WSSD Configuration using iWARP Mesh Topology

Quick Start Guide for Windows

Overview

Chelsio's fifth/sixth generation (T5/T6), high performance iWARP RDMA 10/25/40/50/100GbE adapters enable incremental, non-disruptive server installs, and support the ability to work without requiring any discrete external network switch, delivering a brownfield strategy to enable high performance, low cost, scalable WSSD deployments. Major benefits include cost savings on switches at higher speeds with each deployment. Windows SMB Direct over iWARP RDMA provides higher performance by giving direct access to the data residing on hyper-converged or disaggregated WSSD storage, while the CPU reduction enables a larger number of VMs per Hyper-V server, enabling savings in power dissipation, system configuration and deployment scale throughout the life of the installation. This document provides quick steps to configure WSSD on a 5-node mesh topology using Chelsio iWARP RDMA adapters.

Mesh Topology

This involves connecting each node to every other node. Supported configs using this approach: N ports per node, N+1 node cluster. The below is a 5-node Mesh using 4-port Chelsio T540-CR adapters. NIC ports on each server connected to each other (1<->2, 1<->3, 1<->4, 1<->5, 2<->3, 2<->4, 2<->5, 3<->4, 3<->5, 4<->5).

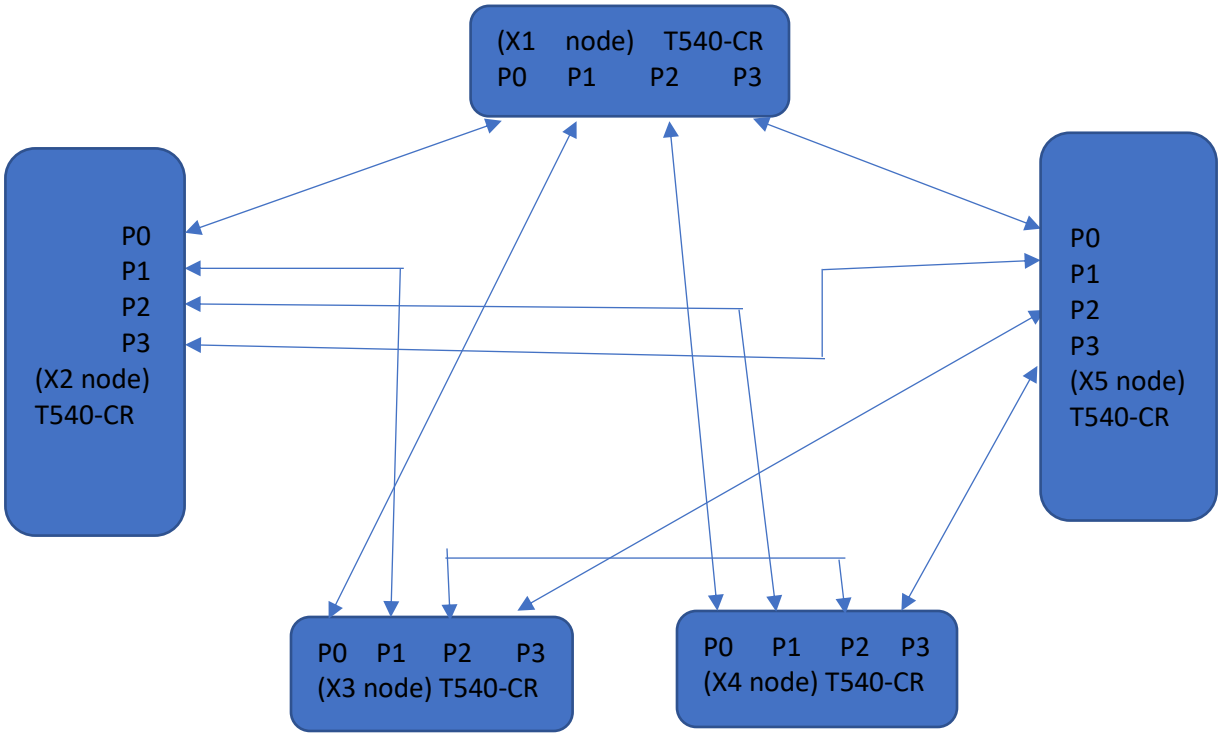

Figure 1 – 5-node Mesh Topology

WSSD configuration with no External Switch

Follow the steps mentioned below to install and configure WSSD using Chelsio Mesh topology:

Prerequisites:

- Please follow the hardware requirements provided by Microsoft for Windows Storage Spaces Direct (S2D).
- For high bandwidth, it is recommended to use NVMe drives for S2D capacity.

i. Connect all nodes in Mesh configuration as shown in Figure 1 (no switches required, saving $$$ for each deployment).
ii. Download the latest Chelsio Unified Wire for Windows from Chelsio Download Center.
iii. Install Unified Wire on all the nodes.

Note: Please refer the support documentation in the package for different installation methods.

iv. Follow the Step 1 of deployment guide for validating network connectivity and configuration.

Note: Step 2 and 3 (VLAN and DCB) in the above network guide are not required to configure with Chelsio IWARP RDMA.

v. On the Chelsio NIC interfaces, leave Dynamic IPv6 configuration (no need to assign any IPV4 addresses).

Note:

- Using IPv6, cluster network validation will be successful without any warnings.
- Using IPv4, cluster network validation will report some warnings/failures, but you can proceed with the configuration. In case of IPv4, if IP addresses in different subnets are used, cluster will create the required number of network subnets for node communication. Microsoft confirmed these network failures are expected with Mesh approach and will address these issues in future releases.

vi. Enable WSSD on the nodes using configuration guide from Microsoft.

Example:

i. Assign IP addresses (IPv4 or IPV6) on interfaces and enable WSSD using below steps:

```
#Test-Cluster –Node –Include “Storage Spaces Direct”,Inventory,Network,”System
  Configuration”
#New-Cluster –Name –Node < host1,host2,host3,host4> –NoStorage -StaticAddress
  -Verbose
#Enable-ClusterS2D -PoolFriendlyName - Verbose
#Get-ClusterNode |% { New-Volume - StoragePoolFriendlyName s2d -FriendlyName
  $_ -FileSystem CSVFS_ReFS -Size 500GB - Verbose }
#New-Volume -StoragePoolFriendlyName s2d -FriendlyName volume1 -FileSystem
  CSVFS_ReFS -Size 100GB -Verbose
```

