Making the right message reach the right customer at the right time: A real case of efficient Data Processing by Indium

Application: Mobile Engagement Platform

Services Offered: Big Data Processing

Background

Customer engagement is a candid future-proof strategy for any business. The importance of captivating customers with the right messaging never goes out of style to build a brand and therefore drive revenue

Client

Our client is one of those pioneers who has developed a Mobile Engagement Platform that enables enterprises to drive their marketing outreach through mobile messaging technology. They are an aggregator in the US with direct connectivity to all major wireless carriers with a best-in-class campaign management platform.

Key Highlights

Key Success:
Manual effort reduced by 50%

Domain:
Telecommunications

Duration:
12 months (ongoing)

BigData Team:
2 Big Data Engineers

1 Senior Big Data Archiect

Technology:
Hadoop Ecosystem, Oozie, Sqoop, Hive, HDFS, HBase, Phoenix

Application Overview

The platform powers mass mobile interactions and time-sensitive campaigns through SMS, authentication codes, notifications, promotional messages and much more. The platform permits billions of daily and monthly mobile SMS transactions to target audience

The average number of message transactions in the platform that happen per day goes up to 150 million. The aggregate messages traded are stored in the data base and the status log of those messages that actually reached/ failed to reach the intended user are mapped against each message. Data thus stored goes through a refined ETL process to derive performance and integrity reports.

Typically, businesses sharpen their marketing efforts in view of getting their SMS as close as possible to actual interaction with their end users by defining various submission-delivery SLAs : message scheduled delivery at a strategic time, defining validity of a message (OTPs for instance) and so on.

In view of this, the reach and responsiveness of the messages (marketing communications) needs to be carried out at lightning pace without giving up on the control of the engagement. There was a pressing need to have flexible room for data storage and processing.

Business Situation

The existing architecture of the client used

 # MySQL (RDBMS) Server for storing messages. This limited insertion rate leading to an IO bottleneck. The system needed a Horizontal scale up i.e. adding more hardware and also configure time sensitive features to meet marketing SLAs.

In spite of handling key functions: Disaster Recovery, Back up and Reporting, the data container was stagnant beyond 100 million capacity leading to process inefficiencies

 # Pentaho Server for ETL Process required heavy IO and computation

ETL process took an average of 11 hours and observed spikes up to 15 hours during promotional days/ special campaigns. The system significantly slowed down the output productivity and efficiency

 # PostgreSQL Server for Reporting and walling Aggregated Logs

Business users were unable to access the reporting data in real time.

Aggregator logs contained only recent 2 month data for reporting; any data requests beyond that would need special data access requests that would typically take 24-48 hours to retrieve

Our Solutions

Indium analyzed the platform from functional and operational perspectives and defined the following goals to be met to make the current system a robust one.

• Lowest Latency

• Superior responsiveness

• Powerful Integration

• Cost effective architecture

To achieve the above, Indium proposed a phased approach to implement a potential Data Processing Solution.

Our Solutions

1. Enhance the ETL Process Solution: Enforcing a centralized data system

The ETL function streamlines aggregated data to MySQL and logs to PostgreSQL. These batch jobs were run on an expensive Pentaho architecture. We migrated the set up to a powerful Hadoop ecosystem. Data is centralized in the Big Data containers offering high flexibility, highly scalability, fault tolerance benefits and cost advantage.

2. Merge log database into Big Data ecosystem

Log insertions into PostgreSQL were tedious taking up about 4-5 man hours. The logs were also stringent on access to end users for viewing the daily/weekly and historic reports.

Indium's hindsight on this pain point resulted in migrating the log database into the Hadoop clusters. This enabled the system with distributed processing and effort would be minimal. The data required for real time reporting was easily generated using Hive tables.

3. Remove the Data Insertion bottleneck by introducing direct inserts into Hbase

There were significant delays in data insertions into the SQL database. The direct consequence of this bottleneck was that the downstream processes were affected i.e. the business users were unable to view real time message delivery status.

Indium's architects introduced HBase on top of HDFS and Phoenix (SQL on Hadoop). This loads real time data into HBase and accessing reports real time is effortless.

While removing the bottlenecks in the process, Indium's clever choices of updated technology and approach to migrate the legacy architecture into a stabilized one, largely influenced the marketing efficiencies of the client. Our Big Data expertise helped crunch the Data processing times from 3days to 3hours!