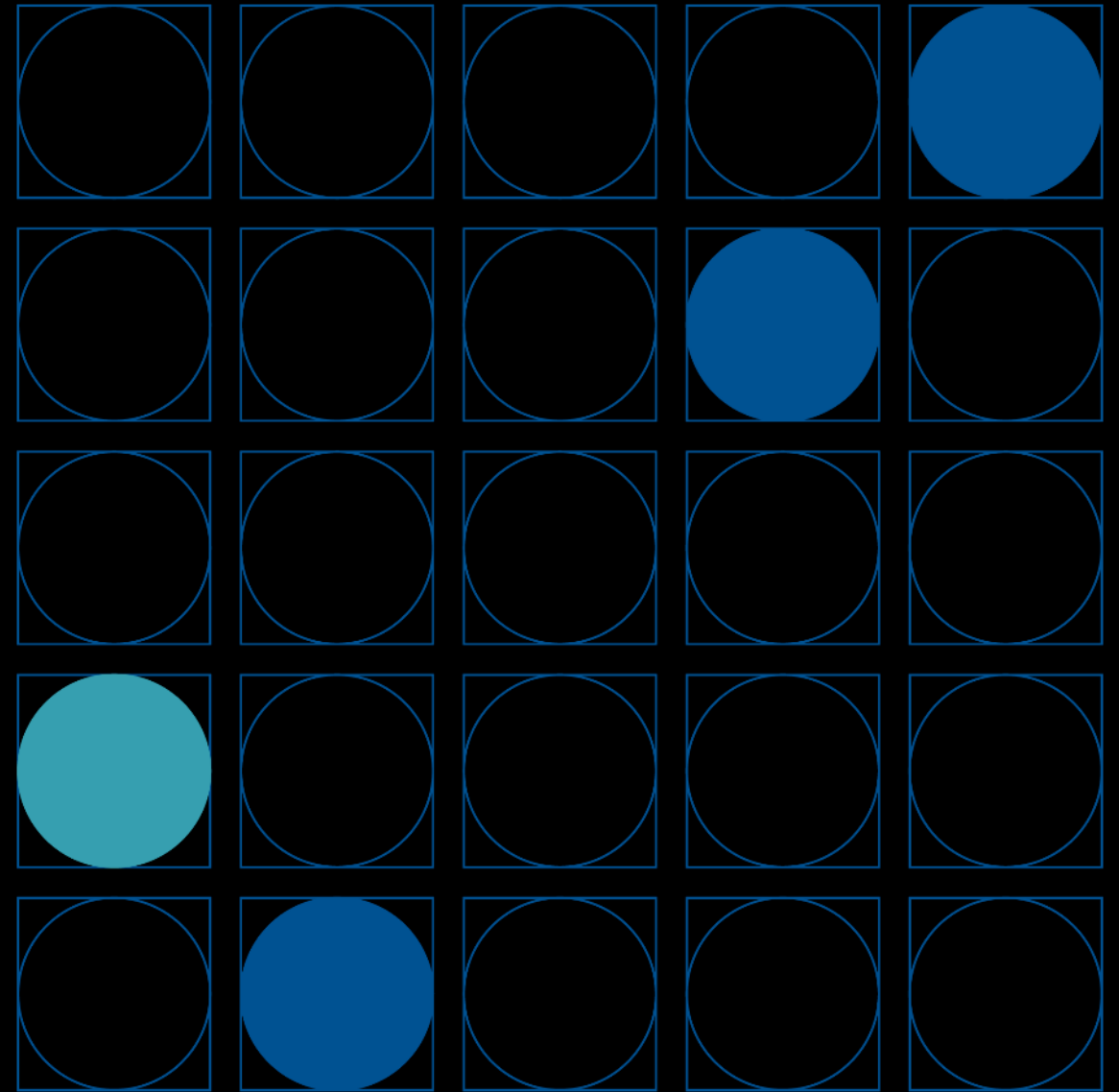


It would take a lot of time to cover every support situation or tools for these technologies.

So, let's look at monitoring and troubleshooting Azure Stack HCI at a high level.



Windows Admin Center Troubleshooting Tools



Diagnostics

Use this tool to collect information for troubleshooting problems with your cluster. If you call Microsoft Support, they may ask for this information.

azstackcluster.liteware.com

Tools

Search Tools

Azure Kubernetes Service

Storage

Volumes

Drives

Storage Replica

Networking

Logical networks

Virtual networks

Access control lists

Route tables

Public IP addresses

Load balancers

Gateway connections

SDN monitoring

Virtual switches

Tools

Azure Monitor

Updates

Diagnostics

Diagnostics

Use this tool to collect information for troubleshooting problems with you

Install diagnostic tools

Hyper-converged diagnostic tools ⓘ

1.1.28

Latest available version ⓘ

1.1.28

Networking diagnostic tools ⓘ

0.0.0.12

Latest available version ⓘ

0.0.0.12

✓ Good news! You're up to date.

Collect diagnostic information

View collected diagnostic information

ⓘ There are no archives—you need to collect diagnostic information first

SDN Monitoring – Data Path Diagnostics

[Summary](#) [Network Controller](#) [Load Balancer](#) [Gateway](#) [Data Path Diagnostics](#)

1 **Diagnostics Scenarios**

2 Capture Conditions

Choose the scenario that you would like to diagnose from of the following scenarios:

WorkLoad Scenarios :

- ☒ Virtual Machine to Virtual Machine Connection ⓘ
- ☐ Outbound NAT Scenario ⓘ
- ☐ Inbound NAT Scenario ⓘ
- ☐ Public IP Scenario ⓘ
- ☐ Public/Private VIP load balancing Scenario ⓘ
- ☐ Internal load balancing CA VIP Scenario ⓘ
- ☐ Gateway Scenario ⓘ

Infrastructure Scenarios :

- ☐ Border Gateway Protocol (BGP) traffic between resources and the switch ⓘ
- ☐ Network Controller and Host Agents communication ⓘ
- ☐ Network Controller and Multiplexers communication ⓘ
- ☐ Network Controller and Gateways communication ⓘ

4

[Back](#) [Next: Capture Conditions](#)

Azure Monitor in WAC

Tools

Search Tools

Storage

Volumes

Drives

Storage Replica

Networking

Logical networks

Virtual networks

Access control lists

Route tables

Public IP addresses

Load balancers

Gateway connections

SDN monitoring

Virtual switches

Tools

Azure Monitor

Updates

Monitoring and alerts with Azure Monitor

2 of 2 nodes in the cluster are connected to [logAnalytic-ws](#)

Azure Monitor connection

Onboard cluster2 items

Node	Onboarding status	Connected to workspace
s2d2.wcurtis.net	Connected	logAnalytic-ws
s2d1.wcurtis.net	Connected	logAnalytic-ws

[How do I stop using Azure to monitor a server?](#)

Alerts created in Windows Admin Center

To get notifications via email or text messages, or to automatically react to events and performance counters, create an alert in Azure Monitor.

