How a DDI Network Services Leader transformed their critical reporting infrastructure with CloudIO FLOW Data Pipelines

Customer Profile

- Silicon Valley-based, VC-owned company.
- Recognized leader in core network services (collectively known as DDI) which includes
    - Domain Name System (DNS),
    - Dynamic Host Configuration Protocol (DHCP)
    - IP address management (IPAM)

Enterprises worldwide rely on this global tech giant to achieve...

- ✓ Seamless network experiences
- ✓ Security from the ground up
- ✓ Reliability and efficiency

...through DDI-focused automation

Scope & Core Challenges

Transformation of Existing Reporting Infrastructure

- o Create a scalable reporting infrastructure on AWS Cloud that can scale with their growing information needs.
- o Unify their highly-siloed corporate information on revenue, discounting and Installer base.

Data Silos in Sales and Marketing Functions

- o Get real-time visibility into the sales and marketing operations, through real-time reporting of key business metrics.
- o Eliminate manual efforts in data integration and reporting

Need for Rapid Turnaround for Time-to-Insight

- o Rapidly integrate and orchestrate data from different types of data sources into a data lake and support different types of data transport protocols.
- o Enable reliable high-volume data movement for real-time streaming analytics.

The existing enterprise reporting system was not reliable. It was developed in silo without considering the information requirements of different business functions.

Given the scope and complexity of the data infrastructure involved, there was a need for a flexible data ingestion framework. But the challenge was the tight timelines, which meant creation of a custom data ingestion framework wasn't possible.

Solution

The customer was looking to create 100+ new reports and dashboards based on cloud-based data lake architecture on AWS. The solutioning depended on an efficient data orchestration framework that could handle the high-volume real-time data flows.

Data Pipeline Workflow

- The CloudIO FLOW accelerator was used for data ingestion and transformation, consuming data from different source systems, and staging the data in Amazon S3.
- Once the data was imported into the CloudIO FLOW pipeline, it was streamed through various stages such as

  ✓ **Data Masking** of some sensitive parts of the data,

  ✓ **Validations** were orchestrated as per requirements

  ✓ **Transformation** of the data as per workflows

  ✓ **Mapping to the output schema** as per requirements and then

  ✓ **Loading into the output data lake** (i.e. Amazon Redshift), for analysis and reporting.

Integrations

- CloudIO FLOW integrates seamlessly with the AWS ecosystem.
- It was also used to transport data from Amazon S3 to Amazon Redshift staging database, which was used as a repository for the data warehouse holding star schema, staging database and source for Enterprise Reporting.
- In addition, Amazon RDS was used as a repository for the ad-hoc reporting needs, hosted on AWS, while leveraging Tableau as the visualization platform for analytics, hosted on AWS.
- Most of the components are hosted in AWS EC2 cloud and AWS S3 buckets.

Business Impact

✓ Accelerated time-to-insight 30% sooner as a result.

✓ Cut down on time and labor costs to build a custom data orchestration framework.

✓ Facilitated reliable movement of data from various source systems into data warehouses.

✓ Enabled data refresh at 30-minute intervals.

✓ Ensured the creation of a unified view of business metrics across different business functions.

✓ Enabled readiness for near real-time reporting requirements.