



Banking Modern Data Fabric In a Box



Infrastructure (Azure)
Data & AI (Azure)
Digital & App Innovation (Azure)
Security



**Anirudh
Kala**

CEO & Co-founder



**Anupam
Gupta**

Co-founder, Head
Corporate Strategy



**Dinesh
Thawani**

Chief Technology
Officer



**Anish
Ashirgade**

Head Celebal Tech.
Singapore



**Mukul
Biswas**

Head Celebal Tech.
UK



**Jayanta (Jay)
Banerjee**

Corporate Vice
President



**Piyush
Gupta**

Global Head, ERP
Modernization Practice



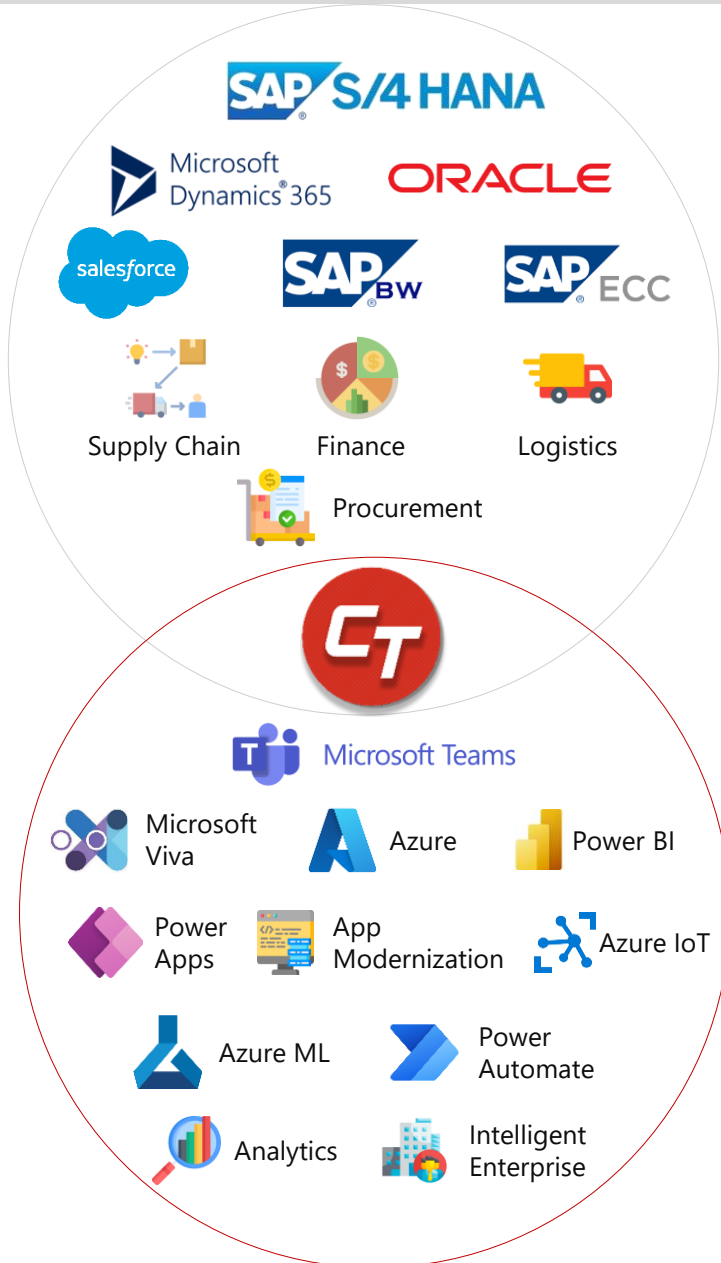
**Abhishek
Goyal**

Solutions Sales
Principal



**Srinivas
Satyanarayana**

BFSI Practice Lead




2000+
Employees




980+
Azure Certifications

Advanced Specialization



- AI & Machine Learning
- Analytics
- Infra and Database Migration to Microsoft Azure
- Kubernetes
- Cloud Security



Global Presence

US, Canada, UK/Europe, Asia Pacific, India, MEA



databricks
Partner Of The Year Award
Asia Pacific and Japan



databricks
Regional System Integration Partner
India
2 YEARS IN A ROW



databricks
Top Consumption Partner
Asia Pacific and Japan

Channels	Mobile/SMS Banking	Internet Banking	ATM/ POS/ Kiosk Banking	Handheld Devices	Call Center			
Channels	Consumer Banking and Wealth Management				Corporate Banking and Trade Finance			
	Savings & Checking	Term Deposits	Retail Loans	Mortgages	Current/ Overdrafts	Notional Pooling	Syndication	Bill Discounting
	Insurance	Mutual funds	Financial planning	Structured products	Export/ Import Financing	Forward Contracts	Documentary Credits	Guarantees
Functional	Card management	Customer Information	DEMAT and online trading	Locker/ Strong Room	Commercial Lending	Limits and Collateral		
	Standing Instructions	Sweeps/ Pooling	Payment Systems	Referral	Bill Payments			
Business Components	Treasury	Asset Liability management	Liquidity Management	Fees/ Charges				
	Interest/Tax	Exchange rates	Bank management	Signature Verification				
	Inventory	Channel Rules	Discounts/ Preferential Pricing	Risk Management				
	Auto-Debit Credit	Business Intelligence	Document Management	Cash Management				
	Legal	Clearing						
Accounting	General Ledger	Multi-Currency	Transactions	Finance and Accounting	Audit and Compliance			
Basic Features	24*7	Multi-lingual	GUI	Single Sign On				
	Integration Framework	Purge	Access Control	Workflow (BPEL)				
	Reporting	Audit	Multi-Calendar	Scheduler				

Data Flow & Operations



Connectivity with Data Sources to facilitate data extraction and transformations with automation and optimized orchestration using cloud native technologies.