0

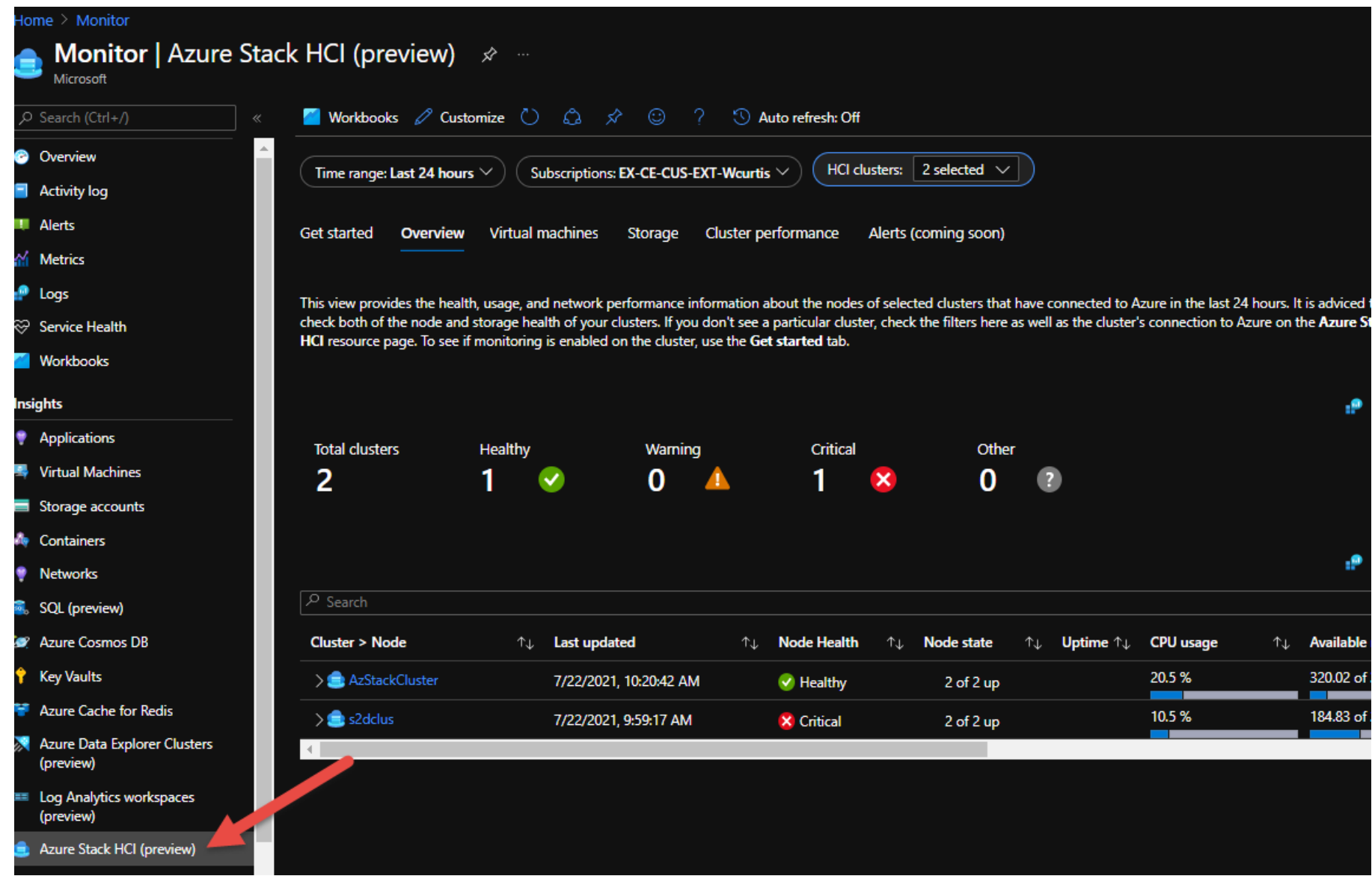
Based on workspace

To view alerts including alerts created in Azure, [visit the Azure portal.](#)

