VIRTUAL INSTRUMENTS

BENCHMARK REPORT

Validating Performance of NexentaStor Software Defined Storage Solution

Nexenta Software Defined Storage All-Flash Performance

AFAs differ from traditional spinning disk arrays in behavior, performance and durability. For example, SSDs write and read in blocks and are limited in the number of writes that can be performed on a particular block. Sophisticated data reduction/On and array-wide wear leveling techniques can dramatically increase SSD durability, in some cases beyond the expected life of a spinning disk drive.

While the performance advantages of AFAs are well documented, much effort has gone into the design of modern AFAs to attack the price premium of flash over disk: sophisticated efficiency techniques promise, for a few workloads (such as full clone VDI), to bring the effective cost of disk arrays and AFAs within striking distance of each other. In addition to the AFA design, one of the potential big benefits of software defined storage (SDS) is a significant decrease is storage costs. But the total cost of storage ownership isn't just about commodity vs. proprietary hardware. It includes the value of all the software and integration and testing work done. To assist the prospective customer, Nexenta has commissioned Virtual Instruments to provide a profile of NexentaStor all-flash array performance.

Storage engineers and architects considering AFAs for their workloads must explore the behavior of these products, and as much as possible, assess them in the context of their expected workloads. With a robust validation process in place, storage engineers and architects can optimally select and configure AFA products for their workloads with potentially considerable impact on both the performance and the cost of their production solutions.

Background & Objectives

Virtual Instruments was chartered with characterizing, benchmarking, and validating the performance of NexentaStor All-Flash arrays. The performance testing workload was representative of higher performance, low latency workloads with both random and sequential I/O profiles. By using Virtual Instrument's workload models, organizations can accurately simulate how well the underlying VDI storage infrastructure will perform based on a variety of real-world testing scenarios.

Virtual Instruments is a leading provider of pre-deployment testing based on production workloads utilizing the proven Load DynamiX Enterprise performance testing platform.

Nexenta is a global leader in Open Source-driven Software-Defined Storage (OpenSDS) with 6,000+ customers, 400+ partners, 42 patents, and more than 1,500 petabytes of storage under management. Nexenta uniquely integrates deep software-only "Open Source" collaboration with one of the largest and most vibrant Open Source communities (46,000 members) and a comprehensive vision around a commodity hardware-centric "Software-Defined Storage" innovation enabling ANY app, cloud platform and protocol

The objectives of the testing were to:

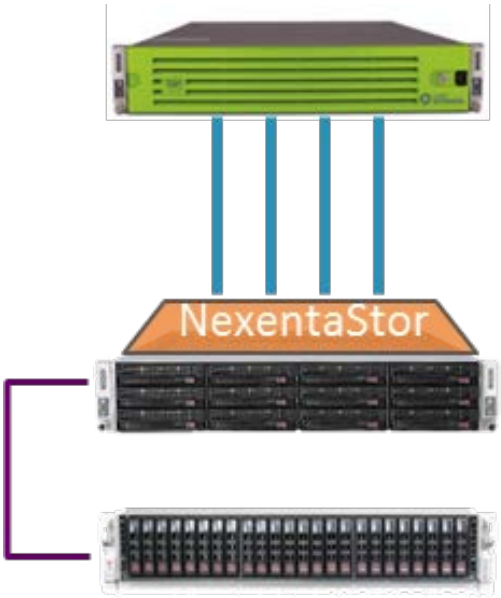
- Characterize and benchmark performance of the NexentaStor all-flash systems using a range of block sizes, access patterns, and mirrors.
- Provide the prospective customer a profile of expected performance characteristics with a range of application workload characteristics.

Table of Contents

Executive Summary

The tests successfully characterized and validated the performance of the NexentaStor All-Flash Array for a range of I/O workload profiles. All latencies and throughputs are at expected levels for the IOPS, I/O sizes, and access patterns configured for this Software Defined Storage solution.

Test Setup and Array Under Test



- Dedicated Load DynamiX Enterprise Appliance, connected via 4x 10GbE to NexentaStor heads
- SMC hardware for All-Flash
  - 256GB of DRAM
  - 24x SSDs, 400GB HGST 1600MM
- 8x NFSv3 shares with 2.4TB working set
  - Each share with 15x files of 20GB each
  - Mimic 120x VM images
- All testing done with lz4 inline compression enabled, on non-compressible dataset (worst case scenario)
- Workload set to 95% Data & 5% Metadata
- Each test run for 10 mins to get sustained numbers