Azure Data Factory/Databricks Notebooks helps in data collection at a standard location, cleaned and processed. Persist the data from the source systems into the landing area.



Raw/unstructured data lands into Staging/Bronze Layer. Further the data is processed and traverse through different layer.



Data models/Business logics will be implemented on Gold Layer (Lakehouse or Azure Synapse) which will act as an Analytical Layer to BI applications or Use-Cases.



Attentiveness on failures and automatic retries of the data pipelines. Integration with Power BI which will provide flexibility to publish dashboards/reports across the enterprise.

Data Lifecycle

Ingest

Classify

Transform

Curate

Consume

Distribute

Archive

Data Types

Streams Messages

Logs

Batch

Structured

Unstructured Media & Text

Data management & Processing

OLTP

OLAP

Columnar

Graph

Document

Index

Time

In-Memory

File System

Data Storage

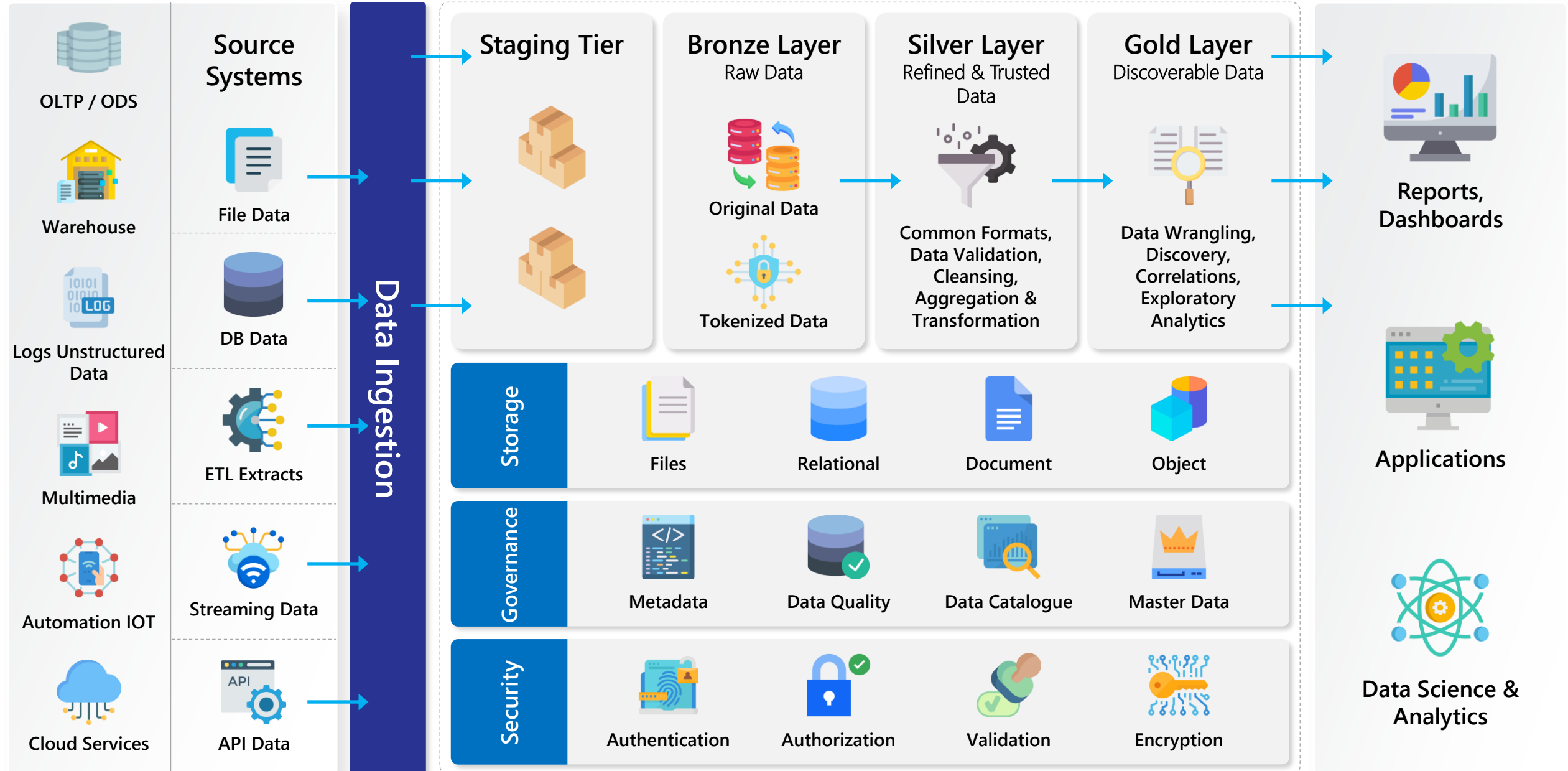
Block Storage

Object Storage

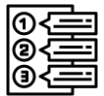
Data Security

Data Governance

Reference Data Architecture



Theme: To construct and automate the Banking Data Lake solution that can be **versatile** with hybrid necessities, **apparent** with regards to data governance and lineage, **adaptable** concerning the immense size and variety of data and able to carry out the state-of-the-art advance analytics. In preference to give a logical approach to business choices, modern data fabric (enable controlled, secure and self-service economy) to deliver data model with business critical KPIs for subdomains of Banking such as – Retail, Operation and Assets



Requirement Gathering and Assessment to understand the current data landscape and to take safeguards in event falling during the phased development.



