ANODE: APPROXIMATE NEIGHBORS ON DISK EXTENDED

Rise of AI

The ever-growing demand for real-time vector similarity search in applications like *image retrieval*, *recommender systems*, and *anomaly detection* pushes the boundaries of existing Approximate Nearest Neighbor (ANN) algorithms. Traditional methods often struggle with high resource consumption, limited scalability, and slow performance on large datasets. This white paper introduces ***ANODE (Approximate Neighbors On Disk Extended)***, a novel indexing scheme developed by **JetEngine** specifically for **PulsejetDB**, our high-performance vector database. ANODE breaks new ground in ANN search, offering unmatched efficiency and flexibility for modern vector applications.

Challenges of new AI age

Current ANN algorithms, such as FreshDiskANN and DiskANN, suffer from significant limitations. Their reliance on sequential indexing leads to concurrency bottlenecks, hindering performance in high-throughput environments. Additionally, these methods require substantial memory resources and often struggle with high-dimensional vectors.

Companies focusing on GenAI solutions

The advent of Generative AI (GenAI) has ushered in a new era of technological advancement, revolutionizing industries from healthcare to entertainment. As companies race to harness the power of GenAI, a critical component has emerged: *vector databases*. These specialized databases are optimized to store and retrieve high-dimensional vectors, the fundamental data structures used in GenAI models.

Introducing ANODE

ANODE is a revolutionary indexing scheme developed by JetEngine specifically for Pulsejet Vector Database. ANODE shatters the limitations of previous methods, offering unparalleled performance, scalability, and real-time capabilities.

ANODE addresses these limitations previously mentioned by offering a compelling alternative for demanding ANN workloads. Here's what sets ANODE apart...

Unparalleled Efficiency & Performance

Pay and get what you deserve, not the limitations.

- **Unparalleled Concurrency** with Concurrent Indexing: Unlike its predecessors, ANODE employs a novel concurrent indexing scheme, enabling simultaneous indexing across multiple cores/threads. This dramatically boosts performance, especially in clustered environments.
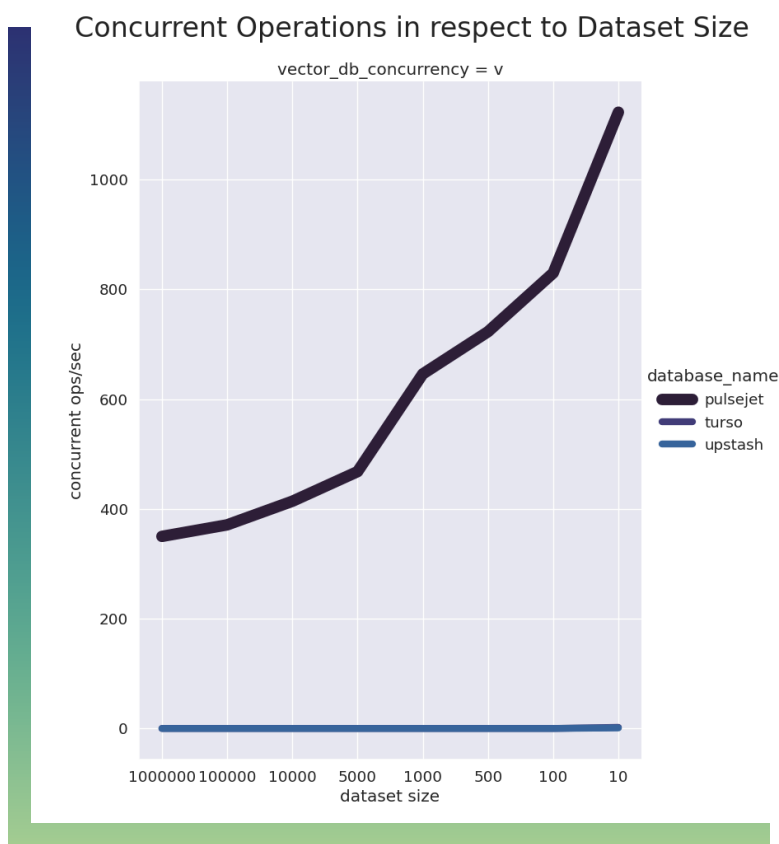
*– Figure 1: Concurrent operations **Pulsejet** vs others (1) –*

The cost-effectiveness of vector databases compared to traditional approaches is a significant factor driving their adoption. GenAI models often require vast amounts of data to train and operate, placing a strain on storage and computational resources. Vector databases, with their specialized indexing and retrieval techniques, can significantly reduce storage costs and improve query performance, making them a more attractive option for businesses seeking to maximize the value of their GenAI investments.

Finally, the flexibility and scalability of vector databases are essential for meeting the evolving needs of GenAI applications. As models become more sophisticated and data volumes grow, databases must be able to adapt and scale accordingly. Vector databases like *PulsejetDB* offer the flexibility to handle a wide range of data types and sizes, ensuring that businesses can leverage GenAI to its full potential without being constrained by technological limitations.

- **Extreme Compression – Up to 95x Smaller Datasets**: ANODE leverages advanced compression techniques, reducing dataset size by up to 95x compared to traditional methods, where traditional methods are actually compressing at most 50x.

- **Minimal Memory Footprint**: ANODE minimizes memory usage during indexing and querying. This allows users to handle massive datasets efficiently, even on resource-constrained machines.

Unleash the Power of Real-Time Vector Search

In the real world, where actions and reactions are immediate, everything operates in real time.

- **Blazing-Fast Loading**: The SIFT1M dataset loads from disk in less than couple of seconds, showcasing the exceptional speed of ANODE.

- **Uninterrupted Querying**: Perform vector queries without interruption even during indexing and optimization.

- **ACID Compliance**: In clustered environments of PulsejetDB, ANODE guarantees ACID (Atomicity, Consistency, Isolation, Durability) properties, ensuring data integrity.

Adaptability and Scalability for Ever-Evolving Data

Move on, change, experiment without boundaries and problem spaces created by tools around you.

- **Dimensionality Agnostic – No Performance Degradation**: Don't get the dimensionality curse, dispel the curse itself! ANODE operates seamlessly across any vector dimension. Whether you're working with low-dimensional or high-dimensional data, ANODE delivers optimal performance without sacrificing accuracy.

- **Real-time Streaming Indexing & Updates**: Queries are seamlessly handled even during indexing and optimization, ensuring continuous availability.

Step up your AI Game

One of the most significant advantages of *PulsejetDB* is its ability to perform seamless querying during indexing. This means that users can continue to access and analyze their data without interruptions, even while the database is being updated. This is a major departure from traditional systems that often require enormous amount of waits and reindexing tasks for changing datasets.

PulsejetDB also prioritizes data consistency and reliability. Its ACID compliance and atomic vector inserts ensure that data integrity is maintained, even in clustered environments. This is crucial for applications that require accurate and up-to-date information.

Finally, *PulsejetDB*'s streaming-friendly design makes it ideal for modern applications that rely on real-time data pipelines. By eliminating the need for expensive daily rebuilds, *PulsejetDB* can adapt to ongoing updates, providing users with the most current and relevant information.

- **Sparse Data Champion**: ANODE excels with sparse datasets, offering faster convergence compared to existing solutions.
- **Exceptional Recall Rates**: ANODE delivers consistently high recall rates, exceeding 0.9 for both 5-recall-rate@5 and 10-recall-rate@10. This translates to highly accurate retrieval of relevant results.
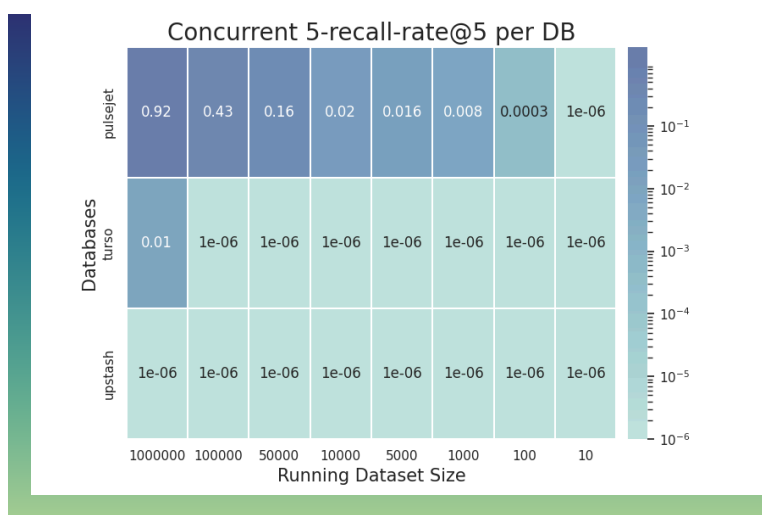
*– Figure 2: Running dataset 5-recall-rate@5 for **Pulsejet** vs others (2) –*

