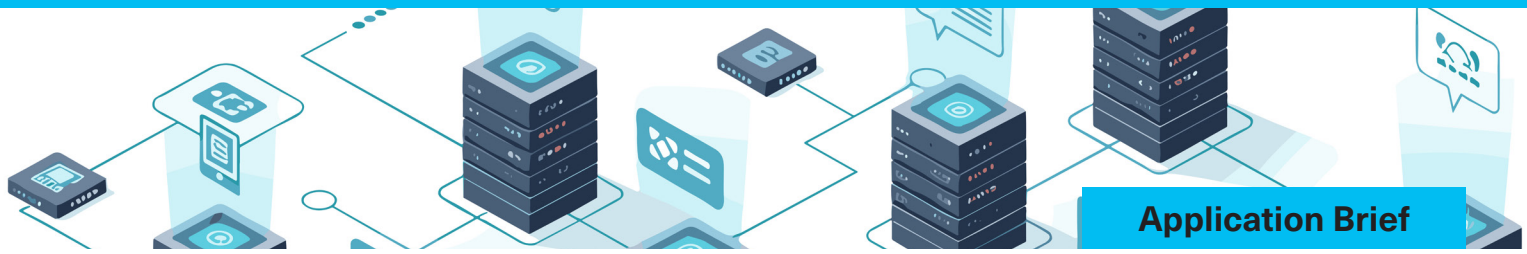
KIOXIA



Application Brief

Using a Disaggregated Storage Architecture as a High-Performance DiskANN[1] Storage Backend for Milvus® Vector Databases

Measured at 1M, 10M and 100M Vector Scale and Tested with 30 TB[2] KIOXIA CM7 Series NVMe™ SSDs in a WD® OpenFlex® Data24 Storage Platform

Introduction

Modern AI applications, from chatbots that retrieve relevant documents to provide additional context to search engines that understand meaning rather than just keywords, rely on vector databases (DBs) to store and query embeddings that capture the semantic content of text, images or other data. The core operation is always the same - given a new query vector, find the stored vectors most similar to it (known as nearest neighbor searches). As the total amount of stored vectors increases and is scaled up, it can become increasingly difficult to support high-query throughput and low latencies, which are necessary to supply accurate information, decrease hallucinations and ensure that application response times are kept low.

One of the most adopted nearest neighbor search algorithms used today is Hierarchical Navigable Small World (HNSW), which builds a layered proximity graph across all vectors enabling fast, accurate searches. Its weakness is that the entire layered proximity graph must live in host DRAM to function. When the math is performed on a modest corpus (such as 100M vectors each with 768 dimensions and stored as 32-bit floats), the raw embedding data alone can consume roughly 307 gigabytes[2] (GB). If the graph structure is added on top of the raw embedding data, most servers will exceed what they can hold in DRAM.

DiskANN, developed by Microsoft® Research, offers another option. It was designed from the ground up to be used with NVMe SSDs as the primary storage tier, keeping only product quantized (PQ) vectors in DRAM. It then pulls the relevant graph edges and full-precision vectors from the NVMe SSDs. Since modern NVMe SSD latency is fast enough for these access patterns, DiskANN achieves recall competitiveness with in-memory HNSW approaches at a fraction of DRAM costs. While there are many DiskANN deployments that use locally-attached storage, there are limitations to usable capacity within a single server that can hinder scalability of higher vector counts.

Disaggregated storage powered by the NVMe over Fabrics™ (NVMe-oF™) protocol is critical for modern vector DBs because it breaks the performance and scaling bottlenecks of traditional direct-attached storage in AI-native infrastructures. Decoupling computation from capacity, organizations can scale storage resources independently to handle massive embedding datasets without being restricted by the PCIe® slots or CPU resources of a single server. Using high-speed networks such as RDMA over Converged Ethernet (RoCE™), the NVMe-oF protocol provides latency and throughput nearly identical to local NVMe drives, which is crucial to achieve high input/output operations per second (IOPS) and low latency required during real-time nearest neighbor searches. This architecture facilitates a 'shared-everything' approach where multiple compute nodes can directly access a centralized, persistent pool of vectors, ensuring consistent performance and simplified management for dynamic, containerized vector DB workloads.