Setting up the Infra with network and security for azure services, laying the foundation for Life Insurance Data Platform as per the designed Solution & Network Architecture.



Deploying Ingestion Framework to construct the pipelines via ADF and further data evolution via Spark Notebooks.



Building Data Platform to stage source data in Raw zone (hierarchical structure), massaged data in Enriched zone and aggregated data as facts & dimensions in Curated zone.



Setting up Purview to cater automated data discovery, classification and end-to-end lineage with prime business glossary thereby enhancing data literacy in the organization.



CI/CD Pipelines deployment and configuration to setup DevOps framework.



Deploying 4 Power BI Reports consisting semantic data model and 15-20 business critical KPIs involved in Banking – Summary, Retail, Operational and Assets

Data integration

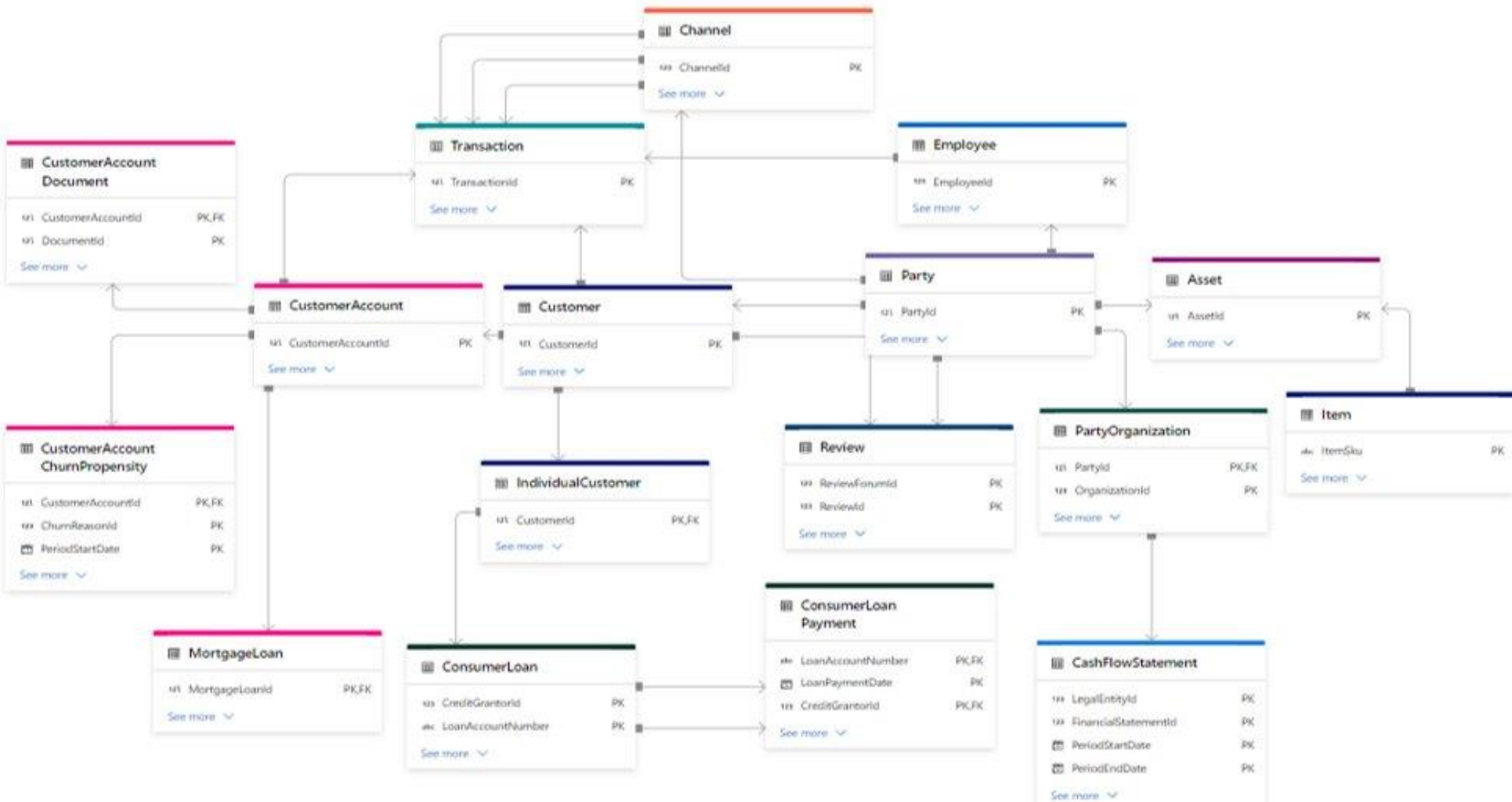
Data warehousing

Big data analytics



S. No.	Dashboard	KPI Group	KPI Name
1	Summary		<ol style="list-style-type: none"> 1. Overall Status 2. Number Of Branches Across Region 3. Total Number Of Customers 4. Fcy 5. Lcy 6. Liquidity ratio 7. Earning Asset (Current) 8. closing Shareprice vs Total Shareprice 9. Npl vs total Loan 10. Employee expense(Actual vs Budget) 11. Capital Adequacy Ratio vs Benchmark 12. Actual Deposite Balance vs Budget (liabilities) 13. Actual vs Budget(operating expence) 14. Total Number Of Echannels Transaction 15. Amount Of Echannels Transaction 16. No of Social Media
2	Retail	Transaction	<ol style="list-style-type: none"> 1. Transaction Count Per Day 2. Transaction Value Per Day 3. Transaction Count by Channels 4. Percentage Of Successful To Failed Transaction 5. Number Of Transaction 6. Commission 7. Amount Transacted

