**MongoDB Financials Use-Case – Consultant Status Summary**

**(Apr 15 – Jun 19 2025)**

### **1. Scope of Work**

Design and implement a MongoDB schema and event-driven processing pipeline mirroring legacy Oracle financial tables while introducing scenario isolation ( *outlook, budget, live* ) and downstream aggregation levels (L4 ➔ L3, L1 ➔ L2).

### **2. Major Contributions & Deliverables**

| **Area** | **Outcome** | **Notes** |
| --- | --- | --- |
| **Schema design** | • Mapped Oracle tables to MongoDB collections: lvl1FinancialsSummary, lvl4CostDetails$<scenario>, lvl3CostDetails$<scenario>, lvl2FinancialsSummary.• Standardised key fields (ipLongId, planId, scenario tags) across levels for join-free look-ups. | Green-labelled entities in the architecture diagram are finalised. |
| **Lookup consolidation** | Combined multiple Oracle reference tables (BU, billing keys, supplier, title) into a single lookup collection, leveraging MongoDB’s flexible schema to reduce cross-collection joins during enrichment. |  |
| **Event ingestion & persistence** | Built ECS Java services that ingest SNS events (IP\_COST\_DETAILS, IP\_COST\_SUMMARY) from **GOS DB** and upsert into level-respective collections.• *L4 ➔ L3 flow*: enrich with lookup data, write to lvl3CostDetails$<scenario>.• *L1 ➔ L2 flow*: enrich with allocations, taxonomy, app mappings, write to lvl2FinancialsSummary. |  |
| **Aggregation / enrichment queries** | Authored and optimised MongoDB aggregation pipelines that:• reduce verbose L4 cost arrays (up to 8 k rows) to BU-level L3 docs.• roll fiscal-year figures and enrichment objects into compact L2 documents. |  |
| **Transactional orchestration** | Wrapped multi-collection updates in withTransaction blocks to guarantee atomicity (upsert + audit) and added timestamped stage metrics for latency tracking. |  |
| **Audit framework** | Designed auditCollection, implemented create/update hooks, and extended schema to log duration, failure reason, and retry metadata. |  |
| **Performance & indexing** | • Replaced unsupported wildcard index outlook.$\*\*.fyCost with outlook.$\*\* and proposed flattening outlook into [ {year, fyCost} ] structure.• Guided team on compound indexes for ipLongId + planId + scenario. |  |
| **Team enablement** | Fielded ad-hoc query-tuning questions, reviewed PRs, and documented best practices for India dev team’s upcoming training sessions. |  |
| **Pull requests** | Raised and merged PRs covering:1. L4 ➔ L3 event processor and aggregation pipelines.2. L1 ➔ L2 enrichment logic and transaction wrapper.3. Audit collection schema and helper utilities. |  |

### **3. Delivered Effort by Sprint**

| **Sprint (Jira 251)** | **Ticket** | **Points** | **Status** |
| --- | --- | --- | --- |
| Apr 15 – May 15 (Design) | **Mongo Queries** | 8 | Complete |
| May 15 – Jun 1 (Execution – Txn flows) | **Mongo Queries** | 8 | Complete |
| Jun 1 – Jun 19 (Execution – Aggregations) | **Mongo Queries** | 8 | Complete |

### **4. Dependencies / Items Outside My Control**

| **Collection (red in diagram)** | **Owned by** | **Impact** |
| --- | --- | --- |
| billingKeyRefData, sdmBuHierarchy, expenseHierarchy, supplier | Source-system teams | Required for full enrichment; awaiting finalized feeds & schemas. |
| ipClosedMonthCostAllocations | Finance Ops | Needed to close the loop on live-scenario reconciliations. |

### **5. Outstanding Work (black labels)**

| **Task** | **Status** | **Next Action** |
| --- | --- | --- |
| Finalise ipAllocatedCost roll-ups into L2 | Design drafted | Align on allocation logic once allocation engine stabilises. |
| Produce downstream dashboards (myInsights) | Schema ready | Coordinate with BI team to wire queries and validate aggregates. |

### **6. Next Steps**

1. **Support dependency teams** to onboard their reference data into MongoDB (schema validation, index suggestions).
2. **Benchmark** end-to-end L4 ➔ L3 ➔ L2 latency and tweak pipelines or indexes as needed.
3. **Cluster sizing** review with MongoDB SA once data-volume projections are final.
4. **Deliver training** session for India dev team on scenario-aware querying patterns and aggregation-framework tips.