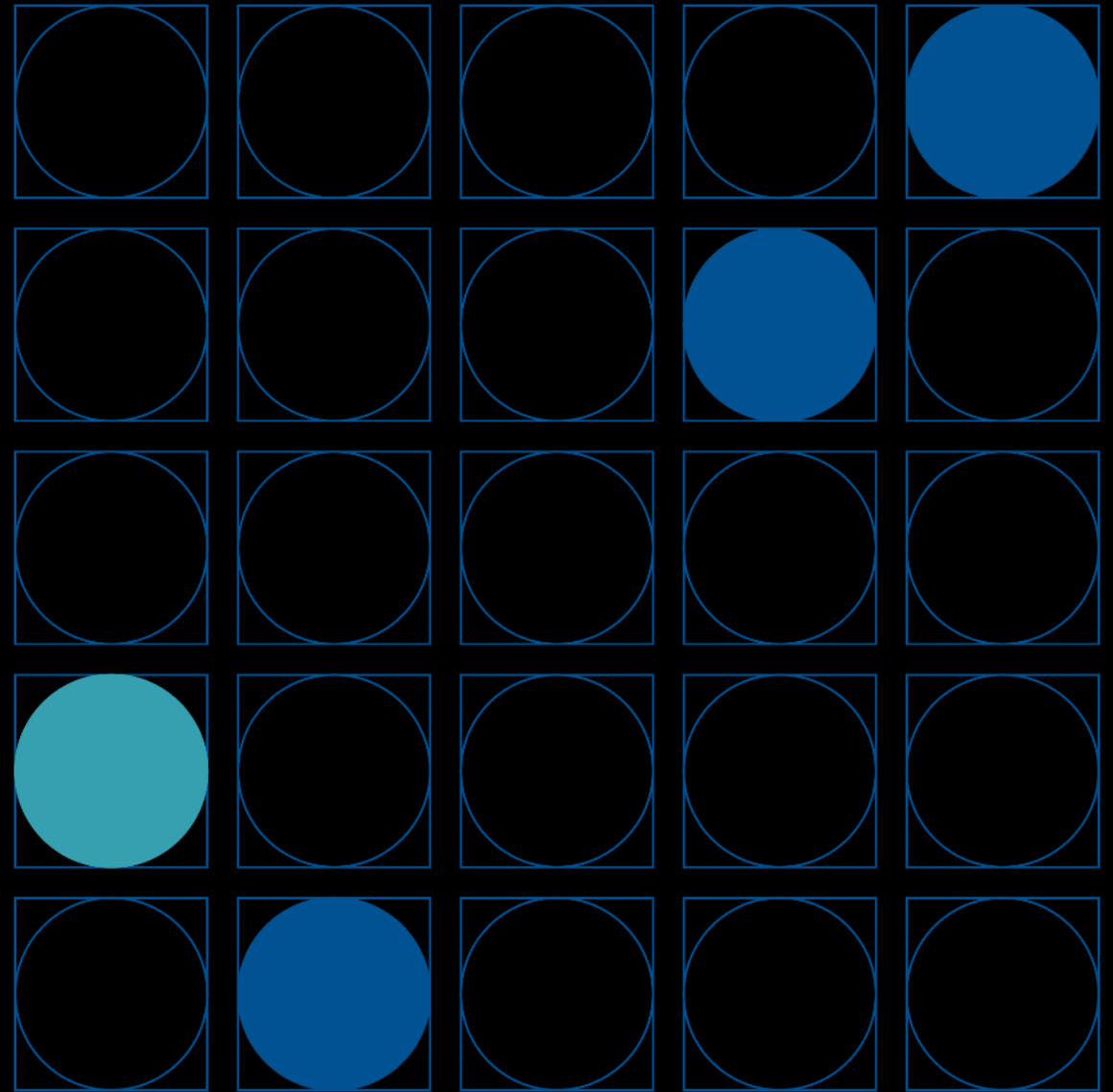
Edit alert rule0 items

Alert name	Resource type	Action group
No records found		

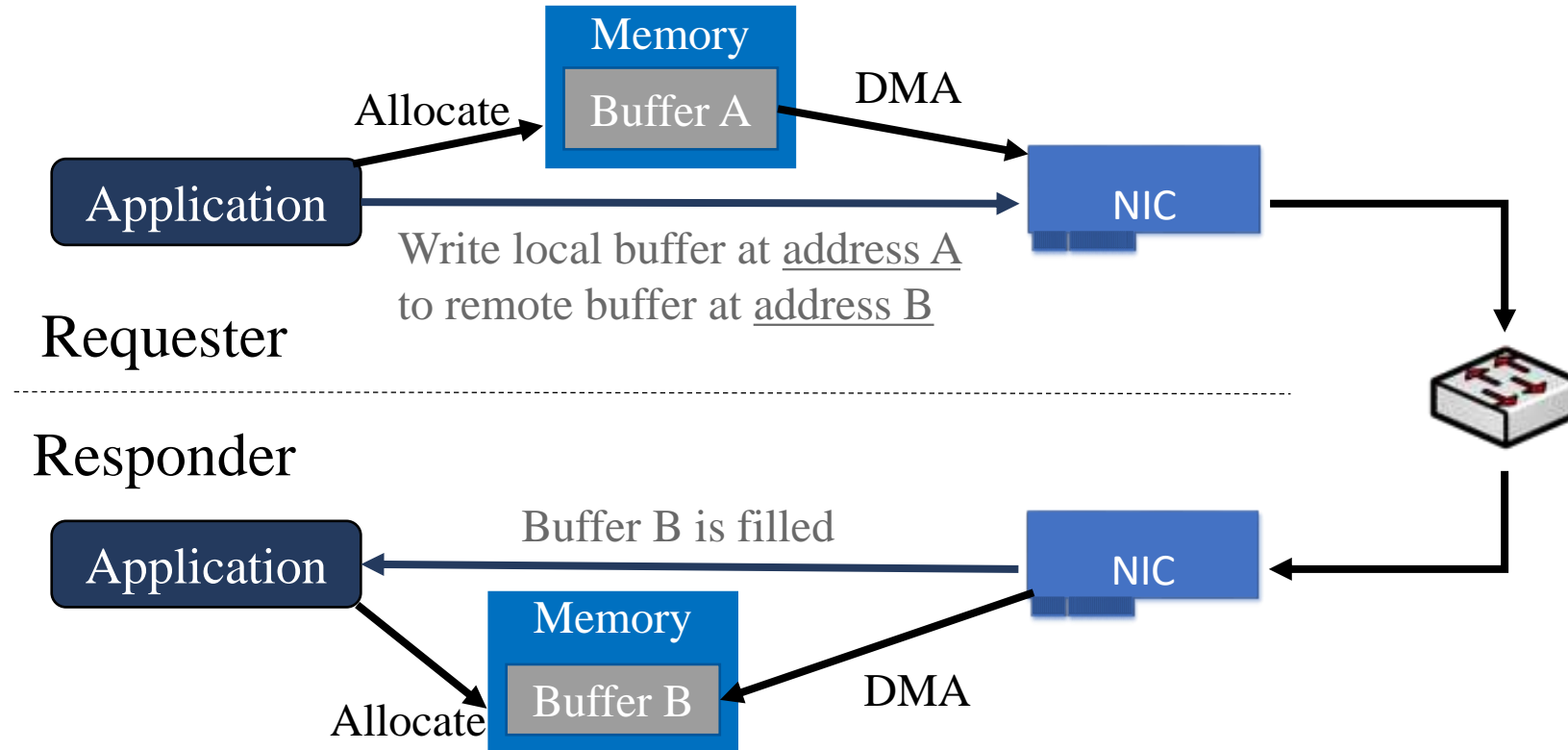
Azure Monitor – Azure Stack HCI and ARC



RDMA Troubleshooting



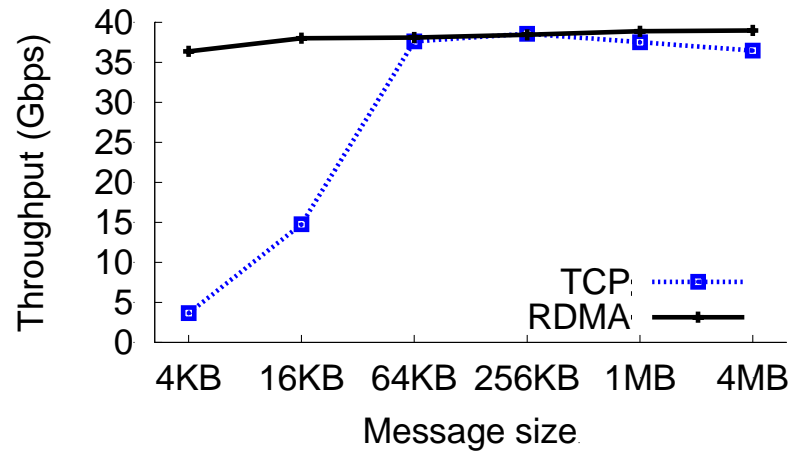
RDMA –Basic flow



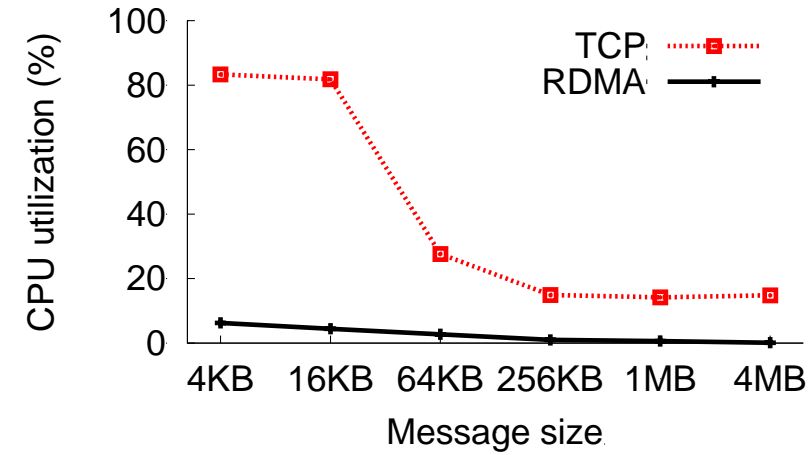
RDMA bypasses host OS stack
→ frees host CPU, lowers latency