This application brief presents performance results measuring a Milvus vector DB using remote SSD DiskANN approaches against an in-memory HNSW baseline. The test configuration included 30 terabyte[2] (TB) KIOXIA CM7 Series NVMe SSDs in a Western Digital® (WD) OpenFlex Data24 4200 EBOF[3] Storage Platform as a disaggregated NVMe-oF storage target over RoCE v2.

VectorDBBench software was used to perform similarity searches that included three dataset scales at 1M, 10M and 100M, each with 768-dimensional vectors. Three remote storage configurations were tested spanning single-path and four-path NVMe-oF access across one and two 200 GbE network interface cards (NICs). Test metrics were gathered for database throughput (queries per second), percentage of recall, 95th percentile (p95) / 99th percentile (p99) latency and load duration.

The test results show the performance ceiling and critical importance of multipath configurations at scale. As a result, a disaggregated storage architecture is a viable choice as a high-performance DiskANN storage backend for Milvus vector DBs. Included is a brief description of each workload test, test results, analysis, the hardware/software test configuration (Appendix A) and the configuration setup/test procedures (Appendix B).

Test Results Snapshot

The following test results measure Milvus using DiskANN approaches (WD OpenFlex Data24 4200 with KIOXIA CM7-R Series SSDs) versus an in-memory HNSW baseline:

Database Throughput

56% more QPS
(1M vectors)
DiskANN 4-Paths / 1-NIC

30% more QPS
(1M vectors)
DiskANN 4-Paths / 2-NICs

19% more QPS
(10M vectors)
DiskANN 4-Paths / 1-NIC

16% more QPS
(10M vectors)
DiskANN 4-Paths / 2-NICs

Percentage of Recall

more than 99%
(1M/10M/100M vectors)
All DiskANN configurations

p95 Latency

63% / 57% ms faster
(100M vectors)
DiskANN 4-Paths / 1-NIC
DiskANN 4-Paths / 2-NICs

p99 Latency

175% / 170% ms faster
(100M vectors)
DiskANN 4-Paths / 1-NIC
DiskANN 4-Paths / 2-NICs

Load Duration

Similar
All DiskANN configurations

Test Scenario

Four configurations were evaluated across three vector scales. The 1M/10M vectors used were Cohere™ datasets, while the 100M vectors used was a Large-scale Artificial Intelligence Open Network (LAION) dataset. All datasets utilized 768 dimensions per vector. Two vector indexes supported by the Milvus® vector DB were tested including HNSW and DiskANN. Tests performed against HNSW were entirely in memory on the client server and used as a baseline for performance. The remote DiskANN configurations featured the Data24 4200 platform via NVMe-oF™ over RoCE™ v2 (NVMe™/RDMA). An NVMe multipath configuration was used in two of the remote DiskANN configurations. The four configurations are summarized in the following table:

Configuration	Index	Storage Path	NVMe Paths	NICs Used	Description
Local HNSW	HNSW	Host DRAM (in-memory)	N/A	1	In-memory HNSW graph. Establishes a maximum queries per second (QPS) ceiling for memory-resident workloads. Requires a full graph in DRAM making it impractical beyond ~50M, 768 dimensional vectors on a 384 GB host.
Remote DiskANN 1-Path, 1-NIC	DiskANN	NVMe-oF via RoCE v2 (Data24 4200)	1	1	Single NVMe namespace path to the Data24 4200 over a single 200 GbE RoCE NIC. It's the baseline for disaggregated DiskANN and exposes single-path bandwidth saturation at the 100M vector scale.
Remote DiskANN 4-Paths, 1-NIC	DiskANN	NVMe-oF via RoCE v2 (Data24 4200)	4	1	Four NVMe namespace paths to the Data24 4200 over a single 200 GbE RoCE NIC to enable multipath IO. It represents a primary production configuration and eliminates single-path latency collapse at large scale.
Remote DiskANN 4-Paths, 2-NICs	DiskANN	NVMe-oF via RoCE v2 (Data24 4200)	4	2	Four NVMe namespace paths to the Data24 4200 are distributed across two 200 GbE RoCE NICs for additional bandwidth headroom and tests the incremental benefit of dual-NIC fabric attach at each dataset scale.

