

BIG DATA INFRASTRUCTURE MANAGEMENT

CASE STUDY

CHALLENGES

A fast-growing Advertising Technology start-up that sells Social Ad Campaign management software wanted its in-house Engineering team to focus on its core product and outsource certain areas of its non-core technology needs. The non-core portion of its technology stack required cutting edge computing skills and entailed creating a Big Data Analytics infrastructure hosted on the Amazon Cloud.

BUSINESS REQUIREMENT

- Design and implement a scalable Crawler to aggregate social data to meet demanding latency requirements.
- Reduce cost for designing and maintaining a highly available and scalable Hadoop based Big Data infrastructure
- Design a robust analytics and predictive modeling backend to meet evolving business initiatives
- Continue to enhance and upgrade the system to the latest stable version of the Hadoop distribution to leverage ongoing Improvement

SOLUTION OFFERED

Serendio was engaged to provide the Ad tech client with a host of Big Data services rendered through our offshore team.

We provided the following

- Virtual team comprised of Team lead/Architect along with 3 experienced developers reporting into the Client's own VP Engineering
- After evaluating multiple **Hadoop** distributions, chose **Hortonworks** distribution as the appropriate platform to build on.
- Improved the scalability of the existing system by re-architecting the Social Data Aggregation and Parsing module through fine-grained **Map-Reduce programming**. The system scaled up well both in velocity and volume.
- **Flume** was architected into the system to improve streaming data aggregation performance and reliability
- Since the business needs mandated flexible data analysis and predictive modeling, an analytics friendly data store including **HBase, MySql, and Amazon S3** was deployed
- An **SNMP** based Application-level Monitoring and Alert system was built to ensure optimal resource utilization and failover
- The entire system was deployed on a **300 node AWS** cluster with flexibility to scale up and down based on data processing volumes
- On-going Big Data thought leadership, 3rd party product evaluation and recommendation

BENEFITS

- Scope and Quality of deliverables increased through leveraging Serendio's expertise in Big Data infrastructure and Social content aggregation.
- Increased Data mining and Predictive modeling capabilities resulted in newer revenue streams
- Savings achieved through reduction in in-house staff and overhead costs

*Created New Revenue Streams
While Reducing Operational
Costs by 60%*

ABOUT SERENDIO

The name Serendio is derived from the expression Serendipitous Discovery.

Unearthing the not-so obvious and seemingly unrelated from data-structured and unstructured, is a hallmark of our technology and our name merely reinforces that.

Our Big Data Science solutions help in driving Decisions and Actions for a wide variety of businesses in Retail, Insurance, Media, Education, and Healthcare. Our mission is to help every Company transform itself into a Data-driven organization.

CONTACT SERENDIO

4677 Old Ironsides Drive
Suite 450
Santa Clara, CA 95054
United States
Phone: +1 408 496 9930
www.serendio.com
info@serendio.com

Big Data Partners

