

CLOUDERA HADOOP FOR ETL

Modern ETL Pipeline

Overview

This is a case study of one of the clients of Pyramid Solutions who migrated from their existing EDW Ecosystem to Hadoop. This case study describes the benefits of migrating to Hadoop, Challenges with existing Ecosystem, Solution for the problem, Migration Methodology, and the security on Hadoop.

Business Objective

- ▶ To reduce Capex and Opex costs.
- ▶ To efficiently utilize resources with ETL.
- ▶ To improve performance and reduce data processing cycle times.
- ▶ To achieve fault tolerance.

Existing system and its Challenges

- ▶ Inefficient resource utilization with ETL.
- ▶ Difficult to process a large volume of data.
- ▶ Incapable of storing and processing unstructured data.
- ▶ Wasted Storage with only ~20% of data in real-time.
- ▶ License Cost and upgrade are expensive.
- ▶ Fault tolerance is not pre-built.

Business Solution

- ▶ Evaluate multiple tools for data warehouse including Hadoop.
- ▶ Develop an end-to-end solution to migrate ETL process to Hadoop.
- ▶ Migrate Data warehouse to Hadoop.
- ▶ Achieve fault tolerance.
- ▶ Reduce storage cost and meet the compliances around online data access.

About Pyramid

Pyramid Consulting Inc. is a global technology and process driven software solutions company offering customer centric solutions. With knowledge and experience of the entire IT lifecycle, we help enterprises streamline core IT processes and augment their competitive advantage. Pyramid's global delivery model harnesses productive tools, agile methodologies, iterative processes and expert frameworks to provide innovative and cost-effective solutions.

Our domain and industry experts engineer robust technology solutions for clients in a wide range of industries including retail, logistics, healthcare, insurance, software development, financial services, media, publishing and telecommunications.

Global Headquarter

Pyramid Consulting, Inc. 11100 Atlantis Place Alpharetta, GA 30022, USA
Phone: 678.514.3500 Toll-free: 877.248.0024

USA

Georgia
New York/New Jersey
North Carolina
Virginia
Texas

Canada

Ontario

Europe

United Kingdom

India

Noida
Hyderabad

Asia Pac

Singapore

Business Benefits

- ▶ ETL Offloading - Offloads data processing workloads to Hadoop to improve performance and reduce data processing cycle times.
- ▶ Augmented data warehouses - Offload high volume storage and processing to Hadoop and deliver ready-to-consume results to a traditional store.
- ▶ Big data warehouses - Built on Hadoop and enable data consumption.
- ▶ Data archival - Data is archived on Hadoop to reduce storage cost and meets the compliances around online data access.
- ▶ DR and Fault Tolerance - Organizations achieve fault tolerance and DR due to Hadoop's distributed nature.

Conclusion

- ▶ Processed a large volume of data.
- ▶ Reduced storage cost.
- ▶ Met the compliances around online data access.

**pyramid**
CONSULTING
SOLUTIONS

For more information:
Visit us at www.solutions.pyramidci.com or
email us at marketing@pyramidci.com