ii. After successfully enabling WSSD on the cluster, you can run traffic by creating virtual disks. Also, you can create or deploy virtual machines. The virtual machine’s files should be stored on the virtual disks.

```
PS C:\Users\Administrator> netstat -xan

Active NetworkDirect Connections, Listeners, SharedEndpoints

  Mode   IfIndex Type            Local Address           Foreign Address         PID

  Kernel      13 Connection      [fe80::c027:49fa:bb9a:1946%13]:445   [fe80::b00b:a382:4c7c:70bd]:13530  0
  Kernel      13 Connection      [fe80::c027:49fa:bb9a:1946%13]:445   [fe80::b00b:a382:4c7c:70bd]:13532  0
  Kernel      13 Connection      [fe80::c027:49fa:bb9a:1946%13]:25181  [fe80::b00b:a382:4c7c:70bd]:445  0
  Kernel      13 Connection      [fe80::c027:49fa:bb9a:1946%13]:25180  [fe80::b00b:a382:4c7c:70bd]:445  0
  Kernel      13 Connection      [fe80::c027:49fa:bb9a:1946%13]:25214  [fe80::b00b:a382:4c7c:70bd]:445  0
  Kernel      13 Connection      [fe80::c027:49fa:bb9a:1946%13]:25215  [fe80::b00b:a382:4c7c:70bd]:445  0
  Kernel      24 Connection      [fe80::8829:ed6d:8e45:406b%24]:25160  [fe80::dd2a:bb85:f3f8:443c]:445  0
  Kernel      24 Connection      [fe80::8829:ed6d:8e45:406b%24]:25161  [fe80::dd2a:bb85:f3f8:443c]:445  0
  Kernel      24 Connection      [fe80::8829:ed6d:8e45:406b%24]:445   [fe80::dd2a:bb85:f3f8:443c]:29528  0
  Kernel      24 Connection      [fe80::8829:ed6d:8e45:406b%24]:445   [fe80::dd2a:bb85:f3f8:443c]:29527  0
  Kernel      22 Connection      [fe80::b4bc:81f5:a258:b917%22]:25143  [fe80::74b1:73b5:8dd3:415b]:445  0
  Kernel      22 Connection      [fe80::b4bc:81f5:a258:b917%22]:25142  [fe80::74b1:73b5:8dd3:415b]:445  0
  Kernel      22 Connection      [fe80::b4bc:81f5:a258:b917%22]:445   [fe80::74b1:73b5:8dd3:415b]:24133  0
  Kernel      22 Connection      [fe80::b4bc:81f5:a258:b917%22]:445   [fe80::74b1:73b5:8dd3:415b]:24134  0
  Kernel      22 Connection      [fe80::b4bc:81f5:a258:b917%22]:25166  [fe80::74b1:73b5:8dd3:415b]:445  0
  Kernel      22 Connection      [fe80::b4bc:81f5:a258:b917%22]:25167  [fe80::74b1:73b5:8dd3:415b]:445  0
  Kernel      13 Listener        [fe80::c027:49fa:bb9a:1946%13]:445   NA                      0
  Kernel      13 Listener        169.254.25.70:445       NA                      0
  Kernel      24 Listener        [fe80::8829:ed6d:8e45:406b%24]:445   NA                      0
  Kernel      24 Listener        169.254.64.107:445      NA                      0
  Kernel      22 Listener        [fe80::b4bc:81f5:a258:b917%22]:445   NA                      0
  Kernel      22 Listener        169.254.185.23:445      NA                      0
PS C:\Users\Administrator> _
```

Figure 2 – netstat showing RDMA connections

\\X2

Processor Information	_Total		
% Processor Time	1.065		
RDMA Activity	**Chelsio Network Adapter #2**	**Chelsio Network Adapter #3**	**Chelsio Network Adapter #4**
RDMA Accepted Connections	43.000	10.000	20.000
RDMA Active Connections	6.000	4.000	6.000
RDMA Completion Queue Errors	0.000	0.000	0.000
RDMA Connection Errors	0.000	0.000	0.000
RDMA Failed Connection Attempts	10,583.000	9,168.000	7,938.000
RDMA Inbound Bytes/sec	68,454.528	0.000	0.000
RDMA Inbound Frames/sec	637.465	0.000	0.000
RDMA Initiated Connections	14.000	10.000	14.000
RDMA Outbound Bytes/sec	278,901.842	0.000	0.000
RDMA Outbound Frames/sec	682.427	0.000	0.000
SMB Direct Connection	**_Total**		
Bytes RDMA Read/sec	0.000		
Bytes RDMA Written/sec	2,606,963.267		
Bytes Received/sec	68,350.615		
Bytes Sent/sec	94,584.589		
Memory Regions	0.000		
RCQ Notification Events/sec	617.482		
RDMA Registrations/sec	592.503		
Receives/sec	636.466		
Remote Invalidations/sec	591.503		
SCQ Notification Events/sec	612.486		
Sends/sec	636.466		
Stalls (RDMA Read)/sec	0.000		
Stalls (RDMA Registrations)/sec	0.000		
Stalls (Send Credit)/sec	0.000		
Stalls (Send Queue)/sec	0.000		

Figure 3 – Performance monitor showing RDMA traffic

Related Links

Deploy Storage Spaces Direct
Storage Spaces Direct hardware requirements
Choosing drives for Storage Spaces Direct