Case Study: Financial Institutions handling 10.5M Customers' data

Introduction

In 2010, A Finance startup needed more flexible and more integrated process to provide the targeted new product and service offerings that their customer seeks. They have vast amounts data of customers' enrolling patterns, transactions, reviews, feedback and social media data.

They also developed a Statement Rewards platform, used by its partner financial institutions with a 10.5M subscriber base. The platform captures credit card transactional data, that is needed to be processed in batch for analytics and business intelligence, to increase loyalty, card spend and top-of-wallet usage by delivering data driven personalized services.

Requirement

- Scalability Challenge: Credit card transactional data is logged in to both traditional RDBMS systems such as MySQL and event logging via restful API.
- Every day there are 20 - 40M events generated and loaded to MongoDB. The data is loaded in to MongoDB in batch every hour in to a stage table.
- After 1 month the MongoDB transaction volume has grown over 1 billion records in stage and the subsequent processing of this stage data in to other collections has started to take its toll on MongoDB.
- Replication being slow, scaling of this DB was not so mature and had many issues with MongoDB V1.x. This has become unmanageable and technically inoperable without a scalable solution.

The other key challenge for them was to figure out

- How to use this massive data for dynamic customer segmentation to aid targeted marketing campaigns and roll out personalized product and service offerings to customers.
- Predict the set of existing customers who will make a new enrollment and the time of the year when the new enrollment may happen
- The banking product they most likely will buy

In addition they needed to know what channels the customer listens to, who are the key influencers in his decision making process, what key life events are happening in his life etc through which the marketing campaigns can be effectively targeted and marketing messages crafted for the particular customer in mind. Understanding the customer's influence quotient also helps in acquiring new leads through his referrals and determining incentive schemes to be given to the customer for referrals.

Solution Implemented

Internet data, customer transaction data, social media data etc are the sources of data were looked at analyzing and developing predictions to ascertain the proclivity of customer to purchase new products, when he would likely buy and who would be key decision makers in the decision to buy. Analyze and deliver predictions using Corpus Big Data Analytics on this huge volume of data pertaining to the customer and develop a purchase profile for the customer along with targeted marketing statements to be used in pitching to the customer at the appropriate time in the right context.

- Tools like SQoop and Flume were used to extract data from MySQL and event API interface on to AWS (S3).
- Using Hadoop (HDFS) cluster loaded data from S3 on to hourly cluster and mapped on to Hive staging warehouse tables.
- The data is then transformed and modeled in to facts and dimensional tables in Hive.
- Data was aggregated using custom MapReduce jobs and PIG scripts.
- The transformed and aggregated data is then loaded into MongoDB as well as on to S3.
- This eliminated the raw data being loaded in to MongoDB, all the processing overhead was pushed on to Hadoop. MongoDB ended up serving only the aggregated and summary results to reporting interface.

Benefit

To gain a full understanding of customers and create the exceptional products and service experience that foster loyalty and long-term relationships, they needed information that is timely, reliable, and robust. With Corpus Big Data Analytics they got the information they needed to make real-time or near-real-time business decisions, minimize customer turnover, detect fraud, and prevent money laundering.

1. Helps make each Customer relationship as profitable as possible
2. Understand Customer Product and Service Preferences and Habits.
3. Customized Offerings to segmented customers
4. Targeted Marketing Campaign to segmented customer
5. Creating Holistic view of Customers for Timely and Innovative Offerings.
6. Prediction of Outcome – Success
7. Focus on profitable customers,
8. Increase customer satisfaction and customer loyalty,
9. Reduce marketing costs and
10. Target product offerings for high margin customer segments.
11. Corpus Big Data Analytics combined with other banking solutions empowers marketing business users to maximize profits with highly tailored product offerings.

About Corpus:

Corpus Software is one of the faster growing IT solution and services company focused on Digital Media Entertainment, Embedded technology and Business Analytics with offices and partners across Americas, Europe, APAC, Middle east & Africa. We work with clients in most emerging technology, that's where we make their business strong and bring in real difference in the way peer operates. A diverse workplace with continues focused towards developing unique ideas and contributions to make our clients business grow, and to keep the momentum going.

Offices: Dallas, London, Singapore, Johannesburg, Hyderabad and Bangalore