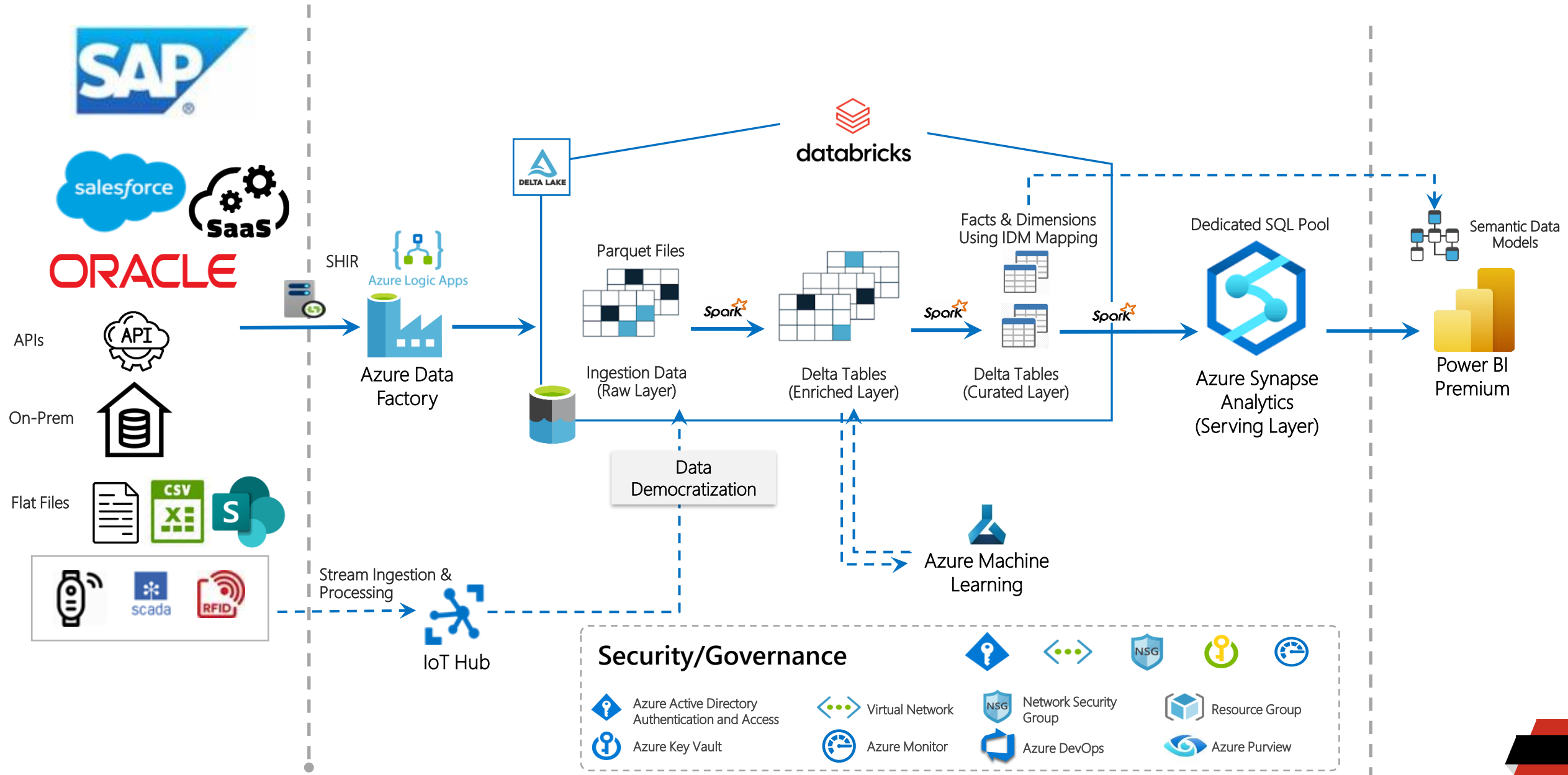
S. No.	Dashboard	KPI Group	KPI Name
3	Operational		<ol style="list-style-type: none"> 1. Number Of Account By Region 2. Number Of New Account by Year 3. Number Of account BY Ageband 4. New Account by gender 5. New Account Per Account Manager 6. Number Of Account by Account Type
4	Assets		<ol style="list-style-type: none"> 1. Daily Analysis Of Earning Asset 2. Earning Asset Ratio By Region 3. Current Day Earning Asset By Region 4. Earning Asset Balance 5. Non Performing loan 6. Recovery 7. NPL vs Total Loan 8. Daily Non Performing Loan vs Total Loan



