Configuring Switchless Ring Backbone

Quick Start Guide for Windows & Linux

The T5/T6 ASIC based adapters support an integrated eSwitch which enables connecting a set of servers, without requiring any discrete external network switch (Switchless Ring Backbone). It has a full suite of L2-L7 features including ACL with support for L2 Ethernet switching, L3 routing, NAT, TCP Proxy, and iSCSI proxy. The Switchless Ring Backbone configuration is supported in Windows Server 2016 (including S2D Clusters) and Linux platforms. This document provides quick steps to configure it on Windows and Linux platforms.
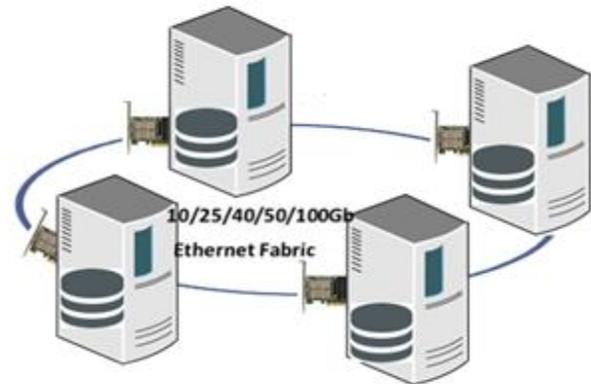
Figure 1 – 4-node setup using Switchless Ring Topology

Windows S2D configuration in Ring Backbone

Follow the steps mentioned below to install and configure a 4-node Storage Spaces Direct (S2D) cluster using Chelsio ring backbone topology:

Prerequisites:

Please ensure that the following requirements are met on each machine in the topology before proceeding with the installation:

- There must be an active management network (apart from Chelsio test network) which enables the machines to be accessed to carry out the required configuration.

- PowerShell Remoting should be enabled using the following command:

```
PS C:Users\Administrator> Enable-PSRemoting -Force
```

- Windows Remote Management (WinRM) should be enabled using the following command:

```
PS C:Users\Administrator> Set-item wsman:localhost\client\trustedhosts - value *
```

- Firewall should be disabled from Control Panel.
- Hyper-V and Failover cluster roles should be installed.
- All the nodes should have the same user credentials.

Ring Installation and Configuration:

i.   Connect the machines in ring backbone. Connect port 1 of one machine to the port 0 of next machine such that a closed ring is formed.
ii. Download the latest Chelsio Unified Wire for Windows from Chelsio Download Center.
iii. Install Unified Wire on all the machines in the ring back backbone.

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

iv. Open PowerShell and run the backbone_cfg.ps1 script on any one of the machines to configure ring backbone.

```
PS C:\Users\Administrator> backbone_cfg.ps1 -Servers <host1,host2,host3,host4> -
Username <username> -Password <password> -BaseIpAddress <test_ip)
```

> **Example:**
>
> 4 nodes N1, N2, N3 and N4 are connected in a ring backbone and configured with network address 102.1.1.0 with netmask 24 (selected by default).
>
> ```
> PS C:\Users\Administrator> backbone_cfg.ps1 -Servers N1,N2,N3,N4 -Username
> administrator -Password windows@123 -BaseIpAddress 102.1.1.0
> ```
>
> N1 will be assigned with IP 102.1.1.1, N2 with 102.1.1.2, N3 with 102.1.1.3 and N4 with 102.1.1.4 respectively.

Note:

- Port 1 should not be configured for any other purpose in the operating system.
- Only single port (port 0) is available to the user after configuring ring backbone.

v. If any of the nodes in the Ring reboots, the ring will be reconfigured automatically. However, if a vSwitch is created or any of the driver parameters is changed (driver will be reloaded) on a node, the ring must be reconfigured manually using below command:

```
PS C:\Users\Administrator> backbone_cfg.ps1 -Servers <host1,host2> -Username
<username> -Password <password> -Reconfigure
```

Note: Only the affected nodes should be mentioned.

S2D Configuration:

i. Before creating the cluster, validate the nodes using the cluster validation tool.

```
PS C:\Users\Administrator>Test-Cluster –Node <host1,host2,host3,host4> –Include
"Storage Spaces Direct",Inventory,Network,"System Configuration"
```

ii. Create a cluster using the nodes validated in the previous step:

```
PS C:\Users\Administrator>New-Cluster –Name <ClusterName> –Node <
host1,host2,host3,host4> –NoStorage -StaticAddress <static_ip> -Verbose
```

Note: Without the –*NoStorage* parameter, the disks may be automatically added to the cluster and you will need to remove them before enabling S2D. Otherwise they will not be included in the Storage Spaces Direct storage pool.

iii. The disks intended to be used for Storage Spaces Direct need to be empty and without partitions or existing data. If a disk has partitions or existing data, it will not be included in the Storage Spaces Direct system.