Test Challenges

Prior to the testing, the following challenges were identified:

- *HNSW index graphs must reside entirely in DRAM, making 100M+ vector corpora impractical on standard server configurations.*
- *Locally-attached NVMe capacity is constrained by server chassis bay count, coupling vector storage scaling to compute node procurement.*
- *A single-path NVMe-oF configuration to a remote target can become bandwidth-saturated at large dataset scales, causing catastrophic latency degradation.*
- *High-recall ANN searches require consistent low-latency storage IO - any queuing on the storage path may directly inflate p95 and p99 query latency.*
- *Vector DB infrastructure must remain operationally manageable to support corpus size scales from millions to billions of embeddings.*

Test Results[4]

Metric 1: Throughput - Queries per Second

This metric measured the number of completed Approximate Nearest Neighbor (ANN) search queries per second as measured by VectorDBBench during the search phase. Higher QPS throughput indicates that the hardware resources are working efficiently and can scan through more vectors in the index, returning the results of the similarity searches faster.

Throughput Results in QPS (higher results are better)

Vector Scale	Local HNSW	Remote DiskANN 1-Path, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 2-NICs	Change vs. HNSW
1M	5,081.6	3,713.9	-36%	7,971.8	+56%	6,624.4	+30%
10M	557.8	539.9	-3%	665.9	+19%	650.6	+16%
100M	77.9	33.3	-133%	67.2	-15%	67.2	-15%

Analysis:

The remote DiskANN 4-Paths / 1-NIC configuration delivered the highest QPS at 1M vector scale with 7,971.8 QPS, or 56% higher than local HNSW. The remote DiskANN 4-Paths / 2-NICs configuration delivered the second highest QPS at 1M vector scale with 6,624.4 QPS, or 30% higher than local HNSW.

At 10M vector scale, the remote DiskANN 4-Paths / 1-NIC configuration delivered 665.9 QPS, or 19% higher than local HNSW, while the remote DiskANN 4-Paths / 2-NICs configuration delivered 650.6 QPS, or 16% higher than local HNSW.

Metric 2: Percentage of Recall

This metric measured the ability of the Milvus® vector DB to find all relevant cases (true positives) within a vector space when conducting a similarity search. The results are calculated as follows:

$$\frac{\textit{Number of True Positives}}{\textit{Number of True Positives + Number of False Negatives}}$$

The recall shows the proportion of actual positives that were identified correctly. From the calculations, the higher the recall, the more precise and similar the results will be when vector data is returned to the vector DB. The results are in % of recall, and the higher results are better.

Recall in % (higher results are better)

Vector Scale	Local HNSW	Remote DiskANN 1-Path, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 2-NICs	Change vs. HNSW
1M	97.99	99.43	+1.44	99.53	+1.54	99.53	+1.54
10M	98.35	99.62	+1.27	99.6	+1.25	99.55	-+1.20
100M	99	99.23	+0.23	99.14	+0.14	99.18	+0.18

Analysis:

All three remote DiskANN configurations delivered a higher percentage of recall versus local HNSW regardless of the vector size. In all cases, the percentage of recall over 99% means that 99% of the correct nearest neighbors were returned. Indexes that trade recall for QPS sacrifice result quality.

