

Capstone Problem Statement

As part of Incedo's hiring process, you are requested to create a presentation for the business case mentioned below. The case is designed to gauge the understanding and approach towards execution in a modern data environment.

In case of doubt, please go ahead and make assumptions keeping in mind the objective and typical business considerations.

Context - Data Platform modernization

- ABC Corp is one of the largest Manufacturing - Consumer Goods Company fully into tool manufacturer worldwide focused on electrical and utility applications.
- They need to address data availability challenges due to Aging Architecture, Operational Silos, Performance bottle neck and Untapped Data with their existing EDW (on-prem)
- They embarked on a Modern Data Platform initiative with the below **objectives**:
 - Develop a scalable Data and Analytics Infrastructure on cloud enabling data availability across different entities of the organization. Solution should have the ability to ingest variety of data sources.
 - Leverage a common Infrastructure to overcome performance bottlenecks, reduce cost, reduce new ETL and BI development effort with increased productivity
 - As a part of reporting, include strategy and tools required to support Operational and BI reporting & dashboards, adhoc reporting and analytical applications needs.
 - In subsequent phase, they intent to build Data Science models across their LOB using the modern data platform. Solution must include details to ingest data from various data available across the enterprise.
 - As part of this program, they intent to have various 3rd party sources (which is not in place in current legacy systems) to carry out various analytics as below
 - ClickStream to Visualize the customer journey - Tie together the clickstream hit from their websites at a user / session level; Identify the events user performed; Compute metrics such as 'Session Duration', 'Time to First Purchase'
 - Social Media Data to increase the business / attract customers (eg. Facebook, Twitter, linkedin) – To understand customer activities on their page, identify frequent visitors & give additional discounts, send flyers etc
 - External HR Data – Productivity Measurement– ABC Corp HR and Payroll being managed by external company – intent to get basic HR data, swipe-in/out, leaves etc.
 - Expectation from business units (Marketing, Sales & Operations) is to have data refreshed within two hours of latency (current system loads daily once as overnight batch)

Current Ecosystem (Tech Stack & Challenges)

- On-Prem Tech Stack:
 - Teradata is used as their EDW solution. 50% of LOB's have built their data pipelines on Teradata via ELT methodology. Some other LOBs, departments and Legacy system have used IBM Data manager as ETL.

- Inventory Management and certain LOBs have used SQL server for building data marts using SSIS for ETL and SSRS for reporting.
- IBM Cognos Report studio is their primary reporting platform. Lot of manual reporting on excel across LOB's.
- Operational metrics of stores & POS across the world is siloed. As are some applications built for regional need.
- There is no single version of truth across their enterprise level.
- There is no exec level dashboards and lot of manual effort involved in downloading few Cognos reports; merging the data to generate excel based dashboards

ABC Corp is engaging with us on the above program, and you have been identified as a Program Manager (that's you 😊 in the context of this case study), who could look at this requirement with a techno-functional lens, to define:

- A high-level Execution Strategy to Modernize the Technology Stack

Please outline your response to the above in a presentation not exceeding 10-15slides, covering.

- Your understanding of the business problem & key objectives
- High level* technical solution, calling out recommended tools / platforms for the data and reporting/consumption layer and key considerations for recommending such choice
- How would you approach this program? Where will you start and how you build a Overall solution and roadmap with stakeholders fully involved in each steps
- Key roles in the team that you would need to manage & deliver this program (You also need to coordinate with an onsite consultant located closer to the client and would be half the globe away with ~12hr time difference)
- Processes & protocols that enable communication while balancing outside hours stretch
- How would the client stakeholders be managed, and what expectations would you set, of them?
- Various key hand-shake points across functional areas, across client vs your org, different internal teams within your org, as well as different internal teams within the client org
- Risks & mitigation plans – from a time / cost perspective.
- Highlight various success criteria for this program – for the initial module as well as ongoing program?

*Please use your knowledge & experience in this space to put out a good outline.

*You can fairly assume any technological Stack as a Solution based on your prior experience and have pointers to substantiate why / how the selected Stack would meet customer expectation and would be able to scale in future.