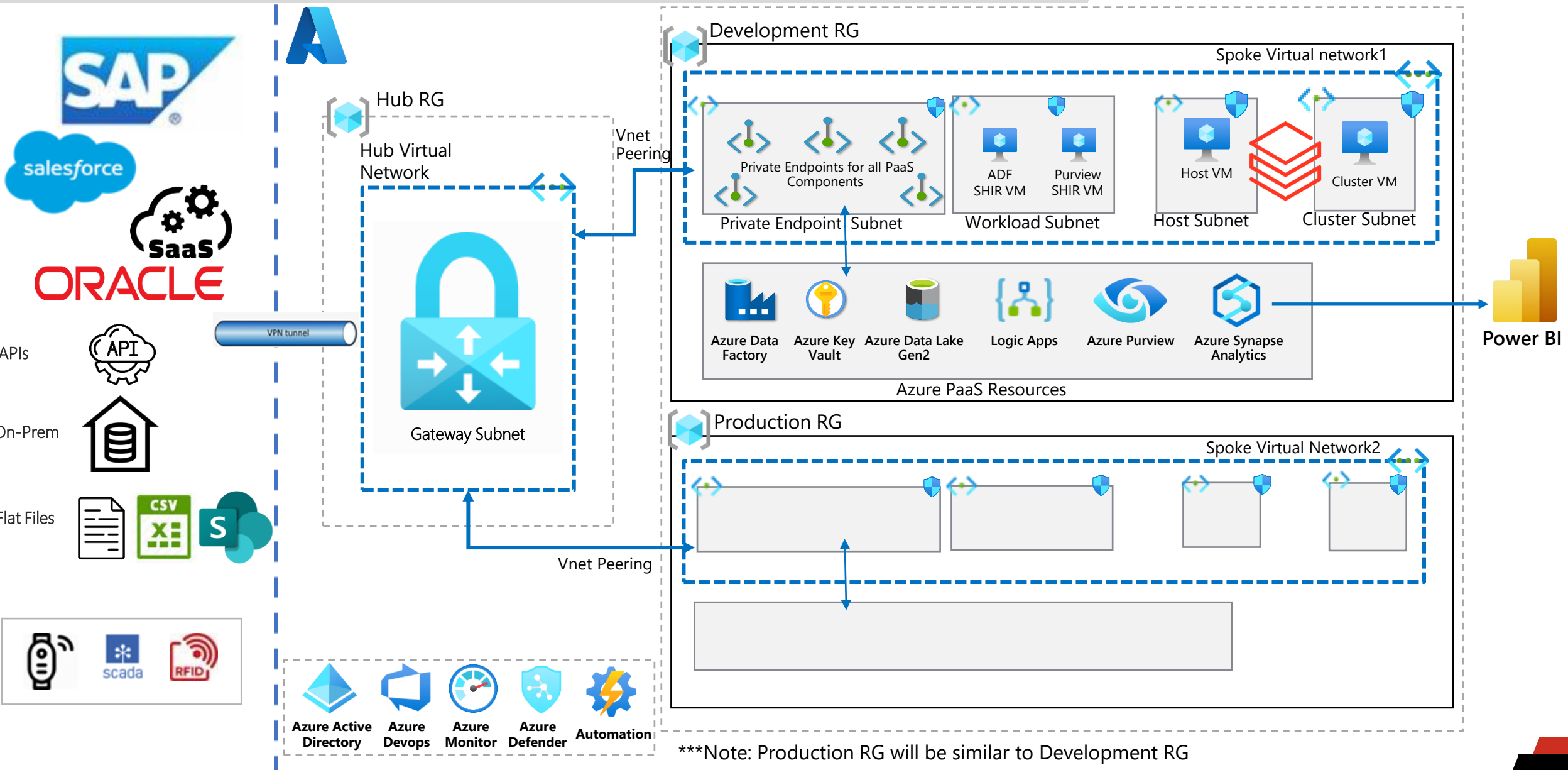
Source Mappings to be provided by Client

Source_Table_Name	Source_Column_Name	Source_Column_Type	Sink_Table_Name	Source_Column_Name	Source_Column_Type
Employee	EmployeeId	string	Employee	EmployeeId	integer
Employee	EmployeeName	string	Employee	EmployeeName	string
Employee_Name	DateOfBirth	string	Employee	DateOfBirth	date
Employee_Name	DateOfDeath	string	Employee	DateOfDeath	date
Employee_Name	EmployeeStandardCostAmount	string	Employee	EmployeeStandardCostAmount	decimal
Employee_Name	EmployeeOvertimeCostAmount	string	Employee	EmployeeOvertimeCostAmount	decimal
Employee_Name	EmployeeHireDate	string	Employee	EmployeeHireDate	date
Employee_Name	HighlyCompensatedEmployeeIndicator	string	Employee	HighlyCompensatedEmployeeIndicator	boolean
Employee_Name	KeyEmployeeIndicator	string	Employee	KeyEmployeeIndicator	boolean
Employee_Name	ExemptEmployeeIndicator	string	Employee	ExemptEmployeeIndicator	boolean
Employee_Name	NonexemptEmployeeIndicator	string	Employee	NonexemptEmployeeIndicator	boolean
Employee_Name	AlienStatusId	string	Employee	AlienStatusId	integer
Employee	SexId	string	Employee	SexId	integer
Employee	GenderId	string	Employee	GenderId	integer
Employee_Name	RacialCategoryId	string	Employee	RacialCategoryId	integer
Employee_Name	EthnicCategoryId	string	Employee	EthnicCategoryId	integer
Employee_Name	EmploymentLeadSourceId	string	Employee	EmploymentLeadSourceId	integer
Employee_Name	SecurityClearanceId	string	Employee	SecurityClearanceId	integer
Employee_Name	SecurityLevelId	string	Employee	SecurityLevelId	integer
Employee_Name	PartyId	string	Employee	PartyId	long

Technical Architecture - Azure Databricks & Synapse



Network Architecture



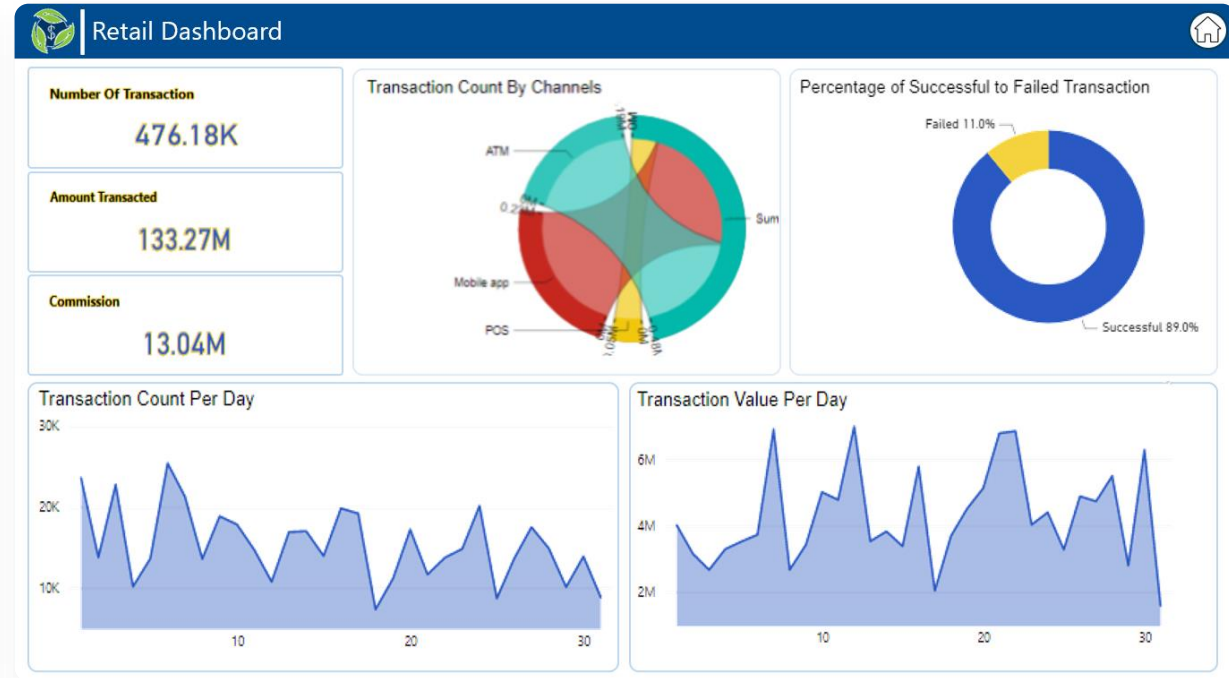
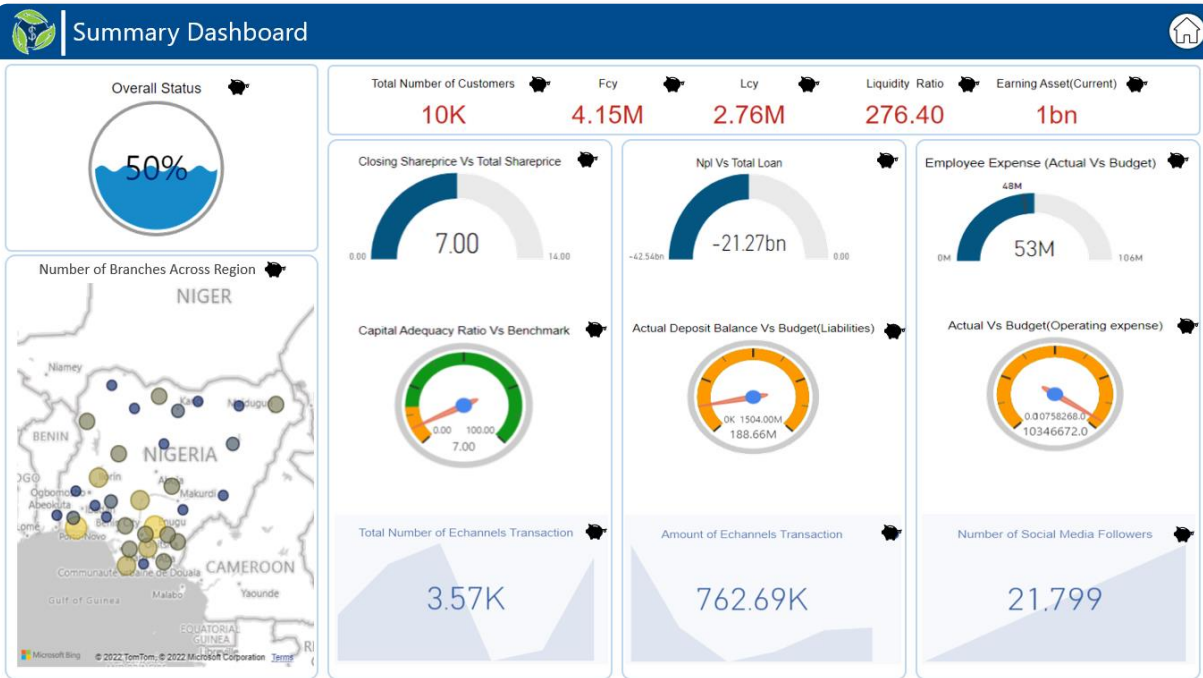
Solution Deployment (Azure x Databricks) - Timeline



Activities	WK1	WK2	WK3	WK4	WK5	WK6
Current Architecture Assessment						
Data Sources, Applications & Landscape Assessment with Workflows KT						
Documentation: Q/A, Assessment Report & Future Solution Approach						
Infrastructure Setup						
Network & Infrastructure setup with Security enablement						
Resource Deployment, Access Management, Connectivity & Throughput						
Data Ingestion Setup						
ADF ARM Templates, Databricks Notebooks Deployment						
Integration check, Execution & Testing						
Data Layer Setup						
Delta Lake Setup (Bronze & Silver Layer)						
Analytical Layer (Gold Layer) Setup						
Data Validation & Monitoring						
BI Reports/Dashboards						
BI Reports/Dashboards Integration & Deployment						
Data Validation & Monitoring						
CI/CD Setup & Governance						
CI/CD Pipeline Deployment						
Setup Purview for Data Cataloguing, Mapping, Lineage, Estate Insights & Access Control						
UAT						
Data Validation & Sign-Off						
Delivery & Sign-Off						
Deliverables demonstration with documentation						

- ✓ The timeline assumed that solution build around set of KPIs/Data Model/LOB will be available to user at deployment completion.
- ✓ Client must have all the necessary approvals before initiating deployment to avoid any latency and impact over the designed timeline. Pre-requisites will be shared by Celebal.

Service Type	Region	Description	Monthly cost
Azure Data Factory	Central India	Azure Data Factory V2 Type, Data Pipeline Service Type, Self-hosted Integration Runtime: 30 Activity Run(s), 100 Data movement unit(s), 200 Pipeline activities, 100 Pipeline activities – External, Data Factory Operations: 50 x 50,000 Read/Write operation(s), 50 x 50,000 Monitoring operation(s)	₹7,664.49
Azure Synapse Analytics	Central India	Tier: Compute Optimized Gen2, Dedicated SQL Pools: DWU 300 x 310 Hours, 0.5 TB of storage with Geo-redundant disaster recovery; Serverless SQL Pools: 1 TB of data queried; Central India Region, 0 GB of data collected per day, 1 days of Hot Cache, 1 days of total retention, 1 times estimated data compression, 0 Hours of 2 x Extra Small (2 vCores) Engine Instances, 0 Hours of 2 x 1 vCore Data Management Instances	₹133,055.49
Storage Accounts	Central India	Data Lake Storage Gen2, Standard, GRS Redundancy, Hot Access Tier, Hierarchical Namespace File Structure, 500 GB Capacity - Pay as you go, Write operations: 4 MB x 10 operations, Read operations: 4 MB x 10 operations, 10 Iterative read operations, 5 Iterative write operations, 10 Other operations. 500 GB Data Retrieval, 500 GB Data Write, 30 GB Meta-data storage	₹2,095.77
Key Vault	Central India	Vault: 1,000 operations, 100 advanced operations, 0 renewals, 0 protected keys, 0 advanced protected keys; Managed HSM Pools: 0 Standard B1 HSM Pool(s) x 730 Hours	₹14.85
Automation	Central India	Process Automation Capability: 500 included minutes and 100 additional minutes, 1 Watchers X 100 Hours	₹33.00
Logic Apps	Central India	Workloads: Standard plan, 1 WS1 (1 vCores, 3.5 GB RAM) x 150 Hours, 10 Standard Connector Calls per day x 1 day, 0 Enterprise Connector Calls per day x 1 day; Integration Service Environment: Premium tier, 0 Base Units x 730 Hours, 0 Scale Units x 730 Hours; Integration Accounts: 0 Standard Integration Accounts x 730 Hours, 0 Basic Integration Accounts x 730 Hours.	₹3,169.10
Azure Databricks	Central India	All-Purpose Compute Workload, Premium Tier, 1 D8DSV5 (8 vCPU(s), 32 GB RAM) x 310 Hours, Pay as you go, 2 DBU x 310 Hours	₹40610.02
Azure Private Link	Central India	5 Endpoints X 730 Hours, 500 GB Outbound data processed, 500 GB Inbound data processed	₹3,835.96
Virtual Machines	Central India	1 D4as v4 (4 vCPUs, 16 GB RAM) x 730 Hours (Pay as you go), Windows (License included), OS Only; 1 managed disk – E10; Inter Region transfer type, 5 GB outbound data transfer from West US to East Asia	₹25,361.88
VPN Gateway	Central India	VPN Gateways, VpnGw1 tier, 730 gateway hour(s), 0 additional S2S tunnels (beyond included amount), 0 additional P2S connections (beyond included amount), 500 GB, Inter-VNET VPN gateway type	₹12,885.52
App Service	Central India	Standard Tier; 1 S1 (1 Core(s), 1.75 GB RAM, 50 GB Storage) x 730 Hours; Windows OS; 0 SNI SSL Connections; 0 IP SSL Connections	₹6,022.04
Azure Monitor	Central India	Log analytics: Log Data Ingestion: 0.05 GB Daily Analytics logs ingested, 0.005 GB Daily Basic logs ingested, 1 months of Interactive Data Retention, 0 months of data archived, 0 Basic Log Search Queries per day with 0 GB data scanned per query, 0 Search job Queries per day with 0 GB data scanned per query; Managed Prometheus: {0} AKS nodes in cluster, {1} Prometheus metrics per node, {2} seconds of Metric collection interval, {3} Average daily Dashboards users, {4} Dashboards, {5} Data samples queried per dashboard, {6} promql alerting rules, {7} promql recording rules; Application Insights: 3 months Data retention, 0 Multi-step Web Tests; 5 resources monitored X 1 metric time-series monitored per resource, 0 Log Alerts at 5 Minutes Frequency, 0 Additional events (in thousands), 0 Additional emails (in 100 thousands), 0 Additional push notifications (in 100 thousands), 0 Additional web hooks (in millions)	₹47.43
Azure DevOps	Central India	5 Basic Plan license users, 1 Basic + Test Plans license users, Free tier - 1 Microsoft Hosted Pipeline(s), 1 Self Hosted Pipeline(s), 2 GB Artifacts, 5 VUMs	₹4,289.67
Microsoft Purview	Central India	Elastic Data Map: 1 Capacity Unit hour, 730 hours, Automated Scanning and Classification: 2 Total scan duration in hours x 32 Total vCores across scans (For other data sources), Other features: 180 Resources Set hours, Microsoft Purview Data Catalog: C0 Service	₹31,195.01
	Total		₹270,280.25





