Fortune 1000 Company: UST Global Proposes Migration of Data Warehouse to Cloud with Dynamic Scaling to Achieve Better Availability and Cost Management

Opportunity
Migrating data warehouse to the cloud

Action
Migrating from Oracle to AWS cloud

Impact
Reduced costs and superior scalability for competitive advantage

Global leader gains flexibility and high efficiencies at reduced costs

Organizations face multiple challenges in data migration and access when embarking on a cloud journey. Moving data warehouses (DW) to the cloud without interrupting business operations, ensuring timely and quality data flow for business users, and integrating multiple applications are some of the critical issues that need to be addressed during the journey.

UST Global's proposal enabled a Fortune 1000 company to significantly reduce their licensing costs of ETL and DW platforms and achieve dynamic scaling of infrastructure requirements, while ensuring meticulous adherence to compliance standards.

Our proposed solution enabled a highly scalable cloud platform and architecture to provide agility and flexibility for scaling data loads. We created the ability to access and use a range of open source tools that reduced their license costs on expensive proprietary software. Our robust solution could pave the way for high availability, fault tolerance and resilience to lend both competitive edge and high return on investment.
Opportunity: Migrating data warehouse to the cloud

The Fortune 1000 client was faced with a critical need to reduce their burgeoning license costs of ETL and DW platforms. They were also looking to dynamically enhance the ability of adding new data sources to the DW and optimize maintenance of ETL jobs, while maintaining strict adherence to SLAs.

The client turned to UST Global based on their extensive expertise in cloud assessment and strategy solutions.

Action: Migrating from Oracle to AWS cloud

After careful evaluation of requirements and various options, we selected Amazon Web Services (AWS) as the cloud platform. Amazon Redshift, a fast and fully managed petabyte-scale data warehouse with its massively parallel processing (MPP) architecture, was deployed to analyze data using existing business intelligence tools. The solution implementation involved the following steps:

- Data extraction from multiple sources to AWS S3 buckets - using an Open Source ETL Tool - Pentaho.
- Data cleansing, ETL steps and application of business logic - using AWS Data Pipeline Service to trigger AWS EMR jobs. For sources that do not support change data capture, a full extract with de-duplication logic was applied in EMR jobs.
- Data loading to Amazon Redshift – after transforming data from all source systems into an intermediary generic data structure.

Impact: Reduced costs and superior scalability for competitive advantage

Our proposal had the following benefits:

- Reduced costs and flexibility: Lowering of licensing costs on expensive proprietary software due to extensive use of open source tools. Amazon Redshift enables starting on a small-scale and subsequently expanding to multi-node clusters based on demand, to further optimize costs.
- Scalability: AWS EMR allows concurrent processing of large data volumes, increasing the throughput of the data load as required. The domain model allows the organization to add new source systems with minimal effort and cost.
- Resilience: High availability and fault tolerance are built into AWS EMR, Redshift and S3 for superior performance.