compassîtes®
digital transformation experts

Case Study

Beverages

Client Profile:

The client is division of the world's largest beverage company. Globally, through the world's largest beverage distribution system, consumers in more than 200 countries enjoy their sparkling beverages, ready-to-drink coffees, juices, juice drinks and beverages at the rate of 1.8 billion servings a day!

Tecnology Used:

Adobe Flex, Adobe Life Cycle Data Services, J2EE, MySQL & Pentaho

Business Situation:

The client sources the market research data from one of the largest market research companies, which provides around 34 trillion data points of relevant market research data, which was made available to client using their proprietary tool. The market research department of client would then work through this data for about two months to manually create a 220 tabbed excel report, after doing analytics and calculations on raw data manually, using tools provided by one of the largest market research companies, and using other tools, to be used by marketers and brand managers to make rightful business decisions.

This was a very inefficient approach since there was loss of valuable time until the report was published. Most of the process was manual; giving scope to human errors. The reports were shared through emails, physical documents creating lack of flexibility and easy availability. Creating new reports was extremely time consuming and an expensive process. The client has to source some of the data from one of the largest market research companies, which usually turned out to be expensive.

Solution Approach:

We got the data from the market research company in the form of raw database files, and processed this data, to generate a macro level database in the form of OLAP cubes, to be served to end users through a web based, advanced data visualization platform called iCharts. Since this was a huge data set that needs to be processed and analyzed, the data was processed in parallel through a MapReduce implementation.

The highly complex, proprietary format raw data files provided by the research companies firm , amounted to around 130GB! We had to deal with the data at the SKU level to get some of the KPIs required by client. The reports to be generated needed at highly complex groupings of products and markets. The technical solution designed was as shown below. The complex data is decoded and loaded by custom written decoders into the tables in the MySQL database in the form of readable raw data. MapReduce was used to split the data into multiple EC2

instances for massive parallel processing, with each EC2 instance executing a pipeline of Linux and MySQL processes on the raw data to reduce it to macro level data, suitable for analytical processing and load it into the star schema. This data is then loaded into OLAP cubes for analytical processing.

Benefits & Results:

¤ Fully automatic process preventing any human errors.

¤ Quick turnaround of reports and the monthly data is now available to managers within 3 days.

¤ Highly organized dashboard, with advanced data visualization capabilities like filtering.

¤ Access controlled dashboard to give selective access to different users.

¤ Individual reports can be downloaded, to be shared with management.