Metric 3 and 4: p95 / p99 Latency

This metric measured the maximum time taken for the top 95% or 99% of vector search requests, respectively. These metrics are crucial for identifying tail latency, revealing how long the worst 5% (1 in 20 queries during the search) or 1% (1 in 100 queries during the search) take, which often indicates system bottlenecks or performance spikes. The results are in milliseconds (ms), and the lower results are better.

p95 Latency Results in ms (lower results are better)

Vector Scale	Local HNSW	Remote DiskANN 1-Path, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 2-NICs	Change vs. HNSW
1M	2.6	4.3	+65%	4.3	+65%	4.9	+88%
10M	6	6.7	+11%	6.7	+11%	7.1	+18%
100M	40.6	2,574	+6,239%	24.9	-63%	25.7	-57%

Analysis:

The single-path remote DiskANN configuration at 100M scale delivered 2,574 ms for p95 latency and is a direct consequence of saturation on a single path at this IO depth. The DiskANN 4-Paths / 1-NIC configurations resolves this entirely as p95 latency drops to 24.9 ms delivering a 103× improvement.

The Data24 4200 platform enables up to 24 SSDs in a compact 2U enclosure with a PCIe® 4.0 backplane and dual-lane, per-drive connectivity. For this test, three KIOXIA CM7 Series drives at 30.72 TB each provided the storage backend and was connected via 12-100 GbE RoCE™ cables. The Data24 4200 presented each drive as an independent NVMe™ namespace, allowing multipath NVMe-oF™ connections from the host server for distributing IO across multiple physical paths to the same or different physical SSDs.

The DiskANN four-path configuration used in testing connected the host to four independent NVMe namespaces on the Data24 4200 over RoCE. The DiskANN index was striped or distributed across these namespaces by the Milvus storage layer so concurrent graph traversal queries could issue IO to multiple NVMe namespaces in parallel without serializing through a single path's queue depth. At 100M vectors, this parallelism is what separates a functional deployment (24.9 ms at p95 latency) from an unusable one (2,574 ms at p95 latency).

p99 Latency Results in ms (lower results are better)

Vector Scale	Local HNSW	Remote DiskANN 1-Path, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 2-NICs	Change vs. HNSW
1M	2.9	4.6	+58%	4.6	+58%	5.3	+82%
10M	6.5	7	+7%	7	+7%	7.5	+15%
100M	73.1	2,653	+3,529%	26.5	-175%	27	-170%

Analysis:

Same analysis as with p95 latency except with different results. For p99 latency, the single-path remote DiskANN configuration at 100M scale delivered 2,653 ms and a direct consequence of saturation on a single path at this IO depth. The DiskANN 4-Paths / 1-NIC configuration resolves this entirely as p99 latency drops to 26.5 ms delivering a 100× improvement.

Metric 5: Load Duration

This metric measured the total time required for the Milvus® vector DB to ingest the full vector corpus, construct the necessary search indexes and commit it to the storage backend, which is crucial for evaluating how quickly the platform can be prepared for production queries. Lower load durations equate to faster times to either construct the vector index initially or reconstruct vector indexes if new additional data needs to be considered. The results are in hours, and the lower results are better.

Load Duration in hours (lower results are better)

Vector Scale	Local HNSW	Remote DiskANN 1-Path, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 1-NIC	Change vs. HNSW	Remote DiskANN 4-Paths, 2-NICs	Change vs. HNSW
1M	0.57	0.61	+7%	0.61	+7%	0.61	+7%
10M	5.84	5.96	+2%	5.96	+2%	5.99	+2%
100M	57.7	58.9	+2%	59	+2%	59.2	+2%

Analysis:

There was no significant difference in load duration between the local HNSW configuration and the three remote DiskANN configurations regarding the total time that was required for the Milvus vector DB to ingest the full vector corpus, construct the necessary search indexes and commit it to the storage backend.

Observations

The following observations were derived from the testing:

QPS: DiskANN on NVMe-oF™ Outperforms In-Memory HNSW at 1M and 10M Vector Scale

The most immediate result of the testing is that disaggregated DiskANN is not a compromise. At both 1M and 10M vector scales, a correctly configured NVMe-oF DiskANN deployment out-queried the in-memory HNSW baseline on the same hardware.

