Western Digital.

Solution Test Report February 2023

Western Digital® OpenFlex™ Data24 and Ondat Storage Platform for Kubernetes Solution Test Report

Abstract

This document provides a record of the testing performed to validate the compatibility of Ondat SDS solutions with the Western Digital OpenFlex Data24.

Table of Contents

Introduction

This Solution Test Report (STR) summarizes the results obtained during the validation of Ondat Storage Platform for Kubernetes with a Western Digital OpenFlex Data24. The objective of this exercise was to verify Ondat Storage Platform properly functions with the OpenFlex Data24 NVMe™ over Fabrics (NVMe-oF™) devices while meeting the reliability, flexibility, performance, and usage patterns of a Kubernetes cluster running common containerized applications. This document includes guidelines for configuring Ondat Storage Platform with an OpenFlex Data24.

Ondat Storage Platform utilizes any and/or all block devices attached to a Kubernetes cluster to create a data mesh while using Kubernetes native contracts to empower stateful applications. Ondat simplifies the integration, allocation, and management of block storage by obfuscating topology and physical architecture while providing advanced storage features such as volume level encryption and replication to ensure a company's data compliance and resilience requirements.

All tests were completed utilizing the Ondat Storage Platform over 100Gb Ethernet interfaces to the OpenFlex Data24. These tests were performed by Ondat from London, UK via remote access to the Western Digital lab in Colorado Springs, CO.

This STR is not an endorsement of Ondat by Western Digital, and no warranty of the product is either expressed or implied.

Objectives

There are several objectives for this test:

1. Verify that the Ondat software can access OpenFlex Data24 NVMe-oF devices.
2. Verify that the Ondat software can integrate, utilize, and deliver storage utilizing OpenFlex Data24 NVMe-oF devices.
3. Verify that the Ondat software can achieve desired or improved performance from OpenFlex Data24 NVMe-oF devices.

Prerequisite Table

Item	Notes
Servers with 100 GbE NIC supporting RoCE v2	Western Digital OpenFlex products require 50GbE – 100GbE network connection supporting RoCE V2
100 GbE Network Switch	Western Digital Storage Servers or Composable Storage require network connectivity ranging from 10 to 100 GbE
5 or more Bare-Metal worker nodes	Ondat recommends 5 or more worker nodes
File System Support: EXT4 & XFS	Ondat creates configuration and blob files related to each persistent volume presented to the cluster.
Increase PID limits to 32768	Ondat storage pods run as part of a PID in cgroup along with other containers which requires an increase in the maximum limit so new pod/container spawning does not fail
Firewall Rules: 2379-2380 : TCP : Two-way 5703-5705 : TCP : Ingress 5710-5711 : TCP + UDP : Ingress 8443 : TCP : Egress 25705-25960 : TCP : Ingress	Ondat requires these firewall rules to be created so communication is successful, specifically etcd, DirectFS, Ondat Data Plane, Ondat REST API, Ondat gRPC API, Gossip Protocol, Ondat Protocol, Shared Filesystems
NTP Clock synchronization	Ondat requires that all nodes within a cluster have their hardware clock synchronized to a reliable Network Time Protocol(NTP) server

Technology Summary

Technology	Version
OpenFlex Data24	Firmware V5.0
Dell® R650 Server	Firmware V5.1
Linux® OS	RedHat® Enterprise Linux® 8.6
Mellanox® ConnectX®-6 DX	Dual Port 100 GbE NIC, PCIe® Gen4.0 x16, Mellanox OFED driver V5.6-1.0.3.3
Mellanox SN2700 Switches	32 port 100 GbE 1U switch
Ondat Storage Platform for Kubernetes	V2.9.1

Solution Overview

Kubernetes is rapidly becoming the de-facto standard for efficient and resilient application deployment. In common with all infrastructure frameworks, Kubernetes needs persistent storage and in particular storage that is managed from the cluster and applications directly. Container-native storage has emerged as a new category of software-defined storage where the traditional features of data storage (resiliency, data protection, scalability) are built into containers running within a Kubernetes cluster itself. This model enables deployments to use locally attached storage resources, including physical storage in on-premises solutions or native storage (including NVMe & NVMe-oF drives) in the public cloud.