Why use RDMA? RDMA Outperforms TCP

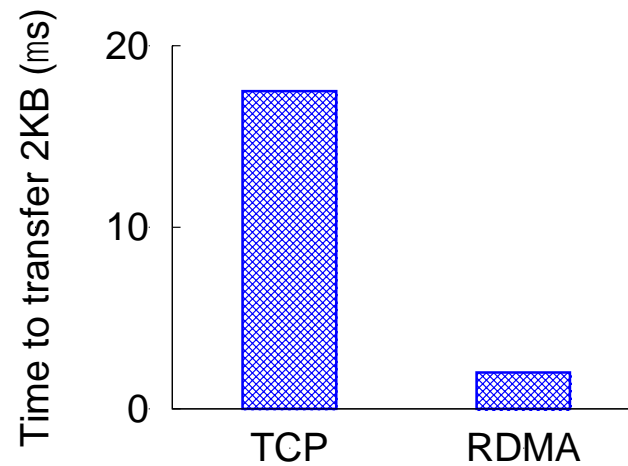
Micro Benchmarks:



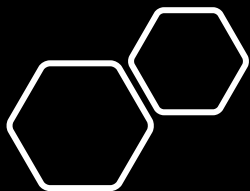
RDMA single thread ~40Gbps



RDMA CPU ~0%



RDMA latency 1~2 μ s



So, what do you do when Storage is Faster with RDMA Disabled?

- Check RDMA Activity Performance Counters
- Check your Switch Configuration (PFC Settings)
- Check your DCB settings on your Hyper-V Host (Validate DCB)
- If possible, switch out RDMA technologies (RoCE to iWARP).

\\s2d1

RDMA Activity	Mellanox ConnectX-3 Ethernet Adapter	Mellanox ConnectX-3 Ethernet Adapter #2
RDMA Accepted Connections	430.000	430.000
RDMA Active Connections	8.000	8.000
RDMA Completion Queue Errors	0.000	0.000
RDMA Connection Errors	0.000	0.000
RDMA Failed Connection Attempts	0.000	0.000
RDMA Inbound Bytes/sec	213,144.457	22,262.655
RDMA Inbound Frames/sec	262.834	41.131
RDMA Initiated Connections	4.000	4.000
RDMA Outbound Bytes/sec	301,813.867	80,130.310
RDMA Outbound Frames/sec	333.057	90.287

\\s2d2

RDMA Activity	Mellanox ConnectX-3 Ethernet Adapter	Mellanox ConnectX-3 Ethernet Adapter #2
RDMA Accepted Connections	4.000	4.000
RDMA Active Connections	8.000	8.000
RDMA Completion Queue Errors	0.000	0.000
RDMA Connection Errors	0.000	0.000
RDMA Failed Connection Attempts	0.000	0.000
RDMA Inbound Bytes/sec	321,163.236	76,312.522
RDMA Inbound Frames/sec	365.036	90.256
RDMA Initiated Connections	430.000	430.000
RDMA Outbound Bytes/sec	224,256.282	23,982.044
RDMA Outbound Frames/sec	264.751	41.117

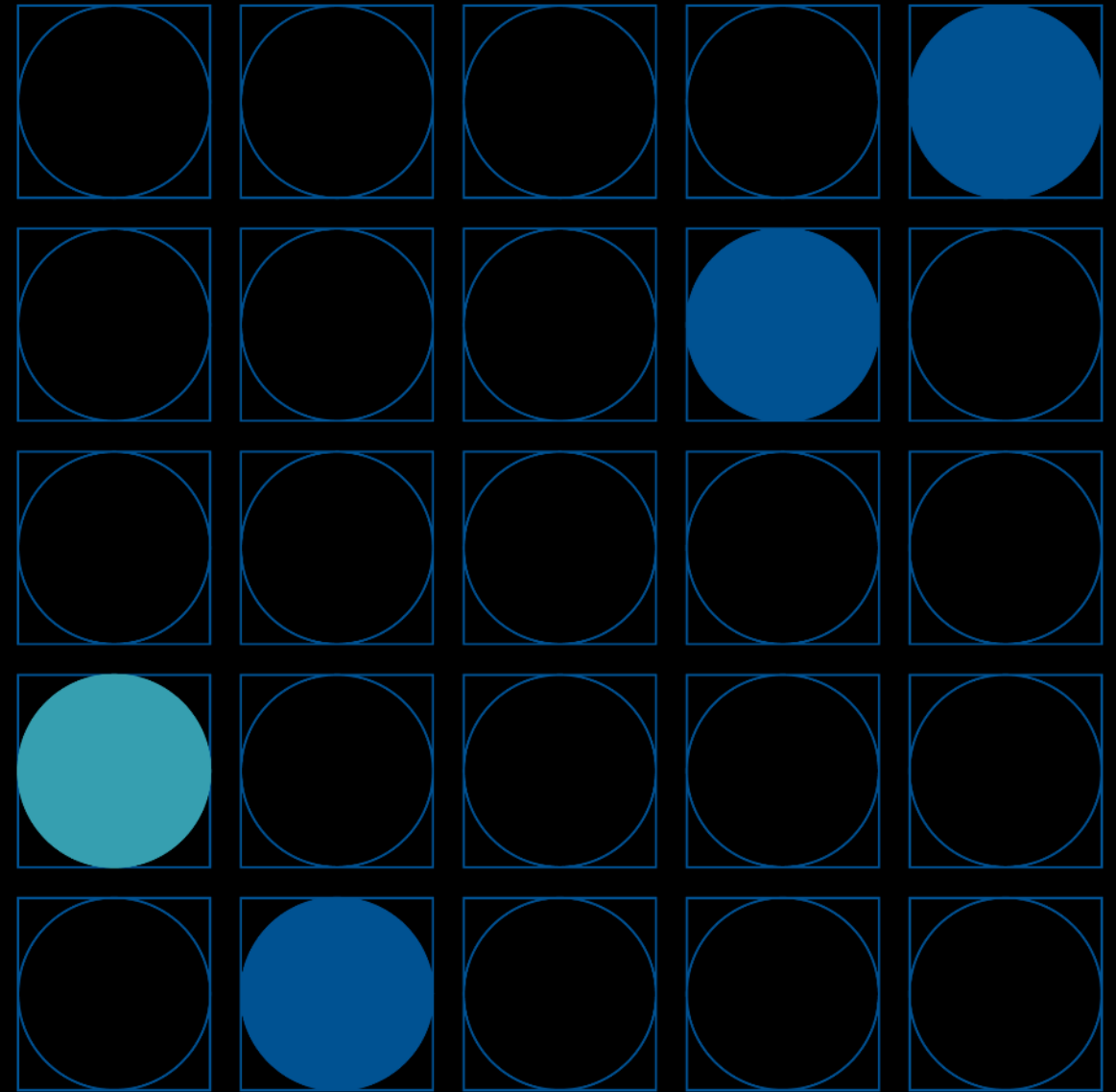
Validate DCB

- Validate-DCB is a module that validates the RDMA and Data Center Bridging (DCB) configuration best practices on Windows.
- Validate-DCB v2.1 is a PowerShell-based unit test tool that allows you to:
 - ✓ Validate the expected configuration on one to N number of systems or clusters
 - ✓ Validate the configuration meets best practices