During our test, we couldn't conclude the test for a few databases, as they were unsuccessful in loading the data. Moreover, they require reconstruction and some of them can't return the inserted value after indexing.
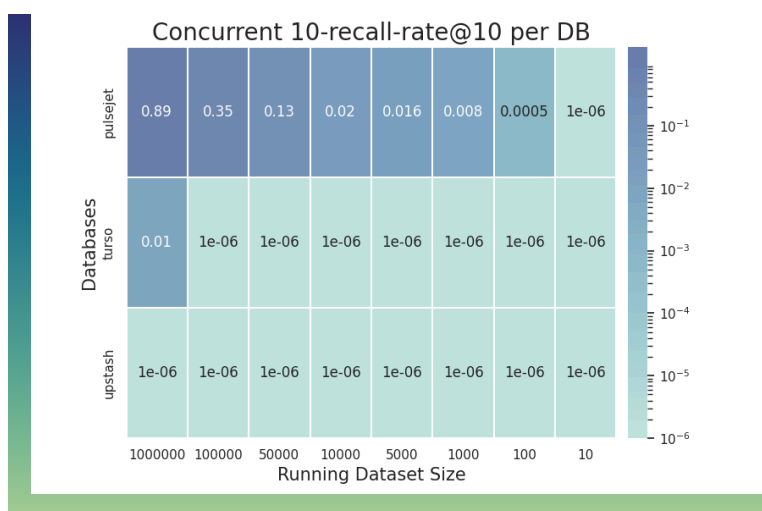
*– Figure 3: Running dataset 10-recall-rate@10 for **Pulsejet** vs others (3) –*

Having billion scale is not possible with existing solutions

One of the most significant challenges faced by existing vector databases is the inability to handle large-scale datasets efficiently. *PulsejetDB*, however, overcomes this hurdle with its advanced compression techniques and minimal memory footprint. This allows users to process massive amounts of data without compromising performance.

Another common issue is the slow indexing speeds of traditional vector databases. *PulsejetDB*'s concurrent indexing scheme addresses this problem by enabling simultaneous indexing across multiple cores/threads, dramatically accelerating the process. This is especially beneficial for clustered environments where performance is critical.

In conclusion, *PulsejetDB*'s **ANODE** indexing scheme represents a *quantum leap* in vector database technology. By addressing the limitations of existing solutions, *PulsejetDB* offers a superior alternative for businesses that demand high performance, scalability, and real-time capabilities.

- **Sparse Data Champion**: ANODE excels with sparse datasets, offering faster convergence compared to existing solutions.

- **Self-Optimizing**: ANODE intelligently adapts to the shape of your data over time, continuously optimizing performance.

> "
> ──────────────────
>
> ANODE revolutionizes indexing by turning scalability challenges into strengths, delivering seamless performance across vast contexts in our advanced RAG without compromising speed, accuracy, or memory efficiency.
>
> —Can Göymen, CEO at Curiosity.tech

Break Through Bottlenecks with ANODE

Because your infrastructure or your hardware shouldn't be the limitation for yourself! On the contrary, it should be game changer and enabler for you.

- **Millisecond Ingestion**: Ingest billions of vectors in just milliseconds, keeping your database up-to-date with the latest data.

- **SSD & NVMe Optimization**: ANODE leverages the capabilities of SSDs and nVME storage, iterating and rewriting itself for long-term storage efficiency.

- **Cloud Storage Integration**: ANODE seamlessly integrates with cloud storage solutions like Google Cloud Storage, scaling your database with ease.

With **ANODE** powering **PulsejetDB**, you gain a powerful vector database solution that delivers unmatched efficiency, scalability, and real-time performance. Say goodbye to limitations and unleash the true potential of your vector data.

Hard tech research, happy customers!

At **JetEngine**, our innovative hard tech research is proving to be a powerful catalyst for customer growth. By investing in cutting-edge technologies and conducting groundbreaking experiments, we are developing solutions that address critical industry challenges. Our commitment to pushing the boundaries of what is possible has attracted the attention of forward-thinking companies seeking to gain a competitive edge. As we continue to deliver tangible results and demonstrate the value of our hard tech research, we are confident in our ability to attract even more customers and establish ourselves as a leading force in the industry.

PulsejetDB with ANODE: A Game Changer

The combination of **PulsejetDB**'s robust architecture and groundbreaking **ANODE** indexing algorithm unlocks a new era of vector search possibilities. With unmatched performance, scalability, and real-time capabilities, **PulsejetDB** with **ANODE** empowers you to:

- *Deliver unparalleled user experiences* with lightning-fast, accurate vector search.
- *Unlock the full potential* of your data by extracting valuable insights from large-scale datasets.
- *Simplify large-scale deployments* with minimal resource requirements and efficient storage utilization.
- *Gain a competitive edge* by harnessing the power of real-time vector search for next-generation applications.

Why settle for less? **PulsejetDB** delivers exceptional performance and minimal resource consumption, unlike other databases that burden you with unnecessary costs. This is where AI industry is evolving and we are inviting you to **PulsejetDB** ecosystem for delivering outcomes.

Ready to unlock the power of ANODE and PulsejetDB? Contact JetEngine today!

Contact us, let's discuss!

Contact us! Let's get in touch, we are looking for transitioning more customers and happy to hear about their success stories. If you want to share your success story with *PulsejetDB* or want to start using it, send us an email at info@jetengine.tech!

If you want to talk with experts in the AI field, come join to our community channels:

Brought to by
JetEngine Engineering Team
info@jetengine.tech