**Nine-Sprint Delivery Summary – MongoDB Financials Use-Case**

**( Sprint 244 → Sprint 252 : Feb 03 – Jun 28 2025 )**

| **#** | **Sprint Window (2 wks)** | **Focus & Outcomes** | **Key Artefacts / PRs** |
| --- | --- | --- | --- |
| **244** | 03 Feb – 16 Feb | **Discovery & scope framing** – mapped legacy Oracle financial tables; captured scenario requirements with Finance & PO; drafted first cut of four-level (L1-L4) document shapes. | Design deck v0.1 |
| **245** | 17 Feb – 02 Mar | **Data-set stimulation** – pulled CSV extracts from Oracle DBA (IP\_COST\_DETAILS, IP\_COST\_SUMMARY, BU, Supplier). Loaded into local MongoDB to validate BSON sizes, array depths, index feasibility. | Mock-data loader scripts |
| **246** | 03 Mar – 16 Mar | **Schema prototyping** – tried three variants (embedded, referenced, hybrid). Benchmarked agg/lookup latency; decided to split **L3 & L4 by scenario** (*outlook, budget, live*) for query locality. | Benchmark report; schema v0.3 |
| **247** | 17 Mar – 30 Mar | **Canonical schema lock-in** – finalised field names (ipLongId, planId, scenario, fiscal-year objects). Designed unified lookup collection combining BU, billingKey, supplier, title. | Schema ERD v1.0; JSON Schema validators |
| **248** | 31 Mar – 13 Apr | **Event-ingest scaffolding** – created ECS Java skeleton; mocked SNS event listener; built idempotent upsert helpers; outlined audit model. | Repo init PR #12 |
| **249** | 15 Apr – 28 Apr | **L4 ➔ L3 pipeline (design sprint)** – wrote aggregation pipeline to condense up-to-8 k cost records, enrich via lookup, and persist to lvl3CostDetails$<scenario>. Added performance counters. | PR #23 – L4toL3 service |
| **250** | 29 Apr – 12 May | **Transactions & audit hardening** – wrapped L4-to-L3 flow in withTransaction; built auditCollection schema, latency stamps, failure logging; load-tested with 1 M docs. | PR #31 – Txn wrapper & audit util |
| **251** | 13 May – 26 May | **L1 ➔ L2 enrichment** – joined allocations, taxonomy, app mappings; produced FY roll-ups; ensured atomic L1-enrich-L2-audit cycle. Provided query-tuning support to dev squad. | PR #38 – L1toL2 processor |
| **252** | 27 May – 09 Jun | **Scenario-aware live handling & clean-up** – distinct path for live events; purge pending approvals; index review (outlook.$\*\* vs flat array {year, fyCost}); zero spill-over declared. | PR #46 – Live scenario handler |

### **Cumulative Impact**

* **Schema Evolution** – from 4 generic collections to 14 scenario-aware, performance-tuned collections (green boxes in diagram).
* **Lookup Consolidation** – collapsed four Oracle reference tables into 1 Mongo collection → removed cross-DB joins during enrichment.
* **Event Pipeline** – end-to-end ingestion, enrichment, aggregation and auditing implemented for both cost-detail (L4→L3) and summary (L1→L2) paths.
* **Quality Gates** – JSON Schema validation, unit + load tests, and transaction-level auditing in place.
* **Team Enablement** – continuous query-design support, PR reviews, and knowledge-transfer sessions scheduled for India dev team.

### **Remaining / External Dependencies**

* Red-label collections (billingKeyRefData, sdmBuHierarchy, expenseHierarchy, supplier, ipClosedMonthCostAllocations) – awaiting upstream teams for feed finalisation.
* Integration of ipAllocatedCost roll-ups into L2 once allocation logic is fixed.
* BI dashboards (myInsights) to be wired after reference data arrives.

All nine two-week sprints have been executed with planned deliverables met and no rollover.