Test-RDMA

- Runs diskspd over network and tests performance
- Located at: [SDN/Diagnostics at master · microsoft/SDN \(github.com\)](#)
- To test RDMA with this script, schedule maintenance or use test-clusterhealth from DISKSPD VM Fleet, [here](#)
- Example: `.\Test-RDMA.ps1 -IfIndex 6 -RemoteIpAddress '192.168.1.2' -PathToDiskspd 'c:\temp\dispd\amd64fre'`

Troubleshooting Tools



Test-ClusterHealth.ps1

- <https://github.com/Microsoft/diskspd/blob/master/Frameworks/VMFleet/test-clusterhealth.ps1>
- Checks Volumes health, SMB Connectivity errors, cluster symmetry.
- Recommended to run even in production environment

Copy/paste to
PowerShell ISE to run
or import as PS
CMDlet

Needs to run from
cluster node.

```
PS C:\windows\system32> C:\Users\stevenek\Documents\test-clusterhealth.ps1
***** Basic Health Checks (23.0s)
All cluster nodes Up
All operational pools Healthy
***** Clusport Device Symmetry Check (24.8s)
***** Total
Pass with 9 per node
***** Disk Type
Pass with 6 per node
***** Solid/Non-Rotational Media
Pass with 3 per node
***** Enclosure Type
Pass with 3 per node
***** Virtual
Pass with 3 per node
***** Enclosure View Symmetry Check (26.5s)
***** Total
Pass with 3 per node
***** Operational Issues and Storage Jobs (29.9s)
No storage rebuild or regeneration jobs are active
***** Physical Disk Health (22.7s)
All physical disks are in normal auto-select or journal state
***** Physical Disk View Symmetry Check (27.3s)
```


Get-ClusterDiagnosticInfo

- You can specify `–verbose` to reveal more info
- This command collects all necessary cluster info from each node
 - Event log
 - Cluster log
 - Systeminfo
 - Basic info in XML files
- Result is stored in ZIP

 HealthTest-S2D-Cluster-20160826-1551 8/26/2016 3:54 PM Compressed (zipp... 10,934 KB

- Usage:
 - `Get-ClusterDiagnosticInfo –Cluster $ClusterName`

Administrator: C:\Windows\system32\cmd.exe

```
PS C:\> Get-ClusterDiagnosticInfo
Writing to path : C:\Users\wcurtis\HealthTest\
Script Version : 0.1
Date of capture : 07/22/2021 07:33:42
```

```
<<< Phase 1 - Cluster Overview >>>
```

Examining each VM

VM 6 of 7

[oo]

