

ETL Deployment using Talend

Large UK Media Group

PROBLEM STATEMENT

Channel4 has recently launched a web portal, one of the first of its kind, for its Agencies, to provide more insight to their data. Agencies registered for the service could get firsthand information with respect to their campaigns and delivery positions, which would help them to charter future course of action.

However, the data that was exposed, was sourced from the existing Data warehousing solution for the OLTP system. Since the source schema was not designed for the new portal, complex logic had to be introduced at application level to present the data. This resulted in not only code maintenance issues but had a severe fall back on the performance of the system as a whole.



SOLUTION

- ❖ NIIT proposed to do away with sourcing the data from the DW system. Instead, the proposed solution was to integrate a replication database to the main OLTP system through Redshift hosted on the AWS Cloud using the replication tool Attunity.
- ❖ Talend was proposed as the ETL layer to extract data from the replicated database, transform them using specific business logic and load them to Dimension and Fact tables in the target Database AURORA, hosted in AWS. Talend jobs were designed to perform both one-time load activity on virgin targets as well as capture and post the delta changes on the OLTP server through Change Data Capture jobs.
- ❖ The processed data in the AURORA databases would then act as the source for the data to be displayed through the portal

BUSINESS OUTCOME

- ❖ Immediate benefits included better code maintainability and ease of deployment. As on date, around 90 Talend jobs have been designed and deployed to extract, transfer and load demand and delivery data in almost real time.
- ❖ The proposed solution offered flexibility for the design as NIIT had the liberty to formulate the ETL rules and the data structures
- ❖ This has provided the agencies with greater visibility on their campaigns and allowed them to fine-tune their directions to meet their delivery targets
- ❖ The chosen solution also offered cost and operational scalability which helped forecast the customers planned IS budgets

