

Top 10 Global Bank Transforms Big Data Infrastructure



A major global bank was grappling with the challenges of modernizing its extensive Cloudera Hadoop infrastructure. With nearly 100 Hadoop clusters spread across its operations, the bank faced significant issues with data silos, inefficient resource utilization, and mounting costs. They required a solution that could consolidate their big data environment while enabling a transition to more modern technologies like Apache Spark, Trino, and other advanced execution engines.

Challenges:

1

Complex and expensive Cloudera/Ozone-based data lake

4

Unable to support downstream AI processing

2

Poor utilization of assets and high licensing/hardware costs

5

Desire to move to modern data pipelines using Trino/Spark and Kafka

3

Multiple siloed data lakes hinder comprehensive data analysis

VAST Data Solution

The global bank chose the VAST Data Platform to consolidate and modernize its big data infrastructure. Key features of the solution included:

1

High-performance S3 platform for consolidating multiple data lakes

4

Resilient system with 99.999%+ uptime

2

Linear scale-out capabilities for both performance and capacity

5

Multi-protocol access to enhance pipeline workflows

3

Built-in data reduction for improved economics and data center space savings

6

Built-in replication for disaster recovery

7

AI/GPU-ready platform to support future initiatives

Implementation

The bank deployed tens of petabytes on VAST Data across three clusters in the UK: production, disaster recovery, and a test environment.

Outcomes and Benefits

1

Performance Boost

3x improvement in overall performance, enabling faster data processing and analytics

5

Capacity Optimization

80% reduced required capacity due to effective data reduction, significantly lowering data platform costs

2

Data Center Efficiency

10x reduction in data center space, leading to substantial infrastructure and energy cost savings

6

Resource Utilization

4x improvement in asset utilization, maximizing the return on investment for the bank's infrastructure

3

Cost Savings

Significant decreases in licensing and hardware costs, improving the overall economics of the bank's big data operations

7

Simplified Management

Consolidate multiple data lakes into a single, manageable platform, reducing operational complexity and overhead

4

Future-Ready Infrastructure

Ability to support AI and GPU workloads for upcoming projects, positioning the bank for advanced analytics and machine learning initiatives

About VAST

VAST Data is the data platform software company for the AI era. Accelerating time-to-insight for workload-intensive applications, the VAST Data Platform delivers scalable performance, radically simple data management and enhanced productivity. Launched in 2019, VAST is the fastest-growing data infrastructure company in history.

Looking to simplify your data center and unlock insights from all of your data? **Contact us at** hello@vastdata.com



[Scan Here](#) learn more about how VAST and Dremio are accelerating and streamlining modern analytics at enterprise scale.

Business Impact

The global bank has successfully transformed its big data infrastructure, setting the stage for more efficient operations and advanced analytics capabilities. The VAST Data Platform has not only solved their immediate challenges but also positioned them to leverage cutting-edge technologies in the future, including AI and machine learning applications.

The consolidation of data lakes and improved performance have enabled the bank to streamline its data pipelines, leading to faster insights and more agile decision-making processes. The significant reduction in data center space and improved resource utilization have also contributed to the bank's sustainability efforts.