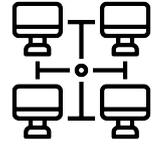


Success Stories

Data from 10 Million IoT Devices into the Snowflake Data Cloud



Client

American Computer
Networking Company

Industry

Utilities

Country

USA

Architecture Services



TECHNOLOGIES LEVERAGED

-AWS: IoT Core, Lambda, S3 and Kinesis Data
-Firehose -IoT Devices -Snowflake Data Cloud
-Tableau



HISTORY OF TRANSFORMATION

Data, an asset that has grown out of control in recent years, is crucial for organizations' futures, especially for our client, a large American computer networking company. Businesses need tools that enable them to not only process the massive amounts of data they generate but also to exploit this data and access information that has real value for the business decision-making process.



CLIENT CHALLENGE

The client wanted to improve service quality and to be able to debug issues down to a granular level utilizing data from ten million IoT devices, which were sending more than 20 different types of real-time data. Furthermore, whenever error rates were detected, those records were destroyed, and the client had no understanding of where or why the errors were occurring. The previous architecture could not handle the high volume resulting in outages and randomly big datasets being sent and sometimes lost. Another issue was geographically dispersed IoT devices across multiple time zones without synchronized clocks. Also, the possibility of IoT devices sending invalid data such as partial records due to power interrupts, or invalid JSON structures. Finally, the client was facing high data

acquisition costs since each IoT message was routed through 20 IoT rules incurring a charge for each rule whether or not it was used.



SOLUTION OVERVIEW

NTT DATA delivered the following solution in order to improve service quality and debug issues down to a granular level leveraging data from ten million IoT devices.

- 1- An AWS Lambda function was designed within the data flow to replace the prior NiFi processing tool in order to avoid reliability issues and reduce costs.
- 2- After analyzing the error rate, it was decided to convert the multi-line JSON into a single-line format in order to understand what the error was, where it occurred, and why.
- 3- To overcome the high ingestion costs, the IoT rules needed to be simplified. For that 20 different IoT rules were merged into one.

Finally, for the deployment of the solution, a parallel flow was implemented, which maintained the old system, which was still sending data in real-time while implementing the new solution, and creating a dummy sink in Snowflake for validation.



BUSINESS VALUE & KP

- Error rates were decreased by 97.6%, and if a device sends a bad request, the record containing the error rate is saved so that it may be located and diagnosed. Because of the reduced error rate, the client gained more reliable and accurate analytics.

With accurate and on-time analytics, the product engineering team was able to use this data for debugging as well as improving the quality of service for devices resulting in large benefits to their IoT device consumers.

- Data acquisition and ingestion costs were reduced by 77.3% (from \$22,000 to \$5,000) due to reducing the number of IoT rules
- The entire solution was delivered in one month to save the client time and give them immediate, cost-saving benefits.

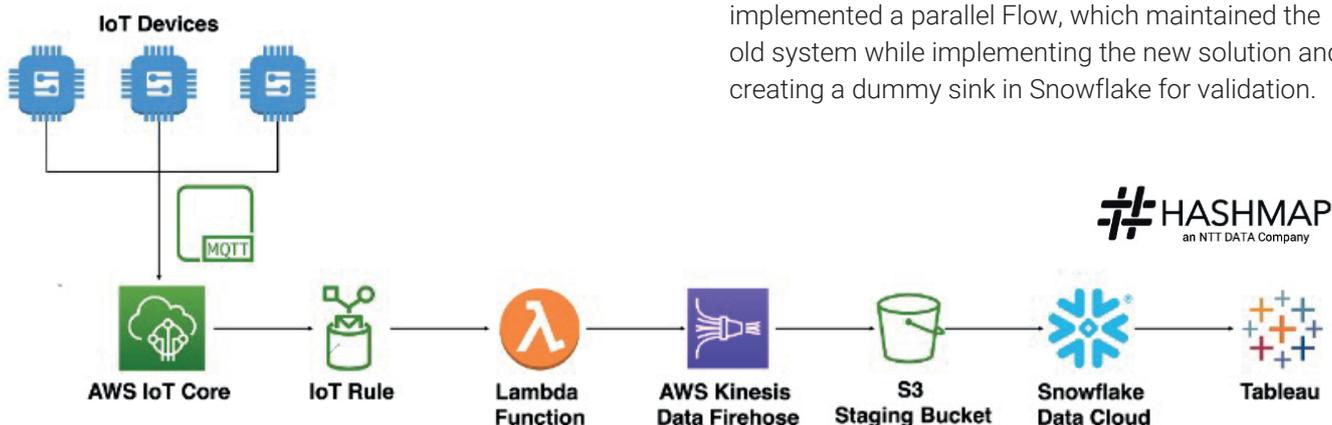


TECHNICAL SPECIFICATIONS

In order to address the client's challenges, which included the ability to debug issues down to a granular level, improve data quality and reliability, reduce error rates and decrease data acquisition and ingestion costs NTT DATA delivered the following solution:

1- Avoid Reliability Issues and Lower Costs

The client was using Apache NiFi for data flow and processing, but it was causing reliability issues and was costly. The Hashmap team decided to replace the NiFi processing by developing an AWS Lambda function, which allowed the client to save money since you only pay for the compute time you use.



KEY SUCCESS FACTORS :

- Our targeted data, cloud architecture, and engineering expertise
- Our fast and methodical delivery practices:
 - Saved the client time
 - Provided them with immediate, cost-saving benefits.

NTT DATA Corporation

Toyosu Center Bldg., 3-3, Toyosu 3-chome, Koto-ku, Tokyo 135-6033, Japan
 Tel: +81 3 5546 8051 Fax: +81 3 5546 2405
www.nttdata.com/jp/

The 10 million IoT devices are now using MQTT, a remarkably lightweight pub/sub messaging protocol, to send telemetry data to AWS IoT core. Further telemetry data is filtered out through IoT rules and preprocessed using AWS Lambda. Data is further buffered with a time and size file limit and compressed by Amazon Kinesis Data Firehose. When either limit is reached first, the file is placed into an S3 bucket to stage into the Snowflake Data Cloud.

2- Error Rate

Another key aspect was the error rate with the previous implementation. The Hashmap team analyzed the error rate and decided to convert the multi-line JSON into a single-line format in order to understand what the error was, where it occurred, and why. This allowed the client to address errors much faster gaining significantly more reliable and accurate analytics

3- Overcome High Ingestion Costs

To address the high ingestion costs, the IoT rules needed to be simplified. For that 20 different IoT rules were merged into one which saved the client a significant amount in costs.

4- Deployment

The last solution to be highlighted in this case is deployment. It is not possible to stop 10 million IoT devices during deployment because devices are still sending data in real-time. Therefore, the team implemented a parallel Flow, which maintained the old system while implementing the new solution and creating a dummy sink in Snowflake for validation.