Test Results Summaries

12x Mirrors, All-Flash, Random I/Os

		12xMirrors - 32KB Record Size - async=4 Random I/Os				
IO Size	Access Pattern - Read %	Avg Total BW (MB/s)	Avg Total Ops / Sec	Avg Latency (ms)	NFS Read Ops/sec	NFS Write Ops/sec
N/A	N/A	2,205	35,133	0.4	-	35,133
4KB	100	1,405	353,539	1.0	335,867	-
8KB	100	2,735	355,941	1.0	338,159	-
16KB	100	3,830	253,534	1.5	240,883	-
32KB	100	3,778	126,145	2.8	119,855	-
64KB	100	3,921	65,731	5.9	62,461	-
128KB	100	3,926	32,972	11.7	31,337	-
4KB	75	1,275	320,384	1.0	228,225	76,153
8KB	75	2,165	281,597	1.1	200,603	66,938
16KB	75	2,536	167,809	1.7	119,531	39,898
32KB	75	2,657	88,705	3.3	63,180	21,108
64KB	75	2,696	45,192	6.8	32,156	10,795
128KB	75	2,742	23,024	13.4	16,359	5,532
4KB	50	1,138	285,676	1.1	135,680	135,735
8KB	50	1,307	169,821	1.5	80,673	80,670
16KB	50	1,376	91,005	2.7	43,224	43,251
32KB	50	1,491	49,762	5.4	23,632	23,658
64KB	50	1,507	25,244	11.8	11,979	12,019
128KB	50	1,571	13,188	23.6	6,251	6,293
4KB	25	930	233,128	1.1	55,431	166,056
8KB	25	974	126,406	1.9	30,073	90,038
16KB	25	1,014	67,026	3.7	15,965	47,730
32KB	25	1,075	35,860	7.7	8,565	25,518
64KB	25	1,066	17,857	18.6	4,273	12,710
128KB	25	1,117	9,364	36.0	2,265	6,647
4KB	0	856	214,334	1.3	-	203,636
8KB	0	858	111,285	2.4	-	105,732
16KB	0	833	55,051	5.0	-	52,314
32KB	0	825	27,529	11.4	-	26,163
64KB	0	833	13,946	26.7	-	13,259
128KB	0	900	7,547	50.6	-	7,179

- All Flash
- 24x SSDs, 12x Mirrors
- 12 SSDs usable capacity
- 100% overhead
- Random I/Os
- High performance, low latency use cases
- Random I/Os, 95% Data, 2.4TB working set
- async=4
- Max perf with 100% Read as we would expect
- Numbers to keep in mind
  - About 320k 4KB IOPS at 75:25
  - About 90k 32KB IOPS at 75:25
  - Max BW of 4GB/s with 128KB 100% Read

12x Mirrors, All-Flash, Sequential I/Os

		12xMirrors - 32KB Record Size - async=4 Seq I/Os				
IO Size	Access Pattern - Read %	Avg Total BW (MB/s)	Avg Total Ops / Sec	Avg Latency (ms)	NFS Read Ops/sec	NFS Write Ops/sec
4KB	0	858	214,997	1.4	-	204,257
8KB	0	869	112,863	2.8	-	107,234
16KB	0	898	59,347	5.3	-	56,395
32KB	0	926	30,889	10.8	-	29,361
64KB	0	915	15,333	24.3	-	14,577
128KB	0	971	8,146	46.9	-	7,750
4KB	50	1,075	269,899	1.2	128,209	128,215
8KB	50	1,311	170,396	1.7	80,924	80,983
16KB	50	1,444	95,558	2.7	45,376	45,429
32KB	50	1,611	53,778	5.2	25,523	25,581
64KB	50	1,599	26,801	11.2	12,717	12,759
128KB	50	1,683	14,123	21.7	6,690	6,743
4KB	75	1,199	301,494	1.1	214,755	71,687
8KB	75	1,775	230,949	1.4	164,487	54,928
16KB	75	2,046	135,419	2.1	96,439	32,233
32KB	75	2,569	85,753	3.3	61,044	20,440
64KB	75	2,558	42,864	6.9	30,505	10,237
128KB	75	2,681	22,513	13.2	15,981	5,423
4KB	100	1,346	338,847	1.1	321,923	-
8KB	100	2,198	286,116	1.3	271,833	-
16KB	100	2,668	176,632	2.0	167,815	-
32KB	100	3,689	123,151	2.9	117,012	-
64KB	100	3,950	66,222	5.8	62,924	-
128KB	100	3,982	33,452	11.4	31,789	-

- All Flash
- 24x SSDs, 12x Mirrors
- 12 SSDs usable capacity
- 100% overhead
- Sequential I/Os

