S2D performance with Chelsio 25GbE

Chelsio T6 iWARP RDMA solution for S2D on Windows Server 2019

Overview

Microsoft **Storage Spaces Direct** (S2D) is a feature which enables building highly available and scalable storage systems by pooling local server storage. You can now build HA Storage Systems using storage nodes with only local storage, which can be disk devices that are internal to each storage node. This not only eliminates the need for a shared SAS fabric and its complexities, but also enables using devices such as SATA solid state drives, which can help further reduce cost or NVMe solid state devices to improve performance. Storage Spaces Direct leverages SMB3 for all intra-node communication, including SMB Direct and SMB Multichannel, for low latency and high throughput storage. This paper presents S2D performance results in Windows Server 2019 using Chelsio iWARP RDMA technology in a hyper-**converged** deployment scenario.
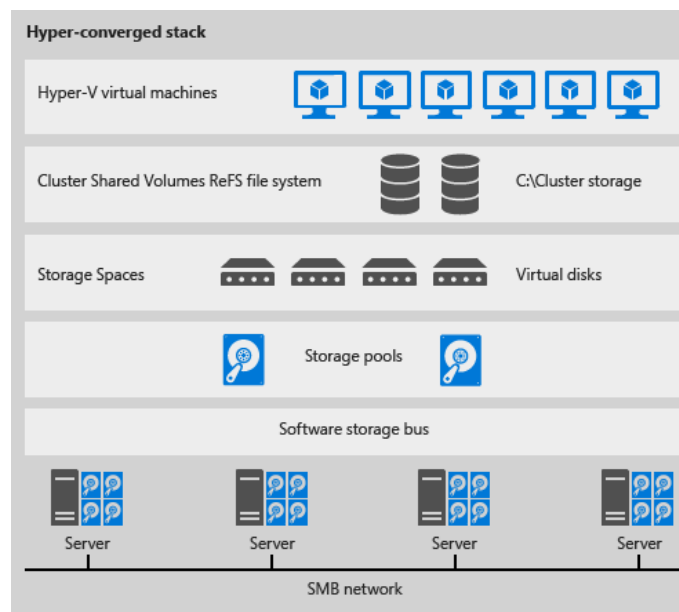

Figure 1 – S2D Hyper-Converged Stack

Why Chelsio iWARP RDMA Solution for S2D

Chelsio's sixth generation (T6), high-performance iWARP RDMA 1/10/25/40/50/100GbE adapters enable incremental, non-disruptive server installs, and support the ability to work with any legacy (non-DCB) switch infrastructure. This enables a decoupled server and switch upgrade cycle and delivers a brownfield strategy to enable high-performance, low cost, scalable S2D deployment.

iWARP has been an IETF standard (RFC 5040) for 14 years, TCP/IP has been an IETF standard (RFC 793, 791) for 40 years. iWARP Inherits the loss resilience and congestion management from

underlying TCP/IP stack and enables a very high performance, extremely low latency, high bandwidth and high message rate solution. iWARP presents no surprises, no fine print, and is a plug and play solution. It is scalable to wherever the datacenter can scale to.

Network QoS is used in **hyper-converged** configurations to ensure that the Software-Defined-Storage system has enough bandwidth to communicate between the nodes to ensure resiliency and performance. Chelsio's iWARP RDMA enabled adapters with enhanced rate-limiting (network QoS) features offload bandwidth allocation to the adapter bypassing the operating system. This eliminates the need for a DCB enabled Ethernet switch to implement Storage Spaces Direct (S2D) in a hyper-converged mode, resulting in reduced total ROI and simplified management.

Microsoft also recommends and prefers to use iWARP RDMA as it is easier to configure/setup, scalable, routable and works with all standard Ethernet switches.
- Microsoft recommends iWARP for S2D
- Microsoft Recommendation on the RDMA alternatives in Windows
- Hyper-converged solution using Storage Spaces Direct in Windows Server 2016

Storage Spaces Direct Performance with iWARP

This paper presents performance results in a 3 node S2D cluster connected using the Chelsio 25GbE T6225-CR/T6225-SO-CR adapters.

