### ****Client Requirements with Informatica ETL Tool****

#### **1. Background and Objectives**

* **Migration Goal**: Transition from Wipro CRM/DMS to Zoho CRM/DMS, including historical and master data migration.
* **Timeline**: Phased go-live starting July 2025, with historical data migrated and tested prior.
* **Informatica Fit**: Informatica PowerCenter or Informatica Cloud Data Integration can extract data from legacy systems (e.g., IBM DB2, SQL Server, Oracle), transform it into Zoho-compatible formats, and load it efficiently, supporting FML’s phased approach.

#### **2. Scope of Work with Informatica**

##### **a) Pre-Migration**

* **i) Data Scoping**:
  + **Informatica Solution**: Use Informatica Data Integration to define data objects (e.g., 177 master data tables, 899 transaction data tables) and set up workflows for bulk and incremental migrations. The tool’s metadata management helps establish a cut-off date and plan phased loads.
  + **Execution**: Create reusable mappings to handle the 1,257 tables and monitor progress via Informatica’s dashboards.
* **ii) Data Profiling**:
  + **Informatica Solution**: Leverage Informatica Data Quality (IDQ) to define data quality rules, assess consistency/completeness, and identify gaps/anomalies (e.g., duplicates, irrelevant data).
  + **Output**: Generate profiling reports for FML review, highlighting inconsistencies across 1 TB of data.
* **iii) Data Integrity Assessment**:
  + **Informatica Solution**: IDQ’s profiling and PowerCenter’s validation capabilities can flag critical issues (e.g., missing keys, orphaned records) pre-migration.
  + **Output**: Deliver data quality reports via Informatica’s reporting tools.

##### **b) Migration**

* **i) Data Cleaning**:
  + **Informatica Solution**: Use Informatica Data Quality to cleanse historical data (e.g., standardize formats, remove duplicates) on a non-production/staging server. IDQ’s cleansing rules ensure Zoho compatibility.
  + **Execution**: Process 650 GB (DMS), 300 GB (Warranty & Audit), and 2 GB (DAR) datasets.
* **ii) Data Mapping & Transformation**:
  + **Informatica Solution**: PowerCenter’s mapping designer creates source-to-target mappings (e.g., Wipro fields to Zoho templates). Transformation rules (e.g., format changes, aggregations) are built and validated with FML using Informatica’s transformation objects (e.g., Expression, Aggregator).
  + **Optimization**: Use PowerCenter’s session tuning to optimize queries and avoid downtime.
* **iii) Data Validation (Pre) & Load**:
  + **Informatica Solution**: Pre-load validation workflows in PowerCenter compare transformed data against FML requirements, allowing iterative adjustments.
* **iv) Data Loading**:
  + **Informatica Solution**: PowerCenter or Informatica Cloud generates CSV files in Zoho’s import utility format. Workflows can be scheduled for full (historical) and incremental (post-cut-off) loads.
  + **Collaboration**: Test loads with Zoho on a non-production server, then deploy to production.
* **v) Data Validation (Post)**:
  + **Informatica Solution**: Use Informatica’s data validation option (DVO) to compare migrated data against transformed datasets. Load test system performance with PowerCenter’s performance monitoring.
  + **Output**: Produce post-load validation reports.

##### **c) Automated Error Handling & Logging**

* **Informatica Solution**: PowerCenter’s error handling logs failed records in real-time (e.g., via log files or database tables). Workflows can include auto-retry logic using session retries or custom scripts.
* **Execution**: Monitor failures across the 1,257 tables and retry partial migrations as needed.

##### **d) Business Sign-Off & Documentation**

* **Informatica Solution**: Generate comprehensive documentation via PowerCenter’s metadata repository (e.g., mappings, workflows, transformation rules). Export reports for FML sign-off.
* **Output**: Deliver cleansed and transformed data documentation.

#### **3. Deployment Model**

* **Preference**: On-premise migration at FML.
* **Informatica Fit**: Deploy Informatica PowerCenter on FML’s on-premise infrastructure (servers, OS, and database connectors).
* **Infrastructure Requirements**:
  + **Servers**: Dedicated Windows/Linux servers with sufficient CPU (e.g., 16 cores), RAM (e.g., 64 GB), and storage (e.g., 2 TB for staging/transformations).
  + **Software**: Informatica PowerCenter (v10.x), Java, and database drivers (IBM DB2, SQL Server, Oracle).
  + **Connectivity**: ODBC/JDBC connectors for source systems.