Note: Storage Spaces Direct does not support disks connected via multiple paths, and the Microsoft Multipath MPIO software stack.

iv. Enable Storage Spaces Direct and create a storage pool.

```
PS C:\Users\Administrator>Enable-ClusterS2D -PoolFriendlyName <pool_name> -
Verbose
```

v. Create virtual disks on the storage pool created in the previous step.

```
PS C:\Users\Administrator>Get-ClusterNode |% { New-Volume -
StoragePoolFriendlyName s2d -FriendlyName $_ -FileSystem CSVFS_ReFS -Size 500GB -
Verbose  }
PS C:\Users\Administrator>New-Volume -StoragePoolFriendlyName s2d -FriendlyName
Collect -FileSystem CSVFS_ReFS -Size 100GB -Verbose
```

vi. Create or deploy virtual machines. The virtual machine's files should be stored on the virtual disks. You can use utilities like VMFleet to deploy and use multiple virtual machines.

Ring Backbone Configuration in Linux

Connect the machines in ring backbone. Connect port 1 of one machine to the port 0 of next machine such that a closed ring is formed. Follow the steps mentioned below on all the nodes to install and configure ring backbone topology:

i. Download the latest Chelsio Unified Wire for Linux from Chelsio Download Center.
ii. Install the Unified Wire.

```
[root@host~]# make install CONF=RING
```

iii. Load network driver (*cxgb4*) with *enable_ringbb* module parameter.

```
[root@host~]# modprobe cxgb4 enable_ringbb=1
```

iv. Bring up port 0 and port 1 interfaces, but assign IP address to only port 0.
v. Disable IPv6 on port 1:

```
[root@host~]# sysctl -w net.ipv6.conf.ethX.disable_ipv6=1
```

Note: Port 1 should not be used for any other purpose in the operating system.

vi. Check dmesg output and note the mirror VF ID:

```
[root@host~]# dmesg|grep -i mirror
```

vii. Check the MPS TCAM entry for port 0 MAC and note the corresponding VF ID.

```
[root@host~]# cat /sys/kernel/debug/cxgb4/<pci_bus_id>/mps_tcam | head -10
```

viii. Apply filters using *cxgbtool*:

```
[root@host~]# cxgbtool <port0_interface> configure_ringbb pf 4 vi <vf_id> mvi
<mirror_vf_id>
```

> **Example:**
> Loading network driver with *enable_ringbb* module parameter and bringing up the interfaces:
>
> ```
> [root@host~]# dmesg | grep -i mirror
> [456586.211088] cxgb4 0000:01:00.4: Port 0 Traffic Mirror PF = 4; VF = 70
>
> [root@host~]# cat /sys/kernel/debug/cxgb4/0000\:01\:00.4/mps_tcam | head -10
> Idx  Ethernet address     Mask       Vld Ports PF  VF
>   0 01:80:c2:00:00:0e ffffffffffff  Y   0x3   7 100
>   1 00:00:00:00:00:00 ffffffffffff  Y   0x3   4  68
>   2 00:07:43:28:cc:68 ffffffffffff  Y   0x2   4  69
>   3 01:00:5e:00:00:01 ffffffffffff  Y   0x3   4  69
>   4 00:07:43:28:cc:60 ffffffffffff  Y   0x1   4  68
>
> [root@host~]# cxgbtool enp1s0f4 configure_ringbb pf 4 vi 68 mvi 70
> Ensure that you have:
> - Copied ring backbone config file in /lib/firmware/cxgb4/
> - Reloaded driver with enable_ringbb=1
> - Brought all chelsio interfaces up
> - Assigned IP address to port0
>
> Press any key to enable ring backbone configuration(Ctrl-C to exit)
> ```

ix. Press any key when prompted to complete the configuration.
x. You should be able to run traffic between the machines/nodes now.

Note: For more ring back bone configuration options, please run *man cxgbtool* and refer the *configure_ringbb* section.

xi. If any of the nodes in the Ring reboots, follow the steps (iii) to (ix) on the affected host to reconfigure the ring.

Related Links
Chelsio Switchless Backbone
High Performance NVMe Over 100G iWARP RDMA
S2D Performance with Network QoS
S2D Performance with iWARP RDMA
iWARP RDMA – Best Fit for Storage Spaces Direct