**EPIC 1 : Report Hub to DCM Metadata Synchronization Engine**

**EPIC Name**

**Report Hub to DCM Metadata Synchronization Engine**

**Description**

Implement a comprehensive data synchronization solution to transfer all reports related metadata from the Report Hub system to the DCM (Data Content Management) system. This initiative will establish a robust, scalable synchronization mechanism that supports both one-time historical data migration and ongoing incremental data capture to ensure data consistency and real-time availability across systems.

The solution will serve as a critical data integration bridge, enabling the DCM system to maintain a complete and up-to-date repository of report metadata, thereby provide catalogs for different reports available in system.

**Acceptance Criteria (EPIC Level)**

**Initial Data Load**

* **Complete Historical Migration**: All existing report metadata from Report Hub is successfully migrated to DCM system with 100% data integrity
* **Data Validation**: Comprehensive validation confirms that migrated data matches source system with zero discrepancies
* **Rollback Capability**: Ability to rollback the initial load if critical issues are discovered post-migration

**Incremental Synchronization**

* **Change Detection**: System accurately identifies and captures all types of metadata changes (CREATE, UPDATE, DELETE operations)
* **Conflict Resolution**: Automated handling of data conflicts with configurable business rules and manual override capabilities

**System Integration**

* **Monitoring & Alerting**: Comprehensive monitoring dashboard with proactive alerting for sync failures, performance degradation, and data quality issues
* **Audit Trail**: Complete audit logging of all synchronization activities for compliance and troubleshooting

**Data Quality & Integrity**

* **Schema Mapping**: Accurate mapping between Report Hub and DCM metadata schemas with data transformation as needed
* **Data Consistency**: Referential integrity maintained across synchronized datasets
* **Error Handling**: Graceful handling of malformed data with detailed error reporting and recovery procedures
* **Data Lineage**: Clear traceability of data origin and transformation history

**Design Details**

**Architecture Overview**

Report Hub → Data Extraction Layer → Transformation Engine → DCM Integration Layer → DCM System

↓ ↓ ↓ ↓

Change Detection → Queue Management → Data Validation → Conflict Resolution

**Data Flow Design**

**Initial Load Process**

1. **Pre-migration Validation**: Verify system connectivity and permissions
2. **Data Extraction**: Full extraction of report metadata from Report Hub
3. **Data Transformation**: Apply schema mapping and data enrichment
4. **Batch Loading**: Load data to DCM in optimized batches
5. **Validation & Reconciliation**: Verify data integrity and completeness
6. **Cutover**: Switch to incremental sync mode

**Incremental Sync Process**

1. **Change Detection**: Identify modified records in Report Hub
2. **Change Capture**: Extract changed records with metadata
3. **Transformation**: Apply necessary data transformations
4. **Conflict Resolution**: Handle any data conflicts
5. **DCM Update**: Apply changes to DCM system
6. **Confirmation**: Verify successful application of changes

**Error Handling Strategy**

* **Retry Logic**: Exponential backoff with 3 retry attempts
* **Circuit Breaker**: Prevent cascade failures during system outages
* **Error Classification**: Categorize errors as transient vs. permanent
* **Notification**: Real-time alerts for critical errors

**EPIC 2: Multi-Tenant Data Catalog Metadata Architecture**

**Epic Name**

**Multi-Tenant Data Segregation for Data Catalog Metadata System**

**Description**

Transform the existing Data Catalog Metadata (DCM) system from a single-tenant architecture to a secure multi-tenant platform that provides complete data segregation between different departments. This restructuring ensures that each department's metadata remains completely isolated, preventing any cross-contamination of data through search results, API responses, UI displays, or any other system touchpoints. The solution will maintain strict security boundaries while preserving system performance and enabling department-specific customization capabilities.

**Epic-Level Acceptance Criteria**

**Data Segregation Requirements**

* **AC-1**: Each department's metadata must be completely isolated from all other departments at the storage layer
* **AC-2**: No search query from any department can return results containing metadata from other departments
* **AC-3**: All API endpoints must enforce department-level access controls and return only authorized department data
* **AC-4**: UI components must display only metadata belonging to the authenticated user's department
* **AC-5**: Cross-department data leakage must be impossible through any system interface (UI, API, batch processes, reports)

**Authentication & Authorization**

* **AC-6**: System must integrate with Bank’s GES policy
* **AC-7**: User access must be restricted to their assigned department(s) only
* **AC-8**: Administrative functions must be scoped to department level with super-admin capabilities for system-wide management
* **AC-9**: All access attempts and data operations must be logged with department context

**Performance & Scalability**

* **AC-10**: Search performance must not degrade compared to current single-tenant system
* **AC-11**: API response times must remain within existing SLA thresholds (<2 seconds for standard queries)
* **AC-12**: Department onboarding must be achievable without system downtime

**Data Migration & Continuity**

* **AC-14**: All existing metadata must be successfully migrated to appropriate department tenants without data loss
* **AC-15**: Migration must maintain all existing relationships and metadata integrity
* **AC-16**: Zero-downtime migration strategy must be implemented for production deployment
* **AC-17**: Rollback capability must be available throughout migration process

**Integration & Compatibility**

* **AC-18**: All existing DCM integrations must continue to function with department-scoped access
* **AC-19**: API contracts must remain backward compatible where possible
* **AC-20**: Third-party tool integrations must be updated to support multi-tenant architecture

**Epic 3 : Tech Debt :   
Already provided you details**

**EPIC: DCM Job Execution Monitoring Dashboard**

**Description**

Develop a comprehensive Streamlit-based monitoring dashboard that provides real-time visibility into DCM (Data Change Management) job executions, tracking job status and quantifying data operations (records added, updated, removed) across configurable time periods. The dashboard will serve as a centralized monitoring solution for data operations teams to track DCM job performance, identify trends, and troubleshoot issues efficiently.

**Business Value**

* **Operational Visibility**: Centralized view of all DCM job executions and their impact on data
* **Proactive Monitoring**: Early identification of job failures or performance degradation
* **Data Governance**: Track data changes and maintain audit trails for compliance
* **Efficiency Gains**: Reduce time spent investigating job status across multiple systems
* **Historical Analysis**: Enable trend analysis and capacity planning through historical data

**Acceptance Criteria (EPIC Level)**

**Core Functionality**

1. **Dashboard displays job execution status for the last 15 days by default**
   * Shows all DCM jobs with status indicators (Success, Failed, Running, Pending)
   * Displays job duration and completion timestamps
   * Provides summary statistics (total jobs, success rate, average duration)
2. **Data change metrics are accurately tracked and displayed**
   * Shows records added, updated, and removed per job execution
   * Provides aggregated totals across selected time period
   * Displays data change trends over time
3. **Interactive date range selection is fully functional**
   * Users can select custom start and end dates
   * Date picker supports past dates with reasonable historical limits
   * Dashboard refreshes automatically when date range changes

**User Experience**

1. **Dashboard is intuitive and responsive**
   * Loads within 5 seconds for default 15-day view
   * Clear navigation and filtering options
2. **Data visualization is meaningful and actionable**
   * Charts and graphs effectively communicate job status trends
   * Color-coded status indicators are accessible and intuitive
   * Export functionality available for reports