Dailyhunt (VerSe) Migrating to ScyllaDB from AstraDB

A Game Changer

About this case study:

Customer Name	: VerSe Innovation Private Limited
Application Name	: Dailyhunt News App
Scope of Work	: Database Migration
Scale	: 2000 Nodes, 35TB Database, 2PB Storage
Timeline	: 10 Weeks

Dailyhunt (VerSe) is a pioneered media unicorn providing news and content aggregation platform in India with over 300Mn monthly active users, relies on a robust database to deliver personalized news in 13 languages. Given the platform's requirement to handle billions of requests daily, efficient database management is crucial.

ACHIEVEMENT

60%
Improvement In Read Latency

90%
DB Cost Optimised

Handled Event of "Ram Mandir Pran Pratishtha"

The database was seamlessly migrated from AstraDB to ScyllaDB with zero downtime, ensuring data integrity while effectively addressing the demands of customer's growing user base. This migration optimized performance by 60% and bring down costs to nearly 90%. Achieved notable improvements in scalability, latency, resource utilization, cost-efficiency, and enhanced monitoring. Impressively, the migration was completed in just six weeks - well ahead of the ten-week timeline - demonstrating both efficiency and enhanced system reliability.

Challenges

- Limitation in Fine Tuning
- Vendor Lock-in
- Complex Migration
- High Maintenance Cost
- Monitoring Limitation
- Lack of Visibility in Credit Consumption
- Prolonged On-boarding

Potential

- Optimized Infrastructure
- Seamless Transition
- Better Uptime & Stability
- Improved Scalability
- Enhanced Flexibility
- Optimized Utilization of Resources
- Improved Performance
- Cost Efficiency

CHALLENGES IN ASTRADB

- **Limitation in Fine Tuning:**
  Fully managed service simplifies infrastructure management but restricts deep customization for performance tuning or network configurations, which can be a drawback for organizations needing fine control.

- **Vendor Lock-in:**
  Using AstraDB can lead to vendor lock-in, making it challenging to switch to another platform due to dependencies on its specific features and APIs, increasing the complexity of future migrations.

- **Complex Migration:**
  Migrating on-premises AstraDB clusters, requires careful planning and can be complex, especially for applications with unique configurations, legacy systems, or specific performance needs.

- **High Maintenance Cost:**
  Serverless architecture can lead to higher costs for small-scale applications or workloads that don't require the extensive scalability or redundancy. Cost can also vary for unpredictable highly variable traffics.

- **Monitoring Limitations:**
  While AstraDB offers built-in monitoring via DataStax Insights, the lack of detailed customization and granularity in metrics can make it harder for enterprises to troubleshoot and fine-tune database performance.

CHALLENGES

- **Lack of Visibility in Credit Consumption:**
  AstraDB uses a credit-based pricing model, where credits are bought for each dollar spent. However, the usage of these credits is not always clear, making it difficult for users to track consumption accurately.

- **Prolonged Onboarding:**
  Despite being a managed service, setting up advanced features like security, multi-region replication, and performance tuning requires a deep understanding of AstraDB, often extending the onboarding process.

POTENTIAL IN SCYLLADB

- **Optimized Infrastructure:**
  ScyllaDB's shard-per-core architecture provides better resource utilization and performance efficiency.

- **Seamless Transition:**
  Zero-downtime migration can be achieved using ZDM Proxy, ensuring no service interruption for end users.

- **Better Uptime & Stability:**
  Low-latency architecture ensures consistent and reliable performance even during high traffic spikes.

- **Improved Scalability:**
  Architecture scales easily, providing better support for Dailyhunt's future growth.

- **Enhanced Flexibility:**
  Giving more control over configuration, performance tuning, and monitoring, addressing AstraDB's limitations.

- **Optimized Utilization of Resources:**
  ScyllaDB efficiently uses hardware, reducing the need for extra computational resources.

- **Improved Performance:**
  Delivering a 60% improvement in read latencies and faster write operations, reducing the 95th percentile write latency from 6ms to 2ms.

- **Cost Efficiency:**
  By switching to ScyllaDB's self-hosted model, Dailyhunt can reduce infrastructure costs by 90%, gaining more control over their expenditure.

AN OVERVIEW

of Customer's Current Environment

A brief insight into the customer's on-premises infrastructure, showcasing their workload distribution, hardware composition, and network architecture across a complex and scalable environment.

WORKLOAD DISTRIBUTION

- Struggling to manage high-volume, fluctuating traffic with billions of daily requests.
- Read/write latencies became bottlenecks as Dailyhunt scaled.
- Serverless pricing model resulted in high costs due to unpredictable traffic spikes.