Results

Below is a screenshot from the VMFleet Watch-Cluster window, which reports IOPS, bandwidth and latency.

CSV FS	IOPS	Reads	Writes	BW (MB/s)	Read	Write	Read Lat (ms)	Write Lat
Total	68,982	68,932	50	36,140	36,140			
azure1	27,722	27,704	19	14,525	14,525		83.244	280.843
azure2	28,346	28,326	20	14,851	14,851		81.491	212.108
azure3	12,914	12,902	11	6,765	6,765		90.208	99.248

Figure 2 – Storage Spaces Direct Throughput Numbers

As you can see, the aggregated bandwidth exceeded **36+ GB/s** for 512KiB random read, which is very impressive for a 3-node cluster.

CSV FS	IOPS	Reads	Writes	BW (MB/s)	Read	Write	Read Lat (ms)	Write Lat
Total	917,089	917,074	15	3,756	3,756			
azure1	306,160	306,157	3	1,254	1,254		0.965	2.453
azure2	297,958	297,952	5	1,220	1,220		0.893	2.713
azure3	312,972	312,965	7	1,282	1,282		0.178	1.497

Figure 3 – Storage Spaces Direct IOPS Numbers

As seen from the above screenshot, this setup was able to demonstrate over **917K IOPs** for 4KiB random read in aggregate using 60 virtual machines, which means **15,200+ IOPs** per virtual machine!

The following is the hardware configuration:

- **3 nodes of Dell R720**
  - 2x E5-2690 v2 @ 3.00GHz 10c/20t
  - 128GiB DDR3 1600Mhz
  - Windows Server 2019 rs5 + Storage Spaces Direct
  - Performance Power Plan
  - 1x Chelsio T6225-CR/T6225-SO-CR (Dual port 25Gb PCIe 3.0 x8)
    - Driver: v6.12.7.0
    - Firmware: 1.22.8.0
    - Dual port connected

- **Total Storage**:
  - 3x 2.0TB INTEL DC P4600 (PCIe 3.0 x4)
  - 5x 2.4TB Micron 9100 (PCIe 3.0 x4)
  - 5x 1.6TB Intel SSD DC P3608 (PCIe 3.0 x8)

A 20 virtual machines per node, for a total of 60 virtual machines configuration was used for this setup. Used A1 default virtual machine size (i.e., 1 vCPU and 1.75GiB RAM). VMFleet tool was used to run DISKSPD in each of the virtual machines with various multi-threaded workloads for IO-sizes 4KiB and 512KiB, 100% random read with queue depth of 8 and 32 per VM.

In addition to S2D, iWARP-enabled Chelsio adapters power other aspects of Microsoft Windows installations such as **Storage Replica** for disaster recovery, **SMB Direct** for high performance file access, **Client RDMA** for bringing RDMA benefits to Windows 10 deployments, **Network Direct** for Windows HPC deployments, hardware offloaded **iSCSI, iSER, NVMe-oF and FCoE initiators** for SAN applications, **d.VMMQ, NVGRE and VxLAN** encapsulation offload, **Guest RDMA, SR-IOV** for Virtual environments.

Summary

Chelsio iWARP RDMA enabled 25GbE Unified Wire and Converged Network adapters (CNAs) deliver a high-performance Storage Spaces Direct (S2D) solution using standard Ethernet infrastructure and enable datacenters to deploy S2D now by leveraging all-inboxed drivers. The ability to work with any non-DCBX switch, enables an immediate plug and play deployment. Moreover, Chelsio CNAs enable a low-cost solution for S2D deployments. Support of iWARP protocol is enabled since Windows Server 2012-R2 release, and in boxed support in Windows Server 2016, has allowed for years of testing for a very robust, tested, and efficient deployment with Chelsio iWARP enabled Ethernet adapters.

Related Links

Storage Spaces Direct throughput with 100GbE iWARP - Microsoft Blog
Storage IOPS Update with S2D - Microsoft Blog