Ondat delivers a data mesh for block storage using Kubernetes native constructs to power stateful applications. Ondat couples any storage to any Kubernetes cluster and, with the simple application of Kubernetes labels, also delivers advanced features such as: Encryption at a per Kubernetes volume level, allowing for safe multi-tenant operations. Topology aware placement of volumes to align with your availability zones and physical architecture to ensure your data compliance. Replication of data at a Kubernetes Volume making sure that the data you need is protected to deliver the business resilience required.

Using Ondat, any storage on any node in your Kubernetes cluster can be delivered to the applications that need it anywhere in the cluster. Intelligent placement makes sure that your workload is always optimized, and by deploying the Ondat data mesh your Kubernetes platforms are responsive to your business applications with compute and storage able to grow independently as your workloads change.

The Ondat Storage Platform delivers the following solution(s):

☐ Backup and Archive ☒ Software Defined Storage ☐ NAS/SAN Services ☒ Storage Platform Manager

☐ Media Asset Management ☒ High Performance Computing ☐Other:

The Ondat Storage Platform is applicable to the following vertical(s):

☒ Media & Entertainment ☒ Life Sciences/HPC ☒ Financials ☒ Enterprise ☒ Government

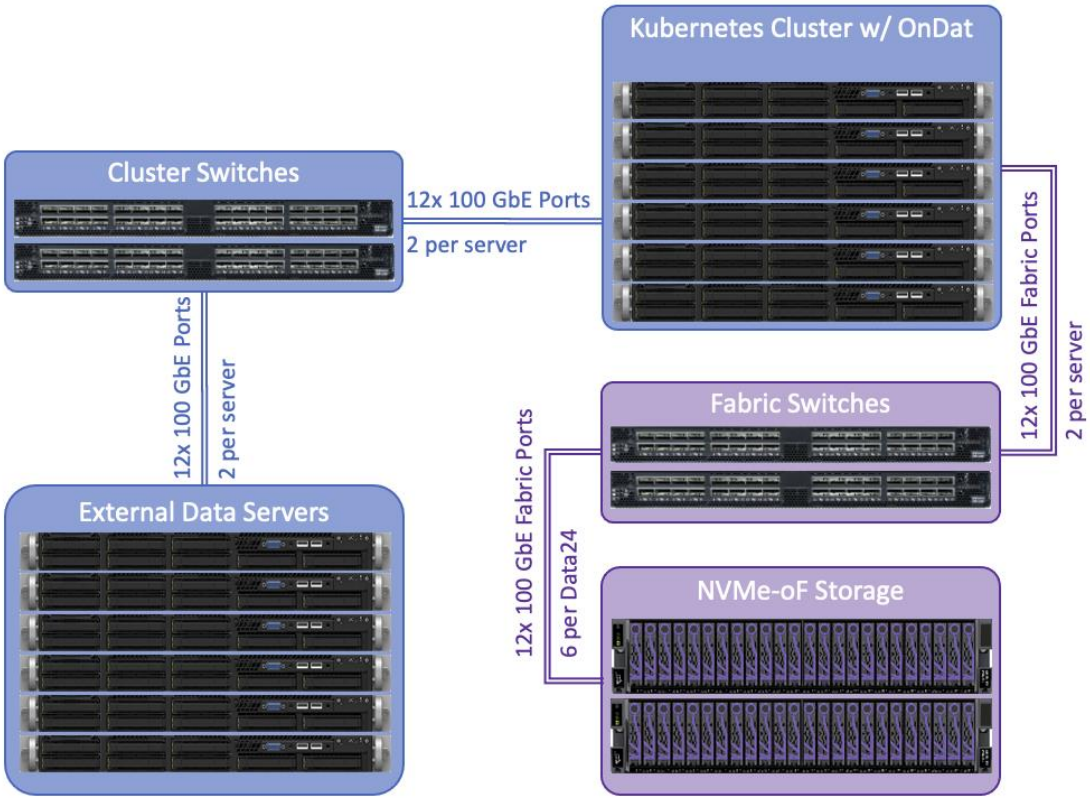
☐ Other:

Architecture

Our co-designed solution architecture was comprised an initial building block of a 6 node Kubernetes cluster with Ondat connected to one OpenFlex Data24, ideally through a pair Mellanox SN2700 fabric switch to ensure fully multi-path support during failure cases. Both the Switch and servers must be configured following the Western Digital Lossless Best Practice Guide as well as the Ondat best practices guide. An addition server was configured to host Rancher for Kubernetes management. Every step was taken to more closely resemble a “real world” deployment of Kubernetes, containers, and workloads generated to utilize the container-storage allocated through Ondat configuration files.

The OpenFlex Data24 can be configured in multiple ways to achieve the desired performance and availability of the NVMe-oF devices and data. This coupled with the Ondat Storage Platform offers both flexibility and scalability to be fully realized to utilize all available hardware and software resources more fully. Configuring the OpenFlex Data24 to use single path access combined with Ondat Hyper-converged Cluster Topology which make persistent volumes accessible to all worker nodes directly maximizing IOPS and Throughput of the OpenFlex Data24. While configuring the OpenFlex Data24 for fully redundant access combined with the Centralized Cluster Topology by limiting volume access to a subset of workers to handle IO intensive workloads while another pool of workers can focus on compute intensive workloads.

The illustration below shows an ideal solution configuration, where all networks are physically segmented for maximum availability and performance.



Network Topology

This solution requires two distinct network segments: cluster intercommunication and storage network. It is heavily recommended to achieve the best possible performance that these are physically segmented but that is not technically require. The cluster network is indented for container intercommunication and IO to and from external data sources. The storage network is intended to only provide performant access to the NVMe-oF storage devices hosted by the Data24.

Testing Environment

Testing environment was designed to emulate real world setups that would include external data sources in addition to the Kubernetes cluster. The OpenFlex Data24 was solely used for Kubernetes container storage which would be configured, managed, and accessed via the Ondat Storage Platform.

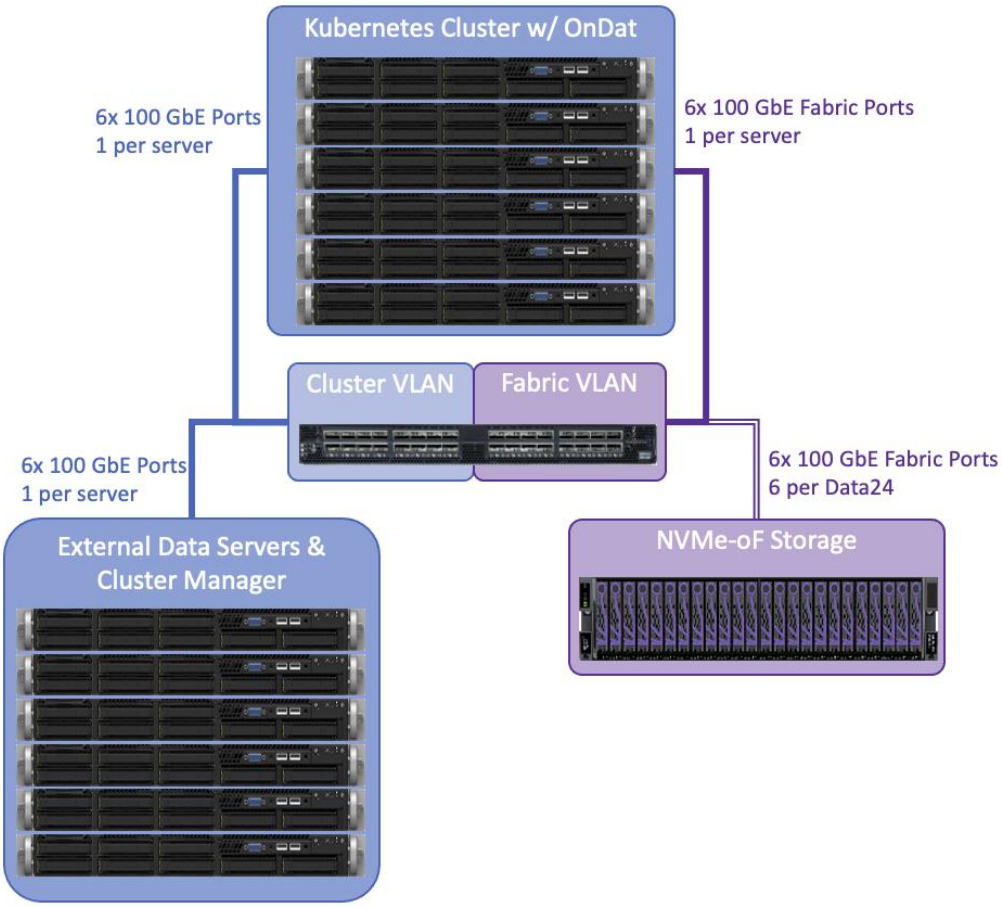
Server Specifications:

- 5x Data Servers & 1x Cluster Manager
  - 2x AMD EPYC™ 7542
  - 512 GiB DDR4 RAM
  - 1x M.2 boot drive
  - OS: RedHat Enterprise Linux 9.0
  - 1x Mellanox ConnectX5 dual port 100 GbE ports
    - Only 1 port cabled to Cluster VLAN
  - 8x Western Digital DC SN640 3.2TB[1] NVMe devices
    - NVMe drives only installed into Data Servers
- 6x Kubernetes Servers
  - 2x Intel® Xeon® Gold 6354
  - 512 GiB DDR4 RAM
  - 2x M.2 boot drives RAID 1
  - OS: RedHat Enterprise Linux 9.0
  - 2x Mellanox ConnectX6 dual port 100 GbE ports
    - 1 port of 1st card cabled to Cluster VLAN
    - 1 port of 2nd card cabled to Storage Fabric VLAN

[1] One terabyte (TB) is equal to one trillion bytes. Actual user capacity may be less due to operating environment.

The External Data resources were provided to supply predetermined datasets for testing as well as generate IOs from an external application. However, to meet the testing timeline we determined this level of testing was not required to prove the solutions capabilities. Therefor we used only one of the servers as the Kubernetes cluster manager server, specifically running Rancher orchestration.

The illustration below depicts the exact environment created and handed over to Ondat to begin testing.



OpenFlex Data 24 and Ondat configuration

1. Configure servers to be on same network as OpenFlex Data24.
2. Deploy and run connect scripts on each server to attach to the desired NVMe-oF devices hosted in the OpenFlex Data24.
3. Verify all volumes are available to the to host OS and that no NVMe-oF device is duplicated on another server.
4. Verify that all lossless settings are properly configured on each server and the storage fabric switch.
5. Verify that all hosts can perform IO to each connected NVMe-oF device hosted in the OpenFlex Data24 within expected performance ranges.
6. Create file systems on each device with mkfs and mount into /var/lib/storageos/data/dev#, sample commands below:

```
mkfs -t ext4 /dev/nvme0n1
mkdir -p /var/lib/storageos/nvme0
mount /dev/nvme0n1 /var/lib/storageos/data/nvme0
```

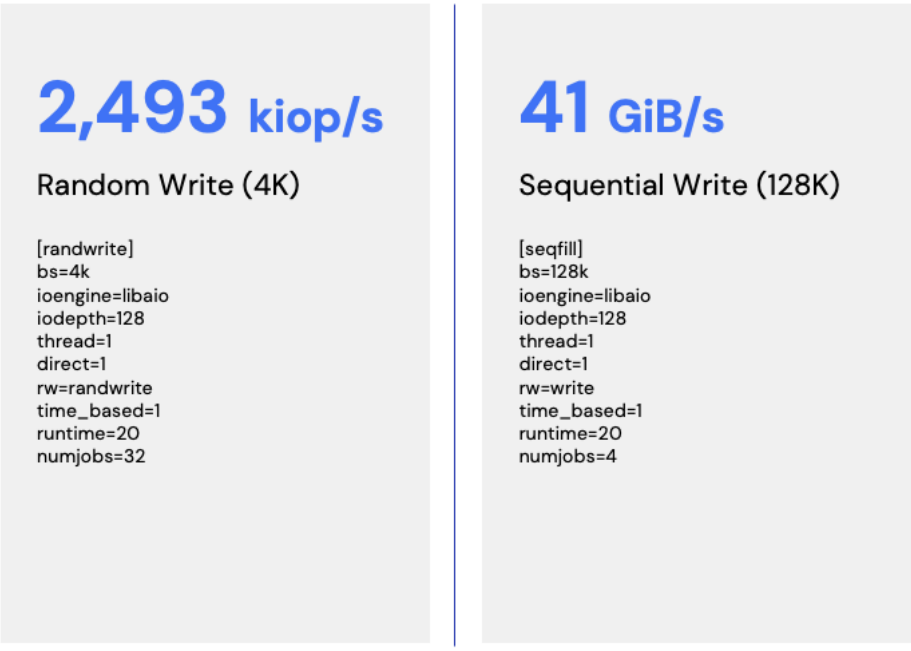
7. Install the Kubernetes distribution of your choice.

8. Install Ondat onto your Kubernetes cluster fallowing the relevant guide.
   https://docs.Ondat.io/docs/install/
9. Create storage configurations for Kubernetes using either the options below
   a. Optionally create storage classes for Ondat Container Storage Interface (CSI) provider.
      https://docs.Ondat.io/docs/operations/storageclasses/
   b. Create a PersistentVolumeClaim with Ondat.
      https://docs.Ondat.io/docs/operations/firstpvc/
10. Verify all containers can identify, access, and perform IO to each PVC defined.

```
storageos get node -ojson | jq -r '.[] | { name: .name, capacity: .capacity.total  }'
```

Results Summary

Testing was performed using Flexible IO (FIO) testing tools for both small block random IO (4KiB) and large block sequential IO (128KiB) while using Ondat Hyper-converged Topology. Small block IO was measured using input/output operation per second (IOPS) to understand performance for Database use cases resulting in an aggregate 2.493 million IOPS. Large Block IO was measured in Gibibytes per second (GiB/s) to understand the performance of image processing use cases resulting in an aggregate of 41 GiB/s. Both results are significantly improved compared to previous performance result published by Ondat.



Testing was conducted by Ondat.

Consideration

Scaling options	Hardware changes	Notes
High-Capacity HA	1x Data24 with 2 Fabric NICs For each server	Each server can achieve same storage performance as tested, but with 24 drives available per server capacity is drastically increased
High Performance HA	1x Data24 with 6 Fabric NICs For every 3 servers	This allows for the realization of nearly the full drive performance as tested above, but with the addition of multiple paths for added resiliency
Minimal Deployment	1x Data24 with 2 Fabric NICs for every 4 servers with 25 Gb NICs or ever 2 servers with 50 Gb NICs	Reduces network port cost, requires break-out cables, lowers CPU requirements, can be replicated to reach capacity & compute targets

Appendix

Contributors

Name	Company	Title
David Tobin	Western Digital	Sr. Technical Marketing Engineer
Chris Milsted	Ondat	Solution Architect

References

Document Title	Date Delivered
OpenFlex Data24 User Manual	9/8/2022
Ondat Storage Platform Solution Brief	9/9/2022
Ondat Storage Platform Best Practices	9/9/2022
Ondat Storage Platform Installation Guide	9/9/2022

Document Feedback

For feedback, questions, and suggestions for improvements to this document send an email to the Data Center Systems (DCS) Technical Marketing Engineering (TME) team distribution list at pdl-dcs-tm@wdc.com.

Version History

Version	Date	Notes
1.0	23 Feb 2023	Initial release

Notices

countries. Mellanox and ConnectX are registered trademarks of NVIDIA Corporation. The NVMe and NVMe-oF word marks are trademarks of NVM Express, Inc. Red Hat and Red Hat Enterprise Linux are registered trademarks of Red Hat, Inc. in the U.S. and other countries. All other marks are the property of their respective owners.

Products may not be available in all regions. Images may differ from actual products.

Western Digital
5601 Great Oaks Parkway
San Jose, CA 95119