1. High performance, low latency use cases
2. Seq I/Os, 95% Data, 2.4TB working set

4xraidz2 (4+2), All Flash, Random I/O

		4xraidz2(4+2) - 128KB Record Size - async=4 - Random I/Os				
IO Size	Access Pattern - Read %	Avg Total BW (MB/s)	Avg Total Ops / Sec	Avg Latency (ms)	NFS Read Ops/sec	NFS Write Ops/sec
N/A	N/A	2,290	36,482	0.4	-	36,482
4KB	100	1,028	258,632	0.9	245,705	-
8KB	100	1,135	147,652	1.9	140,286	-
16KB	100	1,463	96,809	3.3	91,982	-
32KB	100	2,283	76,222	4.6	72,418	-
64KB	100	3,212	53,854	5.8	51,174	-
128KB	100	4,412	37,060	9.8	35,220	-
4KB	75	965	242,532	0.9	172,778	57,642
8KB	75	1,184	153,914	1.5	109,628	36,611
16KB	75	1,353	89,530	3.7	63,753	21,315
32KB	75	1,767	58,971	6.1	41,977	14,061
64KB	75	2,509	42,044	7.7	29,918	10,043
128KB	75	3,291	27,633	11.5	19,647	6,622
4KB	50	979	245,727	0.9	116,731	116,732
8KB	50	1,157	150,283	1.4	71,379	71,402
16KB	50	1,239	81,905	3.4	38,906	38,920
32KB	50	1,400	46,724	5.3	22,190	22,216
64KB	50	1,907	31,947	7.5	15,173	15,194
128KB	50	2,292	19,240	12.2	9,138	9,156
4KB	25	902	225,990	0.9	53,717	160,992
8KB	25	1,074	139,435	1.3	33,164	99,319
16KB	25	1,156	76,432	2.8	18,200	54,431
32KB	25	1,221	40,726	5.1	9,718	28,989
64KB	25	1,614	27,041	7.8	6,463	19,241
128KB	25	1,738	14,588	15.5	3,506	10,368
4KB	0	1,005	251,515	1.0	-	238,943
8KB	0	1,291	167,424	1.3	-	159,069
16KB	0	1,301	85,956	2.6	-	81,672
32KB	0	1,207	40,270	5.4	-	38,271
64KB	0	1,385	23,206	9.7	-	22,054
128KB	0	1,356	11,378	21.8	-	10,820

- All Flash
- 24x SSDs
- 4x raidz2 (4+2)
- 16 SSDs usable capacity
- 50% overhead
- Random I/Os
- High performance, low latency use cases
- Random I/Os, 95% Data, 2.4TB working set
- Numbers to keep in mind
  - About 240k 4KB IOPS at 75:25
  - About 60k 32KB IOPS at 75:25
  - Max BW of 4.4GB/s with 128KB 100% Read

4xraidz1 (4+1), All Flash, Random I/O with only 20 SSDs

		4xraidz1(4+1) - 128KB Record Size - async=4 - Random I/Os				
IO Size	Access Pattern - Read %	Avg Total BW (MB/s)	Avg Total Ops / Sec	Avg Latency (ms)	NFS Read Ops/sec	NFS Write Ops/sec
N/A	N/A	2,328	37,085	0.4	-	37,085
4KB	100	1,040	261,820	0.8	248,737	-
8KB	100	1,228	159,740	1.9	151,761	-
16KB	100	1,693	112,033	2.9	106,452	-
32KB	100	2,445	81,617	4.5	77,553	-
64KB	100	3,402	57,024	6.0	54,187	-
128KB	100	4,418	37,110	9.8	35,268	-
4KB	75	992	249,376	0.8	177,656	59,273
8KB	75	1,202	156,278	1.8	111,311	37,175
16KB	75	1,355	89,660	3.8	63,840	21,355
32KB	75	1,819	60,703	5.9	43,216	14,471
64KB	75	2,624	43,973	7.7	31,298	10,492
128KB	75	3,408	28,613	11.4	20,339	6,862
4KB	50	969	243,237	0.9	115,555	115,536
8KB	50	1,206	156,638	1.7	74,409	74,420
16KB	50	1,322	87,437	3.7	41,527	41,557
32KB	50	1,549	51,684	5.4	24,548	24,569
64KB	50	2,066	34,615	7.3	16,436	16,465
128KB	50	2,461	20,658	12.0	9,804	9,840
4KB	25	968	242,479	0.9	57,626	172,742
8KB	25	1,177	152,753	1.3	36,326	108,809
16KB	25	1,237	81,778	2.6	19,471	58,237
32KB	25	1,325	44,186	4.8	10,549	31,445
64KB	25	1,763	29,530	7.3	7,055	21,015
128KB	25	1,892	15,878	14.6	3,815	11,286
4KB	0	1,041	260,347	0.9	-	247,348
8KB	0	1,321	171,349	1.3	-	162,792
16KB	0	1,375	90,822	2.4	-	86,300
32KB	0	1,297	43,251	5.1	-	41,101
64KB	0	1,515	25,387	8.7	-	24,128
128KB	0	1,479	12,416	19.9	-	11,805