#### **4. Key Deliverables with Informatica**

* **Comprehensive Deployment Plan**: Built in PowerCenter’s workflow manager.
* **Data Profiling and Cleansing Reports**: Generated via Informatica Data Quality.
* **Source-to-Target Mapping Document**: Exported from PowerCenter’s mapping designer.
* **Migration Pipelines and Transformation Rules**: Stored in PowerCenter workflows.
* **Pre-load and Post-load Validation Reports**: Produced via DVO and PowerCenter logs.

#### **5. Implementation Requirements**

* **RASCI Matrix**: Define roles (e.g., FML for validation, Zoho for loading, Vendor for ETL using Informatica).
* **Dependencies**: Informatica requires FML to provide source system access and Zoho to supply target templates.
* **Infrastructure & Access**:
  + **Vendor Location**: Informatica PowerCenter on vendor servers with VPN access to FML systems.
  + **FML Location**: Install PowerCenter on FML servers with admin access for vendor resources.

#### **6. IT / Technical Requirements**

##### **a) Technology Stack**

* **Tool**: Informatica PowerCenter (on-premise ETL) and Informatica Data Quality (profiling/cleansing).
* **Optional**: Informatica Cloud for hybrid scenarios if needed later.

##### **b) Source & Target System Compatibility**

* **Sources**: Fully compatible with IBM DB2 (DMS), SQL Server, and Oracle (Warranty & Audit) via native connectors.
* **Target**: Exports to CSV for Zoho’s import utility.

##### **c) Data Volume Estimation**

* **1 TB Handling**: PowerCenter’s partitioning and parallel processing manage 1,257 tables efficiently.
* **Execution**: Bulk load for 10 years of data, followed by incremental loads.

##### **d) Backup & Disaster Recovery**

* **Informatica Solution**: Pre-migration backups via source system exports. PowerCenter workflows include rollback scripts (e.g., truncate/reload).
* **RTO/RPO**: Configurable in workflows (e.g., RTO: 4 hours, RPO: 15 minutes for staging).

##### **e) Data Anonymization & Masking**

* **Informatica Solution**: Use Informatica Data Masking (add-on) for AES-256 encryption, tokenization, or masking of sensitive data (e.g., customer PII).
* **Compliance**: Supports DPDP, IT Act 2000, and ISO 27001 with key integrity maintained.

##### **f) Data Security During Transfer**

* **Informatica Solution**: TLS 1.2+ encryption for data movement, checksum validation via PowerCenter, and RBAC via Informatica’s user management.

#### **7. Commercial Requirements**

* **Services**: Migration design, ETL development, testing, and support.
* **Product**: Informatica PowerCenter licensing (on-premise).
* **Costs**:
  + **One-time Migration**: License + service fees for 1 TB bulk load.
  + **Delta Migration (5 Phases)**: Incremental service fees per phase.
* **Licensing Model**:
  + **On-Prem**: Perpetual license with annual maintenance; FML provides servers.

#### **8. Vendor Responsibilities with Informatica**

* Assign Informatica-certified project managers and developers.
* Deliver a migration strategy using PowerCenter workflows and risk mitigation via error handling.
* Ensure compliance with FML security policies via Informatica’s security features.
* Provide post-migration support (e.g., 3-6 months).

**Detailed workflow** for migrating Force Motors Limited’s (FML) Wipro-based CRM/DMS to Zoho CRM/DMS using **Informatica PowerCenter** as the ETL tool

### ****Detailed Informatica PowerCenter Workflow for FML Migration****

#### **1. Pre-Migration Phase**

This phase prepares the data and infrastructure for migration, ensuring quality and compatibility with Zoho CRM/DMS.

##### **Step 1: Data Scoping**

* **Objective**: Define the scope of data objects and migration strategy.
* **Tasks**:
  1. **Identify Data Objects**: Map out 1,257 tables (177 master data, 899 transaction data) from source systems (IBM DB2, SQL Server, Oracle).
  2. **Set Cut-Off Date**: Collaborate with FML to agree on a cut-off date (e.g., June 30, 2025) for historical data migration.
  3. **Plan Migration Phases**: Define bulk migration (historical data) and incremental loads (post-cut-off) aligned with Zoho’s July 2025 go-live.
  4. **Create Repository**: Set up Informatica PowerCenter Repository to store metadata for all tables.
* **Informatica Tools**: Repository Manager, Workflow Manager.
* **Deliverable**: Migration scope document with table list and phased plan.

