StorPool System Requirements

2022-11-24

1. Introduction

During the deployment process, CPU cores and RAM are dedicated to the StorPool Storage components installed on physical servers. This resource isolation significantly improves performance maximum and consistency and avoids the adverse effects of CPU saturation and memory congestion. Please refer to the StorPool Storage User Guide for more information on the recommended CPU setup.

Since StorPool Storage ensures resource isolation, customers can buy mix-use servers with enough RAM, CPU, and storage capacity to act as hyper-converged nodes running compute and storage concurrently. This is possible for StorPool Block Protocol client hosts (KVM hypervisors and bare-metal Linux servers) managed with OpenStack, Kubernetes, CloudStack, OpenNebula, OnApp, or custom-built cloud management solutions.

In compute-only hosts and hyper-converged nodes that are StorPool Block Protocol clients, one CPU core is dedicated to the StorPool client driver.

Customers using StorPool Storage over iSCSI or NVMe/TCP can access the targets exposed by their storage system with the standard initiators of their computing platform(s) (e.g. VMware, Oracle VM, oVirt, XenServer, Proxmox, Microsoft Hyper-V, Linux, etc.). In this case, no StorPool Storage components are installed in the compute hosts.

StorPool Storage enables customers with complex environments to build heterogeneous clouds where some servers are hyper-converged, some are storage-only, and some are compute-only. StorPool Storage pools the capacity and performance of the storage drives in hyper-converged and storage-only nodes to serve both compute-only and hyper-converged hosts, using the storage protocol suitable for each use case.

2. Hardware components

This chapter lists hardware expected to work well with StorPool Storage. In most cases, any controller model which uses the same driver (in parentheses) will work well. Components marked in blue are recommended based on experience. StorPool aims for broad hardware support. If your specific model of controller or drive is not on this list, please ask.

CPU:
- Nehalem generation (Xeon 5500) or newer Intel Xeon processor
- AMD Epyc 7000 series
- ARM
- IBM POWER8 or newer

Note: older supported CPUs and server platforms often have severe architectural I/O bottlenecks. We recommend Intel Xeon Scalable, AMD EPYC Rome, or newer CPUs.

RAM:
The typical building block in 2022 comes with 10 x 7.68TB SSDs and 128GB ECC memory (8x16GB memory sticks), of which StorPool Storage uses 60GB. The exact memory requirements depend on the number of drives and their size. The rule of thumb is to plan for 1GB RAM per 1TB of raw storage space. Contact StorPool Technical Account Management for a detailed memory requirement assessment. Non-ECC memory is **not supported**.

HBAs and RAID controllers:
- Intel C200, C600, ICH10(82801JI) (ahci)
- LSI 2308 and 2116-based HBAs and integrated controllers (mpt2sas)
- LSI 3008, 3216, 3224-based HBAs and integrated controllers (mpt3sas)
- LSI MegaRAID controllers with 2108 and 2208 chipsets (megaraid_sas)
- LSI MegaRAID controllers with 3108 chipset (megaraid_sas)
- Dell PERC H730, H730 Mini, H730P Mini (megaraid_sas)
- Intel C600 SCU (isci)

- Primary storage nodes containing HDDs require a power-loss protected write cache as described below in the HDDs section.
- Storage controllers connecting SSDs are using JBOD mode.
- The storage controllers should be supported by the host OS.
- NVMe RAID controllers are **not supported**.

NVMe SSDs:
- Samsung PM1725A, PM963, PM983, PM9A3
- Intel P3520, P3600, P3608, P3700, P4500, P4600, P4510, P4610, P5510, P5520
- Micron 9100 PRO/MAX, 9200 ECO/PRO/MAX, 9300 PRO/MAX, 7300 PRO/MAX
- HPE VO003840KWVMU (PM1723b), VO003840KWVBU (CD5)
- Toshiba PX04PMC, XD5
- Kioxia CD6-R
- HGST/WD Ultrastar DC SN100, SN200

- WD Gold SN600
- Huawei ES300 V3

*Note: See the notes in **Datacenter SATA/SAS SSDs** below.*
*Note 2: Some of the supported NVMe drives in **M.2 form factor** may require additional cooling depending on the specific system design*

Datacenter SATA/SAS SSDs:
- Samsung SV843, SM863, PM863, 883DCT, PM863a, PM883
- Intel DC S3500, S3510, S3520, S3610, S3700, S3710, S4500, S4600, S4510, S4610, S4520, S4620
- Micron M500DC, M510DC, X5100 ECO/PRO/MAX, 5200 PRO/MAX, 5300 PRO/MAX
- Seagate Nytro 1551
- HPE VK000960GWCNR (S3520)
- Toshiba HK4 (Hawk-4) Series
- Toshiba (Kioxia) HK-6R Series
- SanDisk (ex. Smart Storage Systems) CloudSpeed 1000E, Eco, Ascend
- HGST Ultrastar DC SS200
- minimum supported size of pool SSDs is 960GB
- maximum supported size for SATA/SAS SSD is 8TB. For larger SSDs, please check the supported NVMe models
- a minimum number of 8 drives per cluster of the same size is recommended

Note: SSDs on this list have passed validation tests. We have not noticed any performance degradation or general issues for these models. In addition to completing the validation tests, the drives marked in blue have been in production use for an extended period under different workloads. They have not shown performance degradation or other issues.
Note 2: Ensure that the drive is with the latest firmware recommended by the vendor. There are known good versions in the table below.

HDDs:
- Enterprise-grade SAS or SATA hard drives.
- Shingled magnetic recording (SMR) drives are **not supported**.
- We've had a good experience with HGST Ultrastar, WD Re, and Seagate HDDs.
- We've experienced unexpected behavior in some particular series of enterprise-grade HDDs, which led to the whole cluster's reduced performance contact StorPool support for further information.
- Storage nodes containing HDDs must have a power-loss protected write cache using one of the following hardware options:
    - Intel Optane P4801X NVMe drive
    - Partition on large "pool" NVMe drive
    - Intel Optane Persistent Memory 200
    - RAID controller with BBU or CacheVault (not recommended)
- 512n, 512e, and 4kn format drives are supported.
- Maximum size is restricted to 4TB to ensure satisfactory performance. Using bigger size HDDs, up to 12TB, is acceptable only in large backup systems with >100TB usable capacity and really “cold” data.
- For a proof of concept deployment that does not include performance evaluation, a

minimum of 8 HDDs per cluster is acceptable.
- We recommend using 24 or more HDDs per cluster to achieve optimal performance of HDD pools.
- Up to 24 HDDs per storage node are recommended

10/25/40/50/100 Gigabit Ethernet controllers:
- Mellanox ConnectX-3, ConnectX-3 Pro (mlx4_en)
- Mellanox ConnectX-4, ConnectX-5, ConnectX-6 (mlx5_core)
- Intel 82599, X520, X540, X550 (ixgbe)
- Intel X710, XL710, X722, XXV710, XL710 (i40e)
- Intel E810 (ice) - *requires specific restricted configuration in Linux Kernel 5.13 and newer*
- Qlogic QLE3440-CU, QLE3442-CU (bnx2x)
- Broadcom BCM57840S (bnx2x)
- Broadcom NetXtreme-E BCM57414, BCM57508 (bnxt_en)

Ethernet Switches:
- StorPool works with any standards-compliant 10/25/40/50/56/100 GbE switch with Jumbo frames and Flow control
- Mellanox SX1012, SX1016, SN2010, SN2100B, SN2100, SN2410, SN2700
- Dell S4810, S6000

3. Network/fabric architectures

- Flat Ethernet networks
- Routed leaf-spine Datacenter Ethernet networks

Gigabit Ethernet networks are **not supported**.
For deployments utilizing the StorPool native protocol requires 2x network ports per storage node.
Deployments utilizing iSCSI require 4x network ports per storage node.

4. Firmware versions

All network controllers, storage controllers, and SSDs must run the latest known-stable version of firmware available from the vendor. This rule is to prevent the occurrence of known firmware bugs. Note that the latest stable might not always be the latest available version.

Known good versions for some devices:

Broadcom/Avago/LSI 2008,2108,2208,2308 and 2116	19.00.00.00 **
Broadcom/Avago/LSI 3008	12.00.02.00
Broadcom/Avago/LSI 3108	4.650.00-6223
Micron M500DC	0144
Micron M510DC	0013
Micron X5100 Eco 1.92T	D0MU037
Micron 5100 Pro 960GB	D0MU042
Micron 5100 Pro 1920GB	D0MU051
Intel S3500	D2012370
Toshiba (Kioxia) KHK61RSE3T84	1DET6101
Mellanox MT27520 ConnectX-3 Pro	2.42.5000
Mellanox MT27710 ConnectX-4 Lx	14.22.1002
Intel Ethernet Controller X7XX	7.20 0x800079ef 1.2585.0
Intel Ethernet Controller E810-XXV	2.50 0x800077a8 1.2960.0

*** Note that the latest 2x.xx.xx.xx versions are not recommended due to cases with corrupted data with some drives.*

5. Example storage cluster configurations

The following hardware configurations cover the requirements outlined above and are common among StorPool customers.

5.1. Example hardware configuration for dedicated storage - All-NVMe pool

- 3 copies on NVMe SSDs
- Distribute drives evenly across 3+ servers
- 5 000 000 IOPS (random read) and <0.15 ms latency.
- Quoted usable space is before gains from snapshots/clones, thin provisioning, and zeroes detection.

- 50x 7.68 TB NVMe SSDs - 116.4 TB usable (5x 10-bay nodes)
- Other combinations and sizes are possible

Barebone	1 RU Single Socket EPYC server 10x U.2 bays, e.g., SuperMicro AS-1114CS-TNR, HPE DL325 Gen10 Plus v2, Dell R6515	5
CPU	AMD EPYC gen3 Rome 7313P 16C/32T 3.4G 128M	5
RAM	16 GB DDR4-3200 ECC RDIMM	40
Boot drive	M.2 NVMe SSD, e.g., Intel P4101, Toshiba XG6	5
NIC	2x 10/25GbE SFP+ e.g. Mellanox ConnectX-4,5,6, Broadcom BCM57414	5
Pool drive	7.68 TB NVMe U.2 SSD, e.g., Samsung PM9A3, Micron 7300 PRO	50

5.2. Example hardware configuration for dedicated storage - NVMe SSD-Hybrid pool

- 1 copy on NVMe SSDs and 2 spare copies on HDDs
- Total space on HDDs is two times the space on SSDs
- Distribute drives evenly across 3+ servers
- 600 000 IOPS (random read) and <0.15 ms latency. The performance of an all-flash array at a fraction of the cost.
- Quoted usable space is before gains from snapshots/clones, thin provisioning, and zeroes detection.
- 10x 4TB NVMe SSDs, 40x 2TB HDDs - 36.4 TB usable (5x 12-bay nodes)
- Other combinations and sizes are possible

Chassis	2RU with 12x 3.5" bays, 4x U.2 e.g. CSE-826BAC4-R920LPB	5
Motherboard	X11SPH-nCTPF (LSI3008, 2x10G SFP+, 2x OCuLink, 1x m.2)	5
CPU	Intel® Xeon® Silver 4210R - 10 cores @ 2.5 GHz	5
RAM	16 GB DDR4-2400 ECC RDIMM	20
Boot drive	M.2 NVMe SSD, e.g., Intel P4101, Toshiba XG6	5
NIC	2x 10/25GbE SFP+ e.g. Mellanox ConnectX-4,5,6, Broadcom BCM57414	5
Pool drive NVMe	4 TB NVMe U.2 SSD, e.g. Intel P4510	10
Pool drive HDD	2TB Enterprise SATA HDD	40

5.3. Example hardware configuration for Hyper-converged system - All-NVMe pool

- 300 delivered dedicated vCPUs, excluding resources for Host OS and hyperconverged storage system (5 active nodes, spare through increased CPU oversubscription)
- 2048 delivered GB RAM (4 active nodes + 1 standby)
- 3 copies on NVMe SSDs
- Distribute drives evenly across 3+ servers
- 5 000 000 IOPS (random read) and <0.15 ms latency.

- Quoted usable space is before gains from snapshots/clones, thin provisioning, and zeroes detection.
- 50x 3.84TB NVMe SSDs - 58.2 TB usable (5x 10-bay nodes)
- Other combinations and sizes are possible
- Expansion with compute-only or storage-only nodes is possible

Barebone	1 RU Single Socket EPYC server 10x U.2 bays, e.g., SuperMicro AS-1114CS-TNR, HPE DL325 Gen10 Plus v2, Dell R6515	5
CPU alternative	AMD EPYC Rome 7543P UP 32C/64T @3.3G 256M 225W	5
RAM	64 GB DDR4-3200 ECC RDIMM	40
Boot drive	M.2 NVMe SSD, e.g., Intel P4101, Toshiba XG6	5
NIC	2x 10/25GbE SFP+ e.g. Mellanox ConnectX-4,5,6, Broadcom BCM57414	5
Pool drive	3.84 TB NVMe U.2 SSD, e.g., Samsung PM9A3, Micron 7300 PRO	50

5.4. Other configurations

Several other configurations have been omitted from this guide for brevity.

- All-SATA SSD
- SATA SSD-Hybrid
- HDD-Only
- Reverse Hybrid (2 copies on SSD and 1 copy on HDD)

Please contact StorPool support to obtain a complete solution design for your use case.

6. Operating systems

- CentOS 7
- AlmaLinux 8
- Rocky Linux 8
- RHEL (Red Hat Enterprise Linux) 8
- Oracle Linux 8
- Ubuntu 18.04 LTS, 20.04 LTS, 22.04 LTS
- VMware ESXi 6.5, 7 (host only, through iSCSI)
- If you use another Linux distribution, just let us know. We can support StorPool on all Linux distributions with good build and packaging systems.
- x86_64 (amd64), ARM, and Power architectures are supported

7. Cloud management systems and hypervisors

- OpenStack/KVM, OpenNebula/KVM, OnApp/KVM, CloudStack/KVM
- OpenStack/Hyper-V (through iSCSI)
- CloudStack/Xen (through iSCSI)
- VMware Hypervisor with VMFS (through iSCSI or NVMe/TCP)
- Proxmox, XenServer, XCP-NG, oVirt, Oracle VM, Microsoft Hyper-V (through iSCSI)
- Custom cloud management systems through StorPool API

8. Burn-in test

Hardware must pass a burn-in stress test before installing StorPool.

9. Stable server booting

When a server is restarted for whatever reason (e.g., power outage, administrator typed "reboot" at the console), it must get back up without human intervention.

10. Remote console and reboot capability

The server must have an IPMI module, including remote reboot and remote KVM capability.

11. Working kernel crash dump mechanism and debug symbols

If the Linux kernel crashes for whatever reason, we want to be able to investigate. We investigate doing a post-mortem analysis of a kernel crash dump file from /var/crash/. StorPool requires a working kernel crash dump on all servers participating in the StorPool cluster and available kernel debug symbols for the running kernel.

12. Nodes count

StorPool requires a minimum of 3 storage nodes per cluster.
We recommend a minimum of 5 voting nodes and enough free capacity to recover triple replication to guarantee data redundancy in case of a node outage. You can also achieve this reliability by keeping spare hardware or using hardware support service with up to 24h reaction time.

Contacts:

For more information, reach out to us.
support@storpool.com
+1 415 670 9320