**DATA PROJECT**

**Points noted and Research**

Overview: We had meeting with LJ and Andy, leveraging on that meeting in terms of contract which is basically the Rental Man data on AS400; I researched further and found some interesting folders (Data Project) to understand the mindset of the business.

**First, I came up with some business questions / KPIs on Contract, taking into consideration of the following:**

**A. Customer Satisfaction / Cohort Analysis**

**B. Equipment Availability and Status / SKU (Stock keeping unit)**

**C. Equipment Life Cycle at Point of Disposal (Total income / total expenses)**

**D. Income from Equipment-to-Maintenance Ratio (Compare revenue and maintenance cost)**

**E. Equipment Rental Rate (Rental Revenue / number of contracts)**

**F. Financial Utilization of Equipment (Annual Revenue / Total Cost Acquisition)**

**G. Equipment Time Availability and Time Utilized (Days rented / day available)**

**H. Company's Expansion Strategy**

1. Our rental(contract) process. (Reservation, Quote and Quote), data to be captured
2. Which equipment do customers rent frequently?
3. How can we determine the total rental income of an equipment and subtract the original purchase price, any carrying costs and maintenance expenses over its lifetime?
4. Which equipment goes for maintenance frequently?
5. How many hours / days / weeks / months on rent?
6. Which Branch is more efficient?
7. What time of the year do we rent more equipment?
8. Which State / City / County / Region do we need to move more equipment?
9. Do we know our customers?
10. What can we do ensure our equipment rental planning is competitive?
11. How many Contracts do each branch issue daily?
12. How many Contracts are closed daily?
13. Do we meet up with our delivery and pick times?
14. Do we have exact Location of Jobsite, or do we find it difficult locating job sites? How so solve it
15. Do we track our equipment location daily?
16. Status of equipment before and after Jobsites
17. What Brand of equipment do we rent often
18. What is our delivery and pick up processes?
19. Do we have competitive rental rates?
20. How do we track our customers reviews and feedback in real time?
21. How old are some of our equipment
22. What is equipment maintenance cycle?
23. What data do we capture during Rent Reservation and Quote?
24. Do all Reservations and quote get to contract? How do we track why some doesn’t get to?
25. How can we differentiate our rental process from our competitors?
26. How do we track equipment add-ons in contract? Are they returned separately or together?
27. How do we track our market surveys?

Still Developing Questions. All these KPIs target the data from ERP and CRM systems. All these are geared up to having and E-Rental environment or platform.

Integrating Confluent Kafka with Databricks

Diagram

Description automatically generated

Data Flow

1. From Source (Azure SQL, on prem, blob storage etc)

2. Kafka connector connects to source

3. Kafka through ksql create streams (this can be connected to other microservices for Realtime request and monitoring).

3. Kafka connects to Lakehouse (Delta Lakes / Databricks) for analytics and to EDW for storage

Spark Streaming processes data in batches and mostly used for Machine learning and Analysis.

Kafka is a real time message tool. It’s for immediate action, take credit card fraud detection as a case study. It processes data in milliseconds.

**Strategy A**

1. We have the old system writing to AS 400/ IBM, we use Kafka connector on it to bring the data (Transactional) to cloud.

2. We develop new etl pipeline to stream the data at Kafka

3. Then have the data at Streaming (Delta Lake), transform the data with our new table format and load to our EDW, have Data Scientist work on the data as well.

**Strategy B**

We design a new platform using the Azure Database (having the same replica of table transformed in number 3), connect all the new microservices we are developing to the new SQL database.

Connect to Kafka

Connect to Delta

Load to EDW.

Key Business Questions

1. What would we like to be able to do with data/ information that we cannot do today?

2. How would that transform the way we do business?

WHAT IS OUR AVR?

1. Assessment

2. Vision

3. Roadmap.

The AVR seeks to answer:

1. What are we capable of today?

2. What would we like to be capable of tomorrow?

3. How will those new capabilities transform our business?

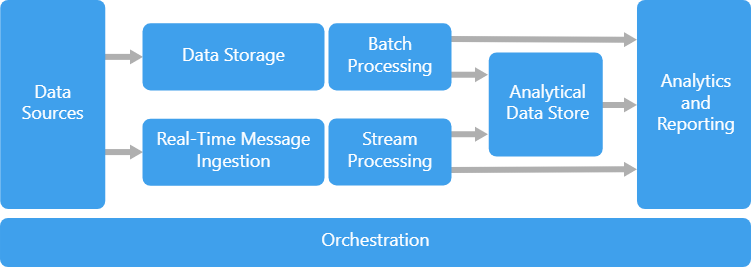
ASSESSMENT

Before you can determine where you want to go or how to get there, you must first understand where you’re coming from. The assessment step of an AVR is an internal analysis that will help paint a picture of your organization as it stands today.