##### **Step 2: Data Profiling**

* **Objective**: Assess and document data quality.
* **Tasks**:
  1. **Connect to Sources**: Configure source connections (ODBC/JDBC) to IBM DB2 (650 GB), SQL Server/Oracle (300 GB), and DAR (2 GB).
  2. **Profile Data**: Use Informatica Data Quality (IDQ) to analyze 1 TB of data for quality, consistency, and completeness.
     + Check for duplicates, null values, and inconsistencies (e.g., mismatched formats in customer records).
     + Identify anomalies (e.g., orphaned transaction records).
  3. **Define Rules**: Set data quality rules (e.g., valid phone number formats, date consistency) with FML input.
  4. **Generate Reports**: Export profiling results (e.g., duplicate count, missing data percentage).
* **Informatica Tools**: IDQ, PowerCenter Designer (source definitions).
* **Deliverable**: Data profiling report highlighting gaps and anomalies.

##### **Step 3: Data Integrity Assessment**

* **Objective**: Identify and document critical issues.
* **Tasks**:
  1. **Run Integrity Checks**: Use IDQ to validate referential integrity (e.g., primary/foreign key relationships).
  2. **Highlight Issues**: Flag critical problems (e.g., missing master data entries linked to transactions).
  3. **Report**: Compile findings into a data quality report for FML review.
* **Informatica Tools**: IDQ, PowerCenter Workflow Monitor.
* **Deliverable**: Data integrity assessment report.

##### **Step 4: Staging Environment Setup**

* **Objective**: Prepare a non-production environment for cleaning and transformation.
* **Tasks**:
  1. **Provision Servers**: Set up on-premise staging servers at FML (e.g., 2 TB storage, 64 GB RAM).
  2. **Install Informatica**: Deploy PowerCenter and IDQ on staging infrastructure.
  3. **Extract Data**: Use PowerCenter to extract 1 TB of historical data to staging tables.
* **Informatica Tools**: PowerCenter Integration Service.
* **Deliverable**: Configured staging environment.

#### **2. Migration Phase**

This phase involves cleaning, transforming, and loading the data into Zoho CRM/DMS.

##### **Step 5: Data Cleaning**

* **Objective**: Cleanse historical data to meet Zoho requirements.
* **Tasks**:
  1. **Develop Cleansing Rules**: Use IDQ to create rules (e.g., remove duplicates, standardize addresses, fix date formats).
  2. **Execute Cleaning**: Run PowerCenter workflows to process 1,257 tables on the staging server.
     + Example: Clean 169 master data tables (<1M records) and 32 transaction tables (>10M records).
  3. **Verify Output**: Sample cleansed data and validate with FML (e.g., spot-check customer records).
* **Informatica Tools**: IDQ (rule definition), PowerCenter (workflows).
* **Deliverable**: Cleansed staging tables.

##### **Step 6: Data Mapping & Transformation**

* **Objective**: Transform data into Zoho-compatible formats.
* **Tasks**:
  1. **Obtain Mapping Document**: Use FML-provided source-to-target field mappings (e.g., Wipro “Cust\_ID” to Zoho “CustomerID”).
  2. **Design Mappings**: In PowerCenter Designer, create mappings for all 1,257 tables.
     + **Transformations**: Use Expression (format changes), Lookup (reference data), Aggregator (summaries).
     + Example: Transform IBM DB2 date field “YYYYMMDD” to Zoho’s “MM/DD/YYYY”.
  3. **Validate Rules**: Test mappings with FML (e.g., run small datasets and review).
  4. **Optimize Queries**: Tune SQL overrides in PowerCenter to minimize performance impact (e.g., parallel processing for >10M record tables).
* **Informatica Tools**: PowerCenter Designer, Mapping Architect.
* **Deliverable**: Source-to-target mapping document and transformation workflows.

##### **Step 7: Data Validation (Pre-Load)**

* **Objective**: Ensure transformed data meets requirements.
* **Tasks**:
  1. **Run Pre-Validation**: Use PowerCenter to compare transformed data against FML criteria (e.g., record counts, key fields).
  2. **Iterate**: Adjust transformations based on FML feedback (e.g., add missing fields).
* **Informatica Tools**: PowerCenter Workflow Monitor.
* **Deliverable**: Pre-load validation report.

##### **Step 8: Data Loading**

* **Objective**: Load data into Zoho CRM/DMS.
* **Tasks**:
  1. **Generate CSVs**: Use PowerCenter to export transformed data into Zoho-compatible CSV files.
     + Bulk Load: Export 1 TB historical data up to cut-off date.
     + Incremental Load: Set up workflows for post-cut-off deltas (e.g., daily updates).
  2. **Test Load**: Work with Zoho to import CSVs into a non-production environment (e.g., sandbox).
  3. **Production Load**: Schedule and execute full load into Zoho production database (July 2025).
* **Informatica Tools**: PowerCenter Workflow Manager, File Writer targets.
* **Deliverable**: CSV files and loaded non-production data.

##### **Step 9: Automated Error Handling & Logging**

* **Objective**: Manage migration errors efficiently.
* **Tasks**:
  1. **Configure Logging**: Set up PowerCenter session logs to capture failed records (e.g., invalid formats, missing keys).
  2. **Implement Retry**: Build workflows with retry logic (e.g., reprocess failed batches up to 3 times).
  3. **Monitor**: Use Workflow Monitor to track errors in real-time.
* **Informatica Tools**: PowerCenter Workflow Monitor, Session Properties.
* **Deliverable**: Error logs and retry reports.

#### **3. Post-Migration Phase**

This phase ensures data accuracy and system readiness.

##### **Step 10: Data Validation (Post-Load)**

* **Objective**: Verify migrated data in Zoho.
* **Tasks**:
  1. **Compare Data**: Use Informatica Data Validation Option (DVO) to compare Zoho data against transformed staging data (e.g., record counts, sums).
  2. **Performance Test**: Simulate peak loads (e.g., 100 concurrent users) and measure response time.
  3. **Generate Reports**: Export validation results for FML review.
* **Informatica Tools**: DVO, PowerCenter Workflow Monitor.
* **Deliverable**: Post-load validation report.

##### **Step 11: Business Sign-Off & Documentation**

* **Objective**: Finalize migration and hand over to FML.
* **Tasks**:
  1. **Compile Documentation**: Export mappings, workflows, and logs from PowerCenter Repository.
  2. **Review with FML**: Present validation reports and cleansed data for sign-off.
  3. **Handover**: Provide FML with documentation and access to workflows for future use.
* **Informatica Tools**: Repository Manager, Workflow Manager.
* **Deliverable**: Comprehensive migration documentation package.

### ****Workflow Execution Flow****

1. **Kickoff**: Repository setup and source connections.
2. **Pre-Migration**: Scoping → Profiling → Integrity Assessment → Staging.
3. **Migration**: Cleaning → Mapping/Transformation → Pre-Validation → Loading → Error Handling.
4. **Post-Migration**: Post-Validation → Sign-Off → Documentation.
5. **Incremental Loads**: Repeat Steps 6-10 for delta migrations (up to 5 phases).

### ****Key Informatica Features Utilized****

* **PowerCenter Designer**: Mappings and transformations.
* **Workflow Manager**: Scheduling and executing bulk/incremental loads.
* **IDQ**: Profiling, cleansing, and quality rules.
* **DVO**: Pre- and post-load validation.
* **Repository Manager**: Metadata storage and documentation.
* **Performance Tuning**: Partitioning for >10M record tables, optimized queries.

### ****Timeline Alignment****

* **March 2025**: Scoping, profiling, and staging setup.
* **April-May 2025**: Cleaning, transformation, and pre-validation.
* **June 2025**: Test loads and error handling refinement.
* **July 2025**: Bulk load into Zoho production, followed by incremental loads.

### ****Sample Workflow (Simplified Example)****

* **Source**: IBM DB2 table “Customer” (1M records).
* **Tasks**:
  1. Extract to staging (PowerCenter Source Qualifier).
  2. Cleanse: Remove duplicates, standardize “Phone” field (IDQ).
  3. Transform: Map “Cust\_Name” to “CustomerName” (Expression Transformation).
  4. Export: Write to “Customer.csv” (File Writer).
  5. Load: Zoho imports CSV.
  6. Validate: Compare record count (DVO).
* **Error Handling**: Log failed records (e.g., invalid “Phone”) and retry.