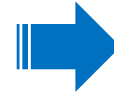
Celebal Team Involved – During Deployment		
Resources/Skill Sets	Number of Resources	Years of Experience
Solution Architect	1	10+
Project Manager	1	8+
Subject Matter Expert	1	10+
Infra Engineer	1	4+
Security Engineer	1	5+
Data Engineer	1	4+
Data Scientist	1	5+
Database Engineer	1	4+
DevOps Engineer	1	4+
Governance Engineer	1	4+
BI Developer	1	4+
QA Engineer	1	3+

Client Team/SPOC Required	
Resources/Skill Sets	Number of Resources
Solution Architect	1
Infra SPOC	1
Security SPOC	1
Data SPOC	1
BI SPOC	1
CloudOps SPOC	1
Governance SPOC	1



Category	Pre-requisites	Requirement
Basic Pre-requisites	VPN Access for all developers	To access On-prem/ Azure Cloud Resources.
Azure Infra Pre-requisites	List of Resources to be added in Azure AD	Celebal team to be added to Customer's Azure tenant for the execution of SoW.
	Contributor Access on Azure Subscription and Owner Access on RG	Owner role to assign access via Managed Identity whereas Contributor role to deploy required Azure services and proceed with the development/testing.
	VNet CIDR range for Azure that must not clash to On-Premises	To establish connectivity among private networks in order to Setup Private Network Topology over Azure.
On-Premise Networking Pre-requisites	Network Engineer required from Client end for Site-to-Site Setup (Max- 2 Days)	To setup on On-prem Firewall for Site-to-Site tunnel.
	DNS Entries Need to be configured on On-premises (DNS Engineer Required for Max-1 Day)	To setup DNS Entries on On-Prem DNS server
SHIR Pre-requisites	SHIR VM Configuration to Access Data Sources	To connect with On-prem resources via Private tunnel
	Required driver installations for connecting data sources to the Data Gateway.	Specific drivers could be required to Install SHIR if the download is blocked on VM
Nomenclature Pre-requisites	Naming convention (if any special requirements).	As per the standards.

1. Use SAP GUI to connect to the SAP System.
2. Go to System -> Status.
3. Check the release of the SAP_BASIS, ensure it is equal to or larger than 701.



Software Compo.	Release	Level	High
SAP_BASIS	730	0008	SAPF

Access approvals & Data Sources Integration will be dealt with high priority requiring Client SPOC from Infra, Security, CloudOps Team & Solution Architect.



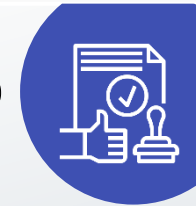
Source field mapping as per the target objects to be provided by the client.



Power BI License to be purchased by Client for developers to deploy BI Reports/Dashboards as per the timelines.



All stakeholders (responsible for report matrices) shall give a sign-off on metrics before Celebal Tech creates the reports.



Access to Azure subscription & Remote database/data sources will be made available for Celebal Tech team



During UAT phase, dedicated SPOC or business user from Client side would be required till Sign-Off, to avoid any delays, cater any issues or revisions which will help expedite solution deployment & delivery.





Assumptions

- ✓ 1 fiscal year data from Data Sources would be considered for loading into Data platform.
- ✓ Source mapping would be provided by client.



Deliverables

- ✓ 4 Power BI reports/dashboards with 15-20 business critical KPIs.
- ✓ Landscape for in-scope Source Systems with ingestion framework and Data Layers.
- ✓ Documentations



Features

- ✓ Power BI model is defined re-usable way for features like Q&A, communication between multiple users via comment. Etc.
- ✓ Self-service features. Multiple data layers can be managed as per the use-cases and user groups.
- ✓ Integrable with other sources like Data streaming, unstructured data, AI&ML capabilities using Data platform

The cost for one solution implementation is INR 10-12 lacs



Assumptions

- ✓ Total data volume to be copy from Source to Staging Layer is ≤ 1 TB(for historical data load and < 10 GB for daily incremental) covering 1 fiscal year data
- ✓ Total number of Source tables considered in scope is between 20-30 covering master and transactions
- ✓ Total 4 Power BI Reports/Dashboards with around 15-20 business critical KPIs.



CELEBAL
TECHNOLOGIES

Thank You

Feel free to
Contact us at

enterprisesales@celebaltech.com

