

WORLDWIDE MEDIA OUTLET TURNS TO PYTHIAN TO OPTIMIZE CASSANDRA FOR COMPILATION OF READER DATA

CLIENT

Daily newspaper

INDUSTRY

Financial news

TECHNOLOGIES

Cassandra

BUSINESS NEED

Capture and access large amounts of data to develop usage profiles in order to understand member interests and to grow advertising revenue.

TECHNICAL CHALLENGE

Bring new servers online in a geographically dispersed infrastructure and improve query performance.

SOLUTION

Pythian's Cassandra experts improved network throttling for easy scalability.

RESULTS

A scalable, resilient, distributed database infrastructure that quickly and efficiently processes data requests:

- Optimized system performance
- Easy scalability
- Enhanced operations

BUSINESS CHALLENGE

One of the world's leading business news organizations set a new standard for news delivery when they implemented a multi-tiered subscription program for web-based access to its news material. With 4.5 million registered users and over 600,000 paying subscribers, the organization relies on its database infrastructure to develop in-depth usage profiles of its membership to understand key interests and to grow advertising revenue.

With business growth so tightly coupled to the performance of its database infrastructure, the organization chose Cassandra because of its strength in supporting geographically dispersed infrastructures and for its ability to scale. They were using in-house resources to design and implement the system, but soon ran into a number of performance issues and realized they lacked the internal knowledge to fully troubleshoot and resolve them.

Cassandra often requires manual effort to ensure data is synchronized across all servers. In addition to performing manual repair operations on a weekly basis, the organization was also having trouble bringing new servers online because data was not being totally transferred to the new servers – in some cases less than 10%.

Compounding the problem were data models that did not properly distribute query requests across all the servers. Instead of the query being routed using the shortest path to the proper server, the query would be sent on a roundabout path that touched almost every server in the network before reaching its intended destination. Performance was severely degraded, and the infrastructure was not able to scale.

SOLUTION

The organization turned to Pythian who could draw upon broad experiences with other Cassandra clients to help troubleshoot and resolve the root cause of the recurring problems.

The first priority was to determine why the new servers were not coming online correctly, and why the replication mechanism wasn't working properly. Cassandra is very sensitive to cloud-based hardware configurations, and spinning up new servers and manual repair processes are extremely taxing on the hardware and network links. After a comprehensive "deep dive" analysis, Pythian determined that certain network conditions caused interruptions within the hardware stack.

Pythian recommended a software upgrade and some changes to the configuration to improve network throttling. Once performance issues were stabilized, Pythian then examined the data modeling architecture and proposed a series of recommendations to fine-tune and optimize system availability and performance.

RESULTS

Recovery from a server failure in Cassandra is extremely difficult, more so than in other environments. Not only did Pythian fix the performance issues, but also provided guidelines to avoid issues in the future and help this client get the most out of its Cassandra infrastructure. A healthy, finely tuned Cassandra environment scales easily even while operating at peak levels by balancing usage across all assets.

"We brought in the Pythian Cassandra experts to help us resolve some very complex performance issues. They worked extremely well with our in-house resources, customizing their support to fit our specific needs rather than forcing a pre-set process on our organization. Their flexibility and broad skill set was beyond our expectations and we learned things about Cassandra and our own operations that will guide us in our future efforts."

ABOUT PYTHIAN

Pythian is a global leader in data consulting and managed services. Since 1997, we have specialized in planning, deploying, and managing business-critical data systems for large and mid-market enterprises. Learn more about Pythian and its elite teams of data experts at www.pythian.com.

OFFICES

New York City, USA
London, England

Sydney, Australia
San Francisco, USA

Hyderabad, India
Ottawa, Canada