Kafka Message
cluster for IoT workloads

The Client

The client is a world leader in advanced positioning solutions known for GPS, laser, optical, and inertial technologies and is renowned for integrating positioning technologies with application software, wireless communications, and services to provide complete commercial solutions. The client's integrated solutions allow customers to manage and analyze complex information faster and easier, making them more productive, efficient, and profitable.

The Problem

The client has many large multi-national customers who use their positioning solutions in the field. Each customer has many hundreds of sensors, IoT devices, and mobile endpoints which generate hundreds of thousands of messages per second on a continuous basis. Each customer also has their own unique requirements of utilizing these messages for business process management. Each customer has a variety of applications and databases that utilize information from these hundreds of endpoints. The client wanted to architect and build a common infrastructure that would be quick to deploy for various customers across various domains and still be scalable to meet the vast variety of requirements. The client was looking for guaranteed message delivery, scalability, durability, and repeatability.

Why AWS

AWS gives an option of flexible services which are designed to enable companies to build and deliver products more rapidly and reliably. This can be done using AWS and DevOps practices.

These services simplify the following:

- Simplify provisioning and managing infrastructure
- Deploying application code
- Automating software release processes
- Monitoring application and infrastructure performance

DevOps is the blend of social methods of insight, practices, and tools that builds an organization's ability to deliver applications and services at high velocity. This helps evolving and improving products at a faster pace than organizations using traditional software development and infrastructure management processes. This speed enables organizations to better serve their customers and compete more effectively in the market.

Why 8KMiles

8KMiles with it's unique combination of Cloud and Analytics expertise was the ideal partner.

8KMiles Differentiators

- **Incumbent Knowledge:** 8KMiles has been servicing various divisions of the client for their various cloud needs successfully and hence had an intimate knowledge of the client's business domain and organizational structure. This enabled 8KMiles to work efficiently and effectively as a partner to the client.
- **CloudEz:** 8KMiles takes a platform approach to every project on the cloud. This CloudEz platform approach always ensures that our solutions bring security, repeatability and the necessary compliance.
- **Analytics:** World class analytics leadership in handling big data, machine learning, prediction, and visualization. In addition to life sciences, we have helped many industries from cyber security to advertising adopt cutting edge analytics technologies.

The Solution

8KMiles recommended big data architecture for stream processing using Kafka on AWS. 8KMiles architected, implemented and is providing managed services for following system:

- Confluent – Kafka platform was configured for processing the messages
- Secure Amazon Virtual Private Cloud (VPC) was instantiated with the necessary capacity to handle the current data sources and scalability for future data sizes running into multiple TB. Entire platform is set on Amazon EC2 and Amazon EBS.
- Kafka Cluster was setup with Multi Availability zone inside Amazon region for High Availability
- Kafka is configured with Zookeeper for
  - Electing a controller. The controller is one of the brokers and is responsible for maintaining the leader/follower relationship for all the partitions. When a node shuts down, it is the controller that tells other replicas to become partition leaders to replace the partition leaders on the node that is going away. Zookeeper is used to elect a controller, make sure there is only one and elect a new one it if it crashes.

  - Cluster membership - to identify brokers that are alive and part of the cluster. This is also managed through ZooKeeper.
  - Topic configuration - identify which topics exist, how many partitions each has, where are the replicas, who is the preferred leader, what configuration overrides are set for each topic
- HortonWorks Ambari is used for Hadoop management, It helps
  - Make Hadoop management simpler by providing a consistent, secure platform for operational control. Ambari provides an intuitive Web UI as well as a robust REST API, which is particularly useful for automating cluster operations. With Ambari, Hadoop operators get the following core benefits:
  - Simplified Installation, Configuration and Management. Easily and efficiently create, manage and monitor clusters at scale. Takes the guesswork out of configuration with Smart Configs and Cluster Recommendations.
  - Centralized Security Setup. Reduce the complexity to administer and configure cluster security across the entire platform.
  - Full Visibility into Cluster Health. Ensure your cluster is healthy and available with a holistic approach to monitoring.

The entire solution was architected to handle around ~ 650K records per second. This was deemed sufficient by the client for most of the use cases and met all their scalability requirements. 8KMiles developed this solution on AWS and continues to manage it with an MSP relationship with the client.

Benefits

With 8KMiles architected and developed solution, the client was able to deploy this messaging as a service solution to their customer within a couple of months. 8KMiles also manages the big data cluster providing operational efficiency and value to the customer.

San Ramon, CA (HQ)

12647 Alcosta Boulevard,
Suite 450,San Ramon, CA 94583, USA
www.8kmiles.com
sales@8kmiles.com
Direct   : 925-270-4800
Toll-Free : 855-856-4537

Chennai, India

Srinivasa Towers
New No.5, Old No. 11,
Cenotaph Road,
Alwarpet,Chennai – 600 018
Phone : +91-44-6602-8000
Fax   : +91-44-4300-9049

Chicago, IL

1827 Walden Office Square
Suite #460
Schaumburg, IL 60173
Phone : 708-289-5111

Ontario, Canada

4 Robert Speck Parkway,
Suite 1500,
Mississauga, Ontario L4Z 1S1
Phone : 416-366-7762

Dallas, TX

17740 Preston Road
Suite #200
Dallas, TX -75252
Phone : 214-272-2404

Sharjah, UAE

Q1-05-109/C
SAIF Zone
PO Box 121213
Sharjah-UAE