- *At 1M vectors, the DiskANN 4-Paths / 1-NIC configuration reached 7,971.8 QPS versus 5,081.6 QPS delivered by local HNSW and provided 56% more throughput from a disaggregated storage index.*
- *At 10M vectors, the DiskANN 4-Paths / 1-NIC advantage narrowed to 19% (665.9 vs. 557.8 QPS), reflecting the higher per-query I/O cost as the graph grows.*

The QPS advantage derives from the DiskANN graph structure, which is optimized specifically for sequential NVMe™ access patterns, combined with the ability of the Data24 4200 to sustain those access patterns over multiple concurrent NVMe paths without contention.

Recall: DiskANN Delivers Higher Accuracy than In-Memory HNSW Across All Vector Scales

DiskANN recall was consistently higher than the in-memory HNSW baseline at every tested scale and every remote configuration.

- *At 1M vectors, the three remote DiskANN configurations delivered faster recall versus 0.9799 achieved by the local HNSW baseline, representing 1.4 to 1.5-point gaps.*
- *At 10M vectors, the three remote DiskANN configurations delivered faster recall versus 0.9835 achieved by the local HNSW baseline, representing 1.2-point gaps.*
- *At 100M vectors, parity is essentially restored within noise, yet the three remote DiskANN configurations delivered faster recall versus 0.99 achieved by the local HNSW baseline, representing smaller gaps and comparable recall. For Retrieval Augmented Generation (RAG) pipelines where retrieval accuracy directly affects generation quality, this can be a very important metric.*

P95 / p99 Latency: Multipath is Mandatory at 100M Scale - Over 100x Latency Difference

- *At 100M vectors, the remote, single-path DiskANN configuration (1-Path / 1-NIC) clearly showed the value of using a multipath approach to NVMe-oF targets. At single path, the test yielded p95 latency at 2,574 ms and p99 latency at 2,653 ms, which is not viable for production use. At this very large vector scale, a single NVMe path cannot handle concurrent DiskANN queries efficiently. Using multiple NVMe-oF paths eliminates saturation, lowering p95 latency to 24.9 ms (103x difference) and p99 latency to 26.5 ms (100x difference), making multipath DiskANN a requirement for production environments at this scale.*

Dual NICs: Headroom With Minimal Incremental Gain at Current Scale

Adding a second 200 GbE RoCE™ NIC (4-Paths / 2-NICs) to the remote DiskANN configuration produces results that are statistically equivalent to a DiskANN 4-Paths / 1-NIC configuration at all three tested vector scales.

- *At 1M vectors, the DiskANN 4-Paths / 2-NICs configuration trailed the DiskANN 4-Paths / 1-NIC configuration in QPS by 16% in QPS (6,624 vs. 7,971), likely due to path balancing overhead or Non-Uniform Memory Access effects at this dataset size.*
- *At 10M and 100M vectors, QPS and latency are nearly identical between the two configurations. The practical implication is that a single 200 GbE RoCE NIC with four-path NVMe-oF is sufficient for the tested scales, and the second NIC provides bandwidth headroom for corpus growth or additional concurrent workloads rather than a latency or throughput improvement at current corpus sizes.*

Conclusions

Three noteworthy conclusions can be drawn from these tests as follows:

1. *Disaggregated NVMe-oF™ DiskANN is faster than local in-memory HNSW at 1M and 10M vector scales, delivering 56% and 19% more QPS, respectively, and with higher recall. Any notion that disaggregated storage carries some inherent performance penalty is falsified given the workloads and vector scales tested.*
2. *At 100M vectors, the choice of path configuration is the single most important deployment decision. Single-path NVMe-oF produces p95 latency of 2,574 ms and shows that incorrectly configured deployments are unusable for any interactive workload. The multipath configuration on the same fabric and hardware produces p95 latency of 24.9 ms and reiterates the key difference between a functional system and a non-functional one. Every deployment targeting 100M+ vectors must be configured with a minimum of four NVMe-oF paths.*
3. *The Data24 4200 with KIOXIA CM7 Series NVMe™ SSDs provides the storage bandwidth and queue depth necessary to sustain DiskANN access patterns at all tested scales on a standard 200 GbE RoCE™ infrastructure. No specialized fabrics, communications standards or proprietary drivers are required while taking operational advantage of disaggregation - independent scaling of storage and compute, shared storage pools, simpler server BOM, and lower total cost of ownership without a performance cost when the system is correctly configured.*

As a result of these conclusions, the Data24 4200 platform deployed with KIOXIA CM7 Series NVMe SSDs as a disaggregated storage architecture is a viable solution for use as a high-performance DiskANN storage backend for Milvus® vector DBs.