- All Flash
- 20x SSDs
- 4x raidz1 (4+1)
- 16 SSDs usable capacity
- 25% overhead
- Random I/Os
- Only 20 SSDs
- High performance, low latency use cases
- Random I/Os, 95% Data, 2.4TB working set
- Max perf with 100% Read as we would expect
- Numbers to keep in mind
  - About 250k 4KB IOPS at 75:25
  - About 60k 32KB IOPS at 75:25
  - Max BW of 4.4GB/s with 128KB 100% Read

4xraidz1 (4+1), All Flash, Random I/O with only 20 SSDs and 100% Data

		4xraidz1(4+1) - 128KB Record Size - async=1 Random I/Os 100% Data				
IO Size	Access Pattern - Read %	Avg Total BW (MB/s)	Avg Total Ops / Sec	Avg Latency (ms)	NFS Read Ops/sec	NFS Write Ops/sec
N/A	N/A					
4KB	100	1,070	257,896	0.4	257,636	-
8KB	100	1,509	187,433	0.6	187,244	-
16KB	100	1,929	121,645	0.9	121,521	-
32KB	100	2,704	85,939	1.3	85,850	-
64KB	100	3,683	58,752	1.8	58,691	-
128KB	100	4,384	35,037	2.5	35,000	-
4KB	75	1,063	255,438	0.4	191,338	63,843
8KB	75	1,498	185,799	0.6	139,160	46,451
16KB	75	1,860	117,250	1.0	87,806	29,326
32KB	75	2,397	76,165	1.5	57,007	19,080
64KB	75	2,742	43,741	2.5	32,692	11,004
128KB	75	3,614	28,885	3.7	21,571	7,282
4KB	50	1,049	251,673	0.5	125,728	125,692
8KB	50	1,410	174,751	0.7	87,290	87,284
16KB	50	1,639	103,233	1.1	51,575	51,553
32KB	50	1,808	57,426	2.0	28,666	28,701
64KB	50	1,979	31,564	3.5	15,726	15,804
128KB	50	2,597	20,760	5.4	10,334	10,404
4KB	25	1,030	246,513	0.5	61,607	184,658
8KB	25	1,409	174,425	0.7	43,611	130,638
16KB	25	1,444	90,945	1.2	22,768	68,084
32KB	25	1,501	47,661	2.4	11,931	35,681
64KB	25	1,640	26,161	4.3	6,550	19,583
128KB	25	1,956	15,636	7.2	3,926	11,693
4KB	0	1,019	243,419	0.5	-	243,174
8KB	0	1,666	206,031	0.6	-	205,824
16KB	0	1,660	104,467	1.2	-	104,362
32KB	0	1,384	43,946	2.9	-	43,900
64KB	0	1,304	20,788	6.1	-	20,766
128KB	0	1,524	12,184	10.3	-	12,170

- All Flash
- 20x SSDs
- 4x raidz1 (4+1)
- 16 SSDs usable capacity
- 25% overhead
- Low Latency
- Only 20 SSDs
- High performance, low latency use cases
- Random I/Os, 100% Data, 2.4TB working set
- Latency measured with client side Queue Depth=1

Conclusions

1. The tests successfully characterized and validated the performance of the NexentaStor All-Flash Array for a range of I/O workload profiles. All latencies and throughputs are at expected levels for the IOPS, I/O sizes, and access patterns configured.
2. With mirrored all-flash, performance for sequential I/Os is very similar to performance with random I/O. For example, for a workload with a 16KB I/O size, 50/50 R/W access, random latency was 2.7ms, sequential latency was also 2.7ms.
3. With RAID 4+2, latency differences between all Read and all Write were surprisingly small, especially when I/O sizes were below 64KB. As expected, significant differences in throughput and bandwidth showed up in workloads where R/W ratios differed a lot. For instance, at 100% Read and 128KB I/O size, bandwidth was 4.4GB/s, while at 100% Write, it was 1.4GB/s.
4. When RAID was changed to 4+1 and the number of SSDs was reduced to only 20, latencies and throughput stayed fairly similar when compared to the 4+2 system and 24 SSDs.
5. With a 4+1 system and client side queue depth = 1, there was a minor drop in total IOPS with 4KB I/O numbers, but the figures are still very good. Note the great latency figures in the in the 500µs to 700µs range for 4 and 8KB IOPS numbers.

VIRTUAL INSTRUMENTS

Sales
Sales@virtualinstruments.com
1.888-522.2557

Website
virtualinstruments.com

©11/2016 Virtual Instruments. All rights reserved. Features and specifications are subject to change without notice. VirtualWisdom®, Virtual Instruments®, SANInsight®, Workload Central and Load DynamiX® are trademarks or registered trademarks of Virtual Instruments in the United States and/or in other countries.