Data Modeling Techniques

# Introduction

The data warehouse was designed to support both analytics and operational use cases. By creating a well-designed model that supports both the dimensional tables and transactional tables. For a time tracking and resource allocation system, the data model has been developed to accommodate both analytical (OLAP) and operational (OLTP) use cases.

# Dimensional Model (Analytical Purposes)

## Design Choices

### Star Schema Implementation

* Two fact tables: ***fact\_time\_tracking*** and ***fact\_allocation***
* Five-dimension tables: ***dim\_client, dim\_project, dim\_employee, dim\_role, and dim\_date***
* Optimized for complex analytical queries and reporting

**Fact Tables**

This contains the incremental tables that will continue to grow on a steady basis. Optimized for analytical queries.

* **fact\_time\_tracking:** Stores actual time entries from ClickUp
* **fact\_allocation:** Stores resource allocation/planning data from Float

**Dimension Tables**

This is information about the fact table containing necessary information used in the modeling process.

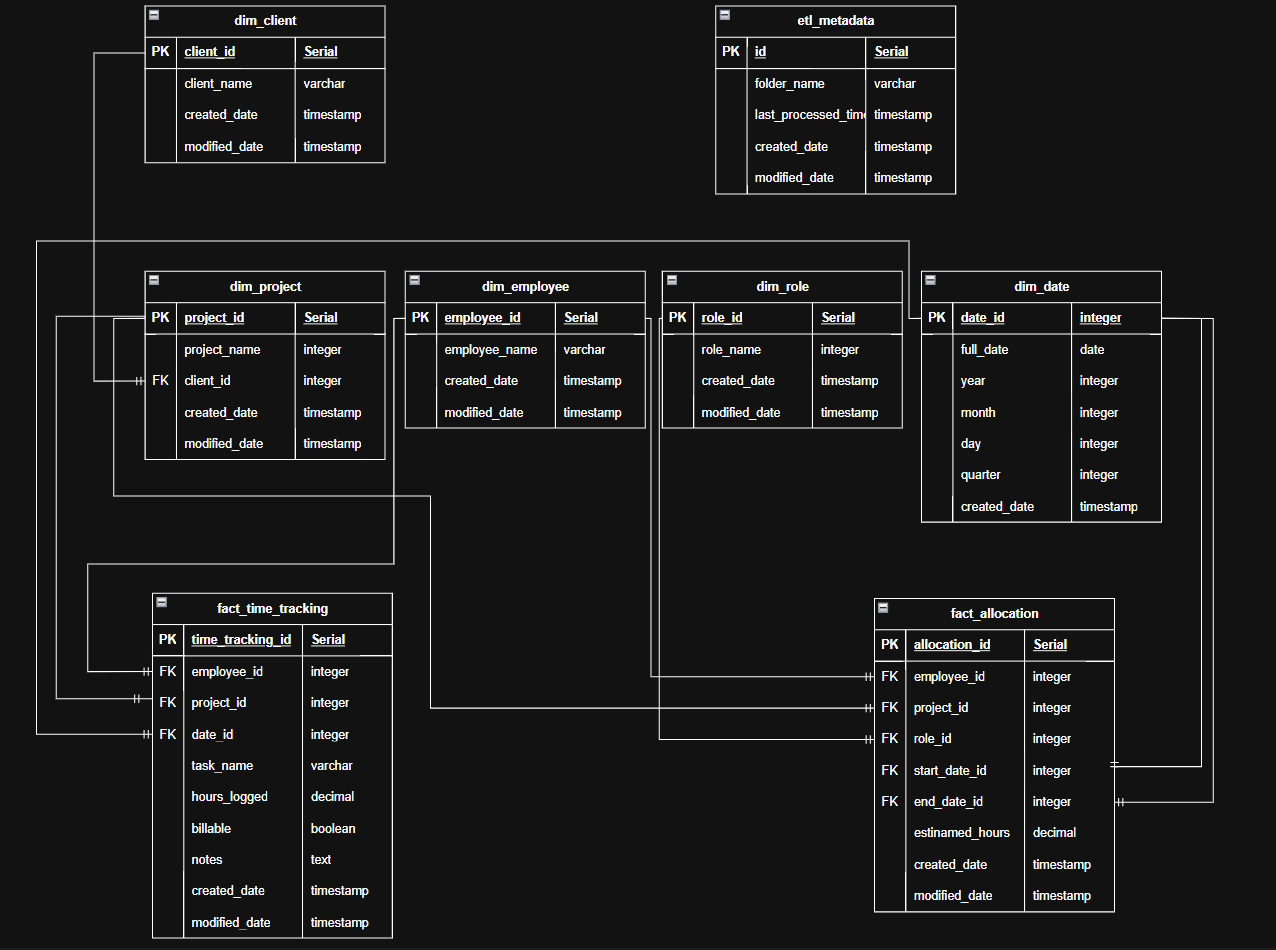
* **dim\_client:** Client information
* **dim\_project:** Project details with client relationships
* **dim\_employee:** Employee information
* **dim\_role:** Role definitions
* **dim\_date:** Date dimension for time-based analysis

**ETL Metadata**

These are standard data used in keeping records and changes.

## Entity Relationship Diagram

Entity-relationship (ER) modeling is a visual approach to data modeling used to represent the structure of a database. It is used to identify the "things" (entities) in a system and how they relate to each other.



*Full ER Diagram & Modeling(Analytics & Operational Purpose)*

### One-to-Many Relationships:

* A client can have multiple projects (1:M)
* A project can have multiple time tracking entries (1:M)
* A project can have multiple allocations (1:M)
* An employee can have multiple time tracking entries (1:M)
* An employee can have multiple allocations (1:M)
* A role can be used in multiple allocations (1:M)
* A date can be referenced by multiple time tracking entries (1:M)
* A date can be the start or end date for multiple allocations (1:M)

### Fact Table Relationships:

**fact\_time\_tracking connects to:**

* dim\_employee
* dim\_project
* dim\_date

**fact\_allocation connects to:**

* dim\_employee
* dim\_project
* dim\_role
* dim\_date (twice, for start and end dates)

**Independent Tables:**

* etl\_metadata is independent and used for ETL process tracking

## Analytical Benefit

* **Performance Optimization:** Denormalized dimension tables reduce JOIN complexity.
* **Business Intelligence Capabilities:** Easy aggregation of hours by various dimensions.
* **Reporting Flexibility:** Multiple grain levels (daily, monthly, quarterly)

# Entity-Relationship Model (Operational Purposes)

## Design Choices

### Normalized Structure

* Clear entity relationships with referential integrity
* Primary and foreign key constraints
* Audit fields (***created\_date, modified\_date***)

### Operational Benefits

**Data Integrity**

* Foreign key constraints ensure referential integrity
* SERIAL primary keys prevent duplicate entries
* NOT NULL constraints where appropriate
* UNIQUE constraints on critical fields

**Transaction Processing**

* Efficient CRUD operations
* Minimal data redundancy

**Operational Features**

* Real-time time tracking entry
* Resource allocation management

# Conclusion

This hybrid design effectively balances analytical and operational needs by ensuring data integrity, enabling efficient reporting, supporting real-time and historical analysis, and allowing for future scalability and modifications.