NETWORK ARCHITECTURE

- Global distribution enabled content delivery across 13 languages.
- Cross-region replication added to operational costs.
- Lack of control over network configurations affected traffic flow optimization.

ScyllaDB - An Introduction

An open-source NoSQL database, capable of delivering high performance results in mission-critical environment. It retains the benefits of other major platforms in the industry, while eliminating the performance bottlenecks associated with Java. ScyllaDB delivers unmatched performance at scale using dramatically fewer while maintaining lower latency, reduced complexity, and lower infrastructure costs.

Why Choose ScyllaDB?

- ScyllaDB stands out due to its consistent low-latency, delivering P99 latencies in single-digit milliseconds even under heavy loads.
- Its high throughput per node and self-optimizing architecture reduce operational complexity, making it simpler and more cost-effective than alternatives like AstraDB, Cassandra etc.
- Ideal choice for data-intensive applications, offering a seamless migration, while cutting operational costs and infrastructure requirements significantly.
- Comprehensive Monitoring stacks enables more control over database granular metrics, allowing for proactive issue resolution and enhanced security management.
- Choose ScyllaDB to scale efficiently, lower total cost of ownership, and achieve high performance without the need for complex tuning or additional caching layers.

Key Features

- Designed for Data-Intensive Applications
- Shard-per-Core
- NUMA Optimized
- I/O Scheduler
- Self-Optimizing
- Full Compatible Ecosystem
- Comprehensive Monitoring Stack
- Low Total Cost of Ownership

ScyllaDB

OUR APPROACH

To address Dailyhunt's challenges and unlock the potential benefits of the cloud, we took the following structured approach:

PREPARATION AND ASSESSMENT

We started with a thorough evaluation of the existing infrastructure to assess hardware, software, and network configurations. This helped us identify key requirements, potential challenges, and the most suitable Google Cloud solutions.

PRE-MIGRATION TESTING

Before starting the migration, a comprehensive Proof of Concept (PoC) was conducted to validate the feasibility of using ScyllaDB. This gave us the confidence that it would meet our performance and scalability needs.

- Schema and Data Migration:
  Migrating from the existing database to ScyllaDB while ensuring data integrity was initially complex. However, after thorough testing and validation, we were able to confirm seamless schema compatibility.

- Performance Testing:
  ScyllaDB's performance was tested under heavy workloads. Although we faced some initial latency issues with configuration tuning, ScyllaDB's shard-per-core architecture eventually showed superior performance and scalability compared to the existing database.

- Application Integration:
  Some adjustments were needed to ensure that our applications worked smoothly with ScyllaDB. After resolving compatibility challenges, we observed improved query response times and smoother application workflows.

- Cluster Setup and Network Sync:
  Ensuring proper connectivity and synchronization between nodes and regions posed some challenges. Once resolved, ScyllaDB demonstrated exceptional ability to handle distributed clusters with minimal latency.

MIGRATION STRATEGY

- Zero Down-time Migration
  Downtime was avoided during the migration from AstraDB to ScyllaDB by implementing a Zero Downtime Migration (ZDM) strategy. We deployed the ZDM proxy using Ansible, which facilitated a seamless transition while keeping Dailyhunt's services online without disruption. Here's how downtime was avoided:

ZDM Proxies:

The ZDM Proxies act as intermediaries between the client applications and both the origin database (AstraDB) and the target database (ScyllaDB). These proxies handle data read and write operations and ensure the system continues to function during the migration. Specifically, they ensured:

Dual Writes:

During the migration, every write request from the application was directed to both AstraDB and ScyllaDB. This guaranteed that any new data was written to both databases, keeping them in sync.

Intelligent Reads:

The proxies intelligently routed read operations to AstraDB at first, and then gradually switched to ScyllaDB once data was migrated and verified.

- Phased Migration
  The migration process was broken down into four phases, ensuring a smooth transition:

Phase 1: ZDM Proxy Setup:

The ZDM Proxy was deployed between the client applications and AstraDB, and client traffic was routed through the proxy. This allowed for traffic interception and routing, which ensured that all reads/writes were properly handled during the migration.

Phase 2: Data Migration:

Tools like the Cassandra Data Migrator (CDM) were used to migrate existing data from AstraDB to ScyllaDB. Since new write operations were being sent to both databases, there was no risk of data inconsistency.

Phase 3: Read Transition:

After verifying the migrated data, read operations were gradually shifted from AstraDB to ScyllaDB. At this point, the ZDM proxies ensured that all reads and writes to ScyllaDB were functioning as expected.

Phase 4: Full Cutover:

After ensuring all data was successfully migrated and tested, the ZDM proxies were bypassed, and client applications directly connected to ScyllaDB, completing the migration without any downtime.

REAL-TIME MONITORING & VALIDATION

During the entire migration process, Prometheus and Grafana were used to monitor both AstraDB and ScyllaDB in real-time. This allowed the team to track any latency issues, verify data consistency, and quickly resolve any performance bottlenecks before fully switching over to ScyllaDB.

THE RESULT

The migration resulted in several measurable improvements for Dailyhunt.

60% Improvement in Latency:

ScyllaDB's latency improvements resulted in faster database operations, reducing read and write latencies by 60%.

90% Cost Reduction:

Migrating to a self-hosted ScyllaDB model allowed Dailyhunt to cut operational costs by 90%, while maintaining high availability.

Seamless Scalability:

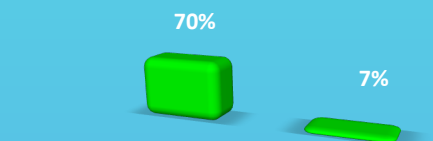
ScyllaDB's architecture allows Dailyhunt to scale their database effortlessly without compromising performance.

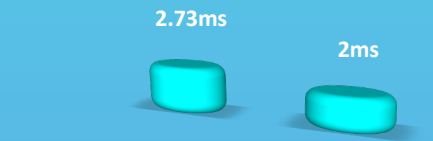
Enhanced User Experience:

End users benefited from a faster, more responsive experience due to reduced latencies and improved uptime.

Robust Security and Monitoring:

With ScyllaDB's detailed monitoring stack, Dailyhunt gained more control over database metrics, allowing for proactive issue resolution and enhanced security management.





AstraDB Performance

ScyllaDB Performance

WHY PARTNER WITH US?

1

Certified and Experienced Team

With over 15 years of experience, our seasoned database professionals have worked extensively with MySQL, PostgreSQL, RocksDB, Redis, TiDB, MongoDB, Aerospike and more. They are supported by a highly skilled team of over 50 certified experts across cloud, networking, security, and DevOps, ensuring specialized, high-quality solutions tailored to your project's needs.

2

Scalable, Proven Solutions

We've successfully handled large-scale deployments supporting millions of users and transform legacy applications into cloud-native architectures, enhancing agility and reducing costs through automated autoscaling and resource optimization.

3

Customer-Centric, Problem-Solving Approach

We start by deeply understanding your business, data flow, and current infrastructure. This includes assessing software, third-party integrations, and licensing constraints, identifying potential issues during migration, and providing tailored solutions. We ensure a seamless transition by aligning our technology recommendations with your specific business needs.

4

Future-Proof Planning

Our approach includes precise capacity planning based on your growth projections. For start-ups and established businesses alike, we propose scalable solutions that avoid unnecessary long-term commitments. Our migration strategy guarantees zero downtime, no data loss, and protects your brand's reputation.

5

Infrastructure as a Code (IaaC)

By utilizing Infrastructure as Code (IaaC), we ensure your infrastructure is not only quickly deployable and consistent but also resilient. When incident strikes, IaaC enables rapid recovery, reducing downtime from weeks to just minutes, ensuring seamless continuity and efficiency.

ANALYSE THE DATA IN REAL-TIME
REAL TIME IS REAL VALUE

”

ABOUT US

Since Our Inception In 2018, We have consistently been at the forefront of technology and innovation, serving prominent clients like Dailyhunt, Josh, PublicVibe, and several respected organizations from the Middle East. Our technological journey has evolved from public to private cloud solutions, establishing a hybrid cloud infrastructure in 2015, and further advancing to a poly-hybrid cloud by 2019. We collaborate with all major hyperscaler's and excel in various hypervisors such as VMware, Nutanix, and KVM. Backed by a robust team of 70 professionals across DevOps, Databases, Networks, CDN, Security, and Sales, we are well-positioned to make a significant impact in the industry.

WHAT ARE WE CAPABLE OF?

100+ TB DATABASE	**50+ TB** DAILY LOGS	**<3MS** RESP. TIME	**<1,5MS** LATENCY	**100000** PODS
10000 END POINTS	**700Bn** TRAFFIC	**50 PB** BANDWIDTH	**300+ Mn** MAU	**1000K** RPS

WHAT VerSe Innovation
SAYS ABOUT US

> Partnering with QuarkWiz proved to be benefcial for us in every way – right from achieving and tracking day-to-day deliverables, to complying with industry best practices. Really appreciate Quark bringing Microsoft to the table. The Google Cloud platform, as provisioned by, and the solution architecture are poised to make a significant impact on our customer service needs. Thanks to the team for its hard work, dedication and enthusiasm.
>
> CTO
> Verse Innovation Private Limited

Ready to elevate your business ?
Schedule a free Consultation

www.quarkwiz.com

Helios Business Park | Bengaluru
+91 80925 30599 | sales@quarkwiz.com