Products Tested:

KIOXIA CM7 Series SSDs

The KIOXIA CM7 Series enterprise NVMe SSDs support EDSFF E3.S and 2.5-inch form factors and are compliant with the PCIe® 5.0 and NVMe 2.0 specifications. These SSDs are available in two configurations: KIOXIA CM7-R Series for read- intensive applications (1 DWPD[5], up to 30.72 TB capacities) and KIOXIA CM7-V Series for higher endurance, mixed-use applications (3 DWPD, up to 12.8 TB capacities). Additional features include a dual-port design for high availability applications, flash die failure protection and security options[6]. Additional KIOXIA CM7 Series SSD information is available here.

KIOXIA CD8P Series SSDs[7]

WD® OpenFlex® Data24 4000 Series Storage Platform

The WD OpenFlex Data24 4000 series NVMe-oF storage platform extends the high-performance of NVMe flash to shared storage. The 4000 series provide low latency sharing of NVMe SSDs over a high-performance Ethernet fabric to deliver similar performance to locally attached NVMe SSDs. The WD RapidFlex® NVMe-oF controllers allow up to six dual-pathed hosts to be attached without a switch. The OpenFlex Data24 4000 series uses three RapidFlex A2000 fabric bridge adapters per input/output module to provide up to 12 ports of 100 GbE that can connect to RDMA and/or RDMA configured host ports. Additional WD OpenFlex Data24 4000 Series information is available here.

WD OpenFlex Data24 4000 Series Storage Platform[7]

Appendix A

Hardware/Software Test Configuration

Server Information	
Vendor/Model	Dell™ PowerEdge™ R6615
No. of Servers	1

CPU Information	
Vendor/Model	AMD EPYC™ 9454P
No. of CPU Sockets	1
No. of CPU Cores	48
CPU Frequency	2.75 gigatransfers per second

Memory Information	
Memory Type	DDR5-4800
Memory Size	384 GB
No. of DIMMs	12 x 32 GiB[8] DIMMs
Total DIMMs Size	384 GiB

OS Information	
OS	Ubuntu®
Version	24.04.3 LTS
Kernel	6.8.0-94-generic

Vector DB Information	
Vector DB	Milvus®
Milvus Version	2.6.9
MinIO® Version	RELEASE.2024-12-18T13-15-44Z
etcd® Version	3.5.25

Load Generator Information	
Load Generator	VectorDBBench
Version	1.0.18

File System Information	
File System	XFS®
Mount Options	noatime, nodiratime

Storage Target Information	
Platform	WD® OpenFlex® Data24 4200 Series
Backplane	PCIe® 4.0
Drive Lane Connectivity	x2

SSD Information	
Model	KIOXIA CM7-R Series
Interface	PCIe 5.0 x4
No. of Drives	3
Form Factor	2.5-inch
Capacity	30.72 TB
DWPD	1 (5 years)

Fabric Information	
Fabric	100 GbE RoCE™
No. of Cables	12
Type	NVMe™/RDMA (NVMe-oF™ over RoCE)

Index Information	
HNSW Parameters	M=30, ef_construction=360, ef_search=100
DiskANN Parameters	Search_list=100, k=10, metric=Cosine

Dataset Information	
Vector Sizes	1M / 10M /100M
Dimensions	768-dimensional float32 vectors (Performance 768D)

Appendix B

Configuration Setup/Test Procedures

Configuration Setup

A Dell™ PowerEdge™ R6615 server was set up with a 200 Gb/s RoCE™-capable NIC with an Ubuntu® 24.04.3 LTS operating system and connected to a 200 Gb/s switch.