The purpose of the assessment is to identify challenges and opportunities by answering questions like, “What data is being held hostage?” and “Are you equipped with the right tools to do your job?”

Step one of the AVR is focused on establishing business relationships and collecting feedback within your organization.

**SUNSTATE BIG DATA ARCHITECTURE**



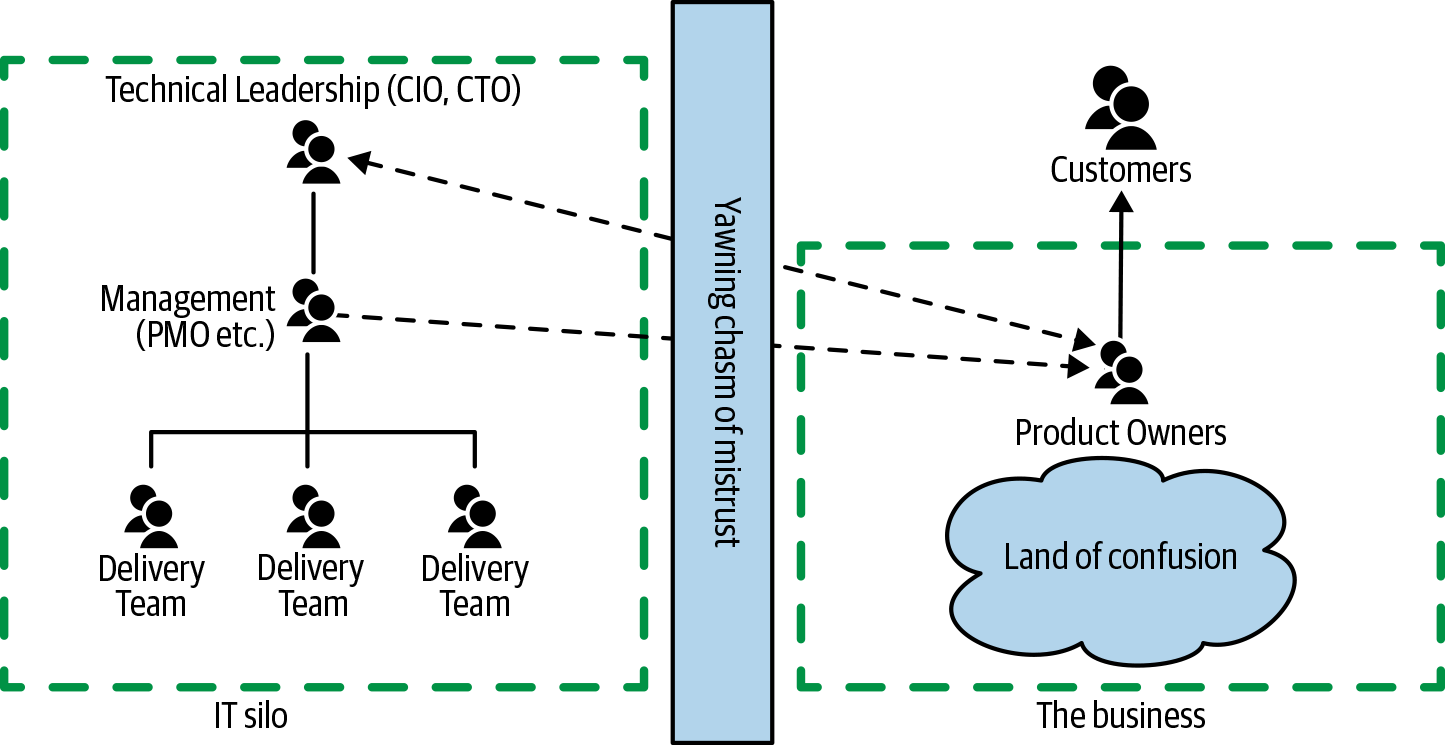
The four concepts of big data architecture are:

1. Batch processing of big data sources at rest.

2. Real-time processing of big data in motion.

3. Interactive exploration of big data.

4. Predictive analytics and machine learning.



Event Driven Architecture

Chart, bubble chart

Description automatically generated

Diagram

Description automatically generated

Diagram

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated