

ETL, DW & BI solution for AES Corporation - a leading US based energy distribution company



BUSINESS NEED

- Global generation of energy across various countries. Needs visibility on a real-time basis.
- Large number of transactions each day.
- There were times when a heavy load on the database caused it to crash, hence abruptly terminating the data loading and fetching process.



ENGAGEMENT

- Experts from Sigma worked closely with the client organization to design an Enterprise Business Intelligence Reporting Solution which enables real-time information to make data-driven decisions.
- The solution removed unnecessary caching problems in Pentaho Data Integration, optimized the database and mondrian application for better performance.
- Extracted data from JIRA & Google APIs, maintained heavy datasets of ~300 gigs.
- # of ETL Jobs written: 200+ (including Verification and Audit ETLs)
- # of Data sources: 40
- Data extracted from:
 - ✓ ISO sites – Url based file downloads, Rest APIs, Scrapping through XPath and Selenium scripts
 - ✓ JIRA
 - ✓ Google Spreadsheets



SOLUTION

- Developed charts, management dashboards, automatic data loading scripts, JIRA Reports, Jobs to fetch JIRA data etc.
- Operational Reports and KPIs for various projects broadcasted on a monthly basis to all the major stakeholders of respective projects.
- Generate OLAP structures (Cubes and fact tables) for client database
- Aggregate the data for various market places, ISOs and services
- Developed Analytical reports and dashboards for market pricing for Ancillary and Energy prices
- Developed the OLAP DB model, ETL and Analytical reports for Project/Fleet data Optimization of reporting and Analytical reports by improvising the performance across the different layers like DB optimization with buffering, query caching, partitioning, indexing, etc. Pentaho Mondrian cache optimization with Smart Refresh Integration with 3rd party products or components



TECHNOLOGY

- Pentaho BI suite
- PDI
- Mondrian
- JIRA
- Crowd SSO
- Amazon RDS – MySQL
- Selenium
- Google Docs
- AWS