A WD® OpenFlex® Data24 4200 Series storage enclosure was set up and connected to a 200 Gb/s switch with 12x connections.

Three 30.72 TB KIOXIA CM7-R Series SSDs were installed into the Data24 4200 storage enclosure.

- *The SSDs were set up with an XFS® file system and mounted with noatime and nodiratime flags once the SSDs were remotely attached to the Dell PowerEdge R6615 server.*

Additional NVMe™ paths were allocated and the NVMe path count per drive was increased from 1 to 4.

The Milvus® vector DB was downloaded and set up so that raw data and the vectors and indexes would be stored onto the SSDs.

The VectorDBBench load generator was installed and used to run a search performance test for various vector DB sizes. Two Cohere™ datasets (1 million and 10 million vectors) and one LAION dataset (100 million vectors) were downloaded and inserted into the Milvus vector DB. All datasets utilized 768 dimensions per vector.

A DiskANN index was created to run the test metrics.

Test Procedures

The VectorDBBench Performance 768D test was run for 1M, 10M and 100M vectors in local in-memory HNSW and in remote DiskANN configurations that included 1-Path / 1-NIC, 4-Paths / 1-NIC and 4-Paths / 2-NICs.

The following metrics were gathered:

- *Database Throughput (in QPS)*
- *Recall (in % of Recall)*
- *p95 / p99 Latency (in milliseconds)*
- *Load Duration (in hours)*

The metrics of the remote DiskANN configurations were compared to the local HNSW in-memory configuration.

NOTES:

[1] The DiskANN repository requests the following citation:
@misc{diskann-github;
authors = Simhadri, Harsha Vardhan and Krishnaswamy, Ravishankar and Srinivasa, Gopal and Subramanya, Suhas Jayaram and Antonijevic, Andrija and Pryce, Dax and Kaczynski, David and Williams, Shane and Gollapudi, Siddarth and Sivashankar, Varun and Karia, Neel and Singh, Aditi and Jaiswal, Shikhar and Mahapatro, Neelam and Adams, Philip and Tower, Bryan and Patel, Yash;
title = DiskANN: Graph-structured Indices for Scalable, Fast, Fresh and Filtered Approximate Nearest Neighbor Search;
urls = https://github.com/Microsoft/DiskANN};
version = 0.6.1;
year = 2023;

[2] Definition of capacity: KIOXIA Corporation defines a megabyte (MB) as 1,000,000 bytes, a gigabyte (GB) as 1,000,000,000 bytes, a terabyte (TB) as 1,000,000,000,000 bytes and a petabyte (PB) as 1,000,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1Gbit = 2^{30} bits = 1,073,741,824 bits, 1GB = 2^{30} bytes = 1,073,741,824 bytes, 1TB = 2^{40} bytes = 1,099,511,627,776 bytes and 1PB = 2^{40} bytes = 1,125,899,906,842,624 bytes and therefore shows less storage capacity. Available storage capacity (including examples of various media files) will vary based on file size, formatting, settings, software and operating system, and/or pre-installed software applications, or media content. Actual formatted capacity may vary.

[3] EBOF = Ethernet Bunch of Flash and a type of storage enclosure or architecture that connects a large number of NVMe™ SSDs directly to a network via Ethernet rather than through a traditional server.

[4] Read and write speed may vary depending on various factors such as host devices, software (drivers, OS, etc.) and read/write conditions.

[5] DWPD: Drive Write(s) Per Day: One full drive write per day means the drive can be written and re-written to full capacity once a day, every day, for the specified lifetime. Actual results may vary due to system configuration, usage, and other factors.

[6] Optional security feature compliant drives are not available in all countries due to export and local regulations.

[7] The product images shown are representations of design models and are not accurate product depictions.

[8] One gibibyte means 2^{30} or 1,073,741,824 bytes.

TRADEMARKS:

DISCLAIMERS: