Breaking Silos through Data Decoupling

Chew Wei Liang
Group Technical Data (GTD)
Project Delivery & Technology (PD&T)

Chua Ket Peng
Group Technical Data (GTD)
Project Delivery & Technology (PD&T)
Breaking Silos through Data Decoupling

1. Problem Statement
2. Solution
3. Summary
Problem Statement

Top 5 Major Problems by Data Related To Digital Transformation

1. **Data not Available**
   Data often not easily accessible when user needed them for work purposes.

2. **No Line of Sight**
   Most of data are available in scattered data source across the network but no line of sight through single source of truth.

3. **No Standardization**
   Data are scattered across the network with different naming and standardization.

4. **Data Security**
   By allowing user to access to the database directly, there is a high risk of data security breach.

5. **Data Quality**
   No metrics available to determine the quality of the data.
Problem Statement

Putting Perspective into Metrics

**Workflow**

52% Non-Productive Time

Business users’ time spent on data related matters such as data gathering and searching, quality control (QC) and loading based on 10 key workflows in Field Development Plan (FDP).

**Data**

17% Time Spent

Everyone who acquires data spends time re-checking it. 17% time spent (equiv to 58 man years) for well header, check shot, deviation and basic logs (find and quality check).

Breaking Silos through Data Decoupling

1. Problem Statement

2. Solution

3. Summary
# SIMPLIFY Data Access through Data Access Robot (DART)

## Corporate Databases Stack

- COS WELL
- PSLG
- PGPS
- PEDMS
- LIMS
- ...

## Document Stack

- UDRS
- ECM

### Characteristics

<table>
<thead>
<tr>
<th>Sequestered</th>
<th>Disconnected</th>
<th>Connected</th>
<th>Integrated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Interface</td>
<td>Request System</td>
<td>Data Portals</td>
<td>Data Gateway</td>
</tr>
</tbody>
</table>

### Advantages

1. Interface coupled to single database
2. Access to all areas of the database
3. Used mostly by data managers

1. Data & report can be requested via ticket request
2. One Stop Centre for data request
3. Satisfies Data on Demand 24/7
4. Web Enabled
5. Standard data output

### Disadvantages

1. Limited Access
2. No direct extraction of data to user environment
3. Long lead time
4. Data Quality Unknown

1. User need to load into interpretation environment
2. Need to access multiple portal for different data types
3. Significance development efforts required to find and transform the data

### DART

<table>
<thead>
<tr>
<th>DART 1.0 (2015)</th>
<th>DART 2.0</th>
</tr>
</thead>
</table>

1. Direct access & immediate to specific database
2. Satisfies Data on Demand 24/7
3. Web Enabled
4. Standard data output
5. Data centric – no longer database centric

1. Limited data access
2. Data Quality Status
3. API for integration

### DART 2.0

- Data Centric
- Data Access Robot (DART)
How do we solve the problems?

#1 Decoupling & Liberating data

Most of the data are tied to application database. Therefore, in order to liberate the data:

1. We design data abstraction layer that is highly tailored to the business needs.
2. The abstracted data are presented into data types.
3. Each data types are carefully designed to determine its mandatory and optional attributes needed for business activities.

Example Data Types:
- Well Schematics
- Checkshots
- Pipelines
- Daily Operations Report (Facilities)
- Deviation Data
- Well Test
- Well Integrity
- Pressure Volume Temperature (PVT)

and 280+ data types...
How do we solve the problems?

1. Data abstraction layer (Decoupling)

- Databases
- Data lakes

Copy only critical data

2. Data Platform (DART)

- Preview
- Download
- API

3. Quality Checking

#2 Data Platform

Liberated data presented by the data types are listed on the data platform. This enable line of sights of the data availability, and enable business users with proper entitlements to quickly access the data.

Every action will be log for audit trail and security purposes.
How do we solve the problems?

1. Data abstraction layer (Decoupling)
   - liberated data will go through the process of data quality checking. The process of measuring the data quality is done through the following step:
     1. A set of business rule is defined based on the input from Subject Matter Expert (SME) or Subject Matter Focal (SMF).
     2. The metrics are defined based on the set of business rule and implemented into the system.
     3. Data quality is presented to the user in different level of granularity, and they can be aggregated into high level dashboard view.

2. Data Platform (DART)
   - Copy only critical data

3. Quality Checking
   - #3

Databases

Data lakes
How do we solve the problems?

Designed with Robustness and Security in Mind

DATA PLATFORM (DART)

Internet

Request to Platform

Data abstraction layer

Databases

Data lakes

Safe Zone

Preview

Download

API

Copy only critical data

Act as layer of firewall

Perform throttle and caching as necessary
How do we solve the problems?

Designed with Scalability in Mind

Routing Mechanism (Load Balance)

Platform (Node 1)

Platform (Node 2)

Platform (Node 3)

Safe Zone

Preview

Download

API

Data abstraction layer

Databases

Data lakes for caching

Auto cache frequent access data
Breaking Silos through Data Decoupling

1. Problem Statement
2. Solution
3. Summary
Data decoupling give the user with proper entitlement the “superpower” ability to be able to access the data anytime, anywhere they needed for operation and analytic purposes.

Database filled with data has no value until it can be easily accessible and turned into valuable insights for the business.

Data quality is measured and the level of quality data is automatically aggregated and tracked in dashboard for further action by data custodian.

Time savings in acquiring data and quality checking translate into millions of dollar in cost saving.
Thank you for your passion!