Got remainder cluster objects & exporting groups resources

Exporting remaining objects

Getting cluster properties

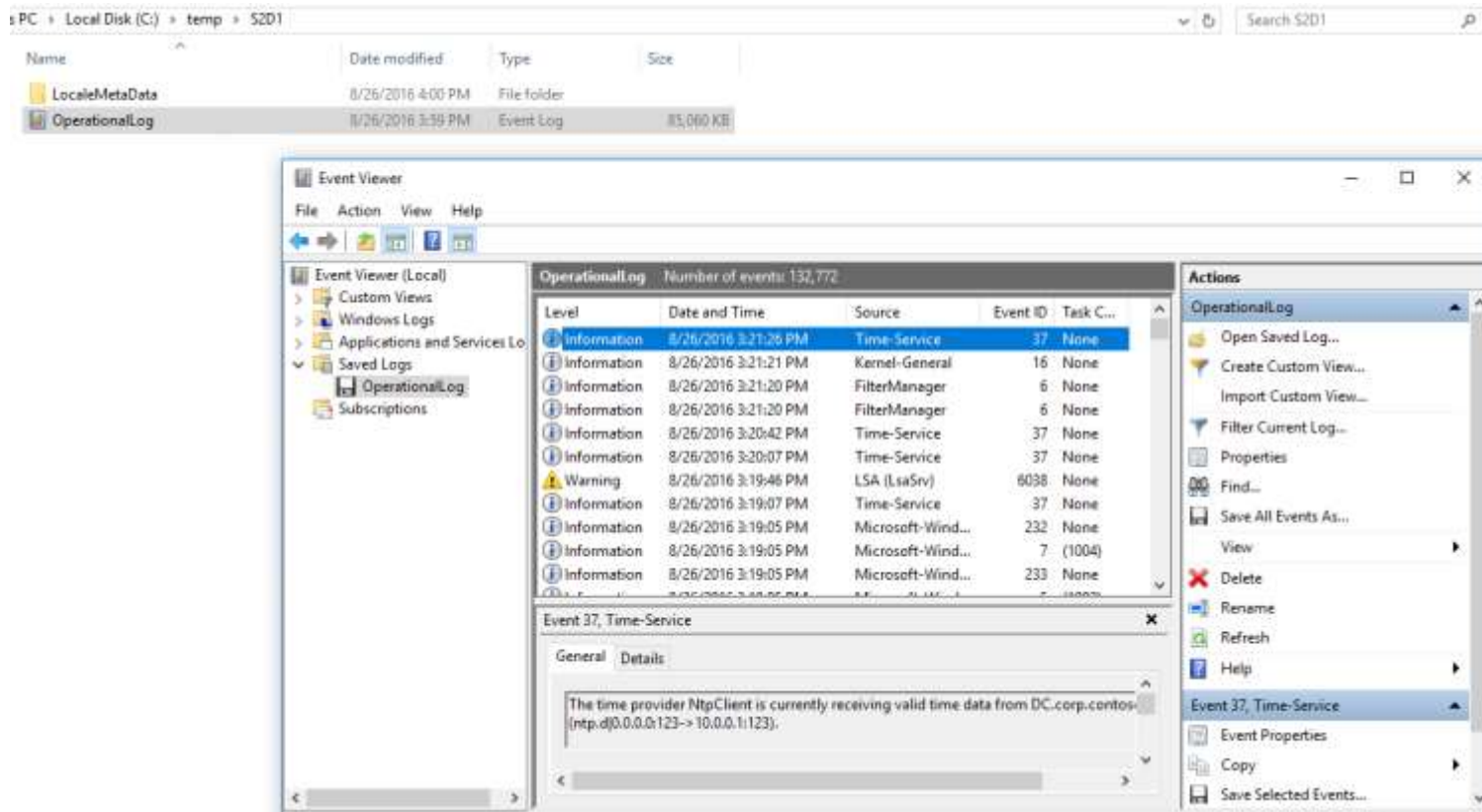
Getting VMs

Exporting vms

The states and counts of vms are

state	count
Running	7

Get-StorageDiagnosticInfo



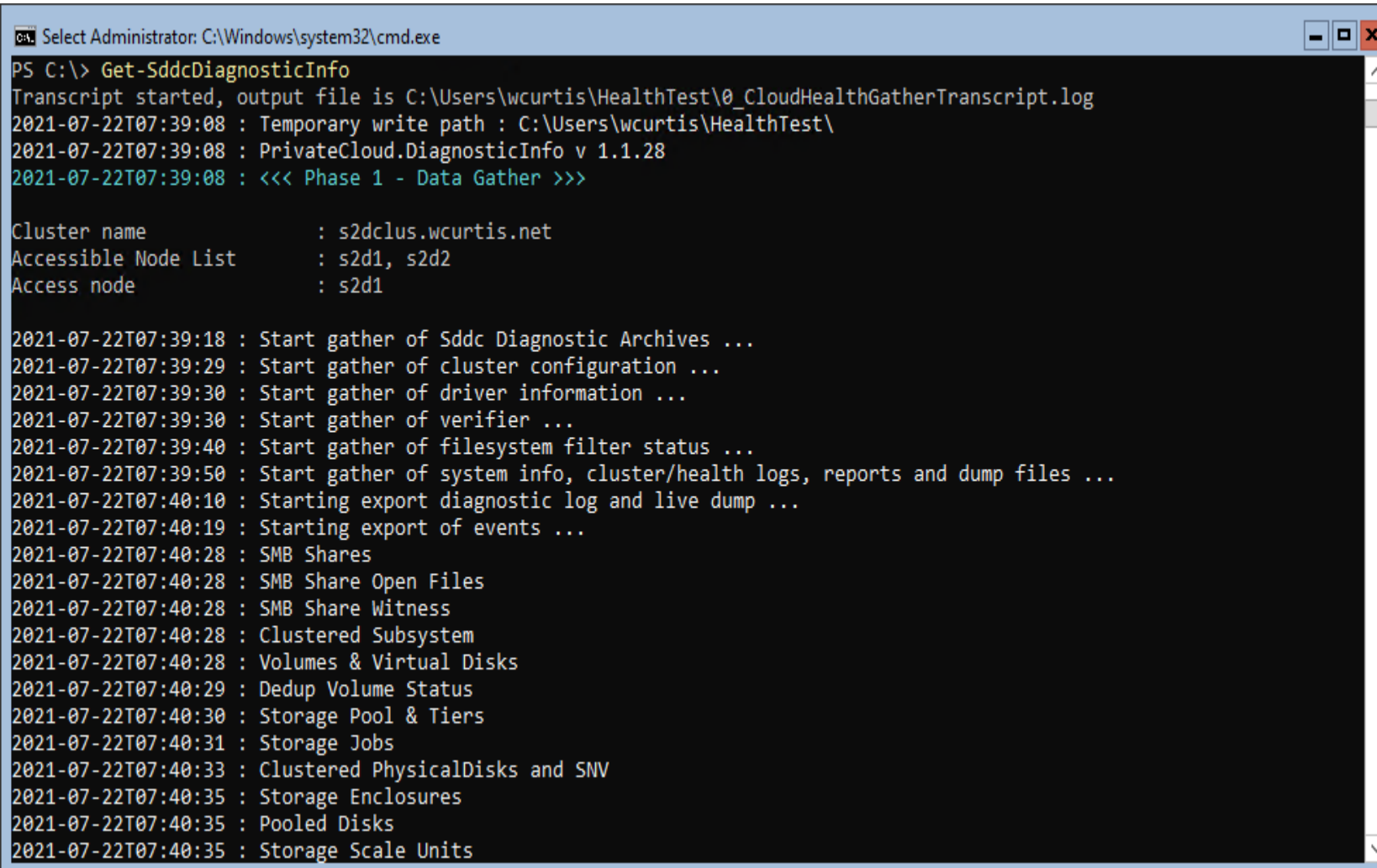
Generates log for each node

Log consists of all storage related info

Usage

- Get-StorageSubSystem - FriendlyName *\$ClusterName | Get-StorageDiagnosticInfo - DestinationPath c:\temp

Get-SddcDiagnosticInfo



```
Select Administrator: C:\Windows\system32\cmd.exe

PS C:\> Get-SddcDiagnosticInfo
Transcript started, output file is C:\Users\wcurtis\HealthTest\0_CloudHealthGatherTranscript.log
2021-07-22T07:39:08 : Temporary write path : C:\Users\wcurtis\HealthTest\
2021-07-22T07:39:08 : PrivateCloud.DiagnosticInfo v 1.1.28
2021-07-22T07:39:08 : <<< Phase 1 - Data Gather >>>

Cluster name           : s2dclus.wcurtis.net
Accessible Node List   : s2d1, s2d2
Access node            : s2d1

2021-07-22T07:39:18 : Start gather of Sddc Diagnostic Archives ...
2021-07-22T07:39:29 : Start gather of cluster configuration ...
2021-07-22T07:39:30 : Start gather of driver information ...
2021-07-22T07:39:30 : Start gather of verifier ...
2021-07-22T07:39:40 : Start gather of filesystem filter status ...
2021-07-22T07:39:50 : Start gather of system info, cluster/health logs, reports and dump files ...
2021-07-22T07:40:10 : Starting export diagnostic log and live dump ...
2021-07-22T07:40:19 : Starting export of events ...
2021-07-22T07:40:28 : SMB Shares
2021-07-22T07:40:28 : SMB Share Open Files
2021-07-22T07:40:28 : SMB Share Witness
2021-07-22T07:40:28 : Clustered Subsystem
2021-07-22T07:40:28 : Volumes & Virtual Disks
2021-07-22T07:40:29 : Dedup Volume Status
2021-07-22T07:40:30 : Storage Pool & Tiers
2021-07-22T07:40:31 : Storage Jobs
2021-07-22T07:40:33 : Clustered PhysicalDisks and SNV
2021-07-22T07:40:35 : Storage Enclosures
2021-07-22T07:40:35 : Pooled Disks
2021-07-22T07:40:35 : Storage Scale Units
```

Best Tool to gather Azure Stack HCI Cluster Information.

Can be downloaded and run manually or run from Windows Admin Center.

....or just use Windows Admin Center!

The screenshot shows the Windows Admin Center interface for a cluster named 's2dclus'. The left sidebar contains a 'Tools' section with a search bar and a list of categories: Drives, Storage Replica, Networking, Virtual switches, Virtual networks, Access control lists, Logical networks, Gateway connections, SDN monitoring, Tools, Azure Monitor, Updates, Diagnostics (highlighted with a red box and a red arrow), Performance Monitor, and Extensions. The main content area is titled 'Diagnostics' and includes a description: 'Use this tool to collect information for troubleshooting problems with your cluster. If you call Microsoft Support, they may ask for this information.' Below this is a section 'Install diagnostic tools' with a table showing the status of diagnostic tools. The table has two columns: the tool name and the latest available version. The first row is for 'Hyper-converged diagnostic tools' with version '1.1.28'. The second row is for 'Networking diagnostic tools' with version '0.0.0.12'. Below the table is a green checkmark icon and the text 'Good news! You're up to date.' Underneath is a section 'Collect diagnostic information' with a large blue button that says 'Collecting diagnostic information. This may take a few minutes.' Below this is a section 'View collected diagnostic information' with an information icon and the text 'There are no archives—you need to collect diagnostic information first.'

Windows Admin Center | Cluster Manager

Microsoft

s2dclus

Tools

Search Tools

Drives

Storage Replica

Networking

Virtual switches

Virtual networks

Access control lists

Logical networks

Gateway connections

SDN monitoring

Tools

Azure Monitor

Updates

Diagnostics

Performance Monitor

Extensions

Load balancers

Public IP addresses

Diagnostics

Use this tool to collect information for troubleshooting problems with your cluster. If you call Microsoft Support, they may ask for this information.

Install diagnostic tools

	Latest available version
Hyper-converged diagnostic tools	1.1.28
Networking diagnostic tools	0.0.0.12

Good news! You're up to date.

Collect diagnostic information

Collecting diagnostic information. This may take a few minutes.

View collected diagnostic information

There are no archives—you need to collect diagnostic information first.

Connecting remote storage subsystem

Useful when working with remote storage. Debug-StorageSubsystem and Get-StorageDiagnosticInfo benefit from this

```
#register storage subsystem
Get-StorageProvider | Register-StorageSubsystem -ComputerName $ClusterName

#unregister storage subsystem
$ss=Get-StorageSubSystem -FriendlyName *$ClusterName
Unregister-StorageSubsystem -ProviderName "Windows Storage Management Provider"
-StorageSubSystemUniqueId $ss.UniqueId
```

Get-StorageReliabilityCounter

```
Administrator: Windows PowerShell
PS C:\Windows\system32> Get-PhysicalDisk | Get-StorageReliabilityCounter | fl

ObjectId           : {1}\JAROMIRK-WIN10\root/Microsoft/Windows/Storage/Providers_v2\SPACES_StorageReliabilityCounter.ObjectId="{f6973183-510b-11e6-933f-806e6f6e6963}:RC:{2885aeb2-bd05-d0f3-7866-8a2145915eae}"
PassThroughClass    :
PassThroughIds      :
PassThroughNamespace :
PassThroughServer    :
UniqueId            : {f6973183-510b-11e6-933f-806e6f6e6963}:RC:{2885aeb2-bd05-d0f3-7866-8a2145915eae}
DeviceId             : 0
FlushLatencyMax     : 106
LoadUnloadCycleCount :
LoadUnloadCycleCountMax :
ManufactureDate      :
PowerOnHours         : 4372
ReadErrorsCorrected  : 0
ReadErrorsTotal      : 0
ReadErrorsUncorrected :
ReadLatencyMax       : 6133
StartStopCycleCount  :
StartStopCycleCountMax :
Temperature          : 0
TemperatureMax        : 0
Wear                  : 0
WriteErrorsCorrected :
WriteErrorsTotal      :
WriteErrorsUncorrected :
WriteLatencyMax       : 6119
PSComputerName       :
CimClass              : ROOT/Microsoft/Windows/Storage:MSFT_StorageReliabilityCounter
CimInstanceProperties : {ObjectId, PassThroughClass, PassThroughIds, PassThroughNamespace...}
CimSystemProperties   : Microsoft.Management.Infrastructure.CimSystemProperties

PS C:\Windows\system32>
```

- Displays SCSI Primary Commands 4 (SPC-4).
- Can display data even with Storage Spaces enabled (unlike SMART)
- Counters
 - Wear
 - SSD Wear level
 - Some disks report 0 as healthy
 - Some disks report 100 as healthy
 - Temperature
 - Errors
 - PowerOnHours

Looking for a Bad Disk

#Look for bad disks (OperationalStatus Not Equal OK)

```
Get-StorageSubSystem -CimSession $Cluster -FriendlyName "Clustered  
Windows Storage*" | Get-PhysicalDisk | where OperationalStatus -NE OK
```

```
PS C:\Users\administrator.WSSDCORP> Get-StorageSubSystem -CimSession $Cluster -FriendlyName "Clustered Windows Storage*" | Get-PhysicalDisk | ? OperationalStatus -NE OK
```

FriendlyName	SerialNumber	CanPool	OperationalStatus	HealthStatus	Usage	Size	PSComputerName
ATA MB6000GEBTP	1EK4Y4WJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4W41H	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4V0XJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MK1200GEYKF	BTHV611401U31P2PGN	False	In Maintenance Mode	Warning	Journal	1.09 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4WMJJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4Y3YJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MK1200GEYKF	BTHV611402J11P2PGN	False	In Maintenance Mode	Warning	Journal	1.09 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4Y6NJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4VV4J	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4WPMJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MK1200GEYKF	BTHV611402Q01P2PGN	False	In Maintenance Mode	Warning	Journal	1.09 TB	QP-HPR01Cluster
ATA MK1200GEYKF	BTHV611402Q81P2PGN	False	In Maintenance Mode	Warning	Journal	1.09 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4Y2RJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster
ATA MB6000GEBTP	1EK4VLDJ	False	In Maintenance Mode	Warning	Auto-Select	5.46 TB	QP-HPR01Cluster

See drives within specific server

```
Administrator: Windows PowerShell
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

PS C:\> Get-ClusterNode

Name                State Type
----                -
261193E04-19-1      Up   Node
261193E04-19-2      Up   Node
261193E04-19-3      Up   Node
261193E04-19-4      Up   Node

PS C:\> Get-StorageSubSystem Cluster* | Get-StorageNode

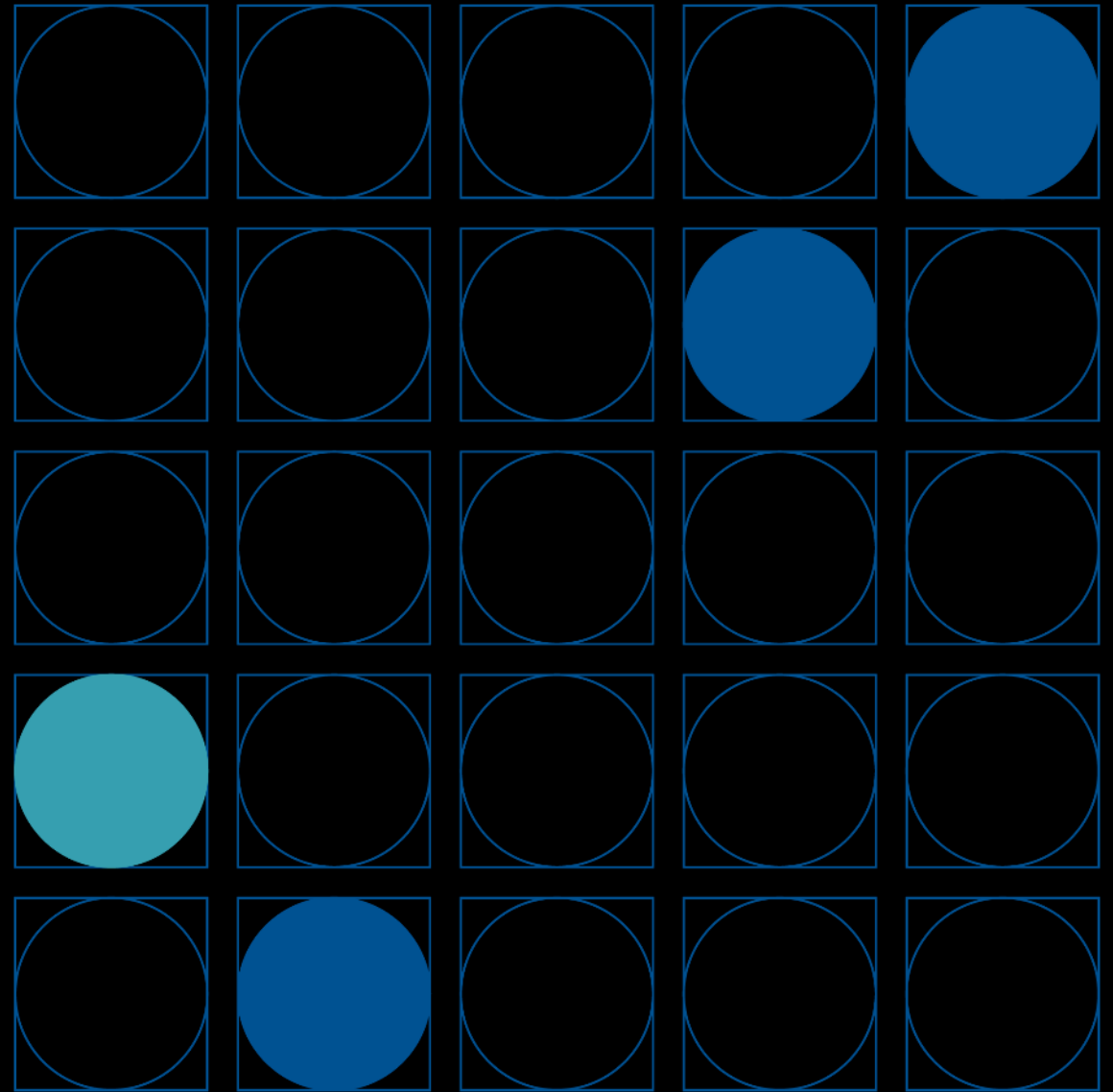
Name                Manufacturer          Model          SerialNumber      OperationalStatus
----                -
261193E04-19-1.cfdev.nttest.microsoft.com Quanta Cloud Technology Inc. QuantaPlex T41SP-2U QTFCKJ60500010 Up
261193E04-19-2.cfdev.nttest.microsoft.com Quanta Cloud Technology Inc. QuantaPlex T41SP-2U QTFCKJ60500010 Up
261193E04-19-3.cfdev.nttest.microsoft.com Quanta Cloud Technology Inc. QuantaPlex T41SP-2U QTFCKJ60500010 Up
261193E04-19-4.cfdev.nttest.microsoft.com Quanta Cloud Technology Inc. QuantaPlex T41SP-2U QTFCKJ60500010 Up

PS C:\> Get-StorageNode 261193E04-19-3.cfdev.nttest.microsoft.com | Get-PhysicalDisk -PhysicallyConnected

FriendlyName      SerialNumber      MediaType CanPool OperationalStatus HealthStatus Usage      Size
-----
ATA SAMSUNG MZ7KM1T9 S2HNNYAG800111 SSD      False OK      Healthy      Auto-Select 1.75 TB
ATA SAMSUNG MZ7KM1T9 S2HNNYAG800105 SSD      False OK      Healthy      Auto-Select 1.75 TB
ATA SAMSUNG MZ7KM1T9 S2HNNYAG800110 SSD      False OK      Healthy      Auto-Select 1.75 TB
ATA SAMSUNG MZ7KM1T9 S2HNNYAG800093 SSD      False OK      Healthy      Auto-Select 1.75 TB
INTEL SSDPE2MD800G4 0000_0000_0000_0000. SSD      False OK      Healthy      Journal 745 GB
INTEL SSDPE2MD800G4 0000_0000_0000_0000. SSD      False OK      Healthy      Journal 745 GB

PS C:\> _
```


Disk Performance



VMFleet



Tool for orchestrating diskspd in multiple VMs in multiple Volumes



Simulates real workload



Creates volumes and VMs



Designed to do tests before going into production



<https://github.com/Microsoft/WSLab/tree/master/Scenarios/VMFleet>

File copy to CSV as a Benchmark....



Always use `\\<ClusterName>\ClusterStorage$`



If you use `\\<NodeName>\c$\ClusterStorage` , it might fail



All IO to a VHD is unbuffered IO, and that is what CSV is optimized for.



With a file copy, you are performing buffered IO

doing a file copy (especially with Explorer) is the worst way to evaluate the performance of the system

XCOPY with the `/J` switch to do unbuffered IO

Diskspd "on CSV" vs "inside VM"

- on the CSV - **5,915 IOPS**

diskspd.exe -Z20M -Z -h -t1 -o8 -b4k -r4k -w0 -W30 -C30 -d300 -D -L
c:\ClusterStorage\collect\testfile1.dat

Total IO thread	bytes	I/Os	MB/s	I/O per s	AvgLat	IopsStdDev	LatStdDev	file
0	7268605952	1774562	23.11	5915.01	1.351	3906.49	3.856	c:\ClusterStorage\collect\testfile1.dat (10240MB)
total:	7268605952	1774562	23.11	5915.01	1.351	3906.49	3.856	

- Inside VM - **84,204 IOPS**

diskspd.exe -Z20M -Z -h -t1 -o8 -b4k -r4k -w0 -W30 -C30 -d300 -D -L
e:\shares\testfile1.dat

Total IO thread	bytes	I/Os	MB/s	I/O per s	AvgLat	IopsStdDev	LatStdDev	file
0	103470358528	25261318	328.92	84204.40	0.094	13324.50	0.304	e:\shares\testfile1.dat (10240MB)
total:	103470358528	25261318	328.92	84204.40	0.094	13324.50	0.304	

S2D Cache performance counters

- How much S2D cache writes to cache devices*:
- Cluster Storage Cache Stores\Update Bytes
- Cluster Storage Hybrid Disks\Cache Populate Bytes
- Cluster Storage Hybrid Disks\Cache Write Bytes
- How much S2D cache writes to capacity devices:
- Cluster Storage Hybrid Disks\Destage Bytes
- Cluster Storage Hybrid Disks\Direct Write Bytes
- *does not contain write amplification for resiliency (3x) and updating the read cache as workload churns ...

